This thesis has been submitted in fulfilment of the requirements for a postgraduate degree (e.g. PhD, MPhil, DClinPsychol) at the University of Edinburgh. Please note the following terms and conditions of use:

- This work is protected by copyright and other intellectual property rights, which are retained by the thesis author, unless otherwise stated.
- A copy can be downloaded for personal non-commercial research or study, without prior permission or charge.
- This thesis cannot be reproduced or quoted extensively from without first obtaining permission in writing from the author.
- The content must not be changed in any way or sold commercially in any format or medium without the formal permission of the author.
- When referring to this work, full bibliographic details including the author, title, awarding institution and date of the thesis must be given.
Study of the N-terminal domains of MDM2 and MDM4, and their potential for targeting by small-molecule drugs

Maria Concepcion Sanchez Perez

University of Edinburgh
Thesis submitted for the degree of Doctor of Philosophy
June 2011
Erase una vez…
Abstract

The MDM2 and MDM4 oncoproteins are both involved in regulating the tumour suppressor, p53. While the MDM2–p53 interface is structurally and biophysically well characterised, the MDM4-p53 interaction has only recently attracted researchers’ attentions. The goal of this project was to establish structural and chemical ground rules for the disruption of the interactions between the N-terminal domains of MDM2/4 and p53, which is an attractive anticancer strategy.

In the current work, successful recombinant production and purification protocols for both the N-terminal domains of MDM2 (i.e. MDM2-N, residues 11-118) and MDM4 (MDM4-N, residues 14-111) have been established, yielding protein in sufficient quantity and quality for analysis using nuclear magnetic resonance spectroscopy (NMR).

Two screening strategies were employed to identify small-molecule antagonists of the MDM2-N:p53 interaction. First, a virtual screening exercise identified several compounds that were shown (by NMR) to bind to MDM2-N with μM $K_{D}$s. Docking studies supported by NMR chemical shift perturbation analysis suggested proposals for binding modes. The results are discussed in relation to the previously reported binding to MDM2-N of well-characterised inhibitors of the MDM2:p53 interaction such as Nutlin-3. Second, a fragment-based library was screened against MDM2-N using TROSY-type NMR spectra to monitor binding. Several hits were identified and the results are discussed with regard to the “druggability” of the MDM2-N p53 interaction.

To better understand the p53-binding groove of MDM4-N, multidimensional NMR was used to investigate the structure and backbone dynamics of double-isotopically labelled samples of MDM4-N, both free (i.e. apo-MDM4-N) and in complexes with a p53-derived peptide or Nutlin-3. The apo-MDM4-N is more conformationally dynamic than MDM2, since it contains unstructured regions. These regions appear to become structured upon binding of a ligand. MDM4 appears to bind its ligand through conformational selection and/or an induced fit mechanism involving reorganization of key sub-sites within the binding groove. This study highlighted
differences between Nutlin-3 and peptide binding that suggest the rational design of specific inhibitors of the MDM4:p53 interaction.
Acknowledgments

First I would like to thank my supervisor Prof. Paul Barlow for the opportunity he gave me some four years ago, and for his help, support and encouragement. I would also like to acknowledge my industrial supervisor Dr Martin Vogtherr for his patience and dedication, and for making my industrial placement in AstraZeneca the highlight of my PhD. And, thanks to Prof. Malcolm Walkinshaw for trusting me with some 33 compounds.

Many people have helped me over my PhD in different ways. People from bioNMR group (University of Edinburgh) had been supportive and patient with me, and also regulars and staff at the PPF and BCF (University of Edinburgh). In particular I would like to thank Dr Schmidt for always having a smile, and Dr Phelan for her encouragement and good energy. And thanks to Juraj for never giving up on me.

During my placement in AstraZeneca, I had a great time. The NMR group and the colleagues at 50F60 made me feel at home, as well as indulging me in any request a PhD student could have. Especially Mr Renshaw, who taught me loads and more, and again to Dr Vogtherr, for many things, but for introducing me to pragmatism. I have also to mention Kevin Embrey, just because.

Last but not least I would like to thank many times my family and friends, for not giving up in my lack of contact or social interest. To my mother; thank you for the regular phone calls, and to my sister for the bundle of pictures sent during this period to keep me up to date on las fantasicas aventuras de Julia y Alvaro. To Zuzana, who is always there. Finally, many thanks to Barry simply because, this PhD could not have happened without you - …and the rest is noise.

And to you for reading. I hope you enjoy.
Declaration

I hereby declare that all the work described in this thesis is my own work unless declared otherwise and that it has not been submitted in whole or partially for obtaining a degree at this or any other University.

Maria Concepcion Sanchez Perez, Edinburgh July 2011
### Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ac3</td>
<td>1-amino-cyclo-propanecarboxylic acid</td>
</tr>
<tr>
<td>Aib</td>
<td>Aminoisobutyric acid</td>
</tr>
<tr>
<td>ARF</td>
<td>ADP-ribosylation factor</td>
</tr>
<tr>
<td>BSA</td>
<td>Bovine serum albumin</td>
</tr>
<tr>
<td>CIAP</td>
<td>Calf Intestinal Alkaline Phosphatase</td>
</tr>
<tr>
<td>CSP</td>
<td>Chemical shift perturbation</td>
</tr>
<tr>
<td>CV</td>
<td>Column volumes</td>
</tr>
<tr>
<td>DM</td>
<td>Doble minute</td>
</tr>
<tr>
<td>DM</td>
<td>Double minute</td>
</tr>
<tr>
<td>DMSO</td>
<td>Dimethyl sulfoxide</td>
</tr>
<tr>
<td>DNA</td>
<td>Deoxyribonucleic acid</td>
</tr>
<tr>
<td>DO</td>
<td>Dissolved oxygen</td>
</tr>
<tr>
<td>DTT</td>
<td>Dithiothreitol</td>
</tr>
<tr>
<td>EDTA</td>
<td>Ethylenediaminetetraacetic acid</td>
</tr>
<tr>
<td>FID</td>
<td>Free-induction decay</td>
</tr>
<tr>
<td>Fig.</td>
<td>Figure</td>
</tr>
<tr>
<td>FPLC</td>
<td>Fast protein liquid chromatography</td>
</tr>
<tr>
<td>G1</td>
<td>Gap1</td>
</tr>
<tr>
<td>G2</td>
<td>Gap2</td>
</tr>
<tr>
<td>HBA</td>
<td>Hydrogen Bond Acceptors</td>
</tr>
<tr>
<td>HBD</td>
<td>Hydrogen Bond Donors</td>
</tr>
<tr>
<td>HINT</td>
<td>Hydropathic interactions</td>
</tr>
<tr>
<td>HPLC</td>
<td>High performance liquid chromatography</td>
</tr>
<tr>
<td>HSQC</td>
<td>Heteronuclear single quantum correlation</td>
</tr>
<tr>
<td>HTS</td>
<td>High Throughput Screening</td>
</tr>
<tr>
<td>IMAC</td>
<td>Immobilized metal-affinity chromatography</td>
</tr>
<tr>
<td>IPTG</td>
<td>Isopropyl β-D-1-thiogalactopyranoside</td>
</tr>
<tr>
<td>ITC</td>
<td>Isothermal titration calorimetry</td>
</tr>
<tr>
<td>Abbreviation</td>
<td>Definition</td>
</tr>
<tr>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>Kan</td>
<td>kanamycin</td>
</tr>
<tr>
<td>KD</td>
<td>Dissociation constant</td>
</tr>
<tr>
<td>LB</td>
<td>Luria Bertrani</td>
</tr>
<tr>
<td>LE</td>
<td>Ligand Efficiency</td>
</tr>
<tr>
<td>LIDAEUS</td>
<td>Ligand Discovery At Edinburgh University</td>
</tr>
<tr>
<td>M</td>
<td>Mitosis</td>
</tr>
<tr>
<td>m/z</td>
<td>Mass/charge ratio</td>
</tr>
<tr>
<td>MDM2</td>
<td>Murine double minute 2</td>
</tr>
<tr>
<td>MLogP</td>
<td>Moriguchi octanol-water partition coefficient</td>
</tr>
<tr>
<td>MW</td>
<td>Molecular Weight</td>
</tr>
<tr>
<td>MWCO</td>
<td>Molecular-weight cutoff</td>
</tr>
<tr>
<td>NCI</td>
<td>National Cancer Institute</td>
</tr>
<tr>
<td>NES</td>
<td>Nuclear export signal</td>
</tr>
<tr>
<td>NLS</td>
<td>Nuclear localization signal</td>
</tr>
<tr>
<td>NMR</td>
<td>Nuclear Magnetic Resonance</td>
</tr>
<tr>
<td>OD</td>
<td>Optical density</td>
</tr>
<tr>
<td>PBS</td>
<td>Phosphate-buffered saline</td>
</tr>
<tr>
<td>PCR</td>
<td>Polymerase chain reaction</td>
</tr>
<tr>
<td>PCS</td>
<td>Poly cloning site</td>
</tr>
<tr>
<td>PDB</td>
<td>Protein data bank</td>
</tr>
<tr>
<td>PEG</td>
<td>Polyethylene glycol</td>
</tr>
<tr>
<td>PES</td>
<td>Polyethersulfone</td>
</tr>
<tr>
<td>PMI</td>
<td>p53-MDM2/MDM4 inhibitor</td>
</tr>
<tr>
<td>Pmp</td>
<td>Phosphonomethylphenylalanine</td>
</tr>
<tr>
<td>res.</td>
<td>Residues</td>
</tr>
<tr>
<td>rf</td>
<td>Radiofrequency</td>
</tr>
<tr>
<td>RING</td>
<td>Really Interesting New Gene 1</td>
</tr>
<tr>
<td>S</td>
<td>Synthesis</td>
</tr>
<tr>
<td>SAR</td>
<td>Structure-Activity Relationship</td>
</tr>
<tr>
<td>SDS-PAGE</td>
<td>Sodium dodecyl sulfate polyacrylamide gel electrophoresis</td>
</tr>
</tbody>
</table>
Abbreviations

SOC  super optimal catabolite suppressor
TAE  Tris-acetic-EDTA
TB   Terrific broth
TCEP tris(2-carboxyethyl)phosphine
TROSY Transverse relaxation optimized spectroscopy
UV/vis Ultraviolet/visible
v/v  Weight/volume ratio
VsL  Virtually screened Library
w/v  Weight/volume ratio
wt   Wild type
Zn   Zinc
Ala  A  Alanine
Arg  R  Arginine
Asn  N  Asparagine
Asp  D  Aspartate
Cys  C  Cysteine
Gln  Q  Glutamine
Glu  E  Glutamate
Gly  G  Glycine
His  H  Histidine
Ile  I  Isoleucine
Leu  L  Leucine
Lys  K  Lysine
Met  M  Methionine
Phe  F  Phenylalanine
Pro  P  Proline
Ser  S  Serine
Thr  T  Threonine
Trp  W  Tryptophane
Tyr  Y  Tyrosine
Val  V  Valine
## Table of Contents

Abstract ................................................................................................................................. i  
Acknowledgments .................................................................................................................. i  
Declaration ............................................................................................................................. ii  
Abbreviations ........................................................................................................................ iii  
Table of Contents ................................................................................................................... vi  
List of Figures and Tables ...................................................................................................... x  

1. Introduction ........................................................................................................................ 1  
1.1. P53 is a key player in cancer biology ........................................................................... 1  
1.2. MDM2-MDM4 coordinate regulation of p53 levels in the cell .................................... 3  
1.3. What is known about MDM2 and MDM4 ................................................................. 5  
  1.3.1. Discovery of MDM2 ............................................................................................. 5  
  1.3.2. Discovery of MDM4 ............................................................................................. 5  
  1.3.3. Domain composition of the MDMs and functional consequences ....................... 6  
1.4. Details of the binding sites within the MDMs ............................................................ 11  
  1.4.1. Details of the p53-binding groove of MDM2 ...................................................... 12  
  1.4.2. Insights into the p53-binding groove of MDM4 ................................................ 17  
1.5. Drug discovery and the modulation of MDM2/4-p53 binding .................................. 21  
  1.5.1. MDM2-N antagonists ......................................................................................... 22  
  1.5.2. Antagonists of MDM4-N ................................................................................... 32  
1.6. Aims of the thesis ......................................................................................................... 37  

2. Materials and Methods ..................................................................................................... 38  
2.1. DNA constructs .......................................................................................................... 38  
  2.1.1. The MDM2-N construct .................................................................................... 38  
  2.1.2. The MDM4 1-134 construct .............................................................................. 39  
  2.1.3. The MDM4-N construct .................................................................................... 39  
2.2. Cloning of MDM4 1-134 ........................................................................................... 39  
  2.2.1. MDM4 1-134 primers ....................................................................................... 39  
  2.2.2. Polymerase Chain Reaction (PCR) ................................................................. 40  
  2.2.3. Agarose gels ...................................................................................................... 41  
  2.2.4. DNA purification ............................................................................................... 41  
  2.2.5. Plasmid amplification and isolation .................................................................. 42  
  2.2.6. DNA quantification ......................................................................................... 42  
  2.2.7. Restriction-enzyme digestion ........................................................................... 43
2.2.8. Dephosphorylation of the open vector ........................................43
2.2.9. Ligation .....................................................................................43
2.2.10. Screening for recombinants and sequencing ................................44
2.2.11. Storage of cells and glycerol stocks ........................................44

2.3. Cell cultures ................................................................................45
2.3.1. Rich media preparation ...............................................................45
2.3.2. Minimal media preparation ......................................................45

2.4. Estimates of protein expression ....................................................46
2.4.1. SDS-PAGE ..............................................................................46
2.4.2. Western blotting ........................................................................47
2.4.3. Estimates of differential protein expression: the soluble and insoluble fractions 48
2.4.4. Estimation of protein concentration ...........................................48
2.4.5. Mass spectrometry ....................................................................49

2.5. Protein expression ........................................................................49
2.5.1. Shaking flasks ...........................................................................49
2.5.2. Fermentation (New Brunswick Bioflow2000) ...............................50

2.6. Protein purification ........................................................................50
2.6.1. Lysis of the cells .........................................................................50
2.6.2. Purification of proteins from soluble fractions ..............................51

2.7. Insoluble fractions: refolding protocol ............................................53

2.8. Buffer screening assays ..................................................................54
2.8.1. Thromofluor stability assay .........................................................54
2.8.2. Dynamic light scattering ..............................................................55

2.9. Isothermal titration calorimetry .......................................................55

2.10. Peptide synthesis ........................................................................55

2.11. Nuclear Magnetic Resonance (NMR) experiments .......................56
2.11.1. NMR toolkit .............................................................................56
2.11.2. Preparation of NMR samples ......................................................57
2.11.3. Screening by NMR: Technical concepts .....................................58
2.11.4. NMR experiments for screening ...............................................59
2.11.5. $K_D$ calculation by NMR ...........................................................60
2.11.6. NMR experiments for assignment .............................................62
2.11.7. Details of NMR experiments ......................................................69
2.11.8. NMR relaxation studies of MDM4 ...........................................69

2.12. Molecular visualization programs and other computational tools .......71
5.1. MDM2-N assignment ................................................................. 186
5.2. MDM4-N assignment ................................................................ 193
  5.2.1. MDMN-N optimization NMR conditions .............................. 193
  5.2.2. MDM4-N backbone assignment .......................................... 200
5.3. MDM4-N chemical shift analysis ................................................ 213
  5.3.1. Introduction ....................................................................... 213
  5.3.2. ¹³C and ¹⁵N shift perturbations upon ligation of MDM4-N ...... 215
5.4. Analysis of the chemical shift perturbations of MDM4-N .......... 222
5.5. MDM4-N relaxation ............................................................... 225
  5.5.1. Introduction ....................................................................... 225
  5.5.2. Datasets ........................................................................... 225
  5.5.3. apo MDM4-N relaxation rates ........................................... 225
  5.5.4. Relaxation rates for MDM4-N:Nutlin-3 complex .................. 228
  5.5.5. Relaxation rates MDM4-N p53 .......................................... 231
5.6. Conclusions ........................................................................... 234

Chapter 6: Conclusions and final remarks ............................................ 240
Bibliography ................................................................................. 243

Appendix A.1. Protparam details........................................................... 255
  A.1.1. hexaHis-Tag MDM2-N ...................................................... 255
  A.1.2. MDM2-N ......................................................................... 255
  A.1.3. hexaHis-Tag MDM 1-134 .................................................... 256
  A.1.4. MDM4 1-134 .................................................................... 256
  A.1.5 hexaHis-Tag MDM4-N ....................................................... 256
  A.1.5 MDM4-N .......................................................................... 257
A1.7 p53 derived peptide ................................................................. 257

Appendix A.2. Sequencing results ....................................................... 258

Appendix A.3. Vitamins and Micronutrients recipe ............................... 260

Appendix A.4. Peptide synthesis methodology ..................................... 260
  A.4.1. Activation of resin ............................................................. 260
  A.4.2. Coupling of Fmoc protected amino-acid ............................. 261

Appendix B.1. VsL details ................................................................. 262

Appendix B.2. Autodock ................................................................ 269
  B.2.1. Grid details ...................................................................... 269
  B.2.2. Docking details ............................................................... 269
  B.2.3. Autodock results ............................................................. 271

Appendix C. NMR details ................................................................. 271
Appendix C.1 .................................................................................................................. 271
Appendix C.2 MDM2-N re-assignment .......................................................................... 272

List of Figures and Tables

Fig 1.1. Schematic representation of p53 as a transcription factor ........................................ 2
Fig. 1.2. MDM2-MDM4: control over p53 ............................................................................. 4
Fig. 1.3. Sequence similarity of the MDMs ........................................................................ 6
Fig. 1.4. N-terminal region of MDM2 ................................................................................ 7
Fig. 1.5. The lid ................................................................................................................... 8
Fig. 1.6. Structures of MDM2-N and MDM4-NTD ................................................................. 10
Fig. 1.7. The p53-binding grooves of MDM4-NTD and MDM2-N ......................................... 11
Fig. 1.8. Topology of MDM2-N .......................................................................................... 12
Fig. 1.9. Surface-mapping of hydrophobic residues of MDM2 ............................................ 13
Fig. 1.10. MDM2 sub-pockets ............................................................................................ 14
Fig. 1.11. Plasticity of the p53-binding site of MDM2-N ..................................................... 15
Fig. 1.12. The role of MDM2 Tyr^{100} ................................................................................ 16
Fig. 1.13. First MDM4-N structures solved by X-ray crystallography ................................ 18
Fig. 1.14 MDM4 and MDM2 binding to p53-derived peptide ............................................. 19
Fig 1.15. Structures of MDM4-N solved in complex with different partners .................... 20
Fig. 1.16. Variation of the structure of the Leu26 sub-pocket within MDM4-N ................. 21
Fig. 1.17. Topology of p53-derived peptide when in complex with MDM2-N ................. 23
Fig. 1.18. Peptides that bind to MDM2-N ........................................................................... 24
Fig. 1.19. Early antagonist of MDM2 NTD-p53 interaction ............................................... 25
Fig. 1.20. MDM2 antagonists from virtual screening projects ........................................... 27
Fig. 1.21. MDM2-N in complex with Nutlin-2 [78]. ......................................................... 31
Fig. 1.22. Indole derivatives that are antagonist of the MDM2-N:p53 interaction ............. 31
Fig. 1.23. PMI peptide in complex with MDM4-N ............................................................ 33
Fig. 1.24. Small organic compound antagonist of the MDM4-N:p53 interaction ............. 35
Fig. 1.25 SJ-172550 compound, MDM4 antagonist ............................................................ 36
Fig. 2.1. MDM2-N protein sequence .................................................................................. 38
Fig. 2.2. MDM4 construct 1-134 ....................................................................................... 39
Fig. 2.3. MDM4 construct 14-111 DNA sequence ............................................................. 39
Fig 2.4 Primers for cloning MDM4 1-134 .......................................................................... 40
Table 2.4. Beer-Lambert equation and extinction coefficients ........................................ 49
Table 2.5. Some optimized parameters for 1-D experiments ............................................ 60
Table 2.6. KD by NMR: A1 fragment ................................................................................. 62
Fig. 2.8. A 3D NMR experiment ....................................................................................... 65
Fig. 2.9. Schematic representation of sequential assignment ........................................... 66
Fig. 2.10. $^{15}$N NOESY experiment ................................................................................. 68
Table 2.11. List of NMR experiments performed for assignment of the different constructs .... 69
Fig. 2.12. Relaxation diagrams ......................................................................................... 70
Fig. 3.1. Growth of bacterial cultures in various media ....................................................... 74
Fig. 3.2. Analysis of proteins produced by cells growing in various conditions ............... 75
Fig. 3.3. E. coli growth in the BioFlow 2000 fermentor as monitored by dissolved O$_2$ levels .... 77
Fig. 3.4. Fermentor to optimise conditions for production of MDM2-N, A ....................... 78
Fig. 3.5. The refolding process .......................................................................................... 81
Fig. 3.6. The results of a refolding trial ............................................................................. 82
Fig. 3.7. Conclusions on the refolding protocol ............................................................... 83
Fig. 3.8 First step in the purification of MDM2-N ............................................................ 84
Fig. 3.9. Thrombin cleavage .............................................................................................. 85
Fig. 3.10. Elution of MDM2-N after thrombin cleavage and loading on to the His-trap column .... 86
Fig. 3.11. Polishing step: gel filtration ............................................................................. 87
Fig. 3.12. MDM2-N LCMS results .................................................................................. 88
Fig. 3.13. 1D $^1$H spectrum of MDM2-N ...................................................................... 89
Fig. 3.14. A. $[^1$H, $^{15}$N] HSQC spectrum recorded for 0.05 mM $^{15}$N-MDM2-N (previous pages) .... 93
Fig. 3.15. HSQC spectra confirm integrity of re-folded protein ......................................... 94
Table 3.16. MDM4 publication up to May 2008 ................................................................. 95
Fig. 3.17. Cloning strategy of the N-terminal domain (1-134) of MDM4................................................. 96
Fig. 3.18. “Miniscale” protein production trial of MDM4 1-134 ................................................................. 98
Fig. 3.19 Precipitation problems with MDM4 1-134................................................................................. 99
Fig. 3.20. Gel filtration chromatography of MDM4 1-134........................................................................... 101
Fig. 3.21. HSQC spectrum MDM4 1-134...................................................................................................... 103
Fig. 3.22. HSQC spectrum of the refolded MDM4-N is near-identical to the HSQC of soluble protein. ................................................................................................................................. 104
Fig. 3.23. Buffer screening conditions for MDM4 1-134 ........................................................................... 106
Fig. 3.24. Structure of N-terminal domain of MDM4 reported in early 2009 .............................................. 108
Table 3.25. Features of E. coli expression host strains ................................................................................. 109
Table 3.26. Grid of expression conditions screened for the gene coding for MDM4-N. (Previous page) .................................................................................................................................................. 111
Fig. 3.27. Results of protein production trial ............................................................................................... 112
Fig. 3.28. First IMAC capture of MDM4-N .................................................................................................... 114
Fig. 3.29. LC-MS of species eluted from the Ni-NTA column. ........................................................................ 116
Fig. 3.30. Last step of MDM4-N purification ............................................................................................... 117
Fig. 3.31. Peptide synthesis ......................................................................................................................... 118
Fig. 3.32. p53 peptide purification ............................................................................................................... 119
Fig. 4.1. Drug development pipeline .......................................................................................................... 121
Table 4.2. Drug-discovery rules .................................................................................................................. 122
Fig. 4.3. SAR by NMR. Structure activity relationship by NMR methodology ............................................. 123
Fig. 4.4. VsL descriptors .............................................................................................................................. 125
Fig. 4.5. Workflow ...................................................................................................................................... 126
Fig. 4.6. Solubility of c17 ............................................................................................................................ 127
Fig. 4.4. DMSO titration ............................................................................................................................ 128
Fig. 4.5. Thermofluor-based stability assay of MDM2-N in DMSO ............................................................. 129
Fig. 4.6. First screening of the VsL - examples. .......................................................................................... 131
Fig. 4.7. Residues selected for analysis of CSP .......................................................................................... 132
Fig. 4.7. Evaluation of compounds in the VsL for binding to MDM2-N ..................................................... 133
Fig. 4.9. HSQC spectra of $^{15}$N- MDM2-N in complex with c29. ............................................................... 134
Fig. 4.10: Solubility test of c29 by NMR .......................................................... 135
Fig. 4.11. Titration of c29 into MDM2-N ............................................................ 136
Fig. 4.12. Poorly water-soluble potential MDM2-N binders ............................. 137
Fig. 4.13. A comparison of the chemical structures of some MDM2-N ligands ....... 138
Table 4.14. Titration details .............................................................................. 139
Fig. 4.15. Estimation of the potential of the hits to become leads ...................... 140
Fig. 4.16. Combined $^1$H and $^{15}$N chemical shift perturbations of MDM2-N amides upon binding to four hits (from the VsL) plotted versus sequence number ......................................................... 141
Fig. 4.17. Differential chemical shift perturbations of MDM2-N Tyr$^{100}$ by four ligands .................................................................................. 143
Fig. 4.18. A series of four MDM2-N binders selected by NMR screening from the VsL .................................................................................. 145
Fig. 4.19. Depiction of c18 and c20 .................................................................. 146
Fig. 4.20 Results of Autodock ......................................................................... 149
Fig. 4.21. Poses for c16 as predicted by Autodock ............................................ 151
Fig. 4.22. Detail of Tyr$^{100}$ side chain displacement upon ligand binding according to the poses created using Autodock ......................................................... 152
Fig. 4.23. Detail of MDM2-N-binding poses represented by conformers 6 and 7, created using Autodock ........................................................................ 153
Fig. 4.24. Comparison of the docked structure of MDM2-N/c16 complex (generated by Autodock) versus the X-ray-derived structure of MDM2 in complex with Nutlin-2 .................................................................................. 154
Fig. 4.25. Detail of putative H-bond formed between MDM2-N and docked c16 ...... 155
Fig. 4.26. Poses of c17 in complex with MDM2-N generated by Autodock .......... 157
Fig. 4.27. Binding modes of c17 conformer 2 vs Nutlin-2 ................................ 158
Fig. 4.28 Detail of the interaction c17 with MDM2-N ......................................... 159
Fig. 4.29. Poses of c21 bound to MDM2-N suggested by Autodock ................. 160
Fig. 4.30. Comparison of the binding mode of c21 (Autodock) vs. Nutlin-2 .......... 161
Fig. 4.31 Poses of c22 predicted Autodock ....................................................... 162
Fig. 4.32 Compound 22 (from VsL) binding to MDM2-N; comparison of poses from clusters 2 and 5 ................................................................. 163
Fig. 4.34. Detail of putative H-bond formed between MDM2-N and docked c22 ...... 164
Fig. 4.35. Detail of the H-bond between Trp23 (from the p53 peptide) and Leu54 (from MDM2-N) ...................................................................................... 165
Fig. 4.36. Proposed structure of MDM2-N binder .............................................. 167

Table of Contents
Fig 5.17. Assigned TROSY spectrum of [1H,15N]MDM4-N in complex with p53-derived peptide

Fig 5.18. Titration of p53 into 15N MDM4-N

Fig 5.19. Overlay of TROSY spectra for apo-MDM4-N vs MDM4-N in complex with p53-derived peptide

Fig 5.20. Assigned TROSY spectra of 15N-labelled apo-MDM4-N and of MDM4-N in complex with a racemic mixture of Nutlin-3

Table 5.21. Assignment of apo-MDM4-N and MDM4-N in complex with Nutlin-3

Fig 5.22. Missing MDM4-N amide resonance assignments

Fig. 5.23. Amide resonances missing from the assignment of the MDM4-N:Nutlin-3 complex

Fig 5.24. Chemical shift perturbations, MDM4-N apo vs MDM4-N Nutlin-3

Fig. 5.25. Combined chemical shift perturbations of apo-MDM4-N vs Nutlin-3 complex

Fig 5.26. Chemical shift perturbations for apo-MDM4-N vs MDM4-N p53

Fig 5.27. Combined chemical shift perturbations of apo-MDM4-N vs p53 complex

Fig 5.28. Secondary structure prediction

Fig 5.29. Details of MDM4-N a-helix number 2

Fig 5.30 Beta-bulge in MDM4-N (3DAB)

Table 5.31. Residues missing or excluded from analysis of 15N relaxation measurements

Fig 5.32 Relaxation data for apo-MDM4-N

Fig 5.33. Relaxation rates mapped onto MDM4-N

Fig 5.34 Relaxation data for MDM4-N in complex with Nutlin-3

Fig 5.35. Surface representation of MDM4-N

Fig 5.36 Relaxation data for MDM4-N in complex with p53 derived peptide. Fig 5.34 Relaxation data for MDM4-N in complex with Nutlin-3

Fig 5.38. Cartoon representation of a model of apo-MDM4-N

Fig 5.39. Cartoon representation of MDM4-N in complex with Nutlin-3

Fig 5.40. Detail of the H-Bonding of Tyr_{100} in MDM2-N

Fig 5.41. Details of H-bonds and WK298 binding to MDM4-N

Fig 5.42. Role of MDM4 Tyr^{66} in p53 recognition
Chapter 1: Introduction

1. Introduction

1.1. P53 is a key player in cancer biology

Cancer is a disease characterised by uncontrolled cell growth. Early studies revealed incidence within diverse tissues. This presented a challenge to the identification of the underlying causes of the disease. With the development of molecular biology, however, tumourogenesis has come to be regarded as a disease instigated primarily by mutations in genes involved in the cell cycle [1, 2].

The cell cycle consists of four stages. First the DNA is duplicated during S (synthesis) phase. This is followed by the G2 (Gap 2) phase during which the cell prepares for entry into the next phase, mitosis (M). During mitosis, the cell divides its contents between two daughter cells, which, in turn enter the G1 (Gap 1) phase. After this period, the cell initiates another cell cycle and divide once more, or enters quiescence [3].

Throughout the cellular cycle, the integrity of the DNA is subjected to constant monitoring. Many abnormalities that may arise in cellular DNA trigger cell cycle arrest by means of the p53 transcription factor. The consequent pause provides the cell with an opportunity to repair any damage incurred by the DNA. If the problem persists, p53 triggers a type of programmed cell death known as apoptosis [4, 5] (Fig. 1.1). This specific mission of p53 as “guardian of the genome” requires that it occupy a central position within a complex network of pathways.

Post-translational modifications of p53 as well as interactions with other proteins have been proven to control activation and inhibition of p53-induced apoptotic activity and hence its role as a tumour suppressor [6-8]. Furthermore, circumvention of p53 pro-apoptotic activities is a key factor in the development of many cancers; for example, p53 is thought to be mutated in half of all incidences of cancer [9-11]. Furthermore, many additional cancers arise from the loss of p53’s tumour-suppression activities due to alterations in its binding partners and in particular in
those components that lie upstream or downstream of p53 on the apoptotic pathway [10, 12].

<table>
<thead>
<tr>
<th>Growth Factor activation</th>
<th>DNA damage</th>
</tr>
</thead>
<tbody>
<tr>
<td>P53 antagonists</td>
<td>P53 activators</td>
</tr>
<tr>
<td>MDM2</td>
<td>Atm, Chk2</td>
</tr>
</tbody>
</table>

Post-translational modifications

![p53](image)

- Ubiquitination
- Proteosomal Degradation of p53
- Transcription Factor Activities

<table>
<thead>
<tr>
<th>Mitosis</th>
<th>Cell cycle arrest</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Apoptosis</td>
</tr>
</tbody>
</table>

**Fig 1.1. Schematic representation of p53 as a transcription factor.** Modulation of p53 activity by post-translational modifications controls the cell cycle. While progression to mitosis is achieved through targeted degradation or inactivation of p53, enhancing p53’s transcription-factor activities after DNA damage will initially arrest cellular growth possibly concluding eventually in cellular death. Adapted from [6]

Consequently, understanding the transcription factor-like activity of p53 has been identified as a critical route to target and defeat cancer [13]. Currently, there are many approaches tailored to expedite specific steps on the pathway leading to p53 restoration. Two of many examples are the efforts to reinstate mutated p53 activity by gene therapy [14] or by using small molecules [15]. In other cases, diverse strategies have emerged to avoid p53 degradation by its major inhibitor, MDM2.
(Murine Double Minute 2) (see below). One route to restrain MDM2 has been inhibition of its expression by antisense oligonucleotides, which showed positive results in breast-cancer cell lines with and without wild type (wt) p53 [16]. Another strategy is to impair the interaction of p53 and MDM2 using small molecules that block the binding between these proteins. Several proven, specific inhibitors of this interaction have been discovered or designed and assessed [17].

1.2. MDM2-MDM4 coordinate regulation of p53 levels in the cell

The p53 protein is a central player in the intricate and interconnecting pathways of apoptosis, where many regulatory mechanisms work synergistically to maintain homeostasis. The current study will focus on inhibition of p53 by MDM2 and MDM4 (also known as MDMX). A comprehensive review can be found elsewhere [18]. Given the amount of publications released periodically on the current subject, this thesis only pertains to papers published before March 2011.

In classic texts, the regulation of p53 is described as an amalgam of post-translational modifications such as phosphorylation, sumoylation [19] and acetylation [6]. On the other hand, MDM2 has emerged as a key player since it physically interacts with p53 and leads to its proteasomal degradation. Moreover, findings of p53-responsive elements in the gene of MDM2 suggested a feed-back loop for regulation and the establishment of a fine equilibrium between the activities of the two proteins [20].

In normal cells p53 is down regulated by MDM2, maintaining a basal concentration of the transcription factor. At the same time, low cellular p53 levels control quantities within the cell of MDM2. DNA damage and other stimuli break this equilibrium via post-translational modifications such as phosphorylation of p53 and/or MDM2 or production of the ADP-ribosylation factor (ARF) protein. ARF is a negative regulator of MDM2 [21], which disassociates the p53-MDM2 interaction, and inhibits MDM2 production. On the other hand, and as mentioned above, p53 induces MDM2 transcription, ensuring that once p53 has accomplished its mission it will be quickly degraded (Fig. 1.2).

In addition, several levels of interaction between p53 and MDM2 have been described, based on different post-translational modifications in both proteins. This
observation supports the findings for the functional significance of different phosphorylation sites situated throughout MDM2, which in turn may lead to attenuation of the p53-MDM2 interaction [9, 22].

The role of MDM4 as a negative effector of p53 was confirmed by the correlation of MDM4 over-expression with several tumour incidences [23, 24]. The fact that MDM4 is an oncoprotein in the p53 pathway together with its striking sequence similarity with MDM2 implicated it in the regulation of p53. Furthermore, several experiments had shown that attenuation of MDM2 does not result in p53 being fully activated [25]. Crucially, knockout mice for mdm2 or mdm4 are lethal, but both can be rescued by deletion of p53 [26], which led to the conclusion that collaboration occurs between MDM2 and MDM4 in the regulation of p53. Therefore, control over p53 has been suggested to be shared and evidence shows that while MDM2 primarily accounts for p53 stability by collaborating in its degradation, MDM4 primarily inhibits the transactivation activity of p53 (Fig. 1.2) [27, 28].

Fig. 1.2. MDM2-MDM4: control over p53. The regulation of p53 is coordinated by MDM2 and MDM4. DNA damage triggers p53 trans-activation (1) genes encoding key proteins involved in apoptosis and cell cycle arrest, and MDM2. Moreover, MDM2 maintains the cellular levels of p53 by ubiquitin-tagging of p53 thus directing it towards degradation by proteasomes, (2) closing a negative feed-back loop. Meanwhile, MDM4 blocks p53 transcriptional functions (3) by physically interacting with p53 and blocking its trans-activation domain. Adapted from [29].

The discovery of heterodimerization of MDM2 and MDM4 RING domain(s) in vivo and the relevance of this phenomenon for control of p53 ubiquitination indicates a sophisticated cooperation between the two MDM proteins [30]. This implicates
MDM4 in the ubiquitination process of p53 [31]. Indeed, while MDM2 homodimerization might effectively target p53 for proteasomal degradation MDM4-MDM2 homodimerization appears to increase the p53-degradation rate in a MDM2 dependent manner [32]. Thus, MDM2 homodimers and MDM2-MDM4 heterodimers control the basal levels of p53. After DNA damage, however, MDM2 homodimers are destabilized by phosphorylation partially increasing basal levels of p53. The heterodimer still exerts control over p53, but MDM2-mediated ubiquitin tagging of both MDM4 and itself will lead to proteosomal degradation of both proteins, finally releasing p53 as a fully entity. Decreasing stress signals, on the other hand, triggers dephosphorylation of MDM2 and MDM4, stabilizing the heterodimers and regenerating the status quo [33].

1.3. What is known about MDM2 and MDM4

1.3.1. Discovery of MDM2

The *mdm2* gene was first isolated from mouse tumour cell line Balb/c3T3 when analyzing the extrachromosomal DNA known as double minutes (DM) [34]. Subsequently, the DM and, specifically, enhanced expression of *mdm2* were associated with tumorigenesis thus branding *mdm2* as an oncogene [35]. A detailed examination of RNA in human sarcomas proved overexpression of the *mdm2* gene in these cell lines supporting the previous data [36]. The expression of *mdm2* yielded a protein – MDM2 - of an apparent molecular weight of 90 kDa. Its interaction with p53 was proven initially by co-immunoprecipitation of the MDM2-p53 complex. Furthermore, transactivation of a reporter gene containing the p53 response element was shown to be impaired by MDM2 [37]. In a key experiment, *mdm2* knockout mice proved lethal in the embryogenic state while *p53* and *mdm2*-null mice developed normally, definitively linking *mdm2* early activities to control of p53 functions [38].

1.3.2. Discovery of MDM4

The protein MDM4 was isolated from a mouse cDNA expression library and identified as a p53-binding protein [39]. The high sequence similarity of MDM4 to MDM2 (Fig. 1.3) prompted research proving that MDM4 is a negative effector of
p53. This was surmised from the lethality in MDM4-knockout embryos that was observed in a p53-dependent manner, while overexpression of MDM4 led to tumour formation by suppressing p53 activities [40]. Moreover, the tumorigenic character of MDM4 was proven after finding MDM4 over-expressed in wt p53 tumours [24]. The findings of embryonic lethality of *mdm2* or *mdm4* knockdown out mice rescued by *p53* knockout- as mentioned above - proved that both proteins collaborate in keeping p53 levels under control and together.

Fig. 1.3. Sequence similarity of the MDMs. Top panel shows an alignment of the human MDM2 and MDM4 highlighting in black identical residues, and in grey, similar residues. The bottom scheme depicts the domain composition of MDM2 and MDM4. Differential presence of NLS (nuclear localization signal) and NES (nuclear export signal) are indicative of different location profiles with MDM4 as a cytoplasmic protein while MDM2 can shuttle from nucleus to cytosol freely [40].

1.3.3. Domain composition of the MDMs and functional consequences

The MDM2 protein is 491 amino acid residues in length and contains three well conserved regions that have been structurally characterised: the N-terminal domain involved in p53 interactions (residues 19-108, labelled MDM2-N in the current
work); a zinc (Zn) finger domain (residues 290-330) and a RING domain with E3 ubiquitin ligase properties (residues 435-482). Within the sequence, various motifs were localized, such as the NLS (nuclear localization signal) (residues 179-185) and NES (nuclear export signal) (residues 190-202); an acidic domain with no sequence conservation (residues 237-288); and a flexible structure at the very N-terminus of the protein known as “the lid” (residues 16-24) which has also been found to be involved in the p53 interaction (Fig. 1.3, bottom) [41-44].

Of the domains in MDM2, the N-terminal region has been most extensively studied. This is partly because it was recognised to be directly involved in the mechanism by which MDM2 binds to p53 and suppresses its activity. A deletion-mutagenesis derived functional map of MDM2 identified the approximate residues involved in engagement with p53 and demonstrated that MDM2 physically blocks the p53 transactivation domain [42, 45]. The determination of the crystal structure of MDM2-N by Kussie et al., [46] revealed two similar sub-domains, each formed by two α-helices supported by β-strands (see Fig. 1.4). These have an axis of pseudosymmetry forming a hydrophobic pocket that harbours the interaction with p53.

**Fig. 1.4. N-terminal region of MDM2.** A. Representation of MDM2-N (red) bound to a p53-derived peptide (yellow) as first proposed by Kussie et al. in 1996 (PDB 1YCR). B. Schematic representation of MDM2-N topology adapted from [46]. Note that the box in the bottom of the picture clarifies the orientation of the viewer.
The extreme N-terminal region of MDM2 corresponds to the aforementioned lid and consists of a sequence that is highly conserved in mammals but appears to be very flexible (Fig. 1.5). The lid appears to have intramolecular affinity for the p53-binding groove of MDM2 (see Fig. 1.4), suggesting an auto-inhibitory role in the interaction with p53. It was proposed that the lid could close over the groove after phosphorylation of MDM2 Ser$^{17}$ [44]. This mechanism falls within the definition of “intrasteric regulation”, which is a novel regulatory system based on conformation changes after post-translational modifications such as phosphorylation [47]. However the role of the lid is not fully understood. Paramagnetic studies on the MDM2 N terminus suggested that the lid prefers a predominantly closed conformation in solution and that the open conformation is poorly represented in the absence of ligand [48]. This agrees with previous structural evidences that the lid has to be displaced from the binding groove of MDM2-N to make room for the p53 binding partner, with an accompanying readjustment of the secondary structure [49].

![Fig. 1.5. The lid. Cartoon representing different conformations of the lid of MDM2-N in reference to the core structure, as obtained from the structure solved in solution by Uhrinova in 2005 (PDB code 1Z1M [49]).](image)
The central region of MDM2 harbours various motifs. The NLS and NES motifs allows MDM2 passage across the nuclear membrane supporting the notion that MDM2 sequesters p53 from the nucleus, thus impairing by means of delocalization its transcription factor activities [41, 50]. Another motif located in the central region of MDM2 is the acidic region. Its sequence does not seem to be conserved amongst orthologues; however the acid nature of this region seems to be a conserved feature in different species. Reports showed a second binding site for p53 in this MDM2 central region [51, 52], and this interaction appears to be modulated by phosphorylation and the tetrameric conformation of p53 [53]. Finally, studies of deleted/mutated forms of MDM2 lacking or impairing the acidic domain demonstrated the importance of this section for the ubiquitination of p53 [54, 55], supporting evidence of cooperative function between MDM2 domains in p53 modulation [55].

The Zn-finger motif next to the acidic domain region (Fig. 1.3) of MDM2 was found to be involved in interaction with the ARF negative effector of MDM2 and various ribosomal proteins [56].

The C-terminal motif in MDM2 is a RING finger domain. Briefly, the RING finger domain was first discovered in RING1 (Really Interesting New Gene 1) protein and proved to be involved in protein-protein interactions. Further screening indentified this domain as an essential component of E3 ligases involved in ubiquitin transfer in the proteasome-mediated degradation pathway [57, 58]. The identification of the RING finger domain in MDM2 shed light on its biochemical function and several subsequent studies proved the ability of MDM2 to transfer ubiquitin residues onto p53 [59]. Mutagenesis studies showed that this domain was essential for trans- as well as auto- ubiquitination and support the evidence that MDM2 acts as an E3 ubiquitin ligase [60]. However, recent studies, focused on the RING finger domain, have found that it is also involved in heterodimerization with MDM4, and this interaction has been proven to be important in ubiquitin ligase activity [30]. The structure of the homodimer (MDM2-MDM2) [61] as well as heterodimer (MDM2-MDM4) [62] of the RING domain of MDM2 implicated E3 domain in the regulation of p53, highlighting the importance of cooperative partnership of MDM2 and MDM4 [63].
The domain composition of MDM4 closely matches to that of MDM2 (Fig. 1.3) with the three main domains distributed in the same fashion i.e. p53-binding region (residues 26-106, MDM4-NTD), acidic domain (residues 246-332) and the RING domain (residues 393-490) [40]. Despite minor differences, the assumption that similar structure equals to similar functions is central in the understanding of MDM4. Initial structural studies on the MDM4-NTD domain were based on models of MDM2-N, using molecular replacement to predict the structure of MDM4-NTD binding a p53 peptide [64]. Initial results predicted a less shallow binding pocket for MDM4 caused mainly by the different orientation of α-helix 2’ (Fig. 1.6). The sequence Pro$^{95}$-Ser$^{96}$-Pro$^{97}$ at the C-terminus α-helix 2’ is displaced 2.5 Å to that of MDM2-N counterpart (His$^{96}$-Arg$^{97}$-Lys$^{98}$), contributing Met$^{50}$ and Tyr$^{96}$ (in zebrafish) or Met$^{53}$ and Tyr$^{99}$ (in human) side chains towards the binding site, changing the binding groove shape and quite drastically diminishing the accessible area of the Leu26 sub-pocket (Fig. 1.7) [65].

![Fig. 1.6. Structures of MDM2-N and MDM4-NTD. Cartoon representation of the MDM4-N (green) in complex with p53-derived peptide (yellow) as solved using crystallography by Popowicz et al. in 2008 (PDB code 3DAB) [65]. Overlay with MDM2-N (in red, PDB 1YCR) clearly shows the difference in α-helix 2’ as highlighted with the dashed lines.](image)
The p53-binding grooves of MDM4-NTD and MDM2-N. Details from surface representations of MDM4-NTD (green, PDB 3DAB), and MDM2-N (in red, PDB 1YCR), highlighting the size difference in the Leu26 sub-pocket. The arrow points out the extra bump within the MDM4 binding groove caused by the orientation of Met$^{53}$ and Tyr$^{99}$ side chains that reduces access to the floor of the groove.

Another difference (versus MDM2) in the structure of MDM4 lies in their acidic domains (Fig. 1.3). While this region is not conserved, an acidic region in the middle part of MDM4 is also detected; however, this is smaller than the MDM2 counterpart. The role of this region on MDM4 has not yet been elucidated. The lack of NES or NLS in this region of MDM4 marks another difference between the proteins; it also suggests a cytoplasmic role for MDM4.

The RING finger domain of both proteins share up to 53% sequence identity [40]. Despite the fact that MDM4 has ubiquitin ligase properties \textit{in vitro}, it has been shown that it does not target p53 \textit{in vivo} [66]. In addition, the RING finger domain was shown to be involved in heterodimerization between MDM2 and MDM4 suggesting MDM4 has the role of a stabilizer of MDM2 and thereby enhances p53 degradation [32].

1.4. Details of the binding sites within the MDMs

Although the interaction of MDM2 or MDM4 with p53 probably extends throughout the length of the protein, most attention, and indeed the focus of this thesis, has been
on their N-terminal domains that carry the well explored p53 binding groove. Therefore, the remainder of this review is concerned mainly with MDM2-N and MDM4-N terminus.

1.4.1. Details of the p53-binding groove of MDM2

Mapping of the interaction of p53 with MDM2 started with the abovementioned crystal structure solved in 1996 [46], which showed there is a specific groove-like contact surface in MDM2 for a helical stretch of p53. This makes it a potential target for drug design. The binding cleft features a 25-Å long and 10-Å wide pocket [49], with two α-helices delimiting the walls of the groove and two β-strands acting as hinges, while two more α-helices close off the bottom of the cleft (Fig. 1.8). The calculated accessible area of MDM2 involved in the interface with p53 is 660 Å while the total buried surface areas of the MDM2-N:p53 (17-29) complex is 809 Å [67].

Fig. 1.8. Topology of MDM2-N Surface representation of MDM2-N with detailed secondary structure represented in cartoons. The binding site is mainly defined by α-helices 2 and 2’ (blue) flanked by two β-strands (light blue) and highly flexible loops (red). The remaining α-helices, 1 and 1’ (yellow), close up the binding groove at the bottom of the structure.
The binding pocket of MDM2-N is mainly composed of hydrophobic residues (Fig. 1.9), where p53 amino acid residues Phe\(^{19}\), Trp\(^{23}\) and Leu\(^{26}\) are buried. Various studies of the dynamics of the docking between MDM2-N and different peptides of p53 have suggested contact between aromatic amino acids as the stabilizing force of the interaction [49, 68, 69], while hydrogen bonds between the backbone amide of p53 Phe\(^{19}\) and O\(^{\#}\) Glu\(^{72}\) of MDM2, and N\(^{\#}\) of p53 Trp\(^{23}\) and the backbone carbonyl of MDM2 Leu\(^{54}\) orientate the peptide in a specific binding pose [70, 71]. Moreover, the backbone carbonyl of Asn\(^{29}\) of p53 and hydroxyl Trp\(^{100}\) of MDM2 also establish an intermolecular H-bond in complexes containing longer-size p53 peptides, which is not essential but presumably stabilises the physiological interaction with p53 [72].

**Fig. 1.9. Surface-mapping of hydrophobic residues of MDM2.** Mapping of the hydrophobicity of residues in MDM2-N (PDB 1YCR) accordingly to the Kyte and Doolittle scale [73]. The colour coding goes from blue for hydrophilic residues to orange for most hydrophobic residues with white as a mid-point. The picture was produced using Chimera [74].

Indeed the MDM2-binding pocket is complementary to the p53 MDM2-interacting region in many aspects besides polarity. Shapewise, both interfaces fit nicely, thereby defining three sub-pockets on the MDM2-interacting surface named after
their corresponding buried p53 amino acid residues. The deepest sub-pocket (Fig. 1.10 B) falls in the middle of the cleft and accommodates Trp$^{23}$, while Phe$^{19}$ and Leu$^{26}$ are anchored in shallower, flanking pockets (Fig. 1.10 A and C, correspondingly).

Both interacting partners bring an element of plasticity to the interface. Experiments with different size p53-derived peptides demonstrated that the peptide does not have a specific conformation in solution, but adopts an extended helical fold when in complex with MDM2. Moreover, the non-liganded (apo-)MDM2-N has been characterized by many researchers as being relatively dynamic and flexible [49, 75] while it becomes more rigid when bound to its natural partner, p53 [76]. Plasticity of the binding groove is indeed evident based on comparisons of the crystal structures of MDM2-N in complexes with different p53-derived partners. The shape and the size of the cleft vary greatly depending of the binding partner (Fig. 1.11) and this has to be taken in consideration in further drug development (see below). Further studies of MDM2-N [48] discounted the possibility of global changes arising from interactions with other domains of the full-length protein. But the discussion is still open, with many strands of evidence suggesting a significant amount of structural and functional cross-talk between the various MDM2 domains [47, 77]
Fig. 1.11. Plasticity of the p53-binding site of MDM2-N. Surface representation, coloured by hydrophobicity (by Chimera) of MDM2-N in complex with (although they are not shown in these pictures) p53-derived peptide (1YCR), Nutlin-2 (1RV1), p53-derived peptide analogue (2GV2) and small-molecule inhibitor (3LBK). The location of Tyr^{100} is indicated in the 1YCR representation, as is the location of the Leu26 sub-pocket. Figure adapted from [70].

The plasticity of the binding site is mainly due to the position of MDM2 residue Tyr^{100}, which is located at the rim of the Leu26 sub-pocket (Fig. 1.12). Residue 100
is tilted towards the binding groove in the *apo*-MDM2-N solved by Uhrinova in a “closed disposition”, anchored by an H-bond with Leu\(^{54}\). It moves away from the groove in an “open disposition” in the crystal structure of MDM2-N complexed with p53 peptide [70, 71]. Moreover, the disposition of Tyr\(^{100}\) changes depending on the ligand so as to accommodate the interaction accordingly (Fig. 1.12).

**Fig. 1.12. The role of MDM2 Tyr\(^{100}\).** Representations of different structures of MDM2-N in which the positions of residue 100 (sticks) are highlighted, so as to show the different dispositions relative to the p53-binding groove. In *apo*-MDM2-N (red, PDB 1Z1M [49]) the Tyr\(^{100}\) side chain is orientated towards the binding groove (“closed”), while in MDM2-N:p53-derived peptide complex (green, PDB 1YCR [46]) it is tilted towards the outside of the binding groove (“open”). MDM2-N in complex with Nutlin-2 (navy blue, PDB 1RV1[78]), p53-derived peptide analogue (cyan, pdb 2GV2 [79]) and a small-molecule inhibitor (purple, PDB 3LBK [80]) display differing positions of the residue 100 in semi-open dispositions.

Due to the strategic location of Try\(^{100}\), it has been proposed that the movements of this particular residue could be modulated by the lid, and therefore when the lid closes over the binding site, allostERIC clashes would force Tyr\(^{100}\) to adopt the closed conformation. Release of the lid by p53 binding would therefore provide more freedom of movement of Tyr\(^{100}\) and this residue can then be displaced by p53 Leu26-occupancy of the sub-pocket [70].
1.4.2. Insights into the p53-binding groove of MDM4

Early understanding of the importance of MDM4 in p53 regulation as well as its high level of similarity with MDM2 prompted investment of effort in structural studies on MDM4-N. Initial reports described the protein as highly unstable, and the first X-ray structure recorded in the literature was that of a zebrafish construct humanized by changing several critical residues (L46V and V95L) to mimic the human p53-binding site [64] (Fig. 1.13). The X-ray structure revealed a similar fold to that of MDM2-N with sheet-helix pairs arranged around an axis of pseudosymmetry defining the hydrophobic surface where the p53-derived peptide is buried (Figs. 1.6 & 1.4). In comparison with MDM2-N, however, the binding pocket of MDM4-N is smaller and shallower. This is caused, mainly, by the presence of several bulky residues orientated towards the binding pocket and occluding the groove (Fig. 1.6).

Fig. 1.13. First MDM4-N structures solved by X-ray crystallography. The top panel presents 3D representations of MDM4-N PDB entries for zebrafish (2Z5S), “humanized” zebrafish (2Z5T) and human MDM4-N (3DAB) respectively. The arrows point to the single-point mutations that converted the zebrafish variant into “humanized” MDM4-N (L46V and V95L). Clearly, the shape of the p53-binding pocket is different, while that of the humanized zebrafish MDM4-N is more similar to the human MDM4-N p53-binding pocket. The nomenclature of the binding pocket is shown on the human MDM4-N structure, 3DAB. The bottom panel shows alignment of the different constructs used for
crystallographic studies of MDM4. The coloured boxes refer to the secondary structure defined in the PDB, thus green for β-sheet and yellow for α-helix. The conservation of the residues is very high in the structural motifs with low occurrence of non-conservative changes (highlighted in black). The figure was generated with Chimera. Adapted from [64].

Indeed the MDM4-N has invariably been described as a challenging system and this is borne out by the inventive techniques used to purify the protein [81, 82]. Eventually NMR studies proved that apo-MDM4-N is in fact loosely structured with evidence of multiple conformations in intermediate exchange [83]. The poorly structured apo state reverts to a structured one when MDM4-N is bound to the p53-derived peptide [83].

The sketch structural details available for apo-MDM4-N make it difficult to itemise the structural differences between liganded and apo-forms of the protein domain. Nonetheless, the several X-ray crystallography-solved structures of MDM4-N bound to different ligands provide valuable information on the binding process. As expected, the interaction with the p53-derived peptide involves three key p53 residues (Phe\textsuperscript{19}, Trp\textsuperscript{23} and Leu\textsuperscript{26}) interacting in with MDM4 in a similar fashion to that seen in the complex of MDM2-N:p53-derived peptide (Fig. 1.14). The interaction is mainly mediated by hydrophobic contacts and secured by p53 Trp\textsuperscript{23}-MDM4 Leu\textsuperscript{54} hydrogen bonding as was observed in the case of the MDM2:p53 interaction [69, 71]. But the orientation of Tyr\textsuperscript{90} (equivalent to Tyr\textsuperscript{100} in MDM2) within the MDM4 p53-binding surface influences the interaction, and the C-terminus of the p53 peptide adopts a more extended conformation, avoiding steric clashes with the rim of the binding groove defined by the protruding side chains of Met\textsuperscript{54} and Tyr\textsuperscript{99} (Fig. 1.14) [84].
Fig. 1.14 MDM4 and MDM2 binding to p53-derived peptide. Both proteins interact with p53 peptides in a similar way. Top panel shows the overlay of the stick representation of the p53 peptide from 3DAB (MDM4-N complex) in blue, and 1YCR (MDM2-N complex) in green. The orientation of Phe\textsuperscript{19}, Trp\textsuperscript{23} and Leu\textsuperscript{26} in both is very similar; however, Leu\textsuperscript{26} does not penetrate as deeply into the MDM4 p53-binding pocket due to the shape of the latter. The bottom panel shows surface representation of MDM4-N (green) and MDM2-N (red) with their respective p53-derived peptides bound. Very clearly, the rim of the Leu\textsuperscript{26} sub-pocket forces Leu\textsuperscript{26}-Pro\textsuperscript{27} to adopt a more extended conformation in order to bypass the protuberance created by the Leu\textsuperscript{26} side-chain. Picture adapted from [84]

By analogy with MDM2-N, the plasticity of the MDM4-N p53-binding site is likely to be of interest. Most of the structures solved for MDM4-N have been in complex with peptides [84, 85] or peptidomimetics [81]. Only one solved structure is of MDM4-N in complex with a small compound [80]. Nonetheless, comparison of the solved structures suggests that MDM4-N has inherently more backbone flexibility than MDM2-N [81, 83]. In particular, there appears to be plasticity within the MDM4-N Trp\textsuperscript{23} sub-pocket and Leu\textsuperscript{26} sub-pocket (Fig. 1.15). Thus, like MDM2-N but probably to a greater extent, MDM4-N can adapt its binding groove to maximise interactions with an interacting partner [81]. The physiological relevance of this observation remains a matter of speculation but it clearly has important consequences for the design of compounds that could block wt MDM2/4-p53 interactions.
Fig 1.15. Structures of MDM4-N solved in complex with different partners. Surface representation of structures of MDM4-N in complex with (although they are not shown in these pictures): p53 peptide (3DAB, [64]), p53 chlorinated peptidomimetic (3FE7, [81]), 12-mer peptide (3JZO, [85]) and with a small-molecular weight inhibitor (3BLK, [80]). The differences in shape and size of Trp23 and Leu26 sub-pockets demonstrate the flexibility of the MDM4 p53-binding pocket and its adaptability to individual partners. Moreover, chlorinated derivates provoke a widening of the Trp23 sub-pocket very evident when comparing the structures on the left (non-chlorinated derivates) with the ones on the right (chlorinated binding partners).
The most striking example of MDM4 p53-binding site changes described so far pertains to the accommodation of chlorinated peptidomimetics in the Trp23 sub-pocket. This deep sub-pocket is modified by movements of the MDM4 Leu<sup>98</sup> side chain in order to accommodate the larger chlorine atom [80, 81]. The chlorine-induced repositioning of the Leu<sup>98</sup> side chain prompts other local conformational changes in the vicinity. Important amongst these is a reorientation of Tyr<sup>99</sup> from closed (with chlorine present) to open (with chlorine absent) dispositions (Fig. 1.16), contradicting previous assumptions of a rigid role for this residue in the conformation of the Leu sub-pocket [81].

![Fig. 1.16. Variation of the structure of the Leu26 sub-pocket within MDM4-N. The residue Tyr<sup>99</sup> adopts different dispositions to accommodate individual ligands in the p53-binding pocket of MDM4. The transition of Tyr<sup>99</sup> between open (blue) and closed (orange) dispositions as been likened to the side-chain readjustments of Tyr<sup>100</sup> in MDM2-N [81].](image)

### 1.5. Drug discovery and the modulation of MDM2/4-p53 binding

A detailed account of the very extensive literature describing efforts to discover or design compounds that efficiently inhibit the MDM2-p53 interaction is beyond the scope of this thesis. Such reviews are already available elsewhere [67, 86-89].
Instead, the aim of this section is to describe the main steps in drug design and to review the most potent inhibitors reported so far (up to March 2011). This will set the scene for the experimental work in the current project.

1.5.1. MDM2-N antagonists

1.5.1.1. First steps in discovery of antagonists

It was understood early on that while the p53-derived peptide is unstructured on its own in solution [75], non-ligated MDM2 retained a well-structured binding pocket. The interaction between MDM2-N and p53-derived peptide thus involves the twisting of the polypeptide so that it can bury itself into the cleft of MDM2. Whether or not the binding site in p53 corresponding to the peptide undergoes a similar transition from disordered to ordered remains unknown, but the behaviour of the isolated peptide is important for designing peptidomimetics and other inhibitors. Indeed a predominant strategy has involved mimicking the relevant region of p53 [42]. This was founded mainly upon structural insights into the MDM2-N:peptide interaction (as described above).

Structurally, the bound peptide adopts an amphipathic helical conformation with Phe\textsuperscript{19} and Trp\textsuperscript{23} side chains orientated towards the p53-binding groove in quasi-parallel directions. The peptide then adopts a type-I β-turn, burying Leu\textsuperscript{26} in the MDM2 binding groove (Fig. 1.17) [46].
Fig. 1.17. Topology of p53-derived peptide when in complex with MDM2-N. The top panels show the sequence of the p53-derived peptide as solved by Kussie in complex with MDM2-N (Kussie 1999), followed by a drawing of the molecule in ChemDraw (Cambridge Soft). The bottom panel is a cartoon showing the structural disposition of this peptide when binding to MDM2-N.

Systematic screening of peptide libraries by phage display identified various peptides with affinity for MDM2-N [90, 91]. As expected, Phe$^{19}$, Trp$^{23}$ and Leu$^{26}$ are conserved as the p53 residues key attaining affinity for MDM2-N, but the length of the peptide could be reduced from 15 residues to 8, with only a small loss of affinity (Fig. 1.18, A) [91]. Studies to improve the potency of 8-mer peptide derivatives followed several paths. Positive entropic contributions were gained after incorporation of unnatural amino acids that are known $\alpha$-helix inducers such as aminoisobutyric acid (Aib) and 1-amino-cyclo-propanecarboxylic acid (Ac3). Gain of affinity was achieved by incorporation of phosphonomethylphenylalanine (Pmp) in position 22 introducing a salt bridge with MDM2 Lys$^{54}$. Adding a chlorine substituent in position 6 of the Trp$^{23}$ side chain improved the occupancy of the Trp23 sub-pocket, increasing significantly the affinity of the resultant peptide (Fig. 1.18)[91].
Fig. 1.18. Peptides that bind to MDM2-N. List of peptides that summarizes the increasing potency achieved by adding unnatural amino acids or by modification of existing residues (boxed). The IC$_{50}$ values reveal increasing inhibition of the p53-MDM2-N interaction. Figure adapted from [91]

Approaches with β-peptides produced satisfactory results, achieving sub-micromolar values for $K_D$. However, these compounds still have a major drawback due to their poor cellular permeability [92]. Research on β-hairpins derived from p53 also produced several promising molecules with increased stability, and $K_D$ in the nM range [93].

Indeed the pace of ligand discoveries for MDM2 NTD has picked up in recent years, and the unique features of the MDM2 binding groove have come to be regarded as a “textbook case” for inhibition of protein-protein interactions by small molecules.

### 1.5.1.2. Small-molecule inhibitors of MDM2:p53 interaction

#### 1.5.1.2.1. Early drug discovery campaigns

The first campaigns to identify small-molecule inhibitors and initial high-throughput screens were focused on natural derivates with previously reported anti-tumour effects. These, however, achieved little success [94, 95]. Subsequently, studies of
peptide:MDM2-N complexes led to interest in terphenyls as scaffolds that could mimic \( \alpha \)-helical structures. Addition of alkyl and aryl substituents to terphenyl was undertaken to emulate the side chains of p53 Phe\(^{19} \), Trp\(^{23} \) and Leu\(^{26} \) [96] (Fig. 1.19). More sophisticated alternative approaches employed Trp\(^{23} \) as a starting point and utilised combinatorial chemistry to yield a library of related compounds. Subsequently, screening for hits revealed compounds with IC\(_{50} \) values in the 0.5 \( \mu \)M range [97] (Fig. 1.19).

**Fig. 1.19. Early antagonist of MDM2 NTD–p53 interaction.** Representations generated with ChemDraw of first inhibitors of MDM2-N adapted from [86]. A. The terphenyl provided a scaffold upon which alkyl or aryl derivates were attached at the \( X_1, X_2 \) and \( X_3 \) positions to mimic key p53 residues. B. A tryptophan derivative obtained by combinatorial chemistry.

### 1.5.1.2.2. *In-silico* screening for antagonists

Methods have been developed for identifying small-molecule inhibitors of the MDM2-p53 interaction based on virtual selection of compounds by *in-silico* screening tools. Virtual screening is gaining increasing importance across the drug-discovery field. In particular, virtual docking of compounds has been used in a pre-selection mode to assess their potential and the prospects of further elaboration and development [98]. Virtual screening involves mapping key atoms involved in the

Chapter 1: Introduction

25
interaction and defining parameters that must be met by a prospective inhibitor. Three different approaches that were implemented using MDM2 as target are described here: hydropathic interactions (HINT) [99, 100], ligand discovery at Edinburgh University (LIDAEUS) [101], and screening using an isoindolinone scaffold [102].

The program HINT creates a pharmacophore based on structural information derived from inspection of MDM2-N in complex with various different peptides [100]. A HINT-based search of the National Cancer Institute (NCI) 3D library yielded a sulfonamide derivative that inhibited in the micromolar range (see Fig. 1.20). The LIDAEUS program, on the other hand, defines an array of interactions based on the X-ray crystal structure of MDM2-N in complex with the p53-derived peptide. This information was then used to dock into the binding site a virtual library of available small molecules, with scoring for energy minimization and pose. The resultant best hits were bisaryl sulfonamide or thiophene-2-sulfonic derivatives with IC$_{50}$ values in the low-micromolar range [86] (see Fig. 1.20). Finally, Hardcastle et al [102] using isoindolinone scaffold in a fusion of in silico screening and combinatorial chemistry yielded isoindolinone Nu8231 (Fig. 1.20) with IC$_{50}$ of 5 µM, highlighting the importance of organic scaffolds in drug discovery, and using in silico tools in rational design of inhibitors. Overall in silico screening proved to be a valuable methodology, especially considering its low cost and high efficiency.
Fig. 1.20. MDM2 antagonists from virtual screening projects. Representation of small-molecule inhibitors of the p53:MDM2 interaction obtained after in silico screening. A. The programme HINT yielded, initially, a pharmacophore that is defined by the distribution of functional groups of three small-molecules [99]. The subsequent screening of a 3D database selected sulphonamide derivatives as MDM2 antagonist. B. Representation of the thiophene sulfonate scaffold selected by LIDAEUS as a candidate to inhibit the p53-MDM2 interaction. R1 and R2 represent phenyl substituents. C. Isoindolinone Nu8231, inhibitor of MDM2-N:p53 interaction identified and using in silico screening for rational design of inhibitors. The figures were adapted from [87] and [103].
1.5.1.2.3. **High-throughput screening campaigns against MDM2**

Some of the major breakthroughs in MDM2-N antagonist discovery to date have emerged from high-throughput screening (HTS) techniques. For example, the most powerful inhibitors known at the time, the Nutlins, were discovered as a result of screening of large libraries of synthetic compounds against MDM2 [78]. These hits were subsequently confirmed by a solution structure solved by NMR [104] and X-ray crystallography. This group of compounds have cis-imidazolines as scaffolds (Fig.1.21) upon which several substituents have been placed so as to project aliphatic and aromatic groups into the MDM2 binding groove, thereby closely mimicking the p53 binding mode. Their potency was established to be in the nanomolar range, and they were shown to have efficacy *in vivo* [78]. An X-ray-derived structure of MDM2 N bound to Nutlin-2, the most potent disruptor of the MDM2-p53 interaction, showed that the imidazoline sits on top of the Trp23 sub-pocket and is solvent exposed, while the 5-chlorophenyl and 4-chlorophenyl substituents penetrated into the Trp23 sub-pocket and into Leu26 sub-pocket, respectively. This leaves the 2-isopropoxy-methoxy phenyl substituent sitting close to the Phe19 sub-pocket but partially solvent-exposed [104]. This moiety also occupies, partially, the Leu22-binding position, thus increasing the affinity as was predicted by Echevarria et al [91]. In general these findings were of broader importance because they successfully demonstrated that small compounds can inhibit protein:protein interactions, and thereby they demonstrated the potential power of organic scaffolds.
Fig. 1.21. MDM2-N in complex with Nutlin-2 [78]. The Nutlins are cis-imidazolines derivates that bind to MDM2 very tightly, providing some of the most efficient known antagonists of MDM2 (left). Their binding closely mimics the pose of p53, in that the para-phenyl derivates are buried in Trp23 and Leu26 sub-pocket, while the more hydrophilic atoms are solvent exposed. This structure was solved by Lukas in 2004 (PDB 1RV1).

Interestingly other HTS-based campaigns succeeded in identifying benzodiazepine derivatives [105] but although the affinity for MDM2-N of these compounds was in the nM-range, they exhibited a low bioavailability and poor cellular absorption.
1.5.1.2.4. De-novo rational design

An independently developed screening methodology involved *de-novo* rational design of molecules with affinity for MDM2-N. Inspired by the approach, described earlier, of using tryptophan derivatives, Ding *et al.* developed a compound based on an oxoindole ring (Fig. 1.22) [106] by modelling and docking of rationally designed compounds. Their findings yielded a compound, MI63, which inhibited p53-MDM2 interaction in the low nM range, but that unfortunately had very poor pharmacokinetic properties [107]. Using MI-63 as a lead, MI-219 was obtained, which inhibited MDM2 at 5 nM but had increased bioavailability [108]. The X-ray crystal structure of MDM2-N in complex with an MI63 analogue showed the indole ring binding in the Trp23 sub-pocket. The fluoroclorophenyl substituent (Fig. 1.22) occupies the Leu26 sub-pocket while the Phe19 sub-pocket is partially occupied by the neopentyl group [80]. This is not in agreement with the structure modelled by Shangary in 2008 of MDM2 with MI-219 [108]. Interestingly, another structure solved by X-ray crystallography of MDM2-N with an imidazole indole scaffold, WK23 (a small molecule similar to the spiro-oxoindole series), shows very similar disposition in the Trp23 sub-pocket, accompanied by a profound reorganization of the Phe19 sub-pocket to compensate for the “empty space” left by WK23 [80].
Fig. 1.22. Indole derivatives that are antagonist of the MDM2-N:p53 interaction. The chemical structures of compounds (right) that share a common indole structure and have affinity for MDM2-N, binding in the same fashion as shown by the surface representations (left) of MDM2-N with MI63-analogue (3LBL) and with WK23 (3LBK). MI-63 is a lead compound for MI-219 currently in pre-clinical phase [109]. Figure derived from [80]
To summarize, MDM2-p53-targetted drug discovery campaigns have provided an excellent platform for development of screening methodologies and assays. This system has provided a proof-of-principle for innovative HTS methodologies and more importantly implementation of structural information into rational design of drugs. Consequently, several compounds are currently in clinical phases, namely, the Nutlins and spiro-oixindole [109].

1.5.2. Antagonists of MDM4-N

As was discussed above, the first structural information for MDM4-N revealed a very similar fold and overall structure to that of MDM2, interacting similarly with p53-derived peptides. However, differences between the binding grooves of MDM2-N and MDM4-N mean that Nutlins do not inhibit the MDM4-p53 interaction to the same degree as they do the MDM2-p53 interaction [25]. This is also true of MI-219 [108]. Initially, the low affinity of MDM4-N for Nutlins was believed to be associated with its narrower Leu26 sub-pocket, resulting mainly from the orientation of α2' and the positioning of the Met53 and Tyr99 side-chains so as to partly occlude this binding pocket [65] (as described earlier). This occlusion was initially regarded as a fixed feature of MDM4, so that while MDM2 exhibits conformational flexibility that adjusts the size of the Leu26 sub-pocket, MDM4 was thought to be more rigid in this respect. Consequently, several groups achieved nM inhibition of the MDM4-p53 interaction by altering residues in a peptide around a Phe19-Trp23-Leu26 triad (Fig. 1.23). More importantly, these studies presented a “closed” conformation brought about by the orientation of Tyr99.

Alanine scanning of potential peptidic ligands proved that while the critical residues for interaction with p53 are Phe19 and Trp23, mutation of Leu26 has only an attenuating effect on the affinity for MDM4 [110]. In addition, phage-display experiments using MDM4 and MDM2 as baits yielded a 12-mer lead peptide that was further optimized against both targets to achieve a low nM $K_D$ for both MDM2 and MDM4 [84]. This dual p53-MDM2/MDM4 inhibitor (PMI) (Fig. 1.22) sits within the binding groove of both MDM2-N and MDM4-N in the same p53-like fashion, and showed how the C-terminus of the peptide has different conformation when bound to MDM2-N or MDM4-N, thus effectively bypassing the structural
differences in the Leu26 sub-pockets of the two proteins. The Pro$^{12}$ of PMI is buried in a different hydrophobic patch, unique to MDM4 [84] (see Fig. 1.23). On the other hand, recent structural findings for a complex of MDM4-N with a p53-peptidomimetic [81] shed doubt on the permanency of the “closed” disposition of MDM4 Tyr$^{99}$ by demonstrating that this residue does after all have a similar intrinsic plasticity as its MDM2 counterpart [80, 81].

![Figure 1.23](image)

**Fig. 1.23. PMI peptide in complex with MDM4-N.** Top - Representation of the PMI peptide sequence. Below - surface representation of MDM4, coloured according to hydrophobicity, in complex with PMI peptide (green), and superimposition on the PMI peptide as it binds to MDM2 (purple). The duality of the binding given by the flexibility of the C-terminal region of the PMI peptide is clearly seen. When bound to MDM4 (green) Pro$^{12}$ makes contact with a unique additional hydrophobic pocket of MDM4 N-terminus (bottom, right). Figure derived from [84].

Thus peptide-based work produced data helping to understand the intricacy of the binding modes of MDM4-N and comparisons with MDM2-N. But the main goal of drug discovery is focused on small-molecule, dual inhibitors. Work performed with
imidazo-indole compounds yielded compound WK298 (Fig. 1.23 B) with an IC$_{50}$ of 20 µM versus the MDM4-p53 interaction [80]. Structural analysis of the interaction by X-ray crystallography showed that the 6-chloro-indole is buried in the Trp23 sub-pocket, while the 4-chlorophenyl ring is in the Leu26 sub-pocket. The Phe19 sub-pocket is occupied by a phenyl ring while the N-N propylamide moiety sits over the Phe19 sub-pocket, shielding several exposed hydrophobic residues (Fig. 1.24 A). Curiously, compound WK23 is very similar to oxoindoles (see Fig. 1.23 B) and structures of the MDM2-N:WK23 and MDM4-N:WK298 complexes are alike (Fig. 1.24 C), suggesting a well-defined scaffold that could be elaborated into a potent dual inhibitor. Compound WK298 has two times more affinity for MDM4 than does WK23. This suggests that the differences in affinity are provided by the different extent of occupancy of the Phe19 sub-pocket by the N-N propylamide. While, as argued above, inhibitors select for MDM2 versus MDM4 according to the shape of the Leu26 sub-pocket, these new studies indicate that successful design of MDM4 inhibitors has to also take into account the Phe19 sub-pocket occupancy [80]. This is in agreement with previously recorded backbone flexibility in the flanking regions of Phe19 sub-pocket β1’-α1’ and α’-β2’ [81, 83].
Fig. 1.24. Small organic compound antagonist of the MDM4-N:p53 interaction. A. Surface representation of MDM4-N in complex with WK298. B. Representation of imidazole-indole derived compounds that are inhibitors of MDM2/MDM4-p53 interactions. C. Detail of a surface representation of MDM4 in complex with WK298 (white) and WK23 (pink). Figure derived from [80].

Other HTS-based work has yielded a powerful inhibitor of the MDM4-N:p53 interaction. The compound SJ-172550 (Fig. 1.25) was identified after an exhaustive HTS which included a primary screening of 300000 compounds against MDM4 by fluorescence polarization assay, followed by cellular tests to confirm the inhibition. The active compounds were clustered by structure and potency, clearly showing phenyl-pyrazole compounds as hits. Ultimately, SJ-172550 was found to have a 4 \( \mu \text{M IC}_{50} \), and is hence the most powerful published small-molecule inhibitor for MDM4 to date. The details of its interaction with MDM4-N have not been elucidated yet since no 3-D representation of the complex has been published. Modelling studies, however, indicated that SJ-172550 does indeed bind to the p53-binding surface on MDM4-N [111].
In conclusion, MDM4 is an attractive therapeutic option in cancer research given the nature of its binding site for p53 and its close relationship with MDM2. The structural similarity between the proteins is favourable for dual inhibition. Recent work has progressed towards rational development of powerful inhibitors whose proven antagonism to MDM2-p53 interactions may be exploited in also developing MDM4-p53 inhibitors. However, relatively little is known about the MDM4-p53 binding pocket and only recently have small-molecule inhibitors emerged to shed more light on the differences between MDM2 and MDM4 and how these can be exploited.

**Fig. 1.25 SJ-172550 compound, MDM4 antagonist.** The chemical structure of compound SJ-172550, the most potent inhibitor of MDM4 NTD recorded up to 2011 [111].
1.6. Aims of the thesis

The overarching aim of this project is to characterize the interaction between the p53 tumour suppressor and its antagonists, MDM2 and MDM4. This will be achieved by using NMR as the primary structural tool.

As part of this aim, this project seeks to explore the utility of the *in silico* screening tool, LIDAEUS, using MDM2-N as a template. The project will also test MDM2’s and MDM4’s druggability potential by screening with the AstraZeneca in-house fragment library.

In addition, NMR will be used to study *apo*-MDM4-N, and MDM4-N in complex with p53-derived peptide or the MDM2 antagonist Nultin-3, to gain insight into the dynamics of these interactions.

In order to achieve this aim, a robust procedure to produce and purify protein has been developed encompassing different methodologies and screening procedures.
Chapter 2: Materials and Methods


All chemicals used in this work were purchased, unless otherwise stated, from Sigma-Aldrich (Gilligham, UK) and Fisher Scientific (Cramlington, UK).

2.1. DNA constructs

2.1.1. The MDM2-N construct

The DNA encoding the N-terminal domain (residues 11-118) of the protein human MDM2 (Q00987) (Fig. 2.1) was provided in the form of recombinant strains of *E. coli* BL21 Star (courtesy of AstraZeneca). The DNA construct was contained in the vector pET28-a (5369 bp) (Novagen), which provides high levels of protein expression under isopropyl β-D-1-thiogalactopyranoside (IPTG) control and kanamycin resistance for selection of clones. At the polycloning site (PCS) unique targets for restriction enzymes are included to allow easy cloning and the coding sequence is flanked by the DNA for hexa-His tags such that the encoded protein can be designed to have affinity tags at either its N-terminus or its C-terminus. In this particular case, the encoding DNA had been codon-optimized for *E. coli* expression, and the construct was designed so as to incorporate a hexa-His tag at the N-terminus of the recombinant protein (Fig. 2.1).

Fig. 2.1. MDM2-N protein sequence. Amino acid sequence for the MDM2-N construct used in this work. Cloning artifacts are highlighted in bold. The red line signals the final construct after removal of the hexa-His tag.
2.1.2. The MDM4 1-134 construct

The DNA for the MDM4 N-terminal region (residues 1-134) was originally cloned into pET28-a as described in Section 2.2. This gave rise to MDM4 residues 1-134 with the hexa-His tag at the N-terminus of the recombinant protein (Fig. 2.2).

![Fig. 2.2. MDM4 construct 1-134. Amino acid sequence of MDM4 1-134 including the cloning artefact (bold). The red line marks the first amino acid residue after hexa-His tag removal.](image)

```
<table>
<thead>
<tr>
<th>1</th>
<th>11</th>
<th>21</th>
<th>31</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGSSHHHHHH SSSLVPGRSH</td>
<td>MTSFSTSAQC</td>
<td>STSDSACRIS</td>
<td>PGQINQVRPK</td>
</tr>
<tr>
<td>41</td>
<td>51</td>
<td>61</td>
<td>71</td>
</tr>
<tr>
<td>GAQGEMFTVK</td>
<td>EVMHYLGQYI</td>
<td>MVQQLDQOE</td>
<td>QHMVYCGGDL</td>
</tr>
<tr>
<td>101</td>
<td>111</td>
<td>121</td>
<td>134</td>
</tr>
<tr>
<td>MLRKNVLTLA</td>
<td>TATTDAAQTL</td>
<td>ALAQPDSMDI</td>
<td>FSQD</td>
</tr>
</tbody>
</table>
```

**Fig. 2.3. MDM4 construct 14-111 DNA sequence.** See legend to Fig. 2.2 for symbols.

2.1.3. The MDM4-N construct

Due to poor stability, the previous MDM4 construct was discarded and a construct derived from published crystallographic studies [112] was used instead. The DNA construct encoding the MDM4 region 14-111 was codon-optimised, synthetically produced and delivered in pET28-a (Novagen) by Geneart.

```
<table>
<thead>
<tr>
<th>14</th>
<th>23</th>
<th>33</th>
<th>43</th>
</tr>
</thead>
<tbody>
<tr>
<td>MGSSHHHHHH SSSLVPGRSH</td>
<td>MDSASRISPG</td>
<td>QINQVRPKLP</td>
<td>LLKILHAAGA</td>
</tr>
<tr>
<td>53</td>
<td>63</td>
<td>73</td>
<td>83</td>
</tr>
<tr>
<td>MHYLGQYIMV</td>
<td>KQLYDQEQOH</td>
<td>MVYCGGDDL</td>
<td>ELLGRQSFSV</td>
</tr>
</tbody>
</table>
```

2.2. Cloning of MDM4 1-134

2.2.1. MDM4 1-134 primers

For cloning of the MDM4 N-terminal region (residues 1-134), the full-length cDNA (O15151) (courtesy of Prof. Hupp) was used as template. Primers were designed to amplify the DNA coding for these amino acid residues as well as incorporating restriction sites for Ndel at the forward primer and XhoI (Fig. 2.4) at the reverse
primer to permit easy incorporation of the construct into the vector pET28-a (Novagen). The reverse primer incorporated a stop codon while the hexa-His tag was fused to the N-terminus of the recombinant protein (Fig. 2.2).

Fig 2.4 Primers for cloning MDM4 1-134. Primers (forwards and reverse) used for cloning MDM4 1-134 incorporated restriction sites for NdeI and XhoI (boxes) used to facilitate the incorporation of the construct into the vector pET28-a. The underlining highlights Met and Stop codons in the forward and the reverse primers, respectively.

2.2.2. Polymerase Chain Reaction (PCR)

PCR was used to amplify the region encoding residues 1-134 of MDM4 N terminus domain. The PCR master mix (50 µL) was set up as follows:

- 10x Pfu buffer (Stratagene)
- 1 unit/reaction Pfu Turbo DNA polymerase (Stratagene)
- 5% (v/v) DMSO
- 0.5 mM PCR nucleotide mix (Promega)
- 0.5 µM primers forward and reverse (Section 2.2.1)
- 100 ng template
- Milli-Q water

The cycling conditions were:

1. 95 °C for 90 s
2. 95 °C for 30 s
3. 54 °C for 30 s
4. 72 °C for 60 s
steps (2) to (4) were repeated 29 times. Then the mixture was incubated at 72 °C for 5 minutes and finally cooled to 4 °C and held at that temperature.

2.2.3. Agarose gels

Agarose gel electrophoresis was used for separating DNA sequences after amplification by PCR or digestion with restriction enzymes. The methodology used to prepare agarose gels was as follows:

The TAE (Tris-acetic-EDTA) buffer contained

\[
\begin{align*}
10 \text{ mM Tris pH 6.8} \\
0.04 \text{ mM Na}_2\text{EDTA} \\
0.08\% \text{ acetic acid (v/v)}
\end{align*}
\]

The gel was prepared by dissolving 0.7% (w/v) agarose powder (Fisher Scientific) in 1xTAE and heated in the microwave to dissolve the agarose powder. Once the solution had cooled down, SYBR-safe DNA gel stain (Invitrogen) was added as the manufacturer recommends and the solution was poured into a gel casting set (BioRad). The gel was let to set at room temperature.

6x DNA-loading buffer composition was

\[
\begin{align*}
0.25\% \text{ bromophenol blue (w/v)} \\
0.25\% \text{ xylene cyanol (w/v)} \\
15\% \text{ (w/v) Ficoll 400 (v/v)}
\end{align*}
\]

For the electrophoresis, the samples were mixed with 6x DNA-loading buffer to a final concentration 1x and the mixture was loaded into the agarose gel, which was run in 1xTAE solution into a BioRad tank at 100 V for one hour. The samples were visualized on a Benchtop 3UV transilluminator (UVP, LLC) and images were recorded using a Polaroid camera.

2.2.4. DNA purification

At various points during the cloning procedure, it was necessary to purify the target DNA product from other DNA, or from enzymes. This was accomplished by gel extraction using the QIAquick kit (Qiagen, according to manufacturer’s conditions).
Alternatively, to stop reactions and remove enzymes, the DNA was ethanol precipitated. For this, 2.5 volumes of 100% (v/v) ethanol were added to the sample and incubated for 40 minutes in the freezer. Subsequently, 0.1 volume of 3 M sodium acetate was added and the solution was spun at 13200 rpm (Microcentrifuge 5415 R Eppendorf), for 30 minutes at 4 °C. The supernatant was discarded; the pellet was re-suspended with 50 µl of 70 % cold ethanol and spun down for > 10 minutes at 4 °C. Finally the supernatant was discarded, the pellet was air-dried and re-suspended in the buffer of choice.

2.2.5. Plasmid amplification and isolation

Amplification of plasmid pET28-a was performed by transformation of Electrocompetent TOP10 E. coli cells (Invitrogen) with pET28-a vector. For this, 1 µl of the plasmid stock (100 ng) was incubated for 5 minutes with 50 µl of electrocompetent cells on ice. Subsequently, the mixture was placed into a pre-cooled electro-cuvette (BioRad) and loaded onto a Gene Pulse II electroporation system (BioRad), and a single 25 µF and 200 Ω and 2.5 kV pulse was applied. The samples were then resuspended in SOC (super optimal catabolite suppressor) media (Invitrogen) without antibiotic and incubated for 1 hour at 37 °C with gentle agitation. Different volumes were plated out onto agar containing 50 µg/ml of kanamycin and incubated in the oven at 37 °C overnight. Subsequently, the plasmid was extracted using the HiSpeed Plasmid Purification Midi Kit (Qiagen) (according to manufacturer’s instructions). The DNA was eluted in Milli-Q water and the concentration was measured using an Eppendorf spectrophotometer.

2.2.6. DNA quantification

DNA was quantified by Epperdorf biophotometer (1-cm path length) by measuring absorbance at a wavelength of 260 nm.
2.2.7. **Restriction-enzyme digestion**

The pET28-a vector and the amplified DNA (encoding MDM4 1-134) were digested with *Nde*I and *Xho*I restriction enzymes (New England BioLabs) to produce complementary ends (“sticky” ends). The double digestion was set up as follows,

- 0.5 units/reaction of both nucleases (*Nde*I and *Xho*I)
- 10x buffer 4 (New England BioLabs)
- 0.1-1 mg/ml Bovine Serum Albumin (BSA)
- target DNA solution (30 µl)
- Milli-Q water to a final volume of 50 µl,

The mixture was incubated for two hours at 37° C with gentle agitation. This was followed by gel extraction of the double-digested DNA (see Section 2.4.4).

2.2.8. **Dephosphorylation of the open vector**

The 5’ phosphates of the open plasmid were removed to prevent re-ligation with itself. For this 30 µl of purified DNA vector was incubated with two units of Calf Intestinal Alkaline Phosphatase (CIAP) (Promega), 10x CIAP buffer (Promega) and Milli-Q water to a final volume of 50 µl for one hour at 37° C. This was followed by ethanol precipitation of the DNA to stop the reaction (see section 2.4.4).

2.2.9. **Ligation**

For the ligation reaction several ratios of vector:insert were used (1:1, 1:3 and 3:3). Each one was incubated with one unit of T4 DNA ligase (Promega), 10x ligation buffers (Promega) and Milli-Q water up to 20 µl, overnight at 16° C.

The end product of the ligation (20 µl) was transformed into TOP10 *E. coli* competent cells by electrocompetent transformation (see section 2.2.5), and streaked onto agar plates that contained 50 µg/ml kanamycin. Several of the colonies were picked in order to amplify and extract the plasmid as described in Section 2.2.5.
2.2.10. Screening for recombinants and sequencing

For sequencing, the putative protein-encoding recombinant DNA construct was amplified by PCR under the following conditions:

- 75-100 ng purified DNA
- 1.6 pmol forward or reverse primer
- Big Dye Terminator v3.1 5x sequencing buffer (Applied Biosystems)
- Milli-Q water to a final volume of 25 µl

The cycling conditions were:

1. 96°C for 1 minute
2. 96°C for 10 s
3. 50°C for 5 s
4. 60°C for 4 minutes

Steps (2) to (4) were repeated 25 times, and cooling to 4°C thereafter. The PCR products were submitted to the sequencing facility at the King’s Buildings campus of Edinburgh University. The results were analyzed using BioEdit version 7.0.0 (http://www.mbio.ncsu.edu/BioEdit/bioedit.html).

2.2.11. Storage of cells and glycerol stocks

In order to make a stock of the recombinant DNA vector pET28 containing the DNA encoding for either MDM2-N, MDM4 1-134 or MDM4-N were transformed into electrocompetent TOP10 E. coli cells (Invitrogen) as described in Section 2.2.5. Subsequently, a single colony was grown overnight in 3 ml LB supplemented with the appropriate amount of kanamycin at 37°C. Aliquots of 750 µl of the overnight culture were mixed with glycerol to 50 % (v/v) in cryotubes and vortexed. Glycerol stocks were stored at -80°C.

Additionally, the DNA vector pET28 containing the DNA encoding for either MDM2-N, MDM4 1-134 or MDM4-N were transformed into electrocompetent expression strains of E coli (DE3) as described in Section 2.2.5 and stored as glycerol stocks.
Prior to production of protein, a small amount of the glycerol stock was picked and streaked on agar plates with antibiotic. The plates were incubated overnight at 37 °C and then stored at 4 °C for future use.

2.3. Cell cultures

2.3.1. Rich media preparation
Different rich medias were used for cellular growth depending on availability. The medias were prepared as follow:

**LB broth**

- 12.5 g of LB powder (Fisher Scientific)
- 500 ml Milli-Q water

sterilized by autoclave and stored at room temperature for further use.

**The LB agar**

- 8 g of agar powder (Fisher Scientific)
- 200 ml Milli-Q water

The media where then sterilized by autoclaving and stored at room temperature for further use. Note that for making agar plates, the solidified agar was melted in the microwave and cooled down to 50 °C in the water bath. Consequently, kanamycin was added to a final concentration of 50 µg/ml and the solution was plated out in Petri dishes and allowed to set. The plates were stored at 4 °C for future use.

Terrific Broth (TB) (Sigma) was prepared as manufacture specifies.

2.3.2. Minimal media preparation
The minimal media for labelled protein expression was prepared as follows:

**5x M9 salts**

- 42 mM Na$_2$HPO$_4$
- 24 mM NaH$_2$PO$_4$
- 9 mM NaCl

Chapter 2: Materials and Methods 45
4 mM \((^{15}\text{NH}_4)_2\text{SO}_4\)

the solution was then autoclaved.

The final media was prepared by mixing the appropriate volume of

1x M9 salt
1 mM Mg\(_2\)SO\(_4\)
2% w/v (\(^{13}\text{C}\)) glucose
50 µM CaCl\(_2\)

Supplements added to increase expression were either

0.5 % (w/v) \(^{15}\text{N}\)-labelled Isogro (Sigma) and 50 µg/ml thiamine and PTM1 salts;

or 0.5 % (w/v) \(^{15}\text{N},^{13}\text{C}\)-labelled Celltone (Cambridge Isotopes Laboratories Inc), 1x vitamins and 1x micronutrient cocktails (see Appendix A.3).

The media was filtered sterilized using 0.45-µm Polyethersulfone (PES) filters prior to use.

### 2.4. Estimates of protein expression

A range of methods was used to estimate and quantify protein expression and concentration and purity. Below is a list of the different techniques used for monitoring protein.

#### 2.4.1. SDS-PAGE

The success of protein purification steps was assessed by sodium dodecyl sulfate-polyacrylamide gel electrophoresis (SDS-PAGE) using NuPAGE Novex, bis-tris mini-gels form Invitrogen according to manufacturer’s recommended conditions. The samples were mixed with an equal volume of 2x Laemmli sample buffer (Sigma-Aldrich) and subsequently loaded onto the gel. Aliquots of 2 µl of protein molecular-weight markers broad-range (BioRad), or SeeBlue Plus2 (Invitrogen) were
included in the first lane of SDS-PAGE gels to estimate the molecular size of the resolved bands.

**2.4.2. Western blotting**

**Towbin-buffer**

25 mM Tris

196 mM glycine

20% (v/v) methanol

**PBST**

1x phosphate-buffered saline (PBS) (Medicago)

0.1% (v/v) tween 20

**Blocking solution**

PBST

5% w/v non-fat dried milk (Marvel)

Western blotting methodology was used to confirm the presence of a His-tag in the MDM4-N construct. For this, the samples were separated by SDS-PAGE and transferred onto a nitrocellulose membrane (BioRad) using a XCell IIBlot Module (Invitrogen) at a constant voltage of 30 V for 1 h in Towbin-buffer. After the transfer the membrane was rinsed 2 times with PBST and incubated with blocking for 1 h with gentle agitation. Subsequently, the primary anti-His antibody (polyclonal, rabbit, Cell Signalling Technology) was added to the blocking solution to a final 1:1000 dilution and incubated overnight at 4 °C with gentle. Afterwards, the membrane was rinsed with PBS and washed twice with PBST for 20 min, and washed with PBS for another 10 min. Incubation with the secondary antibody (goat-anti-rabbit IgG conjugated with horseradish peroxidase for detection via chemiluminescence, Cell Signalling Technology) in a 1:3000 dilution took place for 2 h at 4C in blocking solution. The membrane was then washed as before and detection was performed using the SuperSignal West PICO-Chemiluminescent substrate Trial Kit (Pierce, Thermo Fisher Scientific, Cramlington, UK) which
generates chemiluminescence at 425 nm which could be detected by X-ray film or the SIGMAFAST3,3’-Diaminobenzidine tablets (Sigma-Aldrich), respectively.

2.4.3. Estimates of differential protein expression: the soluble and insoluble fractions

To estimate the protein produced in cellular cultures, 1 ml samples from the cultures prior to and after induction were collected and then spun down for 10 minutes at 4000 rpm (8983 g) and the supernatant discarded. The cellular pellet was re-suspended in 100 µl SDS-sample buffer and incubated for 5 minutes at 90 °C. The mixtures were spun down for 10 minutes at 4000 rpm (8983 g) and 15 µl was loaded onto a SDS-PAGE gel. To estimate the insoluble and soluble proteins expressed in a cellular culture, the cells in 1 ml of the media were spun down for 10 mins at 4000 rpm (8983 g) and the supernatant discarded. The cellular pellet was dissolved in 100 µl of BugBuster (Novagen) and incubated for 30 minutes at room temperature. The solution was then clarified by centrifugation for 20 minutes at 13200 rpm (Microcentrifuge 5415 R Eppendorf). The supernatant containing soluble protein was mixed with 100 µl of 2x Laemmli sample buffer while the pellet (insoluble protein) was re-suspended in 200 µl of 2x Laemmli sample buffer, and incubated for 5 minutes at 90 °C. The mixtures were spin down for 10 minutes at 4000 rpm (8983 g). Finally 15 µl was loaded onto a SDS-PAGE gel and component proteins were resolved by electrophoresis.

2.4.4. Estimation of protein concentration

The concentration of protein in solution was calculated by solving the Beer-Lambert equation, while the estimated protein extinction coefficients (ε) at 280 nm measured in water was obtained from the PropParam Expasy tool (http://expasy.org/tools/protparam.html) as given in Table 2.4 (see also Appendix A.1 for more details). The absorbance at 280 nm wave length was detected using Epperdorf biophotometer (1 cm pathlength) or Nanodrop (1 mm pathlength) (Thermo Scientific) depending on availability.
Table 2.4. **Beer-Lambert equation and extinction coefficients**. \( \varepsilon = \text{extinction coefficients}; l = \text{pathlength}; c = \text{concentration}. \) The values were obtained from the PropParam Expasy tool.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Extinction coefficient (M(^{-1}) cm(^{-1}))</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDM2-N</td>
<td>10430</td>
</tr>
<tr>
<td>MDM4 1-134</td>
<td>7575</td>
</tr>
<tr>
<td>MDM4-N</td>
<td>7450</td>
</tr>
</tbody>
</table>

2.4.5. **Mass spectrometry**

For confirmation of the molecular weight and purity of the sample, mass spectrometry was performed using a DE-STR MALDI-TOF with nitrogen laser (Applied Biosystems). For this purpose, 1 µl samples (100 µM protein) were spotted into a MALDI 96-well sample plate (Applied Biosystems) and mixed with an equivalent volume of sinapinic acid (Sigma Aldrich) prepared according to manufacturer’s recommendations. The sample was analyzed using Voyager DE-STR (Applied Biosystems). The lowest laser intensity that gave clean and well-resolved peaks was used. The instrument was calibrated with external standards including cytochrome c. Alternatively liquid chromatography-mass spectroscopy (LCMS) was performed in AstraZeneca facilities maintained by Mr Andersen using an in-house LCMS 1100 series (Agilent Technologies).

2.5. **Protein expression**

Several protein production methodologies were employed in this thesis.

2.5.1. **Shaking flasks**

For induction of expression, a single colony from an agar plate was picked and incubated in 75 ml of rich media (either LB or TB with 50 µg/ml kanamycin overnight at 37 °C on the shaker (Infors HT) at 160 rpm. Subsequently, this culture was used to inoculate larger volumes of rich or minimal media, with the appropriate
amount of kanamycin, to an initial $OD_{600}$ of 0.1, and incubated at 37 °C with agitation. When the $OD_{600}$ reached 0.3 the temperature was reduced to the temperature to be used for expression, which was specific for each construct. At $OD_{600} = 0.6$, expression of the target protein-encoding DNA was induced with a construct-specific amount of IPTG (see Chapter 3 for details) followed by incubation overnight at the expression temperature.

2.5.2. Fermentation (New Brunswick Bioflow2000)

A starter culture of 200 ml LB containing 50 $\mu$g/ml kanamycin (kan) was inoculated with a single colony and incubated overnight at 37 °C on the shaker. In parallel, 4 L of appropriate growth media were prepared, poured into the fermentor vessel and autoclaved. The dissolved oxygen (DO) probe was charged with the media overnight. Next day, kan was added to the solution to a final concentration of 50 $\mu$g/ml.

The cells of the starter culture were harvested by centrifugation at 4000 rpm (8983 g) at 4 °C for 15 minutes. The pellet was then re-suspended with 50 ml of growth media (containing 50 $\mu$g/ml kanamycin) and this was used to inoculate the 4 L culture in the fermentor. The mixture was incubated at 37 °C until it reached $OD_{600} = 0.3$ when the temperature was reduced to the temperature to be used for expression, which was specific for each construct. At $OD_{600} = 0.6$, expression of the target protein-encoding DNA was induced with a construct-specific amount of IPTG followed by incubation overnight at the expression temperature.

2.6. Protein purification

Following protein production, cells were harvested by centrifugation at 4000 rpm (8983 g) in Avanti J-26 XP centrifuge (Beckman Coulter) for 30 minutes. The supernatant was discarded and the pellet freeze-stored for at least 1 hour.

2.6.1. Lysis of the cells

The cellular paste was lysed in three different ways, depending on equipment availability or size of the cellular pellet. The lysis buffer was as follow

$40 \text{ mM HEPES pH 8}$
0.3 M NaCl
1 mM tris (2-carboxyethyl)phosphine (TCEP) or dithiothreitol (DTT) depending on availability
20 mM imidazole
1 mg/ml lysozyme
2.5 units/ml benzonase

Lysis by sonication: the frozen pellet was re-suspended in 10 ml of lysis buffer per 1 g of pellet. Subsequently, samples were subjected to four cycles of 40-second bursts of sonication (soniprep 150 MSE) on ice.

BugBuster (Novagen): alternatively, the cells were lysed by using the Bugbuster solution (according to manufacturer’s conditions).

Cell disruptor: the cell paste was re-suspended in 10 ml/g of pellet in lysis buffer and passed twice through a Z2 Plus (Constant Systems Ltd) cell disruptor as recommended by the manufacturer.

After lysis by any of the above methods, the lysates were clarified by centrifugation at high speed (16,000 rpm) (30966g) in Avanti j-26 XPI centrifuge (Beckman Coulter) for 45 minutes, followed by filtration through a 0.22-µm PES filter.

2.6.2. Purification of proteins from soluble fractions

Protein purification involved three chromatographic steps. The buffers used in the purification were as follows:

Buffer A

50 mM Hepes pH 8.5
300 mM NaCl
1 mM TCEP/DTT
20 mM imidazole,
(10% (v/v) glycerol when purifying MDM4-N)

Buffer B
50 mM Hepes pH 8.5
300 mM NaCl
1 mM TCEP/DTT
250 mM imidazole for MDM2-N purification or 500 mM for MDM4 1-134/MDM4-N purification

(10% (v/v) glycerol when purifying MDM4-N)

The supernatant from the previous step was sterile-filtered and loaded into a pre-packed Ni\(^{2+}\) NTA or 25 mL Ni\(^{2+}\) NTA agarose (Qiagen) column packed in an XK 16/20 (GE healthcare) column, equilibrated with buffer A, using a low-pressure peristaltic pump at 4 °C or AKTA FPLC system (GE Healthcare) or P-960 pump (GE Healthcare) depending on availability. A protocol was designed using the software UNICORN version 4.1 (GE Healthcare) that included a 10 column volumes (CV) wash step and a gradient step from buffer A to buffer B (from 10 mM to 250 mM imidazole) of 20 CV for protein elution. The elution was collected in fractions of different volumes depending on the size of the column used, with a Frac-900 (GE Healthcare) fraction collector. The column was washed with buffer B for 10 CV and a finally re-equilibrated to Buffer A. The elution was recorded by UPC900 UV unit monitoring the 280 nm trace providing a chromatogram of the run. The peak-corresponding fractions were analyzed by sodium dodecyl-sulfate polyacrylamide gel electrophoresis (SDS-PAGE) on NuPAGE Novex (see section 2.4.1).

Fractions identified as being likely to contain the target protein were pooled, and buffer-exchanged to reduce the concentration of imidazole. The pooled sample was then incubated with 1 unit/mg protein of biotinylated thrombin (Novagen) according to manufacturer’s recommended conditions, overnight at 4 °C. The thrombin was recaptured with 16 µl streptavidin-coated beads (Novagen) per unit of thrombin by incubation for 1 h at room temperature with agitation. The beads were pelleted by centrifugation.

In an alternative procedure, thrombin (GE Healthcare) was added to the pooled fractions as the manufacturer recommends. The thrombin cleavage mixture was then
loaded into a 8000-Da molecular weight cut-off (MWCO) membrane (SpectrumLabs) and dialyzed overnight against 4 L of buffer A at 4 °C.

Following hexaHis-tag cleavage, a re-pass step was performed on the Ni$^{2+}$-affinity column (Qiagen). To accomplish this, the sample was loaded onto the column on an AKTA FPLC (GE Healthcare) using a 5-ml loop, or directly loaded using P-960 (GE Healthcare). The protocol used was as described above, except that the flow-through was collected. The results were analyzed by SDS-PAGE (see Section 2.4.1).

A final gel-filtration step was performed as follows.

Gel filtration buffer

50 mM HEPES pH 7.5
0.1 M NaCl
2 mM TCEP/DTT
0.5 mM EDTA

(10% (v/v) glycerol when purifying MDM4-N)

A Superdex 75 16/60 or 26/60 S200 column (GE Healthcare), on an AKTA FPLC, was pre-equilibrated with gel filtration buffer. The samples were loaded using a 5 mL loop. Five-ml fractions were eluted over 1.2 CV. The peak-containing fractions were analyzed by SDS-PAGE (see Section 2.4.1).

2.7. Insoluble fractions: refolding protocol

The refolding methodology here described was adapted from [113]. The different buffers compositions were as follows:

Guanidinium buffer

6 M guanidine HCl pH 8.0
20 mM Tris HCl
10 mM reduced glutathione

Refolding buffer

50 mM Tris-HCl pH 8.5
10.56 mM NaCl
2.2 mM CaCl$_2$
2.2 mM MgCl$_2$
0.055% (w/v) PEG 4000
0.55 M arginine
0.1 mM oxidised glutathione
1 mM reduced glutathione

The pellet obtained from the clarification of cell lysates containing insoluble material was re-suspended into 20 ml guanidinium buffer per g of pellet and incubated overnight, with gentle rocking, at 4 °C. Afterwards, the re-dissolved samples were clarified by centrifugation at 16000 rpm (30966 g) in an Avanti j-26 XPI centrifuge (Beckman Coulter) centrifuge for 45 minutes. The supernatant was filter-sterilized and loaded onto a 5-ml His-Trap column attached to an AKTA FPLC (GE Healthcare), using a peristaltic pump, at 4 °C. Protein was eluted using in a single step with 700 mM imidazole, collecting 5 ml fractions. The peak-containing fractions were pooled and diluted 10x with refolding buffer and incubated overnight at 4 °C. Low-molecular weight contaminants were removed by centrifugation at 4,000 rpm and the pooled samples were then dialyzed (at 4 °C) in 8000 Da molecular weight cut-off (MWCO) membrane (SpectrumLabs) four times against 4 L of 50 mM Tris-HCl, pH 8.5, 1 mM DTT, with each dialysis step longer than four hours. Precipitated protein were removed by centrifugation at 4,000 rpm (3724 g) in Allegra X-15R Benchtop Centrifuge (Beckman Coulter) for 30 minutes and sterile-filtered.

2.8. Buffer screening assays

Two different biochemical assays were used to help discern the optimal NMR buffer conditions for the MDM4-N construct.

2.8.1. Thermofluor stability assay

A Thermofluor fluorescence-based stability assay was used to screen for optimal NMR buffer conditions. For this, a mixture of 0.2 mg/ml protein in the buffer of
choice (Table 1), and a 1:1000 dilution of Sypro orange dye (Invitrogen) were added in 25-µL aliquots into a 96 well PCR plate (Bio-Rad). The plate was heated using an iCycler iQ instrument (Bio-Rad), with a scan rate of 1 °C/ min from 20 to 90 °C. Data were collected at 0.2 °C increments (Section 5.2.2) or 1 °C (Section 4.2.4), and analysed using GraphPad Prism 4 (GraphPad software).

2.8.2. Dynamic light scattering

Dynamic laser light-scattering measurements were performed using a Wyatt DynaPro Titan plate-reader instrument. For this procedure, MDM4-N protein samples were prepared at both 50 and 100 µM in a range of buffers. Samples were filtered, centrifuged and aliquoted into 384 well plates (Thermo Fisher Scientific). Measurements were made using a laser power of 10%, for ten acquisitions, each of 10 s duration. Data were analysed using the DYNAMICS V6 software.

2.9. Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) data were collected using a Microcal ITC 200 (GE Healthcare) fitted with an autosampler. The titration data were recorded at 25 °C in 50 mM HEPES, pH 7.5, 100 mM NaCl, 2 mM TCEP, 10% (v/v) glycerol, 0.02% (w/v) NaN₃, 2% (v/v) DMSO. Aliquots containing 200 µM ligand stocks in buffer were added to 20 µM MDM4-N in multiple 0.8 µl steps (45 steps in total). Data were analysed using Origin (Microcal).

2.10. Peptide synthesis

The p53-derived peptide (17ETFSDLWKLLP27) was chemically synthesized using solid-phase methodology [114] with the assistance of Dr Stefano Sabatini (Università degli Studi di Perugia). For this procedure the resin, consisting of divinylbenzene cross-linked polystyrene beads with polylinker arm, was activated and Fmoc-Rink linker was coupled and de-protected exposing the amine group. To confirm the success of the reaction, a small amount of de-protected resin was mixed with ninhydrin, which reacts with primary amino groups thus giving a coloured product when positive [115]. Subsequently, amino acids were added sequentially by
debloking-activation and coupling reactions. For each addition (except proline), the reaction was checked with the ninhydrin test. Finally, the peptide was decoupled from the resin while simultaneously deprotecting amino-acid side chains using a piperidine-based protocol [116]. A detailed description of the protocol is given in Appendix A.4.

The resultant product was freeze-dried and dissolved in 10:90 acetonitrile:H₂O (v:v) buffer prior to purification by high-performance liquid chromatography (HPLC) as follows. A Phenomenex C18 column (Prodigy) attached to Agilent 1100 (Agilent Technologies) was pre-equilibrated with buffer C (H₂O, 0.1% v/v trifluoroacetic acid (TFA)) and the peptide was eluted using a gradient from 100% buffer C to 100% buffer D (acetonitrile, 0.1% TFA) 3 ml/min flow rate, monitoring absorbance at 230, 250 and 280 nm. The results were analyzed by Agilent ChemStation software (Agilent Technologies).

### 2.11. Nuclear Magnetic Resonance (NMR) experiments

#### 2.11.1. NMR toolkit.

NMR offers more than a static, freeze-frame picture of a protein structure. In contrast to crystallography, NMR data can normally be collected on samples that are in a physiological-like environment. This results in more reliable information pertaining to molecular events in a biological context. In addition, drug discovery exploits NMR to determine structural information and to provide direct and far-reaching insights into protein-ligand interaction. Such structural information may be usefully deployed in rational drug design, which has a good record of success in terms of hitting the target while minimizing side effects [117].

The NMR experiments described in this thesis were performed using Avance 600 or Avance II 800 NMR spectrometers (Bruker BioSpin) operating at 14.1 or 18.8 Tesla, respectively. The NMR pulse sequences used for screening of small-compound libraries against protein targets, and for measurements of relaxation parameters, had previously been developed by Dr Dušan Uhrín (Edinburgh University) and were performed with the technical assistance of Mr Juraj Bella (Edinburgh University). The NMR pulse sequences used for fragment-based screening, for MDM2-N re-
assignment, and for MDM4-N assignment, were all developed and implemented within the AstraZeneca NMR group at Alderley Park.

The recorded NMR data in the form of free-induction decays (FIDs) were Fourier transformed to obtain NMR spectra using the Topspin interface (Bruker BioSpin) and then imported into the CARA software suite [118] for assignment. Alternatively, the AZARA 2.7 suite (Wayne Boucher, Department of Biochemistry, University of Cambridge, UK) or NMRPIPE [119] scripts were used to process the data that were then imported into the CCPNMR software suite Analysis 2.1.1.

2.11.2. Preparation of NMR samples

For NMR, single-labelled or double-labelled purified protein samples were adjusted to a final concentration ranging from 50-500 µM, depending upon the experiment required and the stability of the construct, using 5,000-10,000 MWCO Centricolumns (Vivaspin). Each protein sample was then buffer-exchanged into the appropriate NMR buffer.

NMR buffer 1 (MDM2-N)

50 mM phosphate pH 6.8
0.1 mM EDTA
2 mM TCEP/DTT
0.02% (w/v) NaN₃.

NMR buffer 2 (MDM4 1-134)

50 mM phosphate pH 6.8
0.1 mM EDTA
2 mM DTT
0.02% (w/v) NaN₃

NMR buffer 3 (MDM4-N)

50 mM BisTris pH 6.5
100 mM NaCl
2 mM TCEP

0.1 mM EDTA

0.1% (w/v) NaN₃

For NMR data collection it is essential to have a deuterium lock signal available in order to correct for inhomogeneity in the magnetic field. Either 10% (v/v) deuterated H₂O (D₂O) or 1% (v/v) deuterated DMSO was included in the final solution to act as a lock solvent. The final sample was transferred to a 5 mm NMR tube (Wilmad High Precision) avoiding bubble formation. All the NMR experiments were recorded at 300 K.

2.11.3. Screening by NMR: Technical concepts

For the purpose of this thesis a protein-based NMR strategy for ligand discovery was adopted. The first step was assign the [¹H,¹⁵N] heteronuclear single quantum correlation (HSQC) spectrum (see below). Chemical shift perturbations could then readily be used for approximate mapping of the ligand-binding site on the protein to specific residues (which was useful for delineating the interaction surface on a 3-D structure). For this purpose, the labelled protein was titrated with non-labelled ligand and the chemical shifts of the residues are analyzed. In the event of weak binding, recording of a series of HSQC spectra showed migration of cross peaks corresponding to amino acid residues involved in the binding. For strong binders, relevant cross-peaks (of the apo-protein) disappeared and new cross-peaks (protein plus ligand) appeared in a different place in the spectrum. Intermediate binders resulted in spectral broadening that can be quite hard to interpret [120].

The ratio of complex to free protein could often be quantified by chemical shift changes [121]. As stated above, slow exchange (relative to the NMR time-scale) manifests as two resonances for the exchanging residue, one for the apo-form and another for the bound form, with different intensities depending on the saturation. Thus slow-exchanges regimes are theoretically analysable by comparing difference in intensities, but signal-to-noise ratios become critical, making it difficult to correctly extrapolate the \( K_D \). Fast-exchange regimes (weak binding) on the other hand, are easily calculated from NMR data. A residue in fast exchange will have a
unique resonance that is an average of the chemical environments between which it is exchanging. Within fast exchange there are two sub-regimes to be considered. In the very-fast exchange the resonances observed are the result of the average chemical shift and average $T_2$ relaxation rates. In fast exchange, relaxation is a separate event; the ratio of relaxation rates is entirely separate from that of chemical shift perturbations, and hence should be included in $K_D$ calculations. The best $K_D$ range for NMR analysis is in the region of $\mu$M to mM. Note that since the protein concentration has to be smaller than the $K_D$ values, $K_{D8}$ in the nM range cannot be accurately measured by NMR [120, 122]. Note also that compound solubility may impose a maximum value on the $K_D$ that can be measured.

2.11.4. NMR experiments for screening

In the screening process two principal experiments were employed, namely variations on $^1$H NMR (1D), and [$^1$H, $^{15}$N] heteronuclear single quantum coherence HSQC [123] or transverse relaxation optimized spectroscopy (TROSY) [124] (both 2D experiments).

Prior to recording an initial 1D $^1$H-NMR spectrum, the sample was locked to the corresponding deuterium signal and the probe was shimmed. For $H_2O$ suppression, the “presat” routine [125] as improved by the Shaka protocol [126], was utilised.

The parameters necessary to record the NMR experiment were optimised while recording 1D spectra and later used within HSQC or TROSY pulse sequences. Consequently, the 90-degree pulse width was optimized for every sample individually since its length is determined by the specific composition of the solution in terms of salt concentration and solvent. Usually, MDM4-N samples had larger pulse widths (13 $\mu$s) for buffer containing 100 mM NaCl, than MDM2-N (11 $\mu$s).

The offset frequency (i.e. the middle of the spectrum) and the spectral width for the indirectly detected dimension were adjusted to delimit the scan area of the experiment, aiming to increase the resolution and decrease the acquisition time. These parameters are also protein dependent, and resonances outside the spectral width defined in the experiment, are “folded” or aliased (i.e. normally negative peaks in the spectra). The number of scans was adjusted to obtain an adequate signal-to-noise ratio depending on the amount of protein in solution (Table 2.5). A few dummy
scans were employed to negate the possible effects of initial sample heating induced by the pulses applied. Once the acquisition had finished, the raw data were processed by Fourier transforms.

<table>
<thead>
<tr>
<th>Construct</th>
<th>Parameter</th>
<th>90-degree pulse width (ms)</th>
<th>Offset frequency (Hz)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDM2-N HSQC (600 MHz)</td>
<td></td>
<td>11.0</td>
<td>7081.96</td>
</tr>
<tr>
<td>MDM2-N TROSY (600 MHz)</td>
<td></td>
<td>12.0</td>
<td>7200.5</td>
</tr>
<tr>
<td>MDM4-N HSQC (600 MHz)</td>
<td></td>
<td>13.1</td>
<td>7049.92</td>
</tr>
<tr>
<td>MDM4-N TROSY (800 MHz)</td>
<td></td>
<td>12.1</td>
<td>9395.35</td>
</tr>
</tbody>
</table>

Table 2.5. Some optimized parameters for 1-D experiments. Note that experiments for MDM4-N at 800 MHz were assisted by an attendant.

2.11.5. $K_D$ calculation by NMR

The recorded spectra were processed by Azara and loaded into CCPNMR-analysis to assess chemical shift perturbations (CSP). The MDM2-N $^1$H,$^{15}$N HSQC spectrum had been previously assigned using peak list 6612 from the BMRB (supplemented with a $^{15}$N NOESY (see Chapter 5 for more details). In order to select the appropriate peaks for $K_D$ estimations, the fast-exchange regime was preferred. Several observation where made in this respect:

i. For each resonance, the volume and line width of each peak to be included in the calculation was examined since they should be within ±10% and ±10 Hz, respectively. Changes in either of these parameters would imply that the nuclei concerned display different relaxation rates during the titration, and this complicates the calculation.
ii. The peak movement should be in a straight line since changes in direction are indicative of multiple events being sensed by the amide i.e. binding of the molecule in several places.

Selected residues were fitted to (2.1) using CCPNMR analysis and plotted against ligand concentration yielding values for A (2.2) and B (2.3). Subsequently, the $K_D$ values were calculated on a residue-by-residues basis by solving (2.3) [120].

\[ y = A \left( (B + x) - \sqrt{(B + x)^2 - 4x} \right) \]  
\[ (2.1) \]
\[ A = \frac{\text{Max shift change}}{2} \]  
\[ (2.2) \]
\[ B = 1 + \frac{K_d}{[P]} \]  
\[ (2.3) \]
Table 2.6. KD by NMR: A1 fragment. The graph represents chemical shift perturbation (ppm) of individual residues versus ligand concentration (mM). The figure was obtained using GraphPad Prism 4 (GraphPad software).

2.11.6. NMR experiments for assignment

2.11.6.1. Introduction to 3D experiments

This section will briefly review the assignment routines used in this thesis. The backbone assignment of MDM4-N in complex with p53 peptide was used to assist the assignment of apo-MDM4-N and MDM4-N in complex with Nutlin-3 as detailed in Chapter 5. Conversely, MDM2-N had been previously assigned and therefore the
experiment used was intended to merely re-assign the HSQC spectrum. This exercise was assisted by the side-chain assignment published by Uhrinova et al. in 2005 as described in Chapter 5.

As a prelude to describing the NMR experiments utilised in this thesis for assignment purposes, several key concepts warrant brief discussion. The NMR sample is placed in a strong, homogeneous, magnetic field ($B_1$). At the beginning of an NMR experiment, nuclei are irradiated by a short pulse of radiofrequency (rf) that matches to their Larmor frequencies. In the case of proteins, $^1$H, $^{13}$C and $^{15}$N are the main nuclei of interest. The rf pulse excites these nuclear spins to a higher energy level. After the rf pulse, the spins revert to equilibrium by various well-understood relaxation mechanisms. The accompanying loss of magnetization over time in the plane perpendicular to $B_1$ is detected by the receiver coil as a “free-induction decay” (or FID) and recorded for further manipulations. Multiple FIDs are routinely summed to increase signal-to-noise. Fourier transformation of the FID (intensity versus time) yields the NMR spectrum (intensity versus frequency, or chemical shift). Extra pulses (making up a pulse sequence) can be incorporated in the experiment. These allow magnetization to be transferred between spins in highly defined ways, through space or through bonds; thus these NMR experiments establish relationships between different nuclei in the protein sequence. A particularly useful pulse sequence in biomolecular NMR is the $^1$H,$^{15}$N-HSQC experiment. In this double resonance experiment, magnetization is transferred between $^{15}$N and a directly attached (through one bond) $^1$H such as occurs within a $^{15}$N-$^1$H (amide) bond of a polypeptide backbone. The experiment is normally recorded in a 2D-format wherein each peak in the 2D spectrum will have the coordinates of the $^{15}$N (y-axis) and the attached proton (x-axis) (Fig. 2.7).
Fig. 2.7. \([^{1}\text{H},^{15}\text{N}]\text{HSQC}\). The red arrows represent the magnetization transfer in the HSQC experiment between \(^{15}\text{N}\) and \(^{1}\text{H}\) (left). The Fourier transform of the FID results in a 2D spectrum, with \(^{15}\text{N}\) (y-axis) and \(^{1}\text{H}\) (x-axis) frequencies. The coordinates of the resonances are then determined by the chemical shift of the N-H pair (green).

The information collected in the HSQC is alone insufficient to assign the spectrum \textit{i.e.} identify which peak corresponds to which amino acid residue. To accomplish this, a third dimension is commonly introduced that permits inter-residue connectivity to be established. For these kinds of experiment, magnetization is selectively transferred via the \(\alpha^{13}\text{C}\) or carbonyl (\(^{13}\text{CO}\)) carbon. Such triple-resonance experiments are generally plotted as a series of planes (each similar to an HSQC) to form a 3D spectrum with (by convention) \(^{15}\text{N}\) in the F1 dimension, \(^{1}\text{H}\) in the F2 dimension and \(^{13}\text{C}\) in the F3 dimension. Further analysis is subsequently performed by extracting cross-peak-containing F2-F3 slices from each plane (\textit{i.e.} matching to a particular \(^{15}\text{N}\) chemical shift) (Fig. 2.8); these are compared to the equivalent regions of other experiments (involving different pulse sequences) (see Fig. 2.8) such as the HNCA so as to find common peaks, and hence organise the slices according to their sequence-positions.
Fig. 2.8. A 3D NMR experiment. The red arrows show the magnetization transfer in $[^{15}\text{N},^{1}\text{H}]$ HSQC and HNCA experiments (top). The spectrum (bottom) is the result of the resonances collected for correlated $^{15}\text{N}$, $^{1}\text{H}_N$ (green) and $^{13}\text{C}_i$ (light blue) and $^{13}\text{C}_{i-1}$ (dark blue) as dictated by the experiment. The 3D spectrum is sliced at the F1 dimension yielding a strip with information on $^{1}\text{H}_N$ and $^{13}\text{C}$ cross-correlated and their connectivity.

2.11.6.2. Backbone assignment of MDM4-N

Sequential assignment of $^{13}\text{C},^{15}\text{N}$-labelled MDM4-N nuclei was performed on the basis of a suite of triple-resonances NMR experiments. The TROSY-type variants [127] of these experiments were employed in the current work that was conducted on a Bruker Avance II 800. In TROSY experiments, two different relaxation pathways cancel each other out, and these pulse sequences were found to improve line-widths and minimise overlap between peaks compared to their non-TROSY equivalents. In an elaboration of the HNCO and HNCA experiments summarised in Table 2.10, key experiments correlated $^{15}\text{N}$-$^{1}\text{H}$ with $^{13}\text{C}_\beta$, as well as $^{13}\text{C}_\alpha$ and $^{13}\text{C}_O$, of $i$ and $i-1$ residue in a complementary manner. For example, from HNCACB, the amide $^{15}\text{N}$
and $^1$H shifts were correlated with $^{13}$Cαs and $^{13}$Cβs within the same residues ($i$) as well as the previous ones ($i$-1); on the other hand HNCA only yielded connectivity for Cα resonances (Fig. 2.9). Thus by comparing these two experiments, the NH chemical shifts of each (non-glycine, non-proline) residue were successfully linked both to its own Cα and Cβ and to the Cα and Cβ of the preceding residue thus allowing sequential assignment. It was found that runs of sequential assignments could be initiated by identifying spin systems of residues with unusual chemical shift patterns such as Gly (lacking Cβ and having upfield $^{15}$N chemical shifts) or Thr/Ser (Cβ downfield resonance values) and that these could be joined together through further inspection of the spectra.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Nuclei Correlation</th>
<th>Connectivity</th>
</tr>
</thead>
<tbody>
<tr>
<td>HNCA</td>
<td>NH-Cα</td>
<td>$i, i$-1</td>
</tr>
<tr>
<td>HNCACB</td>
<td>NH-Cα-Cβ</td>
<td>$i, i$-1</td>
</tr>
</tbody>
</table>

**Fig. 2.9. Schematic representation of sequential assignment.** Top: table illustrates the connectivity of HNCA and HNCACB experiments used for backbone assignment. Bottom: schematic representation of experiments HNCA and HNCACB providing the bases for sequential assignment by connecting cross-peaks between $i$ and $i$-1 residues for Cα (in red) and Cβ (in green). The resonances are residue specific. Gly does not have Cβ while Thr/Ser Cβ chemical shift is above Cα values (bottom).
2.11.6.3. $^{15}$N NOESY of MDM2-N

The $^{15}$N-NOESY experiment exploits the nuclear Overhauser effect of protons by transferring magnetization though space to neighbouring protons (Fig. 2.10) [128, 129]. This 3D spectrum is defined by $^{15}$N in the F1 dimension, $^1$H$_N$ in the F2 dimension and $^1$H in the F3 dimension (Table 2.11 for details). Each strip will then contain NOEs from all hydrogen atoms less than ~5.5 Å in space from the observed NH proton (residue $i$), and the strength of the NOE will depend on the distance between the two protons (Fig. 2.10). By detecting both intraresidue and $i$-$(i+1)$ signals, the NOESY can help in sequential assignment even in the absence of $^{13}$C-spectra. In addition, residues within elements of secondary structure may display backbone-backbone/side-chain NOE cross-peak to residues that are distant in the primary sequence. The re-assignment of MDM2-N could thus be performed reliably on a $^{15}$N-labelled sample by a combination of the HSQC, NOESY and the previous assignment published by Uhrinova et al.
Fig. 2.10. $^{15}$N NOESY experiment. Representation of the magnetization pathway in the NOESY experiment (top). Cartoon representing the sequential assignment assisted by the NOESY where the NOEs cross-peaks were used to prove connectivity between consecutive residues (bottom).
2.11.7. Details of NMR experiments

The detail of the NMR experiments used in this thesis are listed below in Table 2.10, with corresponding details on experiment connectivity and magnetization transfer as well as practical details on numbers of scans (ns) or sweep width (sw) in all dimensions.

<table>
<thead>
<tr>
<th>Experiment</th>
<th>Construct</th>
<th>Mag Trans</th>
<th>Correlation</th>
<th>Connectivity</th>
<th>Number</th>
<th>Nu</th>
<th>Np</th>
<th>Sw (Hz)</th>
<th>Nu</th>
<th>Np</th>
<th>Sw (Hz)</th>
<th>Mix T (ms)</th>
</tr>
</thead>
<tbody>
<tr>
<td>((^1)H,(^1)N) HSQC</td>
<td>MDM2-N</td>
<td>Tb</td>
<td>NH</td>
<td>i</td>
<td>8</td>
<td>(^1)H</td>
<td>2048</td>
<td>8389</td>
<td>(^1)N</td>
<td>1024</td>
<td>7002</td>
<td></td>
</tr>
<tr>
<td>((^1)H,(^1)N) HSQC</td>
<td>MDM4-N &amp; (1-134)</td>
<td>Tb</td>
<td>NH</td>
<td>i</td>
<td>8</td>
<td>(^1)H</td>
<td>2048</td>
<td>8389</td>
<td>(^1)N</td>
<td>1024</td>
<td>7002</td>
<td></td>
</tr>
<tr>
<td>((^1)H,(^1)N) TROSY</td>
<td>MDM2-N</td>
<td>Tb</td>
<td>NH</td>
<td>i</td>
<td>8</td>
<td>(^1)H</td>
<td>2048</td>
<td>14367</td>
<td>(^1)N</td>
<td>64</td>
<td>1824</td>
<td></td>
</tr>
<tr>
<td>((^1)H,(^1)N) TROSY</td>
<td>MDM4-N</td>
<td>Tb</td>
<td>NH</td>
<td>i</td>
<td>8</td>
<td>(^1)H</td>
<td>2048</td>
<td>14367</td>
<td>(^1)N</td>
<td>64</td>
<td>1824</td>
<td></td>
</tr>
<tr>
<td>HNCA</td>
<td>MDM4-N</td>
<td>Tb</td>
<td>NH,Ca</td>
<td>(i, j, k)</td>
<td>32</td>
<td>(^1)H</td>
<td>2048</td>
<td>14367 (^1)C</td>
<td>72</td>
<td>8048</td>
<td>(^1)N</td>
<td>64</td>
</tr>
<tr>
<td>HNCA(C)</td>
<td>MDM4-N</td>
<td>Tb</td>
<td>NH,Ca &amp; Cb</td>
<td>(i, j, k)</td>
<td>112</td>
<td>(^1)H</td>
<td>2048</td>
<td>14367 (^1)C</td>
<td>72</td>
<td>18864</td>
<td>(^1)N</td>
<td>64</td>
</tr>
<tr>
<td>HINC(C)</td>
<td>MDM4-N</td>
<td>Tb</td>
<td>NH-Ca &amp; Cb</td>
<td>(i, j, k)</td>
<td>64</td>
<td>(^1)H</td>
<td>2048</td>
<td>14367 (^1)C</td>
<td>72</td>
<td>18864</td>
<td>(^1)N</td>
<td>64</td>
</tr>
<tr>
<td>HINCO</td>
<td>MDM4-N</td>
<td>Tb</td>
<td>NH-CO</td>
<td>(i, j, k)</td>
<td>16</td>
<td>(^1)H</td>
<td>2048</td>
<td>14367 (^1)C</td>
<td>72</td>
<td>5534</td>
<td>(^1)N</td>
<td>64</td>
</tr>
<tr>
<td>HINCA(C)</td>
<td>MDM4-N</td>
<td>Tb</td>
<td>NH-CO</td>
<td>(i, j, k)</td>
<td>16</td>
<td>(^1)H</td>
<td>2048</td>
<td>14367 (^1)C</td>
<td>72</td>
<td>5534</td>
<td>(^1)N</td>
<td>64</td>
</tr>
<tr>
<td>((^1)H,(^1)N) HSQC NOESY</td>
<td>MDM2-N</td>
<td>Ts</td>
<td>NH</td>
<td>(i, j, k)</td>
<td>2048</td>
<td>(^1)H</td>
<td>2048</td>
<td>8389</td>
<td>(^1)H</td>
<td>264</td>
<td>7002</td>
<td></td>
</tr>
</tbody>
</table>


2.11.8. NMR relaxation studies of MDM4

As mentioned before, NMR experiments are based on irradiation with specific rf pulses. Once the pulse stops, the system goes back to equilibrium by two routes known as longitudinal and transverse relaxation.

Longitudinal relaxation (return to equilibrium along the z axis (Fig. 2.12)) involves energy exchange between the excited spins and the environment, and is an exponential process described by a time constant, \(T_1\) (Fig. 2.12). Conversely, Transverse relaxation does not involve exchange with the environment, but the
relaxation is subjected to effects between neighbouring spins (CSA and dipole-dipole interactions). The exponential magnetization decay in this plane (x,y) is described by a time constant $T_2$ (Fig. 2.12) [130]. Measurement of these relaxation time constants for backbone amide $^{15}$Ns can be used (along with the strengths of the NOEs between $^{15}$N and $^1$H in backbone amides) to derive the correlation times and amplitudes of the internal motions of the protein, providing a wider understanding of the flexibility of the proteins in solution. In this work, such a “ModelFree” analysis [131] was not performed due to data quality but the raw $T_1$, $T_2$ and heteronuclear NOE values are still useful for inferring backbone dynamics.

Fig. 2.12. Relaxation diagrams. A. Diagram depicting the bulk magnetization (M) tilting into the xy plane after a 90° pulse. B. Longitudinal relaxation time ($T_1$) is defined by the time it takes to the magnetization in the Mz plane to go back to its original value. While transverse relaxation time ($T_2$) is governed by the time it takes to the magnetization in the Mxy plane to go back to cero. Adapted from [130]

For MDM4-N relaxation studies, datasets were collected for 100 μM $^{15}$N-labelled apo-MDM4-N, MDM4-N:Nutlin-3 and MDM4-N:p53 at 800 MHz. The relaxation experiments pulse sequences were developed by Dr Dusan Uhrin (School of
Chemistry, Edinburgh University) and were recorded with the assistance of Mr Juraj Bella (School of Chemistry, University of Edinburgh).

The relaxation times were sampled at the following points:

\[
T_1 \text{ (ms): } 51.2, 301.2, 501.2, 701.2, 901.2, 1001.2, 1101.2 \text{ and } 51.2.
\]

\[
T_2 \text{ (ms): } 16.96, 33.92, 50.88, 67.84, 84.8, 101.76, 118.72, \text{ and } 16.96.
\]

The resulting spectra were processed within the program AZARA while the \( T_1 \) and \( T_2 \) values were calculated using a script within the program ANALYSIS 2.1.1. (CCPNMR). For this, the cross-peak intensities were fitted using an equation for exponential decay, \( y = A \exp(-Bx) \). The tolerance in peak picking was adjusted, in the F2 dimension, to 0.1 ppm. Overlapping or very weak peaks were excluded if individual contributions were judged to be unreliable. Expected and experimental values were compared to estimate the error. The relaxation rates constants \( R_1 \) and \( R_2 \) were calculated by inverting \( T_1 \) or \( T_2 \) respectively and the estimate of error was adjusted accordingly.

For the measurement of the heteronuclear steady-state \(^1\text{H}-^{15}\text{N} \) NOE two \(^1\text{H},^{15}\text{N}-\) HSQC spectra were recorded. The saturated spectrum started with a 5 s period which contained a 3 s period of proton spin saturation. In the meantime, a \(^1\text{H}-^{15}\text{N} \) NOE developed, reducing the intensity of each cross peak. The second spectrum, the reference spectrum, contained a 5 s delay with no saturation. Both spectra were processed and the NOEs were calculated using a script within the program ANALYSIS 2.1.1. (CCPNMR) by fitting NOE = I/I(0), where I is the saturated intensity and I(0) is the reference intensity.

### 2.12. Molecular visualization programs and other computational tools

Several molecular visualization programs were used to generate the figures in this thesis. Pymol ([http://www.pymol.org/](http://www.pymol.org/)) [132] was used to display chemical shift perturbations on surface representations of MDM4-N structures. For this, the B-Factor column from the PDB was replaced with chemical shift perturbation values using PDB Editor [133] and a Pymol script “colour_b.py” (Copyright (c) 2004
Robert L. Campbell) was run. Chimera UCSF (http://www.cgl.ucsf.edu/chimera/) [74] was used to generate hydrophobicity colouring of residues accordingly to the Kyte and Doolittle scale [73].
3.1. MDM2-N

3.1.1. Expression in labelled media

Initially, *E. coli* BL21 Star cells, transformed with recombinant DNA encoding residues 11-125 of the human MDM2 sequence (2.1.1), were grown in minimal media as described in Section 2.3.2 of Materials and Methods, to test production levels of MDM2-N prior to producing an isotopically labelled sample. The MDM2-N expression was induced at 0.1 mM IPTG. The amount of soluble MDM2-N protein obtained in minimal media was, however, very poor (Fig. 3.1 B lanes 7 and 8). In order to boost expression of labelled protein, an isotope-enriched supplement known as Isogro was added (0.5 g/L) to the minimal media. As a result the culture reached an OD$_{600}$ of 0.6 more quickly than in conventional minimal media (Fig. 3.1 A). The production of protein (estimated by comparison of lanes 6 vs 8 Fig 3.1 B) was also higher than in minimal media.
Fig. 3.1. Growth of bacterial cultures in various media. A. The data points were acquired by measuring the OD\textsubscript{600} of 1-ml aliquots removed from the culture at time intervals throughout the incubation. Supplementation of the minimal media with Isogro (MM+Isogro, red) caused more rapid growth than incubation in classic minimal media (MM, yellow) (compare red, with yellow) although not as fast as in rich (LB, blue) media. The culture supplemented with Isogro had reached mid-log phase (OD\textsubscript{600} ~ 0.6, line) and was ready for induction after five hours. B. SDS-page of the expression of MDM2-N in different cultures medias. The results were acquired as described in Materials and Methods 2.4.3. The lanes correspond to (L) ladder, (1 & 2) uninduced cells collected after incubation; (3 & 4) LB media incubated cells before and after induction respectively; (5 & 6) MM+Isogro grown cells before and after induction respectively; (7 & 8) MM grown cells before and after induction respectively.
Experimental conditions to increase protein production were explored by increasing the IPTG concentration used for induction from 0.1 mM to 1 mM, or adding 20% w/v glucose at the induction time (Fig 3.2). The results showed that the extra glucose increased the rate of cell multiplication, but analysis by SDS-PAGE confirmed that protein production was increased neither by glucose nor by IPTG (see Fig. 3.2).

<table>
<thead>
<tr>
<th>Number in gel</th>
<th>Media composition</th>
<th>Extra additives</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>MM Isogro 0.5 % w/v</td>
<td>0.1 mM IPTG</td>
</tr>
<tr>
<td>2</td>
<td>MM Isogro 0.5 % w/v</td>
<td>1m M IPTG 20% Gluc v/v</td>
</tr>
<tr>
<td>3</td>
<td>MM Isogro 0.5 % w/v</td>
<td>20% Gluc v/v 0.1 mM IPTG</td>
</tr>
</tbody>
</table>

**Fig. 3.2. Analysis of proteins produced by cells growing in various conditions.** A 25-ml inoculated culture was grown as described in 2.3.2 of Materials and Methods. At the induction time (OD_{600} ~ 0.6) three different procedures were carried out as follow: (1) induction with 0.1 mM IPTG as a control of the experiment; (2) induction with 1 mM IPTG, and addition of glucose to a final concentration of 20% w/v simultaneously; (3) induction with 0.1 mM IPTG also adding glucose to final concentration 20% w/v. The results were acquired by collecting 1-ml sample from the cultures and prepared as described in Materials and Methods Section 2.4.3. A molecular weight marker (L) and un-induced cells were also included in the gel for reference.

Chapter 3: Protein Production and Characterization
In a further attempt to optimise conditions for protein production, a fermentor was employed. The use of a fermentor provides a greater degree of control over growth conditions, since the temperature, agitation, oxygen concentration and pH can all be monitored and adjusted easily. The parameter used to track *E. coli* growth was the dissolved O$_2$ concentration in the media as a function of incubation time. A decrease of O$_2$ implies that the bacteria are metabolically active (and therefore growing or producing protein), while an increase in the concentration of dissolved O$_2$ is a sign of non-growth. From observing the dissolved O$_2$ concentration during a growth in LB media (blue trace, Fig. 3.3), it was noticed that the dissolved O$_2$ level declined within the first two hours of incubation and then remained at a low level. After 10.5 hours of incubation (five hours after induction), the dissolved O$_2$ levels recover gradually, implying that the cells were no longer growing at the same rate. These data demonstrated that the protein-production window is approximately six hours in rich media.
Fig. 3.3. *E. coli* growth in the BioFlow 2000 fermentor as monitored by dissolved O$_2$ levels. Plot of the parameters monitored in MDM2-N expression in *E. coli* using a fermentor. For this experiment, 4 L of LB were inoculated with 50 ml of starter culture at 37 °C. When the OD$_{600}$ reached 0.3, the temperature (green) was decreased to 25 °C and this was maintained for the remainder of the incubation. Agitation (blue) was maintained at 200 rpm. The dissolved O$_2$ level (red) is constantly measured since this monitors the amount of metabolic activity occurring in the vessel. Induction took place after 2 h from the beginning of the run (OD$_{600}$~0.6). The dissolved O$_2$ trace indicates that *E. coli* cells were growing for a period of time during which O$_2$ was being consumed. After 10 hours the cells enter stationary phase and the culture stops growing.

Comparison of data obtained from cell growth in minimal media and Isogro-substituted minimal media, confirmed similar growth/dissolved O$_2$ profiles. To ensure that O$_2$ is not a limiting factor in the experimental setup of the fermentor, a loop was introduced to maintain the dissolved O$_2$ concentration by increasing agitation whenever the dissolved O$_2$ levels fell below 10%. This should prevent the cells from going anaerobic, which would inhibit growth and increase cell death. A similar growth profile to the previous one was obtained (Fig. 3.4 A). But this trace displayed a more abrupt cessation at 10 hours. This was assumed to be brought about by the increased metabolic rate that is a consequence of the higher dissolved O$_2$
availability. According to this scenario, the cells exhaust their nutrients earlier and hence achieve stationary phase more suddenly.

Fig. 3.4. Fermentor to optimise conditions for production of MDM2-N. A. The fermentation was carried out as outlined in the text and described more fully in Materials and Methods 2.5.2. Two traces of dissolved O₂ levels during the fermentation are shown. Minimal media were supplemented with Isogro using the standard protocol (MM + I in red), or an alternative protocol was tried, in which the agitation was increased when dissolved O₂ falls (MM+I+Loop green). The profiles are all similar; however, the abruptness of growth cessation in MM+I+Loop seems sharper that the gradual decrease that occurs in MM+I. B. The amount of soluble and un-soluble expressed protein was estimated by SDS-PAGE. The samples were prepared as described in Material and Methods Section 2.4.3. The SDS-Page gel shows a protein marker (L), non-induced culture (1); the supernatant containing soluble in (2) while the pellet or insoluble protein in (3).
Despite efforts to optimise cell growth conditions in a fermentor, only 1.3 mg of soluble protein was obtained per litre of culture (Fig. 3.4(B)) with the majority of protein being lodged in IBs. Lowering the temperature to 25 °C during the incubation did not help significantly. It was therefore decided to adopt a refolding approach (see section 3.1.2) in parallel with the exploration of alternative supplements for minimal media according to a protocol developed by researchers at AstraZeneca.

Consequently, an attempt was made to produce protein in shake flasks using minimal media supplemented with 15N-enriched Celtone (Cambridge Isotope laboratories) as described in Sections 2.3.2 and 2.5.1 of Materials and Methods. A sample of soluble target protein was purified with a yield of 17.2 mg of soluble protein per litre of culture. The success of the AstraZeneca protocol might be ascribed to the addition to the cellular media of in-house solutions of micronutrients and vitamins (Appendix A.3 for details), whereas previously only PTM1 salts and thiamine had been used. Nonetheless, refolding was attempted in an effort to recuperate the already expressed protein 15N-labelled in inclusion bodies.

3.1.2. Refolding

As outlined above efforts, to enhance 15N -labelled MDM2-N production using Isogro substitution did not yield sufficient soluble protein for NMR-based studies. Therefore, a strategy was developed to recover MDM2-N from IBs. This methodology was also later applied to MDM4 1-134, therefore the refolding presented here is equally representative of the work performed with both proteins.

There are two equally important steps in the recovery of protein from IBs: solubilisation of the protein with a denaturing agent and refolding in the presence of anti-aggregation agents [134, 135].

There are several denaturants described in the literature, of which guanidine hydrochloride (GuHCl), urea and detergents are the most commonly used [135]. GuHCl is a strong chaotrop agent with denaturant qualities at 6 M concentration. Unlike in the case of urea, there is no risk of it interacting covalently with proteins. Therefore, the buffer for solubilisation of MDM2-N expressed in IBs contained 6 M GuHCl buffer (Guanidium buffer). Solubilisation of protein was conducted at room temperature for at least 60 min with constant agitation as described in Section 2.7 of
Materials and Methods. Afterwards, the solution was centrifuged to spin down insoluble remnants that could interfere in the refolding process. Subsequent capture by His-trap and elution ensured the mixture is clean of contaminants such as DNA, and also reduced the volume of the sample for easier handling.

Refolding is the transition from a disordered polypeptide to a flexible intermediate to compactly folded protein (Fig. 3.5). Several different refolding techniques are described in the literature – these may be classified according to how the denaturant concentration is reduced to enable refolding [134, 135]. In the case presented here the removal of the denaturant was performed by a one-step dilution with the refolding buffer detailed in Section 2.7 of Materials and Methods. Specifically, the peak fraction from the His-trap capture (see above) was diluted 10x with refolding buffer (Fig 3.5, see details in section 2.7). The contents of the refolding buffer included arginine, which is regarded as a key component for prevention of self-association and aggregation [136, 137] as is PEG. CaCl$_2$ and MgCl$_2$ are reported to favour folding of proteins [134] (Fig. 3.5).
### Table 3.1: Co-solutes and Their Refolding Effect

<table>
<thead>
<tr>
<th>Co-solutes</th>
<th>Refolding effect</th>
</tr>
</thead>
<tbody>
<tr>
<td>CaCl$_2$</td>
<td>Folding enhancer</td>
</tr>
<tr>
<td>MgCl$_2$</td>
<td>Folding enhancer</td>
</tr>
<tr>
<td>PEG 4000</td>
<td>Aggregation suppressor</td>
</tr>
<tr>
<td>Arginine</td>
<td>Aggregation suppressor</td>
</tr>
<tr>
<td>Redox agent (DTT, glutathione red, glutathione ox)</td>
<td>Redox equilibrium</td>
</tr>
</tbody>
</table>

**Fig. 3.5. The refolding process.** The table on the top lists the effect of the co-solutes usually added in the refolding buffer. On the bottom, there is a schematic representation of the refolding process adapted from Tsumoto, 2004. The protein is unfolded (U) and soluble in 6 M GuHCl in the guanidium buffer. Dilution of the denaturant agent into refolding buffer allows the polypeptidic chain flexibility while arginine and PEG hinder aggregation (A). In these ideal conditions the protein will fold to its native conformation (F).

After incubation overnight in refolding buffer, the solution containing MDM2-N was exchanged by stepwise dialysis against large volumes of Tris-HCl, pH 8.5, 1 mM DTT. Finally, the soluble and putatively refolded protein was separated from insoluble-forms by centrifugation (Fig. 3.6).
Fig. 3.6. The results of a refolding trial. After removal of the refolding buffer by step-wise dialysis, large amounts of protein precipitated. A 10 µl sample of the soluble and putatively refolded protein was mixed with 10 µl of 2X SDS loading buffer and resolved by SDS-PAGE (lane 1) while the aggregates were spun down and re-dissolved in 20 µl 2x SDS loading buffer (lane 2).

To test whether the re-solubilised protein had been successfully folded into a native conformation, a $^{15}$N-labelled sample of the protein was produced, subjected to the procedure described above, and then used to record a [$^1$H,$^{15}$N] HSQC spectrum. This was overlaid with previous spectra of the same protein batch, but collected on a sample of the soluble fraction. The overlay shows very good agreement between spectra collected on different samples, proving that both the natively folded and the refolded proteins have very similar structures (see Fig. 3.15).

Unfortunately, the yield of protein from refolding was lower than had been anticipated, amounting to only 2.5 mg/L of culture of purified protein for MDM2-N. Since enrichment of the minimal media with Celtone had in any case proved more satisfactory for production of soluble protein, no further efforts were made to optimise the refolding conditions.

Further work would be required to establish better conditions for protein recovery from IBs. Methodical exploration of buffer conditions seems a good strategy to improve re-solubilisation.
**Fig. 3.7. Conclusions on the refolding protocol.** The strategies to improve the yields of refolded protein should tackle screening for better buffer conditions tailored individually for the target protein (square).

### 3.1.3. Purification

The first step in the purification of MDM2-N was to break open the cells. When processing a large cell mass, cell pellets were re-suspended in lysis buffer and passed through a French press. Alternatively, for a small volume of cell culture, the pellet was re-suspended in lysis buffer or Bugbuster (Novagen) followed by four rounds of sonication, as detailed in Section 2.6.1 of Materials and Methods. Sonication was performed to assist DNA fragmentation and thereby reduce the viscosity and improve the tractability of the sample.

The hexaHis-tag is a repetition of six histidine residues, which has affinity for divalent metals ions (including Ni$^{2+}$ or Co$^{2+}$) thereby providing the recombinant protein with an affinity-specific structure useful for purification. Consequently, the lysates were clarified by centrifugation and the supernatant was filtered and loaded onto either a 1-ml Ni$^{2+}$ His-Trap column (Qiagen), for small volumes of induced culture, or 20-ml columns for bigger volumes, as detailed in Section 2.6.2 of Materials and Methods. The samples were loaded onto the column with either a 50-ml superloop (GE Healthcare), for up to 50 ml of lysate volume, or with a pump using the AKTA system.
An AKTA based protocol was developed for elution of the protein from a His-trap column using an imidazole gradient. Fractions were collected (see Materials and Methods, Section 2.6.2) resulting in two low peaks of protein (Fig. 3.8). Analysis of the peak fractions by SDS-PAGE showed that the recombinant protein was eluted in the second peak.

**Fig. 3.8 First step in the purification of MDM2-N.** Cells in the induced bacterial culture were lysed by passing through a French press and then clarified by centrifugation. The supernatant was pump-loaded into a 20-ml His-trap column attached to an AKTA system. The strategy for eluting the MDM2-N was an imidazole gradient (from 10 to 250 mM, green line), collecting 10-ml fractions. Two peaks were resolved, peak (a) at 752 ml of elution volume (23.5% of 250 mM) and peak (b) at 810 ml of elution volume (40% of gradient 250 mM). SDS-PAGE analysis of fractions (prepared according to Section 2.4.1 in Materials and Methods) corresponding to peak (b), confirmed the presence of MDM2-N (see inset).

The next stage of the purification was removal of the hexaHis-tag since the effects of such an artefact on the overall fold of a protein and its dynamics, are unpredictable. Consequently, for NMR studies, the hexaHis-tag is generally removed. In the current project, the tag had been fused to the N-terminus of the MDM2-N, since this reportedly improved the protein yield (AstraZeneca, unpublished data). A thrombin-
cleavage site had been engineered into the sequence between the tag and the native residues (Fig. 3.9 A). To remove any imidazole traces on the fractions containing MDM2-N a buffer-exchange step was performed at the same time as thrombin incubation. Following thrombin cleavage (see details in Section 2.6.2 of Materials and Methods) the samples were analyzed by SDS-PAGE that resolves cleaved from un-cleaved (-1900 Da) proteins (Fig. 3.9 (B)).

![Thrombin cleavage cartoon](image)

**Fig. 3.9. Thrombin cleavage.** A. Cartoon representation of MDM2-N with the hexaHis-tag and thrombin-cleavage site. Thrombin cleaves after the Arg residue (in red), thus the final product of the cleavage will have an artefact GSH at the very N-terminus of MDM2-N (Fig 2.1). B. SDS-PAGE of MDM2-N before and after cleavage reaction. Lanes 1 and 3 are duplicates and correspond to MDM2-N after o/n incubation with thrombin, while lanes 2 and 4 contain duplicate samples of un-cleaved versions of MDM2-N. The cleaved construct is 1.9 kDa shorter than the tagged protein.

Once the hexaHis-tag has been removed, the protein no longer has affinity for the Ni²⁺ column. This feature was utilised to separate impurities and un-cleaved recombinant protein from the tag-free MDM2-N. For this purpose, the products of the cleavage step were pump-loaded onto the aforementioned column mounted on an AKTA FPLC, and fractions were eluted as described previously in Section 2.6.2 of
Materials and Methods. The flow-through fractions were collected and taken forward for further purification (Fig. 3.10).

**Fig. 3.10. Elution of MDM2-N after thrombin cleavage and loading on to the His-trap column.**

The protocol followed was the same as used for the initial capture of the protein. The fractions corresponding to the main peak (180 ml of elution volume) are shown in lanes 1-6 of the inset SDS-PAGE. Fractions eluting after initiating the imidazole gradient (lanes 8-11) correspond to un-cleaved protein or bacterial impurities.
Finally, a “polishing” step consisting of gel-filtration chromatography was performed in order to further clean the sample of impurities and possible aggregation products. The chromatographic profile showed a sharp peak at approximately 220 ml of the elution volume. Comparing this information with a standard calibration for the column, confirmed the eluted protein has the expected molecular weight (Fig. 3.11).

**Fig. 3.11. Polishing step: gel filtration.** The last step of the purification involved gel-filtration chromatography. Samples were loaded into a 5-ml loop attached to the AKTA Superdex 26/60 S200 column (GE Healthcare). The sample was eluted over 1.2 column volumes. The UV (280 nm) reading of the chromatogram revealed a sharp peak at about 220 ml. These fractions were collected and the presence of MDM2-N was confirmed by SDS-PAGE (not shown).

To assess the molecular weight and the purity of the sample, liquid chromatography-mass spectrometry (LCMS) was performed. The molecular weight of non-isotopically labelled MDM2-N was estimated (ExPASy tool ProtParam, see Appendix A.1.2 for details) to be 13719.6 Da (including the expression artifact). The molecular weight of a uniformly $^{15}$N-labelled protein is estimated by adding 0.98 Dalton per nitrogen to the estimated molecular weight (the ammonium sulphate used...
for labelling is only 98% enriched). The recombinant MDM2-N has 160 nitrogen atoms and therefore the molecular weight should be a few Daltons less than 13876.4 Da. The spectra obtained were consistent with a pure sample with a molecular weight of 13874.62 Da - this matches the expected weight within the error of the instrument (Fig. 3.12) and allowing for incomplete labelling.

![MDM2-N LCMS results](image)

**Fig. 3.12. MDM2-N LCMS results.** The purity and molecular weight of the protein were verified by mass spectrometry. The inferred molecular weight is 13,874.62 Da. This value is 2 Da larger than the estimated mass of the protein assuming 98% labelling, but lies within the instrument error and uncertainty over the extent of $^{15}$N enrichment.

### 3.1.4. Initial NMR analysis of MDM2-N

Initially, a solution of non-labelled protein was prepared for the recording of a $^1$H-observe one-dimensional NMR spectrum. The pH was kept lower than 7.5. At higher pH values amide protons of the protein are generally in fast exchange with solvent, and hence not easily observable by NMR. Another consideration was that the literature largely documents MDM2-N to be unstable and difficult to handle. Several different NMR buffer and pH conditions that were tested in the current work (not shown) resulted in precipitation of the protein. The NMR buffer in which the MDM2-N presented the greatest stability was 50 mM phosphate buffer, pH 6.8, 0.1 mM EDTA, 2 mM TCEP and 0.02% w/v NaN$_3$ (NMR buffer 1).

---

Chapter 3: Protein Production and Characterization 88
A 1D \( ^1\text{H} \) spectrum (128-scans) was acquired following instrumental optimization as described in Section 2.11.2 of Materials and Methods. The aliphatic region (Fig 3.13, 1-5 ppm) of the spectrum contained sets of very sharp peaks arising mainly from glycerol that originates from careless preparation of concentration membranes. The amide/aromatic region (Fig 3.13, 7-10 ppm) was characterised by a good dispersion of defined peaks, indicative of a folded protein. Several small peaks down-field of the water signal around the water signal (Fig 3.13 a) are characteristic of beta-strands in the protein and thus provide further evidence for the existence of a well-folded protein. The several up-field shifted peaks detected at around 0 ppm (Fig 3.13 b) indicated the presence of a hydrophobic core containing aromatic side-chains, for the signal is arising from shielded methyl groups. The absence of broad resonances and overlapping peaks is inconsistent with self-association and aggregation.

Fig. 3.13. 1D \( ^1\text{H} \) spectrum of MDM2-N. The spectrum was recorded using a Bruker NMR instrument operating at 600 MHz with data acquired and processed using “TopSpin”. The details show (a) alpha protons in beta strands and (b) up-field shifted peaks (normally methyls).

A commonly adopted next step in the NMR analysis of a protein is to record a \([^1\text{H},^{15}\text{N}]\) heteronuclear single quantum coherence (HSQC) spectrum. This
experiment, which normally depends on the availability of $^{15}\text{N}$-enriched protein, provides a 2-D dataset that offers more “frequency space” and hence less overlap of peaks than is obtained in a 1D spectrum. Moreover, the spectrum is highly informative since each cross-peak corresponds to the $^1\text{H}$ and $^{15}\text{N}$ chemical shifts of an amide group within the protein. The HSQC spectrum thereby affords a “fingerprint” of the protein because the positions of peaks are determined both by amino acid identity and by their local magnetic environment (that reflects shielding and deshielding effects of pi-electrons etc.). In addition, the HSQC spectrum can be used to follow perturbation of frequencies due to ligand binding of conformational change, and is therefore a robust underpinning technique for the study of protein-ligand interactions.

By convention in a $[^1\text{H},^{15}\text{N}]$ HSQC spectrum, $^1\text{H}$ chemical shifts form the $x$-axis and $^{15}\text{N}$ chemical shifts the $y$-axis. A good dispersion of similar-intensity cross-peaks is indicative of a compactly folded protein, while the number of peaks reflects approximately the number of amino acids of the protein. The match is imperfect, though, because proline residues have no amide groups while the N-terminal residue has a free amine (whose protons are in exchange with water and generally not observed) rather than an amide. On the other hand, amino acids such as tryptophan, asparagine, glutamine and arginine produce extra peaks due to the N-H groups within their side chains. Where there are fewer cross-peaks than expected, this could imply that the protein has regions that are conformationally mobile on an “intermediate” timescale leading to exchange-broadening effects. Notably, MDM2-N has been reported to have a conformationally mobile N-terminus that is known as the “lid” region as discussed in Section 1.4.1. As expected, therefore, the $[^1\text{H},^{15}\text{N}]$ HSQC of the MDM2-N prepared in the current study contained less cross-peaks than there are amide groups within the protein (Fig. 3.14 (A)).
Chapter 3: Protein Production and Characterization
Fig. 3.14. A. $[^1\text{H}, ^{15}\text{N}]$ HSQC spectrum recorded for 0.05 mM $^{15}\text{N}$-MDM2-N (previous pages). A. This spectrum was recorded at 300K with 50 µM MDM2-N in NMR buffer 1 (2.11.2) on a 600-MHz Bruker spectrometer fitted with cryoprobe. Overall the protein appears to be well folded since a good dispersion (spread over 3.1 ppm in the indirect dimension ($^1\text{H}$), and over 22 ppm in the $^{15}\text{N}$ or direct dimension) is observed and many of the cross-peaks have similar intensity. B. Overlay of the HSQC spectra of MDM2-N produced in this study and the equivalent protein reported by Uhrinova et al. The current MDM2-N spectrum (blue) was recorded at pH 6.8, 300 K and sweep width (Hz): [$^1\text{H}$] 8389.262 and [$^{15}\text{N}$] 1337.614, while the spectrum for the original MDM2-N sample (red) (Residue 1-118) was recorded at pH 7.2, 27 ºC and sweep width (Hz): [$^1\text{H}$] 8012.811 and [$^{15}\text{N}$] 1519.614. The spectra are overall very similar and discrepancies are probably due to the differences in conditions used for recording them. The reduced dimensions in the sweep width of MDM2-N spectrum in the current project prompt resonances to appear as negative peaks (pink, a’, b’ and c’) to remark that they are aliased. The spectrum recorded by Uhrinova has a wider sweep width and the peaks a, b and c appear at its corresponding chemical shift.

Previous work in Prof. Barlow’s lab yielded the NMR assignment of the N-terminal (1-118) domain of MDM2 (Uhrinova et al. 2005). The good overlay of the previously collected HSQC spectrum with the current one for MDM2-N (residue 11-125) suggested that the two proteins are similar in structure as would be expected. Small discrepancies may be due to differences in recording conditions. Note also that a different spectral width was used and thus there are discrepancies due to the way the spectrum is “folded” for display (Fig. 3.14 B).
Fig. 3.15. **HSQC spectra confirm integrity of re-folded protein.** A $[^1\text{H},^{15}\text{N}]$ HSQC of the refolded protein (orange) was overlaid with a previously recorded HSQC of the soluble protein (blue) proving that the refolding protocol produced natively folded protein. Folded peaks of the soluble protein are in pink and those of the refolded protein in blue. The detail shows the centre region of the HSQC spectra (8.7-7.7 $[^1\text{H}]$ and 116-122 $[^{15}\text{N}]$ in ppm) of the refolded protein overlaid with a the soluble protein (left) and of the refolded protein (right).

An $[^1\text{H},^{15}\text{N}]$ HSQC spectrum was also collected for the sample of re-folded protein and this was superimposed on the HSQC spectrum of the soluble protein (Fig. 3.15).
The spectrum of the refolded protein (orange) overlays very well with that of the soluble purified protein (blue).

### 3.2. MDM4 1-134

#### 3.2.1. Construct design and cloning

Only two publications of structural studies were available at the commencement of this work (June, 2008) and they were reviewed to help decide upon the most suitable sequence, from the N-terminus of MDM4, to clone for NMR and other biophysical studies (Table 3.1). The selected sequence was based on that reported by the Holak laboratory [138]. These researchers used human MDM4 residues 1-134 successfully in NMR-based studies. This construct is potentially informative since it includes the very N-terminal region that is equivalent to the “lid” of MDM2.

<table>
<thead>
<tr>
<th>PDB</th>
<th>Reference</th>
<th>Constructs</th>
<th>Details</th>
</tr>
</thead>
</table>

Table 3.16. MDM4 publication up to May 2008. List of the publications on structural studies of the MDM4 N-terminal domain [65, 138] showing information given in the papers about the constructs and the details of purification.

After deciding the most adequate construct, a cloning strategy was designed as follows. The cloning strategy used in the current work involved amplification (from DNA for full-length MDM4, courtesy of Prof. Ted Hupp, Edinburgh) of the DNA encoding the target region (MDM2 residues 1-134) by PCR. Restriction enzyme target sequences were incorporated at both ends of the primers as follows: The forward primer had a restriction site for Ndel nuclease, while the reverse primer had a XhoI restriction site and a stop codon (Fig. 2.4 in Section 2.2.1). The PCR reaction

Chapter 3: Protein Production and Characterization
was performed as described in Materials and Methods Section 2.2.2, and the end product was analyzed on an agarose gel (section 2.2.4) (Fig. 3.17). The upper band was dismissed as a sub-product of the PCR reaction, and the lower band (size = 400 bp) was gel extracted and taken forward.

**Forward primer**

```
CC CAT ATG ACA TCA TTT TCC
```

**Reverse primer**

```
GG CTC GAG TTA CTG TCG G
```

**Fig. 3.17. Cloning strategy of the N-terminal domain (1-134) of MDM4.** Amplification of the gene of interest was performed by PCR. The gel shows a lower band of 400 bp, which matches with the expected size of the insert. The upper band is a by-product of the PCR reaction. The lower band was gel extracted (section 2.2.4). Subsequently, both insert and vector were incubated with the restriction enzymes Ndel and Xhol. After another purification, the insert and the open vector were incubated with a T4 ligase. The end product was used to transform Electrocompetent TOP10 E. coli cells as detailed in Section 2.2.5. Several colonies were picked and sequenced.

The next step in the cloning process was to digest the insert and the vector with Xhol and Ndel. Each endonuclease cut once within the polycloning site of the vector creating the sticky ends required for subsequent reannealing (the use of two restriction enzymes provides directionality to the insertion). The vector was
dephosphorylated to avoid religation (Section 2.2.8). Finally, the ligation reaction was set up with several ratio of X:Y insert to vector (described in detail in Materials and Methods, section 2.2.9). The ligation product was transformed into *E. coli* TOP10 cells for amplification and sequenced. The results of the sequencing are in Appendix A.2.

### 3.2.2. Expression

For expression of the MDM4 1-134 several *E. coli* strains - DE3, Star and pLys - were transformed with the plasmid containing the recombinant DNA. All strains are designed for gene expression by the T7 RNA polymerase under IPTG control and are deficient in *lan* and *ompT* proteases whose target is RNA polymerase [139]. In addition, the pLys strain has tight control over expression, useful for producing toxic proteins, while the Star strain has enhanced mRNA stability due to decreased activity of RNAaseE (Invitrogen).

Consequently, 5 ml protein-production trials were performed as described in Material and Methods, Section 2.5. Unfortunately, SDS-PAGE showed that the MDM4 1-134 protein was insoluble when produced in all of these strains (Fig. 3.18). Judging by the size of the band corresponding to the target protein compared to the background bands, MDM4 1-134 was produced insoluble at the high quantities in *E. coli* DE3 (BL21) Star.
Fig. 3.18. “Miniscale” protein production trial of MDM4 1-134. The gel prepared as detailed in Section 2.4.3, shows that the MDM4 (1-134) has a predominantly insoluble character when produced by these E. coli strains. Three different strains were tested, but material in the soluble fractions (“Sol”) were barely detectable, unlike the large amount of protein evident from correct-sized bands in the insoluble (“insol”) fractions. Further expression of MDM4 (1-134) was undertaken in the Star strain, since there seems to be less contaminating protein bands relative to the target band on the gel shown.

3.2.3. MDM4 1-134 purification

The first approach to purification of MDM4 1-134 was to follow the same protocol used for MDM2-N purification (Section 2.6.2 of Materials and Methods). Unfortunately the protein constantly precipitated, without apparent reason, even at 4 °C (Fig. 3.19 A). Review of the purification protocol revealed that the buffer used for breaking open the cells (Bugbuster) does not include reducing agents. MDM4-N has three free cysteines: two of them in the region of MDM4 that is equivalent to the “lid” of MDM2, and one buried in the structure (Fig. 3.19 B). Addition of DTT to the solution improved the stability of the protein somewhat but it remained prone to precipitation. This meant concentration steps had to be avoided and buffer exchange was instead performed using buffer-exchange columns.
Fig. 3.19 Precipitation problems with MDM4 1-134. A. After loading the soluble MDM4 (1-134) and elution with imidazole-containing buffer, the sample was buffer exchanged into an imidazole-free buffer imidazole by concentration and re-dilution steps. These caused MDM4 1-134 to precipitate. B. Cartoon representing the location of the three free cysteines in MDM4 1-134. Two cysteines are located near the N terminus, in what is expected to be a flexible random coil region. The third cysteine is probably buried in the structure (according to Popowicz et al 2007). C. Desalting of MDM4 1-134. In order to avoid aggregation the MDM4 1-134 sample was buffer exchanged into an imidazole free buffer using desalting columns. The protein (elution volume 3-8 ml) was successfully separated from the imidazole (elution started after 10 ml) (dark blue). The light blue trace is for a solution of imidazole without protein.
Despite all efforts, the yield of soluble purified protein was only 3 mg/L of rich media culture. The precipitation of protein continuously along the purification was detrimental, but eventually a useful amount of protein was finally purified by size exclusion chromatography as described in Materials and Methods Section 2.6.2. This confirmed that the protein was not self-associating and forming high molecular-weight aggregates (Fig. 3.20). To further assess the molecular weight and the purity of the sample of the peak fraction resulting from size exclusion, Fourier Transform Ion Cyclotron Resonance Mass Spectrometry (FTICR-MS) was performed and analyzed by the SIRCAM facility at the University of Edinburgh. The molecular weight of the sample matches that predicted by Protparam of 15018 Da (see details Appendix A.1.4).
Fig. 3.20. Gel filtration chromatography of MDM4 1-134. A, Size-exclusion chromatography is the last step of the purification. The protein elutes at 70-80 ml of the elution volume, in agreement with the expected size of the protein. B, FTICR-MS results for MDM4 1-134 purified sample. The average mass was calculated by fitting (M/z)+z, where M: mass, z: change as showed at the top of the figure. The data was recorded and analyzed by the SIRCAM facility at the University of Edinburgh.
Subsequently, $^{15}$N-MDM4 1-134 was produced in minimal media as described in Materials and Methods Section 2.5 and the soluble fraction was purified producing a very poor yield of only 20 µg of protein from 5 L of expression media. Accordingly a [${}^1$H,${}^{15}$N] HSQC of 25 µM $^{15}$N-MDM4 1-134 was recorded that showed a well-folded protein with good peak dispersion (a spread of 3 ppm in the $^1$H dimension, and 23 ppm in the ${}^{15}$N dimension) (Fig 3.21 A). As outlined above, lack of signals in this spectrum may arise from conformational averaging or “chemical exchange” on an intermediate timescale (compared to the data collection in an NMR experiment e.g. $10^{-6}$-$10^{-8}$ seconds). It is therefore notable that only 80 cross-peaks were detected in the spectra from some 125-130 amides in MDM4 1-134. This deficit, however, is partially consistent with the 34 residues one might expect in the flexible regions at the N-terminus (1-26) and at the C-terminus (126-134) of MDM4 1-134. Additional missing peaks no doubt arise due to MDM4 1-134 being a very flexible protein in the apo-state [140] (Fig. 3.21 B).
Fig. 3.21. HSQC spectrum MDM4 1-134. A. \([^{1}H,^{15}N]\) HSQC spectrum soluble purified 25 \(\mu M\) MDM4 1-134 recorded at 300K in NMR buffer 2 (2.11.2) on a 600 MHz Bruker spectrometer fitted with cryoprobe. Good dispersion of peaks, spreading over 3 ppm in the indirect dimension \((^1H)\), and over 23 ppm in the \(^{15}N\) or direct dimension, shows that the protein is well folded. The cross-peaks arising from side-chains amides are highlighted in the box. B. Secondary structure prediction of MDM4 1-134 as predicted by PSIPRED Protein Structure Prediction Server \([141]\). The square highlights flexible regions of MDM4 1-134.

Furthermore, it is noteworthy that the signal intensity of the cross-peaks corresponding to the backbone is lower than the intensity of the signals corresponding to side-chain amides (Fig 3.21 A). This is a feature of protein molecules that are undergoing slow overall tumbling (while exposed side-chains remain relatively dynamic), therefore behaving as a larger protein.

In an attempt to boost the yield of purified protein, \(^{15}N\)-MDM4 (1-134) was re-folded following the protocol used for MDM2-N. The yield of re-folded protein was very poor, but enough was obtained to perform an \([^{1}H,^{15}N]\) HSQC. Overlaying this with
the HSQC of soluble protein demonstrated equivalence. Indeed, the $[^1\text{H},^15\text{N}]$ HSQC spectrum of the refolded MDM4-N shows the same hallmarks (Fig. 3.22) of a conformationally mobile protein.

Fig. 3.22. HSQC spectrum of the refolded MDM4-N is near-identical to the HSQC of soluble protein. $[^1\text{H},^15\text{N}]$ HSQC spectrum of the re-folded protein (A) was recorded under the same conditions as referred in Fig 3.21. The spectrum was overlaid perfectly with a previously recorded HSQC of the soluble protein (Fig 3.21 A) as shown in the detail of the region (B) $9.0-7.4$ $[^1\text{H}]$ and $115-125$ $[^15\text{N}]$ in ppm (C).
After confirming the integrity of MDM4 1-134 via NMR experiments effort was directed to address precipitation problems during protein purification. It was then decided to explore buffer conditions and additives that might have a positive impact on the stability and solubility of the protein. Increasing salt concentration in the buffer should deter precipitation caused by electrostatic interactions. Certain buffer combinations that do not affect the NMR signal have been proven to minimise aggregation [142, 143]. The use of dynamic light scattering (DLS) as a more convenient technique than NMR with which to screen buffers for promotion of monodispersal protein was also considered. This technique, however, is extremely sensitive to aggregation and requires the sample be filtered prior to measurement. Therefore, a more rudimentary measurement of light scattering was employed, based on the principle that no amino acids or peptide bonds absorb at wavelengths above 310 nm. Thus any reading on a UV/vis spectrophotometer above 320 nm can be considered as due to light scattering or turbidity arising from high molecular-weight particulate matter in solution or aggregates [144].

Thus, 1 ml aliquots of 50 µM MDM4 1-134 were buffer exchanged into Hepes buffer, pH 7.5, with a variety of additives and spectra in the UV/vis region (220-400 nm) were recorded to assess the aggregation of the protein in solution. Instability and self-association of MDM4 1-134 in solution could arise from several possible factors including the free cysteines, unstructured flexible regions at the N-terminus and C-terminus and/or the presence of the hydrophobic p53-binding groove. The “Golovanov mixture” (50 mM arginine, 50 mM glutamic acid) is appropriate for reducing aggregation caused mainly by surface interactions of folded proteins as these chaotopic small molecules purportedly “cover up” or effectively neutralise highly charged or hydrophobic regions in the protein [145]. The role of non-detergent sulfobetaines (NDSB) as mild detergent have proven valuable additives for protein refolding [142, 146] Finally, arginine is a widely-used antiaggregation chaotrope although the precise mechanism for this activity has not yet been revealed [137] (Fig. 3.23 Top).

The recombinant protein in standard buffer supplemented with 50 mM arginine underwent the lowest extent of aggregation (Fig. 3.23, red), based on a relatively low absorbance at 320 nm of 0.11 AU. In contrast, when the buffer was supplemented
with the Golovanov mixture or NDBS, the absorbance at 320 nm increased to 1.5 and 0.38 AU respectively (Fig. 3.23, black), suggesting that the protein was highly aggregated in those conditions.

### Table 3.23

<table>
<thead>
<tr>
<th>Buffer</th>
<th>Additive features</th>
</tr>
</thead>
<tbody>
<tr>
<td>50 mM Hepes pH 7.5, 100 mM NaCl, 1 mM DTT</td>
<td>Standard buffer.</td>
</tr>
<tr>
<td>50 mM NDBS, 50 mM Hepes pH 7.5, 100 mM NaCl, 1 mM DTT</td>
<td>Anti aggregation effects on acidic fibroblast growth factor.</td>
</tr>
<tr>
<td>50 mM arginine, 50 mM glutamic acid, 50 mM Hepes pH 7.5, 100 mM NaCl, 1 mM DTT</td>
<td>Proven anti aggregation in NMR samples due to increasing concentration.</td>
</tr>
<tr>
<td>50 mM arginine, 50 mM Hepes pH 7.5, 100 mM NaCl, 1 mM DTT</td>
<td>Anti aggregation described in many refolding protocols.</td>
</tr>
</tbody>
</table>

### Fig. 3.23. Buffer screening conditions for MDM4 1-134.

**A.** Table listing the conditions screened to increase MDM4 1-134 solubility in solution. All conditions were tried at identical pH and NaCl concentrations. **B.** UV/vis spectra of MDM4 1-134 recorded in different buffer conditions. The standard buffer comprised of Hepes, pH 7.5 and 100 mM NaCl (blue), to which several additives (50 mM arginine in red, 50 mM arginine + 50 mM glutamic acid (Golovanov mixture) in black and 50 mM non-detergent sulfobetaines (NDSB) in yellow) were added. Aggregation was assessed as a function of absorbance at wavelengths above 320 nm.

In conclusion, the efforts made to stabilize MDM4 1-134 were not very satisfactory although this was due partly to the limited array of possibilities screened. It was
decide that rather than launch a major screen of many new conditions it might be preferable to try using a new construct.

3.3. MDM4 14-111 (MDM4-N)

3.3.1. Construct design and cloning

Two crystal structures of the N-terminal domain of MDM4 were published in early 2009 that employed different methodologies to deal with insolubility and aggregation of their respective construct [82, 112]. These constructs differed from MDM4 1-134 in terms of their lengths (Fig. 3.24). Both were shorter with respect to their N and C termini, avoiding unstructured regions and potentially providing more stability. Also Cys\textsuperscript{10} was deleted in both constructs while Cys\textsuperscript{17} was mutated to Ser in PDB id 3FE7, in keeping with reports that removal of these residues increased protein stability [147]. The protein used to solve the PDB id 2VYR structure was also fused to a tag designed to enhance solubility; the strategy in this case involved stabilization of their MDM4 construct by ligand binding followed by cleavage of the tag. On the other hand, 3FE7 was purified in buffer containing 10% v/v glycerol that was found to reduce aggregation and precipitation.

<table>
<thead>
<tr>
<th>PDB</th>
<th>Author, year</th>
<th>Title</th>
<th>Constructs (residues)</th>
<th>Purification strategies</th>
</tr>
</thead>
<tbody>
<tr>
<td>2VYR</td>
<td>Yu, 2009</td>
<td>Structure Of Human Mdm4 N-Terminal Domain Bound To A Single Domain Antibody</td>
<td>16-116</td>
<td>Expression/purification with solubility tag</td>
</tr>
<tr>
<td>3FE7</td>
<td>Kallen, 2009</td>
<td>Crystal Structures of Human MdmX (HdmX) in Complex with p53 Peptide Analogues Reveal Surprising Conformational Changes</td>
<td>14-111</td>
<td>Buffer with 10% (v/v) glycerol</td>
</tr>
</tbody>
</table>
Fig. 3.24. Structure of N-terminal domain of MDM4 reported in early 2009. The table lists the MDM4 N-terminal domain structures submitted in 2009 and a brief description of the most relevant information for the current research, such as construct length and purification strategies. The alignment of MDM4 constructs 2VYR [82], 3FE7 [81] and MDM4 1-134 using BioEdit Sequence Alignment Editor v7. The constructs used for crystallography differ from MDM4 1-134 by truncations of unstructured regions at the N and C termini.

The protein construct used to solve 3FE7 was selected since it had been purified with a more straightforward strategy. Thus, a synthetic gene optimised for E. coli expression and encoding residues 14-111 of MDM4 (Expasy O15151) was subcloned into pET28b between XhoI and NdeI for production of the hexaHis-tag at the N terminus of the protein. Mutagenesis was also performed by Geneart to replace Cys17 by Ser [112] (Fig 2.3). The resultant synthetic gene was used to transform several different E. coli expression hosts by electroporation as described in section 2.2.11.

### 3.3.2. Protein expression

Several conditions and E. coli strains were screened with the aim of high expression of soluble, isotopically labelled MDM4-N. The design of the screening protocol followed basic rules that should enhance soluble expression and yield.

The expression host of choice was E. coli (DE3) BL21 and three strains (DE3, Star, Gold) were investigated [148]. The strain DE3 does not include any added feature. The strain Gold is designed to overcome low success in protein production ascribed to poor transformation. The strain Star was included in the screening for its unique quality to enhance protein translation. A detailed list of features can be found in Table 3.25.
### Table 3.25. Features of *E. coli* expression host strains.

Different *E. coli* strains were selected for MDM4-N expression trials.

Other important factors affecting protein production that were varied are IPTG concentration (0.05, 0.5 and 1 mM) and the temperature of incubation (18, 25 and 37 °C). These tests were designed to minimise formation of inclusion bodies [149]. In addition to MDM4-N expression in minimal media (M9 supplemented with Celtone) and Terrific Broth (TB) rich media were used as positive controls. Finally, another well-characterised (MDM2-N) construct, with known high rates of expression, was added to the screening as positive control. The extent of the screening is detailed in Table 3.26.
<table>
<thead>
<tr>
<th>Strain</th>
<th>STAR</th>
<th>GOLD</th>
<th>DE3</th>
<th>STAR</th>
<th>GOLD</th>
<th>DE3</th>
<th>STAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp (C)</td>
<td>37</td>
<td>37</td>
<td>37</td>
<td>37</td>
<td>37</td>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td>Media</td>
<td>TB</td>
<td>TB</td>
<td>TB</td>
<td>TB</td>
<td>TB</td>
<td>TB</td>
<td>TB</td>
</tr>
<tr>
<td>IPTG (mM)</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strain</th>
<th>GOLD</th>
<th>DE3</th>
<th>MDM2</th>
<th>STAR</th>
<th>GOLD</th>
<th>DE3</th>
<th>STAR</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp (C)</td>
<td>37</td>
<td>37</td>
<td>37</td>
<td>37</td>
<td>37</td>
<td>37</td>
<td>37</td>
</tr>
<tr>
<td>Media</td>
<td>TB</td>
<td>TB</td>
<td>Celtone</td>
<td>Celtone</td>
<td>Celtone</td>
<td>Celtone</td>
<td>Celtone</td>
</tr>
<tr>
<td>IPTG (mM)</td>
<td>1</td>
<td>1</td>
<td>0.1</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strain</th>
<th>DE3</th>
<th>Gold MDM2</th>
<th>STAR</th>
<th>GOLD</th>
<th>DE3</th>
<th>STAR</th>
<th>GOLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp (C)</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
</tr>
<tr>
<td>Media</td>
<td>TB</td>
<td>TB</td>
<td>Celtone</td>
<td>Celtone</td>
<td>Celtone</td>
<td>Celtone</td>
<td>Celtone</td>
</tr>
<tr>
<td>IPTG (mM)</td>
<td>0.05</td>
<td>0.05</td>
<td>0.1</td>
<td>0.05</td>
<td>0.05</td>
<td>0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strain</th>
<th>DE3</th>
<th>STAR</th>
<th>GOLD</th>
<th>DE3</th>
<th>MDM2</th>
<th>DE3</th>
<th>STAR</th>
<th>GOLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp (C)</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>25</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>Media</td>
<td>Celtone</td>
<td>Celtone</td>
<td>Celtone</td>
<td>Celtone</td>
<td>Celtone</td>
<td>TB</td>
<td>TB</td>
<td></td>
</tr>
<tr>
<td>IPTG (mM)</td>
<td>0.1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>0.1</td>
<td>0.05</td>
<td>0.05</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strain</th>
<th>DE3</th>
<th>STAR</th>
<th>GOLD</th>
<th>DE3</th>
<th>MDM2</th>
<th>DE3</th>
<th>STAR</th>
<th>GOLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp (C)</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>Media</td>
<td>TB</td>
<td>TB</td>
<td>TB</td>
<td>TB</td>
<td>TB</td>
<td>TB</td>
<td>TB</td>
<td></td>
</tr>
<tr>
<td>IPTG (mM)</td>
<td>0.05</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strain</th>
<th>MDM2</th>
<th>DE3</th>
<th>STAR</th>
<th>GOLD</th>
<th>DE3</th>
<th>STAR</th>
<th>GOLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp (C)</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
</tr>
<tr>
<td>Media</td>
<td>Celtone</td>
<td>Celtone</td>
<td>Celtone</td>
<td>Celtone</td>
<td>Celtone</td>
<td>Celtone</td>
<td>Celtone</td>
</tr>
<tr>
<td>IPTG (mM)</td>
<td>0.1</td>
<td>0.05</td>
<td>0.05</td>
<td>0.05</td>
<td>0.1</td>
<td>0.1</td>
<td>0.1</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Strain</th>
<th>STAR</th>
<th>GOLD</th>
<th>DE3</th>
<th>MDM2</th>
<th>DE3</th>
<th>GOLD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temp (C)</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td>18</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Media</td>
<td>Celtone</td>
<td>Celtone</td>
<td>Celtone</td>
<td>Celtone</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IPTG (mM)</td>
<td>1</td>
<td>1</td>
<td>0.1</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Chapter 3: Protein Production and Characterization

110
Table 3.26. Grid of expression conditions screened for the gene coding for MDM4-N. (Previous page) The recombinant protein was produced in different *E. coli* strains (STAR, GOLD and DE3) at different temperatures (Temp (C)) and different IPTG concentrations (IPTG (mM)), in two different medias (TB and Celtone). MDM2-N was also expressed, in randomised conditions, as a positive control.

After gene expression and protein production, the cells were harvested by centrifugation and the cell pellets were lysed and clarified by centrifugation as described in Materials and Methods Section 2.6.1. Subsequently, 1 ml of soluble fraction was loaded into a 96-well plate for high throughput analysis. The Phynexus system (PhyNexus Inc.) provides small-scale chromatographic separation by affinity-capture of recombinant hexaHis-tagged proteins. The soluble fractions were processed with the Phynexus as the manufacturer recommended, and the elution fractions were collected and analyzed by SDS-PAGE.

To analyze the data, the bands from the SDS-PAGE corresponding to MDM4-N were scored with values 1 to 5 depending on their intensity. This exercise revealed that *E. coli* Star is probably (for reasons unknown) not suitable for expression of MDM4-N since no expression was detected in the SDS-PAGE gels. However, the scatter of production levels proved that DE3 and Gold are equally suitable for MDM4-N expression (Fig. 3.27 A). Further up-scaling of the process to 1 L cell cultures showed better yield of purified protein from DE3 over GOLD (Fig. 3.27 B). Finally, the best conditions for production of MDM4-N in *E. coli* DE3 cells were found to be: induction of the culture at 0.6 OD_{600} by 1 mM IPTG o/n at 18 °C (Fig. 3.27 C).
Chapter 3: Protein Production and Characterization

Fig. 3.27. Results of protein production trial. A. Scatter of the scored MDM4-N production yields of *E. coli* strains DE3 vs GOLD. The scatter of MDM4-N production yields from DE3 and Gold do not demonstrate a clear advantage for either strain. B. Analysis of the expression of MDM4-N in *E. coli* strain DE3. Different expression conditions were compared to optimise IPTG concentration and incubation time of *E. coli* DE3 cells transformed with MDM4-N. C. Scale-up of production. After screening for MDM4-N production in 5-ml aliquots, the most suitable conditions were tested in bigger volume of cell culture. Subsequently, expression of 1 L of cell culture and purification of MDM4-N proved that expression of the recombinant MDM4-N was favoured in *E. coli* strain DE3.

<table>
<thead>
<tr>
<th>Strains</th>
<th>T (°C)</th>
<th>IPTG (mM)</th>
<th>Incubation time (h)</th>
<th>Purified protein (mg)</th>
</tr>
</thead>
<tbody>
<tr>
<td>GOLD</td>
<td>18</td>
<td>1</td>
<td>16</td>
<td>1.5</td>
</tr>
<tr>
<td>DE3</td>
<td>18</td>
<td>1</td>
<td>16</td>
<td>2</td>
</tr>
<tr>
<td>DE3</td>
<td>37</td>
<td>0.05</td>
<td>4</td>
<td>-</td>
</tr>
</tbody>
</table>
3.3.3. Protein purification

The first attempt to purify MDM4-N was made by following the methodology described in Kallen et al (2009). This paper, however, neglected to clarify the pH of the buffers, and hence attempts to reproduce their methodology proved very time consuming.

Since MDM2-N purification had proved to be both simple and successful, the same methodology was initially adopted for MDM4-N, but changing the pH of buffer A and B form 8.5 to 7.5, and adding 10% (v/v) glycerol to increase solubility as detailed in Section 2.6.2. Subsequently, following gene expression, protein production and harvest from a 4 L culture, the cellular pellet was re-suspended in lysis buffer and passed through a French press twice as described in Materials and Methods Section 2.6.1. The lysate was clarified by centrifugation and the supernatant was loaded into a 20 ml Ni$^{2+}$-NTA pre-packed column. The protein was eluted with an imidazole gradient and the peak fractions (Fig. 3.28 A) were analyzed by SDS-PAGE (Fig. 3.28 B lanes 1 and 2). In the corresponding peak, fraction 2, two bands appeared at MW of about 11 KDa, consistent with the expected molecular weight as estimated by Proparam (Appendix A1.5). Western blot (Materials and Methods Section 2.4.2) was performed on the samples from the capture step using anti-His tag antibody (Fig. 3.28 C), and MDM2-N with its intact hexaHis-tag as positive control.
Fig. 3.28. First IMAC capture of MDM4-N. A. After clarification of the lysate by centrifugation and filtering, the solution was loaded into a Ni\textsuperscript{2+}-NTA pre-packed column and the protein was eluted with an imidazole gradient. The chromatogram resolved two elution events at 16% and 38% of the imidazole gradient. B. SDS-PAGE of the peak fractions. The peak fractions from the previous step were analyzed by SDS-PAGE. Lane 1 corresponding to peak fraction 1 are mainly *E. coli* contaminants, while the expected recombinant protein appeared mainly at the peak fraction 2. C. Western blot of hexaHis-MDM4-N construct. The Western blot of the peak fractions confirmed that two species of MDM4-N have been co-purified. Detection of the tag by anti-His tag antibody confirmed that the lower band in SDS-PAGE corresponds to a truncated form of hexaHis-MDM4-N. A control was added (C+) which corresponded to hexaHis-MDM2-N construct.
Both bands were positive for hexaHis-tag after performing the western blot suggesting that the lower band is a truncation of hexaHis-MDM4-N. To determine the difference in size, LC-MS was used (see Section 2.4.5), confirming co-existence of two constructs in solution: hexaHis-MDM4-NΔ (9132 Da) and hexaHis-MDM4-N (13185 Da). Thus the truncated construct is 5 KDa smaller than hexaHis-MDM4-N (Fig. 3.29 A). Subsequently, the tag was removed by thrombin cleavage and re-passed through the Ni²⁺NTA column. The flow-through was analyzed by mass spectrometry revealing that the truncation was, as expected, smaller after tag deletion (7381 Da, Fig. 3.29 B). It co-purified with MDM4-N, portending future difficulties to separate the truncated form the intact construct.
Fig. 3.29. LC-MS of species eluted from the Ni-NTA column. A. Two species co-purified. B. LC-MS after His-tag-removal and re-pass over Ni-NTA column. Mass spectrometry revealed that the tag had been removed and the resultant proteins no longer adhere to the resin.

Finally, a gel-filtration step was used to purify the desired product (MDM4-N) from the truncation (MDM4-NAΔ), successfully separating MDM4-N from impurities. Mass spectrometry and SDS-PAGE confirmed the presence of highly purified MDM4-N (Fig. 3.30 B and C), yielding 2.14 mg of protein per L of cellular culture.
Fig. 3.30. Last step of MDM4-N purification. A. The final purification step for MDM4-N was size exclusion chromatography performed as described in Materials and Methods Section 2.6.2. The main peak eluted at 220 ml in accordance with its size. B. SDS-PAGE of peak fractions from size-exclusion chromatography. The peak fractions were analyzed by SDS-PAGE to confirm the size and purity of the peak samples. C. LC-MS of MDM4-N. Finally, the purity of the sample was confirmed by LC-MS, where a single peak corresponding to a molecular mass of 11433 Da was detected confirming that the sample purified is MDM4-N.
3.4. Peptide synthesis

The synthesis of the p53-derived eleven-residue peptide (residues 17-27) was performed by classic solid-phase synthesis [114] (see details in Section 2.10 of Materials and Methods). This methodology relies on the use of aminomethyl-polysytyrene resin providing the support from which the peptide is extended by stepwise addition of amino acids via repeated cycles of de-blocking/activation and coupling reactions. In this procedure, the free amino substituent of the resin is coupled with the activated $\alpha$-carboxyl group of the incoming amino acid forming a peptide bond. This is followed by removal of the 9-fluorenymethyl carbamate (Fmoc) protecting group on the amino terminus of the growing peptide, thus enabling the next coupling reaction (Fig. 3.31). The completion of each cycle may be confirmed by a ninhydrin (2,2-Dihydroxyindane-1,3-dione) test which should be negative after addition of a new amino acid (no change of colour) but positive (blue) after removal of Fmoc.

![Fig. 3.31. Peptide synthesis. Schematic representation of the steps followed in peptide synthesis. 1. The amino substituent of the peptide is protected by Fmoc (red) while the sidechain is protected differently depending upon the nature of the side-chain. The amino acid is activated (green) at the carboxylic end. 2. The acceptor amine moiety is deprotected by removing the Fmoc group. 3. The new amino acid is coupled, forming the peptidic bond. 4. The addition is confirmed by ninhydrin test and the cycle is repeated.](image-url)
Following completion of eleven cycles in the synthesis of the p53-derived peptide the side chain protections were removed and the peptide was cleaved from the resin. The end product was freeze-dried to remove any organic solvent. For further purification, the peptide was submitted to a preparative reverse-phase chromatographic step, and eluted as a sharp peak corresponding to near-homogenous product (Fig. 3.32 A). Subsequently the product was analyzed by MALDI mass spectrometry (section 2.4.5) to confirm its molecular weight and purity (see Fig. 3.32 B).

**Fig. 3.32. p53 peptide purification.** The end product of the peptide synthesis purified by HPLC (top) and confirmed by MALDI (bottom).
Chapter 4: Screening for hits against MDM2 and MDM4

4. Screening for hits against MDM2 and MDM4

4.1. Introduction to screening

4.1.1. The history of screening and some emerging concepts

Blocking protein-protein interactions is an appealing prospect in many therapeutic strategies. The goal is to interfere selectively with target biological processes that depend on specific protein-protein contacts to communicate information [150]. This goal is highly relevant to the development of anti-cancer therapies. Protein-protein recognition is key to the regulation and promulgation of apoptotic pathways. The disruption of the apoptotic signalling cascade is one of the most widespread and effective mechanisms utilised by cancer cells to avoid cell death [151]. In response, the pharmaceutical industry has developed strategies to identify small compounds that act to restore apoptic competency in potentially tumour-forming cells by specifically blocking the relevant protein-protein interactions.

The availability to researchers of high-throughput screening (HTS) techniques and combinatorial chemistry in the 1990s, led to the ability to probe a range of targets using massive compound libraries [152]. Compound collections are based on clusters of chemical entities that share common core structural features but differ in substituent moieties [153]. Typically, HTS involves screening $10^6$ candidate hits against a single protein target in a search for high-affinity ligands. The generally acknowledged low success rates of early implementations of this strategy was due in part to a disregard of emerging structural information, since library design had often been based on historical targets; furthermore, a perceived need to maximise the library size and enrich the portfolio drove the pharmaceutical industry to acquire enormous numbers of drug-like compounds with little regard for coverage of chemical space [154].
Procedures involving HTS are placed towards the beginning of the workflow, enabling identification of “hits” or compounds with good affinity for the target protein. This is followed by characterization of the interaction and analysis of its physicochemical basis and its specificity. Compliance with specificity and physicochemical criteria upgrades a hit to a lead (Hit-to-Lead phase). Next, the lead is optimized in terms of selectivity and potency. Finally clinical trials assess how realistic are the possibilities to achieve a marketable drug [155] (Fig. 4.1).

Fig. 4.1. Drug development pipeline. Cartoon representation of the steps from target identification and hit discovery to drug development, adapted from [156].

Following the disappointing results of early HTS campaigns, retrospective analysis of successful existing drugs prompted the adoption of a set of ground-rules in drug discovery that are now almost universally accepted. The parameters that determine the possibility of a compound being suitable to take directly to the clinic from a screening process were re-assessed. This led to the practice of reviewing libraries of compounds according to their physicochemical properties as encapsulated in Lipinsky’s rule of five [157]. Furthermore, other concepts were redefined and notions of “lead-like” and “drug-like” compounds were clarified [158, 159] (see Table 4.2).

Alternative strategies were developed to re-address the problem of coverage of chemical space. The biggest drawback in HTS is the limitation in the number of distinct chemical entities that can be explored. In order to increase the proportion of the chemical universe screened, compound sizes were reduced to smaller blocks, of 100-250 Da, thus giving rise to “fragment libraries”. This policy both improved the hit rate and decreased the size of libraries that needed to be screened, albeit at a potential cost in terms of affinities [160].
To estimate the quality of hits from a library, a new vocabulary has been established. The quantification of the potential usefulness of hits is addressed in terms of “ligand efficiency” (LE), based on a consideration of affinity in terms of energy of binding per heavy atom [161]. Therefore, a compound with higher ligand efficiency will have a higher probability of being developed into a larger, better binding compound while still observing Lipinsky’s rules. Moreover, a review of fragment leads that had been optimized to produce successful drugs yielded, by analogy with Lipinsky’s rule of five, the rule of three [162] (Table 4.2).

<table>
<thead>
<tr>
<th>Descriptor</th>
<th>Lipinsky (Drug-like)</th>
<th>Actox (Fragment)</th>
<th>Oprea (Lead-like)</th>
</tr>
</thead>
<tbody>
<tr>
<td>MW (Da)</td>
<td>350-500</td>
<td>&lt;300</td>
<td>450</td>
</tr>
<tr>
<td>HBD</td>
<td>5</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>HBA</td>
<td>10</td>
<td>3</td>
<td>8</td>
</tr>
<tr>
<td>CLogP</td>
<td>5</td>
<td>3</td>
<td>4.5</td>
</tr>
<tr>
<td>NROT</td>
<td>10</td>
<td>3</td>
<td>-</td>
</tr>
<tr>
<td>PSA</td>
<td>140</td>
<td>60</td>
<td>-</td>
</tr>
</tbody>
</table>

Table 4.2. Drug-discovery rules. Guidelines have been devised over the years, based on several descriptors, to identify promising compounds during drug discovery efforts. The underpinning data were obtained from reviews of databases of successful drugs in the market place. The goal has been to reduce attrition rates during later stages of drug discovery and clinical trials. A compound is assessed in terms of molecular weight (MW), hydrogen bond donor (HBD), hydrogen bond acceptor (HBA), partition coefficient in octanol/water (CLogP), number of rotatable bonds (NROT) and polar surface area (PSA). According to this strategy, a compound is pursued further in a drug-discovery campaign only if it lies within the parameters for drug-like, lead-like or fragment candidates.

Hand-in-hand with the development of fragment libraries has been the implementation of an “NMR-toolkit” for quick identification of hit-fragments [163]. The [1H,15N] HSQC provides an effective route to relevant hits since chemical shift is sensitive to binding and even weak binders produce detectable perturbations. This ability to detect weak binders is a key advantage of NMR in respect of fragment-based screening. Moreover, the specific effects of binding upon chemical shifts can be easily correlated to structural information with regard to the protein target, the ligand, or both. This engendered the concept of structure-activity relationships (SAR) by NMR [121] (Fig. 4.3). This entails successful identification of fragments
binding to different but adjacent pockets of the target protein by qualitative analysis of an assigned \[^{1}H,^{15}N\] HSQC of the target protein (of known structure) in the presence of fragments. It may then be possible to chemically link the fragments together in an appropriate way and hence create a compound that should bind significantly better than either of the individual fragments [163].

Fig. 4.3. SAR by NMR. Structure activity relationship by NMR methodology. A series of fragments are screened against the target protein, to identify two fragments that bind to different subpockets of the protein. Finally a complex compound is obtained by linking the fragments, whose resulting energy of binding is theoretically equal to the sum of the individual fragments. Adapted from [151]
4.2. The use of a virtually screened library

4.2.1. Introduction and acknowledgments

A virtually screened library (VsL) was kindly provided by Professor Malcolm D Walkinshaw (Centre for Translational and Chemical Biology, University of Edinburgh). Dr Steve Shave (University of Edinburgh) was also involved in this project. He used the program LIDAEUS (for Ligand Discovery At Edinburgh University tool) [164] to dock compounds selected from the large database EDULISS (for EDinburgh University LIgand Selection System), based on various descriptors, into the peptide-binding groove of the N-terminal domain of MDM2 PDB id 1YCQ [165]. The final shortlist of potential ligands were selected by Professor Walkinshaw and Dr Liz Blackburn (University of Edinburgh) and dissolved in dimethyl sulfoxide (DMSO).

4.2.2. Evaluation of compounds

The VsL created for MDM2-N contained 34 small organic compounds in stock solutions arranged according to their solubility in DMSO. The physicochemical descriptors provided by the manufacturing company (Maybridge or Specs depending on availability) were MW (molecular weight), MLogP (Moriguchi octanol-water partition coefficient), HBA (number of hydrogen bond acceptors) and HBD (number of hydrogen bond donors). These descriptors were organized in scatter plots of MW vs MLogP and HBD vs HBA as show in Fig. 4.4. The objective was to decide upon which of the groups within the drug discovery rules each compound belongs helping to decide on the expected $K_D$. The scatter plots show that almost all of the compounds complied with Lypinsky’s rule of five, i.e. they have a MW average of $371 \pm 53$ Da and MlogP values of $3 \pm 0.5$. On the other hand, the values for MW are too high to meet the Astex criteria for fragment libraries (Fig. 4.4 A) but they do fit to the definition of good lead-like compounds (see Table 4.2) [158]. Furthermore, the HBA and HBD satisfy the rule of five (Fig. 4.4, B). Overall, the compounds were mostly within the boundaries of drug-like molecules, and the $K_D$ values expected for them will be in the low $\mu$M range. All the compounds contain aromatic rings and
alkyl substituents and a list of the library with corresponding details can be found in Appendix B.1.

**Fig. 4.4. VsL descriptors.** In order to classify the compounds as drug-like, lead-like or fragments, they were correlated in terms of MW vs MLogP (A) and HBA vs HBD (B). The smaller squares define the area of fragment-like molecules, the middle, lead-like and the larger, drug-like compounds, all as defined by the rules of drug discovery [158]. Although some compounds could qualify as fragments in terms of size and solubility, the hydrogen-bond distribution places most of them within the category of lead-like compounds with the expectation of $K_D$ values in the low µM range.
4.2.3. VsL Workflow

The workflow adopted for this project involved several steps. First, preliminary experiments are needed to estimate compound solubilities in water and stability of MDM2-N under the conditions of the experimental setup (Fig. 4.5, Solubility Test and Controls). Next, the compounds are assessed for their ability to bind MDM2-N, based on a two-point titration monitored by $\left[{\text{H,}}_{15}\text{N}\right]$ HSQC spectra of the $^{15}\text{N}$-labelled protein or binding assay (Fig 4.5). Combining information about solubility and capability of binding helps to exclude some compounds prior to the next stage, which involves a full titration of ligand into the protein sample, monitored by a series of HSCQ spectra as outlined above.

![Fig. 4.5. Workflow.](image)

Fig. 4.5. Workflow. Layout of the plan for screening VsL compounds at Edinburgh University. See text for details.

4.2.4. Controls

A common problem in ligand discovery is that most biophysical techniques for screening for hits require highly aqueous-soluble compounds. Unfortunately, the aqueous solubility of an organic compound is not easily predicted with confidence. This means there is a reliance on experimental data. Therefore, the first step in the screening of the VsL was to determine the solubility of each of its members under the experimental conditions to be applied during the titration. For this purpose, aliquots of the compounds from DMSO stock solutions were added to the aqueous
sample buffer (NMR buffer 1, Materials and Methods Section 2.11.2) resulting in a 5% DMSO (v/v) concentration and two final concentrations of compound at 100 µM and 200 µM. Subsequently a visual inspection was performed to detect and record the presence of any visibly precipitated matter and this was followed by the recording of 1D ¹H-observe NMR experiments on each sample. The intensities of resonance in the 100-µM and 200-µM samples were then compared. Since signal intensity is directly related to the concentration of compound in solution, the ratio in this case should be 2 if there are no solubility issues (Fig. 4.6).

Fig. 4.6. Solubility of c17. The solubility of the components of the VsL was estimated by eye and by [¹H] NMR. The different compounds were dissolved in NMR buffer 1 (see 2.9.2) and 5% v/v DMSO at two different concentrations, and spectra were recorded using a 600-MHz spectrometer. Using the program TOPSPIN (Bruker) the intensities resulting form the recording of the different solutions were superimposed, scaling up the signal from the 100-µM sample so that it matched that of the 200-µM sample. If this ratio worked out to be 2.0 the compound was judged to be fully soluble.

The outcome of solubility tests was that 50% of the organic compounds in the VsL were soluble at the tested concentrations. Most of the remaining compounds did not yield any NMR signal, and this correlated with the cloudiness of the mixture. Others were judged to have reduced solubility on the basis that the intensity of the NMR signals did not correlate with concentrations. All of them were, nonetheless, employed for the binding assay by NMR. This decision was taken on the basis that while poor solubility will likely compromise the assay, and this needs to be
accounted for in data interpretation, any compound remaining in solution is still worth exploring as a potential ligand.

Regarding the target protein, controls were performed to discern the stability of MDM2-N in increasing DMSO concentrations (since DMSO is required as a co-solvent for the potential ligands). Experiments were also undertaken to assess the sensitivity of resonances to very small pH changes that might inadvertently accompany compound addition. The spectral width and number of points in the indirect ($^{15}$N) dimension of the [$^1$H,$^{15}$N] HSQC were also optimized in order to achieve maximum resolution in minimal time under experimental conditions, even though this could result in “folding” of peaks in the spectrum (see details in 2.11.4).

![Fig. 4.4. DMSO titration. To evaluate the effect of DMSO on MDM2-N, [$^1$H,$^{15}$N] HSQC spectra were recorded for the protein under experimental conditions and with 0 (green) and 5 % DMSO (v/v) (red) and overlay using the program ANALYSIS (CCPN). The correlation of the peaks is very good, meaning that DMSO does not affect the structure of MDM2-N.](image)

In addition, a thermal stability assay was performed to elucidate the stability of MDM2-N in the NMR buffer with 5% DMSO v/v. For this, 20 µM of MDM2-N in phosphate buffer with increasing concentration of DMSO were incubated with Sypro orange as described in Materials and Methods Section 2.8.1. The assay relies on the
quenched fluorescence of Sypro orange in aqueous solution while the dye fluoresces in hydrophobic environments. Thermal unfolding of the protein will expose hydrophobic residues buried in the native conformation, causing the dye to bind to them and fluoresce [166]. The melting temperature is identified as the inflexion point in the fluorescence curve, usually represented as the minimum point in the differentially transformed fluorescence plot [167]. Additives or other substances that increase the stability of the protein will increase the melting temperature of the system under study, while substances detrimental to protein stability will effectively decrease the melting temperature of the system. The assay showed that the MDM2-N melting temperature (~55 °C) was not altered by addition of DMSO to 5% v/v (Fig. 4.5), confirming good stability of MDM2-N under screening conditions.

![Fig 4.5. Thermofluor-based stability assay of MDM2-N in DMSO. The melting temperature of the protein in the control conditions (blue stars) was the same when 5% v/v DMSO was present in the solution (red circles).](image)

4.2.5. Binding assay

4.2.5.1. Details of the binding assay

An initial filter of compounds from the VsL was needed to rapidly discriminate binders from non-binders. Consequently, a first screening was performed in an “all-or-nothing” fashion by mixing 20 μM MDM2-N with 200 μM compound in 50 mM phosphate buffer, pH 6.8. A [1H,15N] HSQC spectrum of the apo MDM2-N (Fig 4.6 A) and of the putative complex (Fig 4.6 B and C) were recorded using a 600-MHz spectrometer (Bruker), as described in Materials and Methods 2.11. Overlays of the spectra recorded with and without the potential ligand were inspected to identify
peak movements and/or peak broadening. It is worth noting that changes in the spectrum can arise both from conformational changes and from proximity to a binding ligand.

Chapter 4: Screening for hits against MDM2 and MDM4

130
Chapter 4: Screening for hits against MDM2 and MDM4

Fig. 4.6. First screening of the VsL - examples. A tenfold excess of each VsL compound was added to $^{15}$N-MDM2-N in NMR buffer 1 (see Section 2.9.2) and 5% v/v DMSO. The reference HSQC spectrum of MDM2-N (no ligand) is shown in black (A). Compounds that do not bind to MDM2-N (e.g. c11, spectrum shown in orange B) produced little or no changes in the spectrum, while others (e.g. c17, in green, and c22 in blue C) caused detectable chemical shift perturbations and are considered “hits”.

4.2.5.2. Binding assay results

To avoid over-interpretation and false-positives, several precautions were adopted such as allowing for small changes in temperature, pH or instrumental set-up (e.g. shimming or tuning that will not affect shift but could affect line-width or intensity of peaks). Several isolated, non-overlapped, peaks were selected to assess spectral perturbations upon ligand binding (i.e. chemical shift perturbations (CSP)). The chemical shifts were extracted and $^1$H and $^{15}$N frequencies were combined with a weighting used to allow for their different gyromagnetic ratios as shown in (4.1) and (4.2) [168]
\[
\text{CSP} = \sqrt{(\Delta \delta_H)^2 + (\Delta \delta_N \cdot \alpha_N)^2} \quad (4.1)
\]

\[
\alpha_N = \frac{\gamma_N}{\gamma_H} \quad (4.2)
\]

where $\Delta \delta_x$ is the chemical shift change in the corresponding reference and $\gamma_x$ is the corresponding gyromagnetic ratio as sowed in Appendix C.1.

The selection process was influenced by structural significance i.e. Phe$^{55}$ and Tyr$^{104}$ were observed closely since they are in $\alpha2$ and $\alpha2'$ helix (Fig. 4.7), respectively, forming the walls of the binding site. In addition Phe$^{55}$ chemical shift is affected by the proximity to Tyr$^{66}$, magnifying the CSP effect by ring current. Also, Asp$^{68}$ and His$^{73}$ (Fig. 4.7) are in the loop between $\alpha2$ and $\beta1'$ conforming Phe19 sub-pocket walls. Equally important to the selection was the isolation within the spectrum of a cross peak. The Gly$^{12}$ resonance is a good example of a cross peak that is well isolated from all other cross peaks (Fig. 4.7).

![Fig. 4.7. Residues selected for analysis of CSP. Surface representation of MDM2-N (PDB id 1CYR) with the residues used for residues used for CSP calculations highlighted in red. The picture was obtained using Chimera.](image)
The resultant CSP were averaged to yield a global score for each compound [169] (Fig. 4.8). Thus compounds 16, 17, 21 and 22 were observed to induce significantly greater perturbations than the average (a mean difference of 0.023±0.015 ppm), and are thus considered hits (Fig. 4.8).

![Graph showing CSP per compound and average](image)

**Fig. 4.7. Evaluation of compounds in the VsL for binding to MDM2-N.** Chemical shift perturbations (CSP) are plotted for each of 34 compounds. All the HSQC spectra used to derive these data were recorded in identical conditions, and several controls were included. The weighted, combined amide chemical shifts of residues Gly$^{12}$, Phe$^{55}$, Asp$^{68}$, His$^{73}$ and Tyr$^{104}$ were used to calculate the average MDM2-N chemical shift perturbation for each compound. Hits were then defined as those compounds causing chemical shifts perturbations higher than the average plus one standard deviation.

In addition, other qualities in the spectra were assessed, such as peak broadening or disappearance of resonance. For example, following addition of c29 to $^{15}$N-MDM2-N, the number of amide peaks in the spectrum declined. This implies some form of interaction has occurred but presumably this is on the intermediate timescale leading to broadening out of some peaks (Fig. 4.9).
Fig. 4.9. HSQC spectra of $^{15}$N- MDM2-N in complex with c29. Overlap of HSQC spectra of apo-MDM2-N (in black) and the same sample but with a tenfold excess of c29 (red). The number of peaks decreases when adding c29, indicating that an interaction occurs between compound and protein that leads to intermediate time-scale conformational exchange.

Indeed solubility test by NMR of c29 indicated poor solubility since the intensities have not increased in proportion to the concentration of compound in solution (Fig. 4.10), although no precipitation was observable. Possibly, the compound is forming micelles or suspension when dissolved in aqueous solution at the given DMSO percentage (5% v/v) and it is the transient association of MDM2-N with the micelles that is the source of the conformational exchange.
Fig. 4.10: Solubility test of c29 by NMR. Proton NMR spectra of c29 at concentrations of 100 and 200 µM NMR buffer 1 (see 2.9.2) and 5% v/v DMSO showed that c29 is not fully soluble under these conditions since a scaling factor of 1.22 was required to overlay the spectra.

Titration experiments (see Fig 4.14 and text) revealed that upon addition of c29 \(^{15}\)N MDM2-N in solution, a concentration-dependent disappearance of protein resonances occurred. This is characteristic of an increase of the correlation time (see Fig. 4.11). One possibility is that despite the signal-to-noise test applied in preliminary work described above (Section 4.1.4), c29 does form micelles or (more likely) other heterogeneous aggregates in solution. The resultant particles presumably remain in suspension (despite spinning) in solution and MDM2-N might thus bind to a very high-molecular weight structure causing it to tumble more slowly in solution, with associated line broadening and loss of signal-to-noise.
Chapter 4: Screening for hits against MDM2 and MDM4

**Fig. 4.11. Titration of c29 into MDM2-N.** In the screening series of VsL compounds, the c29:MDM2-N complex had attenuated solubility from an NMR perspective, since the first titration point (1:8 ratio of MDM2-N:c29 – orange) resulted in disappearance and broadening of nearly all HSQC peaks in comparison with the reference HSCQ (black). Successive titration steps (leading to lower ratios because of the reverse mode in which titrations were performed in the current work) proved that c29 does indeed interact with MDM2-N, for peak emerged correlating the c29 concentration.

In addition, while c32 and c28 did not qualify as hits on the basis of the chemical shift perturbation criteria outlined above (Fig. 4.8), obviously detectable peak movements and peak broadening was detected when overlays of the spectra recorded with and without c28 and c32 ligand were inspected (Fig. 4.11). That these effects did not correlate with increasing ligand concentration could be explained by the poor solubility of these compounds in aqueous solution, since visible precipitation was noticed by eye even at 100 µM, and therefore the ratio 1:10 protein to ligand was not fulfilled. In addition to peak broadening, key residues’ resonances were shifted after adding c28 or c32 to the \(^{15}\text{N}-\)protein (Fig. 4.12). One such is Phe\(^{55}\), whose presence
in the $\alpha2'$ helix (Fig. 4.7) makes it very sensitive to conformational changes in the MDM2-N backbone when accommodating a binding partner into the hydrophobic pocket [170].

![Figure 4.12](image)

**Fig. 4.12. Poorly water-soluble potential MDM2-N binders.** The [$^{1}\text{H},^{15}\text{N}]$ HSQC spectrum of MDM2-N (red) was compared to those of the protein with ten time excess of c28 (green) and c32 (blue). Several resonances have been perturbed by the presence of the ligand, including Phe$^{55}$ and Ile$^{19}$ (detail).

The chemical structure of c28 (Fig. 4.13) has a long aliphatic tail, while c32 is rich in aromatic rings, which could explain the poor aqueous solubility. In addition, c32 has a chemical structure that resembles Nutlin-3 (Fig. 4.13), with a scaffold with several phenyl substituent projecting out in the same orientation. By analogy with Nutlin-3, these could be buried in the hydrophobic pocket of MDM2-N while the more polar pyrazolo structure will be solvent exposed. On the other hand, c28 has a methyl-naphthyridine scaffold (Fig. 4.13). The binding mode of c28 is difficult to infer since the structural similarity with Nutlin-3 is not as obvious as with c32. Other, more extended ligands for MDM2-N such as the sulphonamide derivates [171] and isoindolinone [103] could help to understand the binding mode of c28 (Fig. 4.13).
Chapter 4: Screening for hits against MDM2 and MDM4

4.1.3. A comparison of the chemical structures of some MDM2-N ligands. All these ligands have multiple rings. The structural similarity amongst these binders (discovered in this work) versus published ligands, Nutlin-3 and isoindolinone presumably explains their mutual ability to bind to the hydrophobic groove of MDM2-N.

The overall hit rate of the compounds screened was 20% including c28 and c32 that are only vary sparingly soluble. The other binders (c16, c17, c21 and c22) seemingly have no solubility issues. These were therefore carried forward to the next stage.

4.2.6. $K_D$ estimates of the VsL binders by NMR

With regard to the estimation of $K_D$ values by NMR for the interactions between compounds in the VsL and MDM2-N, the need for stock solution of ligands to be dissolved in DMSO was a concern. To accurately estimate the $K_D$ it is necessary to maintain a uniform total DMSO concentration (5% v/v) throughout the titration. It was therefore expeditious to perform the titration steps “in reverse”, i.e. by starting from the highest concentration point (a 16-fold or 8-fold excess of compound over
protein) and then diluting the ligand concentration by adding protein solution in the same buffer (containing 5% v/v DMSO, see Table 4.1).

<table>
<thead>
<tr>
<th>Compound Concentration (mM)</th>
<th>Protein concentration (mM)</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>50</td>
<td>50</td>
<td>1x</td>
</tr>
<tr>
<td>100</td>
<td>50</td>
<td>2x</td>
</tr>
<tr>
<td>200</td>
<td>50</td>
<td>4x</td>
</tr>
<tr>
<td>400</td>
<td>50</td>
<td>8x</td>
</tr>
</tbody>
</table>

Table 4.14. Titration details. See text for details.

Given that resonances in fast exchange were selected for estimations of \( K_D \), as detailed in Materials and Methods Section 2.11.5, values in the low \( \mu \)M range were expected. Consequently, by the end (or beginning if the titration is done in reverse) of the titration (1:8 ratio of MDM2-N:compound), only \( \sim 60\% \) of the protein would be expected to be in complex with ligand, as seen in the graph in Figure 4.15; to obtain a 90% saturation of the protein, and given the \( K_D \) ranges of the compounds, the molar ratio of compound to protein would need to be 35 (see Fig. 4.15 A). Another test to corroborate the \( K_D \) values was therefore deemed desirable. This would obviously need to take account of the aforementioned compound ligand solubility issues. Competition assays have been applied successfully to screening of poorly soluble compounds, and the IC\(_{50}\) values obtained are regarded as acceptable indicators of the binding affinity of a compound. Unfortunately, lack of time prevented to carry them out.

An important frame of reference in drug discovery is provided by the concept of ligand efficiency. As was described above, Ligand efficiency (LE) is a descriptor of the energy of binding of a compound per atom, calculated as free energy of binding (\( \Delta G \)) divided by the number of non-hydrogen atoms (N) (Fig. 4.15 (4.4) and (4.5)). A value of LE above 0.3 Kcal/mol is taken as evidence for a hit that has the potential to
be developed into a good drug and is inherent to compliance with the aforementioned rule-of-five [161, 172]. The average LE for the compounds that were found to bind to MDM2-N in the current study is ~0.2 Kcal/mol per non-hydrogen atom (Fig. 4.15 B). Further development of these hits into drugs would depend upon improving the hit’s potency to achieve a LE of 0.30 Kcal/mol per non-hydrogen atom.

\[
\text{Sat}_i = \frac{\text{Sat}_{\text{max}} ([P] + [L] + K_d)}{2} \sqrt{([P] + [L] + K_d)^2 - (4[P] + [L])} \quad (4.3)
\]

\[
K_d = 200 \mu M
\]

\[
\Delta G = R T \ln (K_d) \quad (4.4)
\]

\[
\text{LE} = \frac{\Delta G}{N} \quad (4.5)
\]

<table>
<thead>
<tr>
<th>Compound</th>
<th>$K_d$ (µM)</th>
<th>SD (µM)</th>
<th>$\Delta G$ (Kcal/mol)</th>
<th>LE</th>
</tr>
</thead>
<tbody>
<tr>
<td>c16</td>
<td>218</td>
<td>100</td>
<td>-5.02</td>
<td>0.2</td>
</tr>
<tr>
<td>c17</td>
<td>200</td>
<td>44</td>
<td>-5.08</td>
<td>0.2</td>
</tr>
<tr>
<td>c21</td>
<td>330</td>
<td>53</td>
<td>-4.78</td>
<td>0.19</td>
</tr>
<tr>
<td>c22</td>
<td>99</td>
<td>17</td>
<td>-5.49</td>
<td>0.22</td>
</tr>
</tbody>
</table>

Fig. 4.15. Estimation of the potential of the hits to become leads. A. The relatively high $K_d$ values typical in ligand screens, combined with limited solubilities of the screened compounds, results in titration curves that may not reach more than 60% saturation. The simulation shown was made using the law of mass action shown in (4.3). B. $K_d$ values for complexes of MDM2-N and compounds from the VsL. The errors are as expected given the failure of the fitted binding curves to reach saturation. The ligand efficiency (LE) was calculated to estimate the potential of the compounds as future drug-leads, by calculating the free energy of binding ($\Delta G$; (4.4)) and normalizing the value by the number of non-hydrogen atoms in the compound (4.5).

4.2.7. Interpretation of NMR data

Note that additional information can be extracted from the NMR-monitored titration exercise. This is a key advantage of NMR over many rival techniques for screening and hit discovery. Provided the HSQC spectrum has been assigned and the protein’s
structure determined, the binding site can be mapped onto the 3-D structure [121]. To do this in the case of MDM2-N, the amide chemical shift perturbations induced by binding were calculated (as explained above) and plotted in histogram format as a function of residue number (Fig. 4.15). The histogram was evaluated to identify the residues undergoing the most substantial ligand-induced chemical-shift perturbations. Chemical-shift movements lower than 0.5 ppm were excluded from consideration during attempts to map the binding surface in order to avoid over-interpretation. For example “allostERIC movements” may occur at sites in the protein other than the binding site [103, 170].

Fig. 4.16. Combined $^1$H and $^{15}$N chemical shift perturbations of MDM2-N amides upon binding to four hits (from the VsL) plotted versus sequence number. This allows identification of the residues whose amides are significantly more perturbed than the average and hence are likely to be the ones that participate most strongly in the binding event. The surface representation of MDM2-N (1CYR) coloured by hydrophobicity (blue hydrophilic residues to orange for most hydrophobic residues with white as a mid-point) and labelled according to secondary structure (see schematic under the plot) is included in the diagram, for reference (picture generated with Chimera).
Examination of the amides that consistently experience the greatest chemical shift perturbations upon binding of ligands, highlighted in Figure 4.16, are residues along $\alpha_2$ and $\alpha_2'$ of MDM2-N which form the walls of the peptide-binding pocket. On the other hand residues within helices $\alpha_1$ and $\alpha_1'$, which have mostly scaffold functions, are perturbed to a lesser extent (see Fig. 4.16). Interestingly, addition of c17 and c22 (but not c16 or c21) provoked significant perturbation in the resonances arising from residues in the flexible N terminus that includes the lid of MDM2-N (Ile$^{19}$ to Glu$^{23}$) [173]. This observation suggested differences, upon binding to c17 and c22 compared to c16 and c21, in the behaviour of the lid region of MDM2-N. Also, based on amide chemical shifts, Tyr$^{100}$ of MDM2-N appears to have a different conformation when c16 and c22 are bound compared to when c17 and c21 are bound. Therefore a pattern seems to emerge involving the role of Tyr$^{100}$ flexibility that is consistent with the ideas discussed in Introduction Section 1.4.1 (and see Fig. 1.12). It seems likely that binding of compounds 16 and 22 to MDM2-N occurs predominantly in the Leu26 sub-pocket, consequently displacing the ring of Tyr$^{100}$ and leading to an “open” conformation. On the other hand, c17 and c21 presumably bind less deeply into the same sub-pocket (Fig. 4.17), and hence perturb Tyr$^{100}$ to a lesser extent.

C16
Fig. 4.17. Differential chemical shift perturbations of MDM2-N Tyr\textsuperscript{100} by four ligands. Addition of c16 and c22 had a bigger effect on the chemical shift of the Tyr\textsuperscript{100} amide during this overlaid titration series from 0 (black) to 8x (blue) of compound to protein, compared to addition of c17 and c21. The cross peaks of Leu\textsuperscript{57} are included for comparison.
4.2.8. **Structural description of soluble binders**

Compounds c16, c17, c21 and c22 have some structurally similarities (see Fig. 4.18). They share a core structural scaffold in the form of triazole with a 3-sulfonyl acetamide group and a phenyl-methyl (in c17) moiety at position 4. At position 5, c16, c17 and c21 each have an N-substituted aminomethyl group, while c22 has an N-substituted 5-(1-aminoethyl) group. The N-substituents on the alkylamino groups are variously para-functionalised phenyl rings (or anilino rings if the substituted amine is considered to belong to this ring); 3-trifluoroanilino in c16, 4-fluoroanilino in c17, 4-chloroanilino in c21 and 4-bromoanilino in c22 (Fig. 4.18). Thus this set of compounds may be considered to belong to the same series, which allows inferences to be made regarding the binding mode of the compounds. A review of the literature undertaken to find similar compounds that act as MDM2-N binders revealed similarity with Nutlins [78].
Fig. 4.18. A series of four MDM2-N binders selected by NMR screening from the VsL. All four share a triazole scaffold with various substituents at the 3, 4 and 5 positions.
Other compounds (c18 and c20, see Fig. 4.19) with identical triazole scaffolds were observed to bind to MDM2-N but only very weakly, based on the lack of chemical shift perturbations during an NMR titration of greater then 0.25 ppm (Fig. 4.7) at the highest ratio of compound to protein. Hence no $K_D$ calculations were attempted. Nonetheless, this poor binding affinity of compounds with limited structural differences to c16, 17, 21 and 22 provides grounds for discussing the mode of binding for hits.

![Fig. 4.19. Depiction of c18 and c20. These very weakly binding compounds share the triazole scaffold with the series of MDM2-N binders in Figure 4.18.](image)

### 4.2.9. Docking

Docking was performed using the program Autodock [174]. The Autodock program explores different conformational dispositions of the compound in the binding site, defines poses and assigns them to clusters that overlay with a root-mean-square deviation $< 2$ Å. This exercise was informed by experimentally solved structures of complexes of MDM2-N with small compounds or the p53-derived peptide. Knowledge of the binding region as indicated by chemical shift perturbations was also utilised. Prior to docking, solved MDM2-N structures were scrutinized to discern the flexibilities of individual MDM2-N residues upon binding small organic
molecules. Several residues have been reported to differ in conformation between the apo- and bound states. This might have a significant effect on affinity and potency. Of particular importance is the aromatic ring of Tyr\textsuperscript{100}, which moves by more than 5 Å to accommodate the p53 peptide [173].

The PDB files for the compounds were generated using a Gaussian energy minimization [175] and the protonation state was simulated using Open Babel [176]. Unfortunately, use of the apo structure of MDM2-N solved by NMR (1Z1M) [177] posed a problem in that the orientations of several residues are incompatible with accommodation of ligands in the binding site. The main goal is to interrupt the p53-MDM2 interaction and therefore it was decided to use for docking the PDB entry 1YCR [165] that corresponds to the crystal structure of the MDM2-N:p53-derived peptide complex. This coordinate file was modified in Molprobity [178] to delete the p53-derived peptide, add hydrogen atoms and confirm good geometry.

Autodock parameters were set to identical values for all compounds to facilitate comparison of results. Default parameters were used with the exception of the size and location of the grid (which defines the three dimensional space of MDM2-N explored in the docking), that was adjusted to accommodate knowledge of the binding site from the literature and from NMR studies (see Appendix B.2). In addition, the Tyr\textsuperscript{100} side-chain was set to be fully flexible. The number of runs can vary depending on the number of torsional degrees of freedom of each compound but in this work all compounds have 7-8 rotatable bonds, and a total of 100 runs per compound was deemed sufficient. Also, to explore the energy landscape thoroughly, for every pose 25 x 10\textsuperscript{6} energy minimizations were performed.

The output of Autodock is primarily a classification of the poses in terms of pose clusters (see Fig. 4.20, histograms) and the mean binding energies for the set of poses in each cluster (see Fig. 4.20, red diamonds). Subsequently, the clusters are ranked on the basis of these average binding energies. The most populated cluster with lower energy of binding is then proposed as the most likely binding pose prioritizing the cluster population over the energy of binding. Of the compounds docked, only c22 has an energetically favoured pose cluster that is more populated than the rest; this is a 60-member cluster with an average of -6.5 Kcal/mol energy of binding (see Fig. 4.20 D). The other compounds yield a less clear-cut result in Autodock; none of
the well-populated pose clusters differed from other well-populated pose clusters by more than 1 Kcal/mol of binding energy. Such differences lie within the resolution of the program; it cannot resolve differences amongst conformers that differ by < 2.5 Kcal/mol binding energy.

A

C16 Number of clusters C16 Mean Binding Energy

B

C17 Number of clusters C17 Mean Binding Energy
Fig. 4.20 Results of Autodock. The docking exercise provides poses for each docked compound. The poses are clustered on the basis of similarity (the size of each cluster is indicated by the histograms) and ranked according to lowest binding energy. The program also outputs the associated mean binding energies in cluster (red diamonds).
4.2.9.1. Analysis of docking results

Clusters of poses, obtained by Autodock, from docking of c16, c17, c21 and c22 into the binding cleft of MDM2-N, were selected on the basis of cluster population and average energy of binding. Poses were also inspected manually to help assess the validity of the results and to compare with the inferences from the NMR chemical shift perturbation results mapped onto the 3D representation of MDM2-N. The PDB files for the complexes of MDM2-N:ligand resulting from Autodock, can be found by following instructions in Appendix B.2.3. The remainder of this discussion focuses on the clusters of poses that have the highest populations.

4.2.9.2. Analysis of c16 Autodock results

Analysis of the results by Autodock showed that “pose cluster 2” (mean binding energy of -5.98 Kcal/mol) has the most members (18 in total) closely followed by conformer number 7 (-5.84 Kcal/mol), with 16 members. Both proposed binding poses were visualised by PYMOL [132], using 1YCR as template and the PDB of the compounds generated by Autodock. The chemical shift perturbations were calculated as described above and used to colour code (from white to blue) the amides most affected by binding (see Fig. 4.19). Both poses bury the 4-phenyl substituent of the triazole in the Trp23 subpocket of the MDM2-N binding groove, while the more hydrophilic triazole scaffold is exposed to solvent. In the cluster 2 pose (Fig. 4.21, green), the p-trifluoromethylphenyl (or anilino) group is buried in the Phe19 subpocket and the 3-sulfonyl acetamide substituent of the triazole is in the Leu26 subpocket. In terms of the position of Tyr\(^{100}\), the pose seen in cluster 2 agrees with a “closed” conformation, which is similar to that screen in apo-MDM2-N; in other words the Tyr\(^{100}\) side chain has not been displaced from the binding groove.
Fig. 4.21. Poses for c16 as predicted by Autodock. The amide $^{15}\text{N}$ and $^1\text{H}$ (combined) chemical shift perturbations were mapped on a surface rendition of 1YCR (minus the p53-derived peptide) using Pymol. Residues not affected by the interaction are in white, while increasing degrees of perturbation are colour coded from light blue to dark blue. The biggest amide chemical shift perturbations occurred in residues close to the hydrophobic groove of MDM2 as previously discussed. The two most populated poses were cluster 2 (green) and cluster 7 (red) (see Fig. 4.20).

This contrasts with the cluster 7 pose (see Fig. 4.21, red) in which the trifluoroethylphenyl group occupies the Leu26 sub-pocket, displacing Tyr$_{100}$ from the pocket, by 4.6 Å, to form an open conformation (Fig. 4.22).
Fig. 4.22. Detail of Tyr<sup>100</sup> side chain displacement upon ligand binding according to the poses created using Autodock. Conformers in both the top-ranking clusters for the complex of c16 with MDM2-N show a displacement of Tyr<sup>100</sup> by 4.6 Å. The sub-pockets on MDM2-N binding groove are labelled as Phe19, Trp23 and Leu26 correspondingly.

Both clusters of conformations (for the complex with c16) generated by Autodock are in agreement with the NMR data. The chemical shift perturbation of Tyr<sup>100</sup> is significant, at 0.061 ppm, but it is difficult to judge without detailed calculations of chemical shift whether this is as large as might be expected if the side-chain is displaced (as predicted by pose 7 rather than pose 2).

The c16 pose represented by cluster 6 (six members, -5.9 Kcal/mol, Fig. 4.20 A) also places Tyr<sup>100</sup> in an ‘open’ conformation (Fig. 4.23). Yet the pose of cluster 6 is opposite to that of cluster 7, with the p-trifluoromethylphenyl moiety buried in the Phe19 sub-pocket, and the 3-sulfonyl-acetamide substituent on the triazole extending into the Leu26 sub-pocket. However, since cluster 7 is more populated than cluster 6, the former is accepted as the more likely mode of c16 binding to MDM2-N.
**Fig. 4.23.** Detail of MDM2-N-binding poses represented by conformers 6 and 7, created using Autodock. The cluster 6 pose (yellow) is consistent with NMR data showing displacement of a Tyr for the Leu26 sub-pocket but its orientation relative to the sub-pockets is the reverse of the pose represented by cluster 7 (red).

Chemical shift perturbation-mapping of those residues most affected by c16 binding (Fig. 4.23, blue) were also compared to the published NMR data for the complex of MDM2-N with Nutlin-3 (Fig. 4.24, gray) [179]. It is apparent that both c16 and Nutlin-3 affect mainly residues in helix α2, β1 and β1’. The effect of Nutlin-3 on residues in the loop connecting α2 and β1’ (Fig. 4.23 α2 and β1’) is higher due to the sterically greater occupancy of the Phe19 sub-pocket by the Nutlin-3.
Fig. 4.24. 

Comparison of the docked structure of MDM2-N: c16 complex (generated by Autodock) versus the X-ray-derived structure of MDM2 in complex with Nutlin-2. The surface of MDM2 was coloured for chemical shift perturbations produced by adding c16 (blue) (this work) or Nutlin-3 (grey, [62]) to MDM2 -N. Nutlin-2 (black) is shown in the binding site since the NMR-derived structure of MDM2 in complex with Nutlin-3 lacks the most flexible/solvent exposed region of the Nutlin-3 and therefore not used in these figures [104]. Structural details (α2 and β1’) are marked on MDM2-N surface for easy identification.

A comparison of the MDM2-N:Nutlin-2 complex crystal structure (1RV1) [180] (See Fig. 4.24, black) with that of the MDM2-N:c16 complex suggested by Autodock cluster 7 (Fig. 4.24, red), reveal similarities; the triazole of c16 is solvent exposed as is the cis-imidazoline of Nutlin-2. The p-trifluoromethylphenyl moiety and the 4-phenyl substituent of the triazole in the c16 cluster 7 pose are orientated in similar ways to the bromophenyl rings of Nutlin-2 (Fig. 4.25). Conversely the sulfanyl acetamide of c16 is buried is in the Phe19 sub-pocket while Nutlin-2 projects its piperazin-2-one and isopropoxy-methyl phenyl groups out into the solvent [104], burying only its ethyl-ether group within Phe19 sub-pocket (Fig 4.25) [78, 104].

To sum up, the data are consistent with c16 binding to MDM2-N in a similar fashion to the Nutlins. The occupation of the sub-pockets is stabilised by hydrophobic interactions of which the most important involves the Trp23 sub-pocket occupied by the 4-phenyl moiety. The triazole on the other hand, is orientated towards the
solvent. Furthermore, the c16 sulfonyl acetamide could form an H-bond with MDM2-N Gln\textsuperscript{72} providing a polarity (or directionality) to the otherwise ambiguous pose (Fig. 4.25) in a similar fashion to that seen for the MDM2-N:Nutlin interaction.

**Fig. 4.25.** Detail of putative H-bond formed between MDM2-N and docked c16. The pose proposed by Autodock, backed up by NMR studies, for c16 includes H-bond formation between the terminal amide of c16 and the carbonyl oxygen of Gln\textsuperscript{72} (A). Nutlin-2 forms a H-bond with the O\varepsilon of Gln\textsuperscript{72} of MDM2-N (B).
4.2.9.3. Analysis of c17 Autodock results

C17 cluster 3 (with 28 members, see Fig. 4.20 B) has a binding energy of -5.74 kcal/mol (Fig. 4.26, green) ranks top for Autodock-generated complexes of c17 with MDM2-N. However, other clusters also have numerous members and must be also considered as good candidates. Of all the high-ranking clusters, only cluster 2 has a closed conformation for Tyr\textsuperscript{100}. This cluster is poorly populated, but it agrees with the NMR data since no chemical shift perturbation for Tyr\textsuperscript{100} occurred upon c17 binding (Fig. 4.26, red). The pose of conformers in cluster 2 does not allow any H-bonds to form with MDM2 but relies mostly on hydrophobic interactions.
Fig. 4.26. Poses of c17 in complex with MDM2-N generated by Autodock. The chemical shift perturbations were mapped onto the surface of 1YCR as described above. Cluster 3 (green) was the most populated, but the open disposition of Tyr^{100} does not match with the lack of chemical shift perturbations for this residue upon binding to c17. Therefore, the pose of cluster 2 (red), suggesting a closed conformation of MDM2-N Tyr^{100} is preferred.
The binding mode of conformers in cluster 2 buries the \( p \)-fluorophenyl (anilino) group in the Phe19 sub-pocket, extending the sulfonyle acetamide towards the Leu26 sub-pocket, while the 4-benzyl substituent of the trizole is buried in the Trp23 sub-pocket, and the more hydrophilic region of the triazole is solvent exposed. This arrangement is the inverse of the pose of Nutlin-2 in the MDM2-N binding cleft, since the more polar region of Nutlin-2 is oriented to the shallower Phe19 sub-pocket (see Fig. 4.27).

**Fig. 4.27. Binding modes of c17 conformer 2 vs Nutlin-2.** The structures were aligned in Pymol. The surface of MDM2-N is colour-coded for chemical shift perturbations produced by c17 (blue) or Nutlin-3 (gray) binding.

The interaction of c17 with MDM2-N is mainly hydrophobic as proposed by Autodock. The extra methylene of the 4-benzyl substituent (compared to the phenyl substituents of the other ligands) allows better occupancy of the Trp23 sub-pocket. While the Phe19 sub-pocket apparently is maintained by stacking between the fluorophenyl (anilino) group of c17 and Try\(^{67} \) of MDM2-N (see Fig. 4.28). The relative low occupancy of the Leu26 sub-pocket is restricted by torsion of the sulfonyle acetamide chain.
**Fig. 4.28 Detail of the interaction c17 with MDM2-N.** The c17 cluster 2 pose could be stabilized upon MDM2-N binding by stacking between the $\text{p-fluorophenyl (anilino)}$ group and the phenyl ring of Tyr$^{67}$.

### 4.2.9.4. Analysis of c21 Autodock results

Cluster number 1 (34 members) has -6.29 Kcal/mol binding energy, while cluster 4 (14 members) has a binding energy of -5.88 Kcal/mol (see Fig. 4.20 C).

In this case, the two most populated clusters correspond to poses that exhibit different occupancies of the Leu26 sub-pocket, and different conformations of the Tyr$^{100}$ side-chain. Cluster number 4 (Fig. 4.29, red) has the $\text{p-chloroanilino}$ buried in the Leu26 sub-pocket and therefore displacing Tyr$^{100}$, while cluster 1 (Fig. 4.29, green) buries 3-sulfonyle acetamide in the Leu26 sub-pocket, and Tyr$^{100}$ has a ‘closed’ conformation. Since the NMR data does not show chemical shift perturbations for the Tyr$^{100}$ amide, it seems likely that the binding pose is better reflected by the conformations in cluster 1.
Fig. 4.29. Poses of c21 bound to MDM2-N suggested by Autodock. Shown are conformers in cluster 1 (green) and cluster 4 (red). The chemical shift perturbations were mapped onto the 1YCR surface as described above.
Mapping chemical shift perturbations onto a surface representation of MDM2-N again shows some similarity in the pattern of affected residues for complexes with c21 and Nutlin-3 (see Fig. 4.30). Comparing the interaction with Nutli-2:MDM2-N complex crystal structure in this case, cluster 1 indicates a higher degree of occupancy of the Phe19 sub-pocket, for it is effectively burying its hydrophobic $p$-chlorophenyl (anilino) ring and minimizing solvent exposure. Nutlin-2 only buries its ethyl-ether group in the Phe19 sub-pocket with lesser occupancy of the available space on the Phe19 sub-pocket. In general, the conformation of bound c21 predicted by Autodock is apparently sustained by hydrophobic interactions, for this pose does not seem to involve H-bond formation with MDM2-N.

Fig. 4.30. Comparison of the binding mode of c21 (Autodock) vs. Nutlin-2. The surface of MDM2-N is colour coded for chemical shift perturbations produced by c21 (blue) or Nutlin-3 (gray).

### 4.2.9.5. Analysis of c22 Autodock results

The top-ranked cluster 2 from the Autodock output for the MDM2-N:c22 complex (-6.52 Kcal/mol, 60 members) (see Fig. 4.20 D) is consistent with a ‘closed’ conformation of Tyr$^{100}$. On the other hand, the Tyr$^{100}$ amide chemical shift was perturbed upon binding c22, which suggests its side chain is displaced and in an ‘open’ conformation. The next most populated clusters were numbers 1 (-6.86
Kcal/mol) and 5 (-6.01 Kcal/mol), with 9 and 10 members respectively (see Fig. 4.20 D). Both of these predict similar open conformations for Tyr$^{100}$, consistent with the chemical shift perturbations (see Fig. 4.31).

**Fig. 4.31 Poses of c22 predicted Autodock.** Chemical shift perturbations were mapped onto the MDM2-N (1YCR) surface as described above. Cluster 2 (green) was the most populated, but the disposition of Tyr$^{100}$ does not agree well with the NMR data. The poses in clusters 1 (orange) and 5 (red) differ but each predicts an open conformation of the MDM2-N Tyr$^{100}$ side chain that is consistent with the NMR chemical shift perturbation data.
However, the poses of conformers in categories 1 and 5 are very different in orientation. In cluster 1 poses, the $p$-bromophenyl (anilino) moiety projects towards solvent, while in cluster 5 poses this group buried within the Leu26 sub-pocket and the sulfon酰 acetamide is orientated towards the Phe19 sub-pocket (Fig. 4.32).

Fig. 4.32 Compound 22 (from VsL) binding to MDM2-N; comparison of poses from clusters 2 and 5. The binding pose is very different for poses 5 (red) and 1 (orange) although both are consistent with NMR data.

The chemical shift perturbation data, mapped onto the surface of the MDM2-N structure, provides additional grounds upon which to decide the preferred propose. The cluster 1 pose is not fully accommodated within the binding groove and therefore this would not explain chemical shift perturbations around the Phe19 sub-pocket. In contrast, the cluster 5 involves quite deep penetration into the binding groove and could be accompanied by small structural adjustments of the groove-flanking residues, consistent with the wide distribution of chemical shift perturbations seen in the NMR spectra when MDM2-N was titrated with c22 (see Fig. 4.32).
Fig. 4.33. **Comparison of c22- and Nutlin-3-induced chemical shift perturbations.** Upon binding to MDM2-N, the chemical shift perturbations associated with c22 binding (blue) are different from those associated with nutlin3 (gray) binding.

One may also compare the chemical shift perturbations induced by c22 (in the current work) with those reported for Nutlin-3. Note that c22 is much smaller than Nutlin-3 and consequently has a lower occupancy of the Phe19 sub-pocket. This appears to be consistent with the differences in chemical shift perturbations for the β1’ region (Fig. 4.33).

Fig. 4.34. **Detail of putative H-bond formed between MDM2-N and docked c22.** The pose proposed by Autodock, backed up by NMR studies, for c22 includes H-bond formation between the terminal amide of c22 and the carbonyl oxygen of Gln^{72}.
In overall, the pose proposed for c22 is very similar to that of c16, sustaining the same occupancy of the sub-pockets, and more interestingly, the H-bond between c22 and MDM2-N (Fig.4.34).

4.2.9.6. Conclusions

In summary, most of the docked compounds share a common property of burying the phenyl (or benzyl) group at the 4-position of the triazole into the Trp23 sub-pocket, while the triazole scaffold itself is oriented towards the solvent. Substitution of a bulkier p-chlorophenyl group for the phenyl moiety at position 4 might increase the affinity of the compounds for MDM2-N by promoting better occupancy of the Trp23 sub-pocket following a similar strategy that increased the affinity of peptides for the MDM2-N binding groove [91].

The lower affinity of c20 might be due to the insertion of a polar ester group at the para-position of the phenyl attached to position 4 of the triazole (Fig. 4.19). Interestingly part of the energy for binding of the p53-derived peptide to the MDM2-N binding groove derives from an H-bond between the epsilon nitrogen of Trp23 of p53 and the MDM2-N carbonyl oxygen of Leu54 (Fig. 4.35). Intuitively, the failure of c20 to achieve this H-bond, thereby failing to recapitulate the p53-binding mode in the deepest sub-pocket, may account for the inability of c20 to bind to MDM2-N.

Fig. 4.35. Detail of the H-bond between Trp23 (from the p53 peptide) and Leu54 (from MDM2-N). See text for more discussion.
The difference in occupancy of the less critical sub-pockets could arise from the bulkiness of the atoms attached to the terminal phenyl group and the associated flexibility of the Leu26 sub-pocket. The conformation of Tyr\textsuperscript{100} modifies the accessibility of the Leu26 sub-pocket, allowing larger groups to be buried, while the Phe19 sub-pocket has more limited space. Thus the large \( p \)-bromo- substituent on the anilino group of c22 may be buried in the Leu26 sub-pocket only upon displacement of Tyr\textsuperscript{100}. The same argument might apply to the \( p \)-fluoromethyl- substituent of c16. The \( p \)-fluoro- of c17 and \( p \)-chloro- of c21 are the smallest substituents and are preferably buried in Phe19.

Regarding c16 and c22, their Autodock-predicted poses look similar with regard to the Phe19 sub-pocket since both are predicted to form an H-bond with Gln\textsuperscript{72} of MDM2-N. Nutlin-2 does something very similar. Clearly, this intermolecular H-bond is important for providing polarity to the pose, and the lack of binding of c18 (as shown by NMR) could be ascribed to the substitution of the sulfonyl acetamide with sulfonyl acetic acid at position 3 of the triazole. This prevents c18 to establish the H-bond with Gln\textsuperscript{72} of MDM2-N.

Consequently, the affinity of MDM2-N for c16 and c22 should be explored further by addition of a bulkier \( para \)-chloro substituent to the phenyl attached to the 4-position of the triazole to optimise its enthalpic contributions to the binding by fulfilling the available space in the MDM2-N Trp23 sub-pocket (Fig. 4.36 \( a \)). It also might be advantageous to explore \( para \) and \( meta \) anilino substituents to achieve better occupation of the Leu26 sub-pocket (Fig. 4.36 \( b \)). Finally reduction of rotatable bonds by introduction of double bonds could induce rigidity to the molecule and this might potentially benefit affinity by reducing entropic penalties upon binding (Fig. 4.36 \( c \)).
Fig. 4.36. Proposed structure of MDM2-N binder. Suggested improvements of the VsL binders would be optimization of enthalpic contributions by exploring other substituents in a, and b. Also, entropic penalties could be reduced by adding double bonds (c) reducing the flexibility of the compounds.
4.3. Fragment library

4.3.1. Acknowledgment and introduction to fragment library

The fragment library (AZ-fragments) under discussion is the intellectual property of AstraZeneca. The structure of several compounds cannot be revealed due to confidentiality issues, and will be omitted from structural interpretations.

The aim of this part of the project was to assess the “drugability” of MDM2-N using NMR as a basis upon which to consider initiating a new hit-discovery campaign. The initial plan also involved assessment of the potential for extrapolating from MDM2-N-binding fragments to MDM4-N-binding ones and hence to judgements of the drugability of MDM4-N.

4.3.2. Workflow

The initial step of the screening process itself consisted of the addition of six fragment mixtures to a 50 µM MDM2-N sample in NMR buffer 1 (see Section 2.11.2). The addition of mixtures saves on data acquisition time as well as on the quantity of labelled protein expended. Subsequent analyses of TROSY spectra allowed identification of cocktails whose members include one or more candidate binders (Fig. 4.37, NMR screening). The candidates were then individually added to a fresh MDM2-N sample to identify the binder(s) (Fig. 4.37, Deconvolution of Mixtures). Note that the individual candidate hits from an identified mixture are added sequentially to the same sample of 15N MDM2-N (to save protein). This practice suffers the disadvantage of only identifying the first binder tested from the mixture, and ignoring any other binders that might have been present.
Chapter 4: Screening for hits against MDM2 and MDM4

A

Fragment library: 162 Compounds

27 Mixtures: 6 compounds

50 μM MDM2

NMR screening: TROSY

15 Mixtures: No significant shifts

12 Mixtures: Significant Shifts

Deconvolution of Mixtures: addition compounds sequentially recording a spectrum each time

Compounds with no effect on protein.

Compounds with effect: titration and Kd calculation.

B

<table>
<thead>
<tr>
<th>Plate mixture</th>
<th>Hit</th>
<th>Deconvolution</th>
<th>Fragment Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>6</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>✓</td>
<td>11</td>
<td>A1</td>
</tr>
<tr>
<td>18</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>24</td>
<td>✓</td>
<td>19</td>
<td>A2</td>
</tr>
<tr>
<td>30</td>
<td>✓</td>
<td>FALSE</td>
<td></td>
</tr>
<tr>
<td>36</td>
<td>✓</td>
<td>36</td>
<td>A3</td>
</tr>
<tr>
<td>42</td>
<td>✓</td>
<td>45</td>
<td>A4</td>
</tr>
<tr>
<td>48</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>54</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>60</td>
<td>✓</td>
<td>64</td>
<td>A5</td>
</tr>
<tr>
<td>66</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>72</td>
<td>x</td>
<td></td>
<td></td>
</tr>
<tr>
<td>78</td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>84</td>
<td>✓</td>
<td>85</td>
<td>A6</td>
</tr>
<tr>
<td>90</td>
<td>✓</td>
<td>92</td>
<td>A7</td>
</tr>
</tbody>
</table>
Fig. 4.37. Workflow for fragment-based screening. A. Schematic representation of the screening plan designed for this project. B. Table showing the work progress on the screening. The hits from the initial screening with mixtures are highlighted in green, while mixtures without hits are in orange. The deconvolution, identified several fragments that were used to calculate the $K_D$ by NMR. The structure of the fragments highlighted in red will not be discussed in this thesis for confidentiality issues.

The screening was performed by recording transverse relaxation optimized spectroscopy (TROSY) on the $^{15}$N-labelled protein samples. In essence, TROSY cross-peaks correlate the amides $^{15}$N and $^1$H nuclei to produce a spectrum that looks very like an HSQC. However, the TROSY pulse sequence overcomes ($T_2$-linked) line-broadening effects characteristic of larger, or self-associating, and therefore more slowly tumbling protein molecules. It was found that TROSY spectra were superior to HSQC spectra for MDM2-N (Fig. 4.38).
A

Chapter 4: Screening for hits against MDM2 and MDM4
Fig. 4.38. $^{15}\text{N} \text{MDM2-N HSCQ vs TROSY}$. Comparing HSQC spectrum of MDM2-N in red (A) and TROSY (B), the latter has better dispersion of peaks.

Out of a total of 162 screened fragments, nine were confirmed to bind to MDM2-N by NMR, for a ~5% hit rate. Of these, two were “out of stock” at AstraZeneca meaning that they could not be further investigated within the time constraints of the project. Of the remainder, fragments A2, A8 and A10 structures cannot be revealed since they are private property of AstraZeneca. The others - A1, A3, A5 and A7 - will be further discussed in terms of their structures below.
4.3.3. Estimates of $K_D$ by NMR

The titrations for $K_D$ estimates were performed according to the range of $K_D$ values expected for fragments (in the mM range). The final point of the titration was set as high as the solubility of the fragments allowed, so as to achieve the highest extent of saturation of the protein.

A

<table>
<thead>
<tr>
<th>Fragment Concentration (µM)</th>
<th>Protein concentration (µM)</th>
<th>Ratio</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td>50</td>
<td>0</td>
</tr>
<tr>
<td>100</td>
<td>50</td>
<td>2x</td>
</tr>
<tr>
<td>300</td>
<td>50</td>
<td>6x</td>
</tr>
<tr>
<td>700</td>
<td>50</td>
<td>14x</td>
</tr>
<tr>
<td>1300</td>
<td>50</td>
<td>26x</td>
</tr>
<tr>
<td>2100</td>
<td>50</td>
<td>42x</td>
</tr>
</tbody>
</table>

B

<table>
<thead>
<tr>
<th>Fragment</th>
<th>$K_d$ (mM)</th>
<th>Error (mM)</th>
<th>LE</th>
</tr>
</thead>
<tbody>
<tr>
<td>A1</td>
<td>3.27</td>
<td>0.37</td>
<td>0.19</td>
</tr>
<tr>
<td>A2</td>
<td>5.77</td>
<td>1.12</td>
<td>0.23</td>
</tr>
<tr>
<td>A3</td>
<td>0.94</td>
<td>0.07</td>
<td>0.32</td>
</tr>
<tr>
<td>A5</td>
<td>2.96</td>
<td>0.43</td>
<td>0.29</td>
</tr>
<tr>
<td>A7</td>
<td>1.75</td>
<td>0.31</td>
<td>0.23</td>
</tr>
<tr>
<td>A8</td>
<td>8.2</td>
<td>2.65</td>
<td>0.16</td>
</tr>
<tr>
<td>A10</td>
<td>2.63</td>
<td>0.66</td>
<td>0.20</td>
</tr>
</tbody>
</table>

Fig. 4.39. $K_D$ values and ligand efficiencies of some compounds in the fragment library. A. Titration series to calculate $K_D$ by NMR. B. The dissociation constant for a fragment:MDM2-N complex was calculated by fitting plots of TROSY-derived chemical shifts versus concentration to equations (2.1) to (2.3) using the “Graph-pad” software as described in Materials and Methods 2.9.3.2. Some of the derived $K_D$ values. The ligand efficiencies (LE) were calculated according to equation (4.5) in Fig 4.15.
A combined chemical shift was calculated for each ligand concentration, as described in Materials and Methods Section 2.11.5. Subsequently, multiple plots (i.e. one plot for each residue followed) of chemical shift perturbation versus ligand concentration were fitted to equations (2.1) to (2.3). Finally, an average was calculated of the $K_D$ values derived from the multiple residues followed (Fig. 4.39).

The $K_D$ values obtained for fragments in the library were in the mM range as expected. Fragment A3 was the tightest binder. The LE scores were low (~0.2) for all fragments except A3, thus implying poor-quality ligands. Attempts to develop them into commercial drugs would be likely to violate the Lispinsky rule-of-five. On the other hand, fragment A3 has a good LE score of 0.3, implying that this fragment has more promise than the others and might be worth taking to the next stage.

### 4.3.4. Structural description of AstraZeneca fragment library binders

The fragments shown in this work to bind to MDM2-N have the desired characteristics of fragments according to the Astex rules. Whereas fragment A5 has a urea core, the core structures for fragments A1, A3 and A7 contain pyrazole, tetrazole and triazole-thione, respectively. These five-member heterocycles are popular for the versatility they offer with regard to expansion of a fragment into a lead [120]. The substituents attached to the core structures are mainly six-membered ring derivatives that could insert into a typical hydrophobic pocket (Fig. 4.40). Interestingly, the fragment-hits identified by screening in the current work share similarities with compounds already known to bind MDM2-N; specifically, they resemble the Nutlins.
4.3.5. Results of screening the fragment library for MDM2-N binders

4.3.5.1. Fragment A1

From inspection of the NMR data it is evident that binding of fragment A1 to MDM2-N does not follow a classical 1:1 model. During titration of MDM2-N with this fragment, the perturbed resonances did not track across the spectrum in the normal unidirectional manner. Instead some residues exhibited an alteration in the responsiveness of their chemical shifts to ligand addition after titration point 4 (i.e. a 26-fold ratio of ligand to protein) (shown in Fig. 4.41). A change in the directionality...
or sensitivity of chemical shift perturbations during the course of a titration often indicates more than one distinct ligand-induced event occurring within the protein or at the binding site [181]. The fact that the MDM2-N sub-pockets are very similar to one another, in terms of both their physicochemical natures and their shapes, is likely to contribute to an increased promiscuity of fragment poses associated with a heightened possibility that two (or more) molecules of a fragment bind simultaneously to different pockets. In the analysis of non-typical perturbed resonances, such as that of A1, it is very difficult to deconvolute the two (or more) contributing events (Fig. 4.41). For these reasons the $K_D$ value calculated for fragment A1 must be treated with caution since it contains information for two binding events; it may be more appropriate to consider it as an apparent dissociation constant.

Subsequently, the chemical shift perturbations caused by fragment A1 were processed as described in Section 4.2.5.2, and mapped onto the surface of the structure of MDM2-N (1CYR) modified by removing the coordinates of the p53-

**Fig. 4.41. HSQC-monitored MDM2-N titration with A1 fragment.** Changes in the directionality of the chemical shift perturbations might be accounted for by a second component to the binding event that affects differently the MDM2-N residues involved. Highlighted with red lines are the movements of resonances that change direction, while in green are highlighted residues that exhibit changes in the line-width or the responsiveness of the perturbation.
derived peptide) to visualise the distribution of the residues involved in the interaction (Fig. 4.42). Those residues whose amides experience chemical shift perturbations > 0.05 ppm (when adding fragment A1) are displayed on the protein surface (Fig. 4.42, red) indicating that A1 prefers these regions for binding to the protein.

Fig. 4.42 Chemical-shift perturbations accompanying interaction of MDM2-N with fragment A1. The $^{15}$N plus $^1$H combined chemical shift perturbations for each residue are plotted as a function of residue number. Mapping of the bigger movers (i.e. > 0.05 ppm) in red on the crystal structure of MDM2-N (1CYR) reveals that most perturbed residues are found in and around the binding cleft. The sub-pockets are also indicated in this figure marked as Phe19, Trp23 and Leu26.
4.3.5.2. Fragment A3

Analysis of the set of MDM2-N TROSY spectra collected during titration with increasing concentration of fragment A3 proved more satisfactory than was the case for A1. The progressive perturbation of resonances as a function of the concentration of A3 is linear, indicating that the binding between A3 and MDM2-N is a simple two-state (bound or unbound) process. In addition, resonances displaying chemical shift perturbations > 0.05 ppm lie in the vicinity of the Trp23 sub-pocket (see Fig. 4.43).

Fig. 4.43 Plot of chemical shift perturbations versus residue number for the interaction of MDM2-N and fragment A3. Mapping of the bigger movers (> 0.05 ppm) in yellow on the surface of MDM2-N (PDB id 1CYR) reveals that they are localised into the binding groove and are particularly well represented in the Trp23 sub-pocket.
4.3.5.3. Fragments A5 and A7

As with the A3 titration, titration of fragments A5 or A7 into MDM2-N resulted in linear changes in chemical shift perturbation as a function of fragment concentration. In these cases, however, residues displaying the largest chemical shift perturbations did not fall into a pattern that helped to identify a distinct binding site (see Fig. 4.44). This could be a consequence of the low affinity of these fragments.

![Diagram of MDM2-N with residues Phe19, Leu26, and Trp23 highlighted.]

![Bar graph showing CSP (ppm) for residues 11 to 115.]

Chapter 4: Screening for hits against MDM2 and MDM4

179
Fig. 4.44 Plots of chemical shift perturbations versus residue number for the interactions of MDM2-N and fragments A5 and A7. Plots representing the chemical shift perturbations of MDM2-N when titrated with fragment A5 (previous page, green) or fragment A7 (black). Mapping of the bigger movers (> 0.05 ppm) onto the surface of MDM2-N (1CYR) surfaces for A5 in green (previous page) and A7 in black.
4.3.5.4. Fragment-based screening of MDM4-N

Those fragments that displayed affinity for MDM2-N were titrated into a sample of $^{15}$N-labelled MDM4-N in NMR buffer 3 and chemical shift perturbations were monitored as described in Section 4.3.3.

As discussed later (see Chapter 5) the conformational flexibility of MDM4-N in solution was readily apparent from its 2-D spectra. Many amide resonances were missing or overlapped (see Fig. 4.45), and therefore the TROSY spectrum could be only partially assigned (see Chapter 5). Consequently, these titrations did not provided as much information as in the case of MDM2-N.

![Fig. 4.45. Titration of fragment A5 into MDM4-N.](image)

Details from two TROSY spectra recorded for apo-MDM4-N (red) and for the last titration point with fragment A5 (blue). The apo-MDM4-N spectra (red) showed a high degree of overlap. This complicated the interpretation.
Given the large amount of cross-peak overlap and the poor assignment of *apo*-MDM4-N, many resonances had to be excluded from the analysis, and the results are therefore based on a relatively small pool of data. Nonetheless, mapping of the chemical shift perturbations onto a surface representation of the MDM4-N structure (PDB 3FE7) showed that the interaction with fragments occur in the vicinity of the binding groove of MDM4-N (Fig. 4.46).

**Fig. 4.46. Comparison of chemical shift perturbations induced by a selection of fragments binding to MDM4-N.** In the surface representation of MDM4-N missing assignments are highlighted in white while those residues affected by ligand binding (chemical shift perturbations above average) are highlighted in red. In green are highlighted assigned residues whose amide resonances are not perturbed by these fragments. Lower panel: Plot of the chemical shift perturbations caused by fragments A1 (black), A3 (cyan), A5 (orange) and A7 (purple) for non-overlapping and unambiguously assigned amide resonances in the [\(^1\)H,\(^{15}\)N] TROSY/HSQC of MDM4-N. The horizontal line (in black, Ave) represents the average for the chemical shift perturbations.
4.4. Overall conclusions

MDM2-N has a very well defined binding groove, but the similarity amongst its sub-pockets is a hindrance to interpretation of screening results. This is because the fragments most likely to bind will be hydrophobic compounds with limited structural diversity. This is evident from inspection of very successful MDM2-N ligands reported in the literature such as the Nutlin-3 and MI-219 [109]. In these compounds, para-halogenated phenyl or chloro-oxindole rings are inserted into the Leu26 and Trp23 sub-pockets. This could readily explain the abnormalities observed in A1 titration, consistent with more than one pose for the fragment within the binding groove of MDM2-N.

Speculation as to the binding configuration of the fragments can be based on a comparison of the structural information available for others, known, binders e.g. from the similarities of A1 with Nutlin-3. The biphasic response of chemical shifts to titration with fragment A1 could, in this light, be the result of two poses that flip due to both pyridine and amino-phenyl substituents of A1, having similar (quite poor) affinities for the two adjacent sub-pockets (Fig. 4.47).
Fig. 4.47. Proposed poses for fragment A1 bound to MDM2-N. The occupancy by Nutlin-3 of binding sites on MDM2-N (above) was used to formulate the hypothesis for flipped poses of fragment A1 (and the biphasic chemical shift perturbation pattern) shown below.

On the other hand, fragments A3, A5 and A7 are very likely to occupy the Trp23 sub-pocket with the corresponding p-chloropheny, p-trifluoromethyphenyl and p-benzyl correspondingly, while the tetrazole substituent in A3 and A7, or the urea moiety in A5, remains solvent exposed, consistent with the several perturbed shifts for residues within the vicinity of this pocket (i.e. Phe55). The differential perturbations occurring in Leu26 and Phe19 sub-pockets (Figs. 4.43 and 4.44) indicate that there is a different element in the binding, and so while A3 and A5 core moieties of tetrazole and urea correspondingly tilt over Phe19 sub-pocket of MDM2-N, the triazole-thione of A7 will towards Leu26 sub-pocket (Fig 4.48).
With regard to the MDM4-N work, it is very clear that the information to be extracted by NMR screening is very limited. The missing assignments hinder the interpretation of the resonances movements that accompany titration with the fragments and deeper interpretation of the data could lead to wrong conclusions. So far, the chemical shift perturbation mapped onto the 3D representation of MDM4 (Fig. 4.6) seems to highlight the association of the compound with Trp23 sub-pocket. But, perturbation of chemical shift of residues Val$^{27}$ and Thr$^{48}$ (Fig. 4.48) are very difficult to justify unless the binding occurs outside the boundaries of p53 binding groove.

Fig. 4.48. Proposed docking for fragments A3, A5 and A7. See text.
Chapter 5: NMR of MDM2/4 – assignment, structure, dynamics

5. NMR of MDM2/4 – assignment, structure, dynamics

Assignment of protein NMR resonances is desirable for NMR-based drug discovery. It allows chemical shift perturbations to be mapped onto the three-dimensional structure of the protein. An extended discussion on NMR theory and pulse sequence development and design is beyond the scope of this thesis; these topics are reviewed elsewhere e.g. [120]. Several concepts are, however, briefly explained below to facilitate understanding of the current study and its outcomes.

5.1. MDM2-N assignment

The interaction of the N-terminal domain of MDM2 with p53 has been studied extensively. Several NMR-based studies of this domain have yielded resonance assignments (Fig 5.1). The Biological Magnetic Resonance Bank (BMRB) contains backbone and side-chain assignments for four proteins of varying lengths that correspond to the N-terminal region (roughly 120 amino acid residues) of MDM2 (see Fig. 5.1). Of these, the two datasets for the apo form of the protein were employed to help re-assign the $[^{1}\text{H},^{15}\text{N}]$ HSQC spectrum collected on the MDM2-N protein used in the current study.
A

<table>
<thead>
<tr>
<th>BMRB id</th>
<th>Author</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>2410</td>
<td>Stoll</td>
<td>Apo-protein pH 7.4; Temp. 300 K [76]</td>
</tr>
<tr>
<td>6612</td>
<td>Uhrinova</td>
<td>Apo-protein pH 7.3; Temp. 288 K [49]</td>
</tr>
<tr>
<td>6248</td>
<td>Fry</td>
<td>Holo-protein pH 70; Temp. 293 K [104]</td>
</tr>
<tr>
<td>15945</td>
<td>Riedinger</td>
<td>Holo-protein pH 6.5; Temp. 293 K [182]</td>
</tr>
</tbody>
</table>

B

2410  MCTMNMSVPTDGVTSQPASEQETLVRKPLLKLKSVGAKDYTEVMKELFVLYQY
6612  MCTMNMSVPTDGVTSQPASEQETLVRKPLLKLKSVGAKDYTEVMKELFVLYQY
MDM2-N  ------GSHMGRGATTSQPASEQETLVRKPLLKLKSVGAKDYTEVMKELFVLYQY
        ------------------------------------------

2410  IMTKRLYDEKQHIVYCSNDLLGDFGVPSFSVKEHRIYTMRIYRNVQESDSST-
6612  IMTKRLYDEKQHIVYCSNDLLGDFGVPSFSVKEHRIYTMRIYRNVQESDSST-
MDM2-N  IMTKRLYDEKQHIVYCSNDLLGDFGVPSFSVKEHRIYTMRIYRNVQESDSGT
        ********************************************

2410  ------
6612  ------
MDM2-N  SVSENLE

Fig 5.1. BMRB entries. A. List of entries in the BMRB related to proteins representing the N-terminal domain of MDM2. B. Alignment of the Stoll (BMRB id 2410) [76] and Uhrinova (BMRB id 6612) [177] sequences along with the MDM2-N sequence used in the current study.

The conditions used to record the NMR spectra of the two constructs, BMRB id 6612 and BMRB id 2410, differed (see Fig 5.1 A). Because resonances are sensitive to conditions such as temperature, salt concentration and pH, the peak lists from the two BMRB entries did not match precisely with one another, and neither spectrum overlaid well with the \[^{1}H,^{15}N\] HSQC acquired in the current study of MDM2-N. For some of the more isolated cross peaks it was relatively straightforward to infer corresponding resonances across the three spectra (see Fig. 5.2, A). In more crowded areas, however, this was not possible (see Fig. 5.2, B).
Fig 5.2. MDM2-N assignment based on previously assigned spectra. Shown are two sections from the \([^{1}H,^{15}N]\) HSQC spectrum of MDM2-N. Also shown are crosses and their alleged assignments corresponding to the \([^{1}H,^{15}N]\) HSQC peak lists downloaded from the BMRB (2410 in blue; 6612 in green). A. Region of the spectra containing isolated cross-peaks. B. Crowded region.
It was decided to use the peak list and assignments deposited in BMRB id 6612 as a template for re-assignment of MDM2-N on the basis that this list is more complete than the one deposited in BMRB id 2410 (a total of 92 NH resonances assigned in BMRB id 6612 vs. 83 NH resonances assigned in BMRB id=2410). A three-dimensional $^{15}$N-NOESY [128, 129] spectrum recorded for MDM2-N, as reported in Section 2.11.6.3, was exploited to assist in the transfer of assignments between proteins. Each “strip” of peaks from the $^{15}$N-NOESY spectrum contains NOEs from all protons closer than ~5.5 Å to the “root” NH group. Thus the side-chain resonance assignments gleaned from the BMRB id 6612 deposition could readily be used to check the assignments inferred in the current work for MDM2-N (Fig. 5.3).
**Fig. 5.3 Use of $^{15}$N-NOESY to confirm reassignment of MDM2-N.** The $^1$H-$^1$H strips correspond to residues 99-104 and 107 of MDM2-N, as indicated (the $^{15}$N shifts are not shown). The horizontal lines connect the strong “diagonal” signal with the weaker sequential $^1$HN$_i$-$^1$HN$_{i+1}$ cross-peak, while the crosses (in red) are the assignment list for BMRB 6612, used to assist the re-assignment of MDM2-N. As expected, there is no NOE between residues 104 and 107.

The MDM2-N construct used in this work, corresponding to residues 11-118 of the mature MDM2 protein, contains 107 amino acid native residues (excluding the cloning artefacts (see Fig. 2.1)) and the four proline residues present in MDM2-N leaves 103 expected backbone amide resonances. Ultimately, only 76 resonances
were unambiguously assigned in the [\( ^1\text{H}, ^{15}\text{N} \)] HSQC spectrum of MDM2-N (see the chemical shift list in Appendix C.2) corresponding to \(~74\%\) of the total. The assignments are summarised in Figure 5.4 and shown on the HSQC spectrum in Figure 5.5. The assigned resonances included all those in the peptide-binding groove of MDM2-N (\( \alpha_2 \) and \( \alpha_2' \)), and should therefore report on occupancy of this binding site by ligands.

Several factors prevented a full assignment of the MDM2-N HSQC. The BMRB id 6612 deposition lacked assignments for several resonances (i.e. 1-6, 105-106 and 111-114). Those regions are located towards the termini of the protein used to acquire the data and missing resonances (in both the previously collected spectra and the ones collected in the current work) presumably arise primarily due to conformational flexibility on an intermediate timescale. More importantly, the pH of the NMR buffer for MDM2-N was set to 6.8 (found to be optimal, data not shown), while a pH value of 7.2 had been used for recording the spectra in BMRB id 6612. This difference in pH was sufficient for some resonances to migrate too far from their positions in the original spectra for them to be traced easily and reliably in the new spectra, even with the help of the NOESY data. These were not pursued exhaustively since sufficient assignments had been made to enable subsequent ligand-screening experiments.
5.2. MDM4-N assignment

5.2.1. MDM4-N optimization NMR conditions

With the objective of assigning MDM4-N resonances from scratch it was important to optimise sample conditions for NMR data acquisition. The aim was to ensure good resolution and maximise protein stability while observing the technical requirement of NMR (and particularly of cryoprobes) for low conductivity (equating roughly to ionic strength) and also maintaining neutral or acidic pH to minimise amide proton exchange with aqueous solvent [122]. Given these multiple requirements and that MDM4-N is notoriously unstable, a screening assay was developed to identify the most suitable conditions.

It was initially observed that MDM4-N displayed a tendency to aggregate and precipitate in a range of conditions. Concentration of MDM4-N to 1 mM in the presence of 10% (v/v) glycerol showed visible signs of aggregation. This ruled out
the use of standard high-throughput automated procedures, as used routinely to screen for conditions favouring crystallisation, since these require ~10 mM stock solutions of protein. For the current work a thermofluor-based stability assay was performed, as described in Section 2.8.1, to explore the stability of MDM4-N in various buffer conditions. Triplicate data points were recorded over a range of pH values from 4.1 to 8.2 (using appropriate buffers) and in the presence of two different salts - NaCl and (NH₄)₂SO₄ – at concentrations of 10 mM versus 155 mM (the overall range of conditions is summarized in Table 5.6).

<table>
<thead>
<tr>
<th>Buffers</th>
<th>PH ranges</th>
<th>Salt concentration</th>
</tr>
</thead>
<tbody>
<tr>
<td>MES</td>
<td>4.4-6.4</td>
<td>10, 155 mM</td>
</tr>
<tr>
<td>NaCit</td>
<td>5.8-7.8</td>
<td>10, 155 mM</td>
</tr>
<tr>
<td>NaPhos</td>
<td>5.8-7.8</td>
<td>10, 155 mM</td>
</tr>
<tr>
<td>MOPS</td>
<td>6.2-8.2</td>
<td>10, 155 mM</td>
</tr>
</tbody>
</table>

Table 5.6. Buffer conditions explored for MDM4-N NMR data collection. List of the range of buffers and salt concentrations used to screen for the most suitable buffer for MDM4-N NMR studies. For each line in the table, both NaCl and (NH₄)₂SO₄ were explored at 10 mM and at 155 mM.

To analyze the results, the apparent melting temperatures (Tₘ) of MDM4-N under the various conditions were graphed as scatter plots to identify trends (see the panels of Fig. 5.7). Inspection of such plots clearly suggested that MDM4-N benefits from high salt concentrations on the basis that Tₘ values were higher at 155 mM salt, with this trend being more marked for (NH₄)₂SO₄ than for NaCl (Fig. 5.7 B).
Fig 5.7. Salt dependencies of apparent $T_m$ for MDM4-N. Scatter plots of MDM4-N $T_m$ when the protein is in NaCl (A) or (NH$_4$)$_2$SO$_4$ (B)-containing buffers at different concentrations (10 mM, 155 mM) showed that MDM4-N is more stable in buffer containing higher salt concentration.
From the pH perspective, the plot shown in Fig 5.8 suggests better stability of MDM4-N at higher pH, with the trend showing an increase of $T_m$ at pH values $> 6.2$.

**Fig 5.8. MDM4-N, pH-stability assay.** Plot of MDM4-N $T_m$ over a range of pH values with different salts in the buffer (NaCl shown in blue; (NH$_4$)$_2$SO$_4$ in black). The trend suggests that MDM4-N is more stable at around neutral pH than low pH.

In general, all the collected data suggested MDM4-N was more stable in 150 mM salt, preferably in (NH$_4$)$_2$SO$_4$, at a pH above 6.2. But while the thermofluor-based stability assay reports on the stability of the protein under study, it does not indicate the aggregation state of the protein in solution. Therefore dynamic light scattering (DLS) measurements were performed on 50 µM and 100 µM MDM4-N in 150 mM (NH$_4$)$_2$SO$_4$, pH 6.5 (Fig. 5.9 A) or 150 mM NaCl, pH 7.2 (Fig 5.9 B), as described in Section 2.8.2. Both plots in Figure 5.9 show evidence of very extensive self-association with particle sizes around 300 nm – a single protein molecule would have a diameter of about 5 or 6 nm. The size distribution profile (*i.e.* hydrodynamic radius according to percentage total intensity) of MDM4-N in 150 mM (NH$_4$)$_2$SO$_4$ was very
broad, suggesting the presence of a mixture of aggregates in solution (Fig 5.9 A); on the other hand a tighter distribution of MDM4-N particle sizes, centered on 100 nm was observed in 150 mM NaCl (Fig 5.9 B).

Fig 5.9. DLS-derived histogram showing particle-size profiles of MDM4-N in two different buffers. A. 150 mM (NH$_4$)$_2$SO$_4$, pH 6.5. B. 150 mM NaCl, pH 7.2.
The previous experiment defined the best buffer conditions for MDM4-N and these were used in initial NMR experiments. Thus, a $[^{15}\text{N},^1\text{H}]$ HSQC spectrum of 50 µM MDM4-N in 150 mM NaCl at pH 7.2 showed relatively good-quality spectra. To increase the number of detectable cross-peaks, more acidic conditions were explored with a view to slowing down the amide exchange with solvent proteins or deuterons. This comparison is clear in the overlay of HSQC spectra of MDM4-N at pH 6.5 (Fig. 5.10, cyan) and pH 7.2 (Fig 5.10, green). More and better resolved, sharper, peaks were obtained for MDM4-N at the lower pH.

![Figure 5.10](image)

**Fig 5.10. HSQC MDM4-N at different pH.** HSQC of 50 µM $^{15}$N MDM4-N at pH 6.5 (cyan) and pH 7.2 (green).

Finally, to improve the resolution of the NMR spectrum, the $[^1\text{H},^{15}\text{N}]$ HSQC experiment was compared with a TROSY-based experiment. Typically molecules with high molecular weight (or molecules that are self-associating) tumble slowly in solution and events such as dipole-dipole interaction, and chemical shift anisotropy provide a source of rapid $T_2$ relaxation. This effect manifests as broad peaks, and poor spectral resolution. The TROSY-based approach exploits the mutual cancellation of dipole-dipole interactions and chemical shift anisotropy that occurs in one of the components of the multiplet that arises, due to J-coupling, in a non-decoupled $[^1\text{H},^{15}\text{N}]$ HSQC experiment [120]. Overlay of the MDM4-N TROSY (Fig. 5.11, red) with the MDM4-N HSQC (Fig. 5.11, green) shows that the TROSY yields
superior results in the case of MDM4-N under the optimised conditions tested. For example comparing central regions of the spectra (Fig. 5.11: [\textsuperscript{1}H] 7.5-8.5 ppm; [\textsuperscript{15}N] 115-125 ppm) shows significantly less overlap in the TROSY spectrum compared to the HSQC spectrum. That an improvement is obtained with TROSY suggests that MDM4-N is undergoing transient self-associations and therefore “behaving” like a larger protein molecule. This is in agreement with the DLS data.

![MDM4-N TROSY vs MDM4-N HSQC](image)

**Fig. 5.11. MDM4-N TROSY vs MDM4-N HSQC.** Two consecutive heteronuclear 2-D NMR experiments were recorded for 50 µM of \textsuperscript{15}N MDM4-N in the same buffer. The TROSY (red) experiment proved better resolution that the HSQC (green) experiment. The apparent “off-set” of the TROSY spectrum (when compared to the HSQC) is due to the selection of the singlet where both sources of \( T_2 \) relaxation are mutually cancelled (dipole-dipole interactions and chemical shift anisotropy). The spectra were collected at 300 K. For the TROSY experiments: number of scans = 8; number of points = 64. For the HSQC experiments, number of scans = 8; Number of points = 1024. For more details, please refer table 2.10.

Further work confirmed the best buffer conditions for recording TROSY spectra on MDM4-N. It was eventually decided to use 50 mM Bis-Tris, pH 6.5, with 100 mM NaCl, 2 mM TCEP, 0.1 mM EDTA and 0.1% v/v NaN₃ (NMR buffer 3). A maximum protein concentration of 100 µM was maintained due to the tendency to
aggregate. Such a low concentration necessitates the use of a cryoprobe, with an attendant need to keep to low salt concentrations (highly conducting solutions reduce cryoprobes sensitivity).

5.2.2. MDM4-N backbone assignment

5.2.2.1. Preliminary work by NMR

The sequence of MDM4-N is 102 residues if the cloning artefact left over after thrombin cleavage (Fig. 2.3) is included. The native sequence of interest consists of 98 residues, five of which are prolines, leaving 93 expected backbone amide peaks in a [\(^{1}H,^{15}N\)] TROSY experiment. An initial inspection of the heteronuclear TROSY spectrum of apo-MDM4-N revealed only 66 peaks (Fig 5.10). This was an early indication of the challenging task of studying the apo-protein, and therefore alternative strategies were sought.

Fig 5.12. MDM4-N TROSY 1. The spectrum was recorded on 100 \(\mu\)M \([^{13}C,^{15}N]\) apo-MDM4-N in NMR buffer 3 (section 2.11.2).
Considering that all the existing structures solved for the MDM4 N-terminal domain were in complex with ligands, samples of MDM4-N in complex with the p53-derived peptide (17ETFSDLWKLLP27) and MDM4-N in complex with Nutlin-3 were considered. To help design these experiments, the $K_D$ of these interaction were estimated by isothermal titration calorimetry (ITC). The experiments were carried out as described in Section 2.9 of Materials and Methods. The $K_D$ values were then calculated by inverting $K_A$ (Fig. 5.13) and found to be $1 \pm 0.1 \mu M$ and $39 \pm 8.1 \mu M$ for the MDM4-N:p53, and MDM4-N:Nutlin-3 interactions, respectively. Given the poor quality of the data obtained from ITC experiments on the MDM4-N:Nutlin-3 interaction (see Fig 5.13, B), the latter value has to be regarded as an approximation. The poor quality could be ascribed to the poor solubility of Nutlin-3 in aqueous solution (Note: the buffer contained only 2% DMSO).
Chapter 5: NMR of MDM2/4 – assignment, structure, dynamics
Fig 5.13. MDM4-N, ITC results. Raw data and integrated peaks (binding curve) obtained after ITC experiments carried out as detailed in section 2.9 A. The p53-derived peptide was titrated into MDM4-N. B. Nutlin-3 was titrated into MDM4-N. Analysis of the binding curve provides values for the number of binding sites (N), binding affinity ($K_a$), enthalpy ($\Delta H$) and entropy ($\Delta S$) of the binding. The $K_D$ can be then calculated by inverting $K_a$. 

Chapter 5: NMR of MDM2/4 – assignment, structure, dynamics
Addition of 200 μM non-labelled p53-derived peptide (\(^{17}\text{ETFSDLWKLLP}^{27}\)) to 100 μM \(^{15}\text{N}\) MDM4-N brought about an increase in the number of observable peaks in the TROSY spectrum, consistent with the 93 resonances expected from the protein sequence (Fig. 5.14).

**Fig 5.14.** MDM4-N TROSY II. The spectrum was collected on 100 μM \([^{13}\text{C},^{15}\text{N}]\) MDM4-N to which had been added 200 μM p53 in NMR buffer.

A TROSY spectrum was also recorded for MDM4-N mixed with Nutlin-3, yielding about 80 cross-peaks (Fig. 5.15). Attempts to calculate the \(K_D\) by NMR were hindered by the slow exchange rate, from the NMR point of view, of the interaction between MDM4-N and Nutlin-3. Titration of increasing concentrations of Nutlin-3 into \(^{15}\text{N}\)-MDM4-N and subsequently recorded \([^{15}\text{N}-^1\text{H}]\) TROSY spectra showed that the 80 cross-peaks present at 1:1 Nutlin-3:MDM4-N (as shown in Fig. 5.15A (blue)) did not change at a 2:1 Nutlin-3:MDM4-N ratio. Therefore, initially Nutlin-3 was added to a final concentration of 200 μM into a solution of 1.5 ml 50 μM \(^{15}\text{N}\)
MDM4-N in NMR buffer 3, \textit{i.e.} a final ratio 4:1 Nutlin-3:MDM4-N. This mixture was then concentrated ~two-fold using a 10,000-MWCO Centricon (Vivaspin) to 100 µM final concentration of $^{15}$N-MDM4-N. The sample was then transferred to a microcentrifuge tube, and again Nutlin-3 was added to a final concentration of 200 µM. The sample was spun down to remove any precipitate and used to record the TROSY spectrum shown in Fig. 5.15, B.
5.2.2. Assignment MDM4-N in complex with p53 derived peptide

Since MDM4-N with p53-derived peptide in excess provided the best signal-to-noise ratio, this sample was used for subsequent backbone assignment. For the assignment several pairs of TROSY-type pulse sequences were adapted for standard triple resonances experiments (see Materials and Methods Section 2.11.6).

Sequential backbone assignment was initiated by using HNCA, HNCCAB and HN(CO)CACB, which correlate backbone NH resonances of residue $i$ with $\text{C}_\alpha$ and $\text{C}_\beta$ nuclei. While HNCA and HNCCAB spectra contain cross-peaks corresponding to $^{13}\text{C}$ of both $i$ and $i-1$ residues, HN(CO)CACB contains only carbon resonances of $i-1$ residues. Subsequent overlay of the spectra therefore helps to distinguish between $i$ and $i-1$ resonance and to establish connectivity of the residues as shown in Figure 5.15.
5.16. Finally, the HNCO and HN(CA)CO pair of experiments was employed to assign the corresponding CO and confirm the reliability of the assignment.

Fig. 5.16 Sequential assignment of MDM4-N in complex with p53. Example of strips as used for backbone assignment of MDM4-N-p53 derived peptide. Detail of HNCACB analyzed with CARA as the Cα peaks (green) or Cβ peaks (blue) are linked (black line) to establish connections between residues.
This approach ultimately yielded a near-fully assigned TROSY spectrum for MDM4-N in complex with p53 as shown in Fig. 5.17. The complete peak list can be found in BMRB id = 16900.

Fig 5.17. Assigned TROSY spectrum of $[^1\text{H},^{15}\text{N}]$MDM4-N in complex with p53-derived peptide. A. The assignment of resonances in a TROSY spectrum to specific residues of MDM4-N, in the presence of an excess of p53-derived peptide. B. Details of the central region ($[^1\text{H}]$ 8.8-7.2 ppm and $[^{15}\text{N}]$ 124-114 ppm) of the spectrum above.
5.2.2.3. Assignment of apo-MDM4-N and in complex with Nutlin-3

An assignment of apo-MDM4-N based on direct transfer of assigned resonances from previously assigned $^1$H,$^{15}$N spectra was not possible. NMR titrations in which increasing concentrations of (unlabelled) p53 were added to $^{15}$N MDM4-N proved that the interaction of MDM4-N with p53 peptide is in the slow-exchange regime in the NMR timescale as expected for a tight binder. Therefore, depending on the ratio of peptide:protein, resonances arising from the same residue may be observed simultaneously for both apo-MDM4 and MDM4-N bound in complex with p53 (Fig 5.18 A), albeit with different intensities (Fig 5.18 B).

![Fig 5.18. Titration of p53 into $^{15}$N MDM4-N. A. Titration of increasing concentrations of p53 into $^{15}$N MDM4-N was monitored by acquiring a series of TROSY spectra. Upper panel: Illustration of slow exchange based on an isolated peak, in this example Met$^{46}$, for which the amide resonance assignment could be confidently transferred form MDM4-N in complex with p53 (blue) to the apo-MDM4-N (red). Lower panel: The peak volume for the amide cross peak of Met$^{46}$ was plotted against the ratio of p53 to MDM4-N; red = apo, blue = bound.](image-url)
Therefore, titration of increasing concentrations of non-labelled p53 into $^{15}\text{N}$-MDM4-N caused some cross-peaks to disappear and new cross-peaks to appear. In many cases these could not be traced to the original apo-MDM4-N TROSY spectrum (Fig. 5.19).

Fig 5.19. Overlay of TROSY spectra for apo-MDM4-N vs MDM4-N in complex with p53-derived peptide. Assignments could not be transferred confidently between the fully assigned TROSY spectrum of the MDM4-N:p53-derived peptide complex (blue) and that of apo-MDM4-N (red).

Subsequently, HNCA, HNCACB and HNCO triple-resonance NMR experiments were recorded for apo-MDM4-N and for MDM4-N in complex with Nutlin-3 (see details in Section 2.11.6). While amide shift assignments between ligand-bound and apo-forms of MDM4-N had not been transferrable (amides are highly sensitive to the chemical environment), $^{13}\text{C}$ assignments proved, as expected, to be easier to translate from the assigned complex (with p53-derived peptide) to the apo- and Nutlin-3-bound forms. Fig. 5.21 shows the amide assignment of apo-MDM4-N and for MDM4-N in complex with Nultin-3, and the complete peak list can be found at the BMRB whose ids are 16893 (apo MDM4), 16894 (MDM4-Nutlin-3 complex).
Chapter 5: NMR of MDM2/4 – assignment, structure, dynamics
Fig 5.20. Assigned TROSY spectra of $^{15}$N-labelled apo-MDM4-N and of MDM4-N in complex with a racemic mixture of Nutlin-3. A. Assigned TROSY spectrum of apo-MDM4-N (red). B. Assigned MDM4-N in presence of excess of Nutlin-3 (black). A detail of the middle region for each spectrum is highlighted ([$^1$H] 8.8-7.2 ppm and [$^{15}$N] 124-114 ppm).
Finally, the extent of the assignment for *apo*-MDM4-N and for MDM4-N:Nutlin-3 complex is shown in Table 5.21.

<table>
<thead>
<tr>
<th>Atom</th>
<th>Number Assigned</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>MDM4-N Apo</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td>54/98</td>
<td>55.1%</td>
</tr>
<tr>
<td>Cα</td>
<td>84/98</td>
<td>85.6%</td>
</tr>
<tr>
<td>Cβ</td>
<td>61/90</td>
<td>67.8%</td>
</tr>
<tr>
<td>NH</td>
<td>66/93</td>
<td>70.9%</td>
</tr>
<tr>
<td>MDM4-N Nutlin3</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CO</td>
<td>71/98</td>
<td>72.5%</td>
</tr>
<tr>
<td>Cα</td>
<td>98/98</td>
<td>100.0%</td>
</tr>
<tr>
<td>Cβ</td>
<td>75/90</td>
<td>83.0%</td>
</tr>
<tr>
<td>NH</td>
<td>80/93</td>
<td>86.0%</td>
</tr>
</tbody>
</table>

Table 5.21. Assignment of *apo*-MDM4-N and MDM4-N in complex with Nutlin-3. Only partial assignments could be achieved for these species due to the poor quality of spectra and in particular the loss of cross-peaks relative to the MDM4-N:p53-derived peptide complex.

5.3. MDM4-N chemical shift analysis

5.3.1. Introduction

The variation in the number of amide signals observed in NMR spectra of the different MDM4-N samples (*apo*, or complexed with peptide or Nutlin-3) is likely due to line-broadening effects. Numerous conformations inter-converting on the ms-timescale probably account for the lack of observable signal in *apo*-MDM4-N. That more signals appear in the TROSY spectrum when the MDM4-N is in complex with Nutlin-3 or p53 is consistent with many regions of *apo*-MDM4-N being conformationally mobile in the *apo* state. Interestingly, the amide signals missing in the *apo*-MDM4-N TROSY spectra are mostly those arising from the segment of the
protein between α2 and α2’ helices (see Fig. 5.22, red), which is composed mostly of β-strands connected by loops forming the walls of the Phe19 sub-pocket (Fig 5.22). However, α2 and α1’ residues are also missing from the apo-MDM4-N TROSY spectrum. This is in agreement with observation that the highest crystallographic backbone atom B-factors [65, 81] occur in β3-α1’ and α’2-β2’, drawing attention to the flexibility of the Phe19 sub-pocket.

Fig 5.22. Missing MDM4-N amide resonance assignments. Upper panels: Surface representation of MDM4-N with underlying secondary structure labelled (in left-hand view). The Phe19 sub-pocket is also labelled. Lower panel: Summary of secondary structure. In red are highlighted those residues for which amide assignments are missing in the case of apo-MDM4-N.

Addition of Nutlin-3 to MDM4-N caused extra peaks to appear in the TROSY spectrum, but the total number of signals still did not match to the expected 93 amides of MDM4-N; several residues presumably remained conformationally mobile. Mapping of those missing assignments on the structure of the protein (Fig. 5.23) indicates that MDM4-N gains rigidity upon Nutlin-3 binding in the region β3-α1’-β2’, but that the connecting loops remain flexible. On the other hand, resonances
in α2 and α2′ helices disappear in the transition from apo-MDM4-N to the Nutlin-3 complex. These residues form parts of the walls of the binding groove.

Fig. 5.23. Amide resonances missing from the assignment of the MDM4-N:Nutlin-3 complex. Surface representations of MDM4-N revealing the underlying secondary structure (upper); and secondary structure summary (lower). The secondary structure and the Phe19 sub-pocket are labelled in the left-hand view. In blue are highlighted those residues for which amide resonances are missing from the assignment of the MDM4-N:Nutlin-3 complex.

Overall, it is clear that the Phe19 sub-pocket is very flexible in apo-MDM4-N and to a smaller extent in the Nutlin-3 complex when comparing Fig. 5.22 with Fig. 5.23. Surprisingly, in the MDM4-N:Nutlin-3 complex α2 appears to be in a disordered state.

5.3.2. $^{13}$C and $^{15}$N shift perturbations upon ligation of MDM4-N

To further understand the flexibility and ligand-binding properties of MDM4-N, both the $^{15}$N and $^{13}$C chemical shift perturbations were analyzed. Changes in the chemical shift of $^{13}$Cα and $^{13}$CO arise mainly from changes in secondary structure and backbone dynamics whereas perturbations of $^{13}$Cβ are more likely to reflect side-chain readjustment and the proximity of the ligand. The chemical shift perturbations
for each assigned carbon was adjusted by its gyromagnetic ratio and then a list was processed to provide a set of normalised values. The extent of the analysis of the data is hindered by missing resonances, e.g. from residues 86-94.

### 5.3.2.1. apo-MDM4-N vs MDM4-N:Nutlin-3

When comparing chemical shifts of *apo-MDM4-N* versus MDM4-N:Nutlin-3, the largest $^{13}$C$_{\alpha}$ resonance differences occur in the vicinity of $\alpha$2 (res. 30-40), $\beta$3 (res. 73-75) and $\alpha$2’ (res. 95-105) as may be seen in Figure 5.24 A, consistent with the reorganization of the binding site upon binding. The $^{13}$CO perturbations are generally small while the observable $^{13}$C$_{\beta}$ chemical shift perturbations are in the vicinity of 0.1 ppm. By way of comparison the $^{15}$N,$^1$H combined chemical shift perturbations (Fig 5.24 B) occurred, as expected, mainly in the vicinity of the binding groove, $\alpha$2 and $\alpha$2’, while $\beta$3 resonances are lacking for apo-MDM4-N and the perturbation could not be estimated.
Fig 5.24. Chemical shift perturbations, MDM4-N apo vs MDM4-N Nutlin-3. Plot of A, $^{13}$C chemical shift perturbations and B, $^{15}$N chemical shift perturbations, for MDM4-N in complex with Nutlin-3 compared to apo-protein.
It may be inferred that $^{13}$C$\alpha$ chemical shift perturbations of residues in $\alpha$2 and $\alpha$2' arise from structural reorganization of the binding site of MDM4-N upon binding to Nutlin-3. In particular the Met$^{53}$, Gly$^{57}$ and Met$^{61}$ residues of $\alpha$2 experience the largest perturbations (Fig. 5.25). The $^{13}$C perturbations in the $\beta$3 region are to be expected, with the most perturbation observed in the case of Tyr$^{75}$, for this goes from disordered in apo-MDM4-N (lacking amide signal) to ordered state when MDM4-N is bound to Nutlin-3. Also, the chemical shift perturbation of Gln$^{71}$ $^{13}$C$\alpha$ is significant, for this is a key residue in the MDM4-N p53 interface [65] since it establishes an H-bond with CO of Phe19 in the peptide.

It is interesting that both $^{13}$C (Fig.5.24 A) and amide chemical shifts (Fig. 5.24 B) of Tyr$^{99}$ exhibit perturbations upon binding; this indicates that Tyr$^{99}$ is affected by the interaction with Nutlin-3 and contradicts some previous reports [64]. The data here presented clearly shows a reorganization of the Leu26 sub-pocket of MDM4-N (Fig 5.25) as a consequence of interaction with Nutlin-3.

![Combined chemical shift perturbations of apo-MDM4-N vs Nutlin-3 complex.](image)

**Fig. 5.25.** Combined chemical shift perturbations of apo-MDM4-N vs Nutlin-3 complex. Chemical shift perturbation mapped on the MDM4-N surface (7FEA)) colour coded from red (largest perturbations) to yellow (no perturbation). In white are indicated residues with missing resonances. Residues with the largest perturbation are labelled.
5.3.2.2. apo-MDM4-N vs MDM4-N:p53

The overall pattern of $^{13}\text{C}$ chemical shift changes for apo-MDM4-N versus the MDM4-N:p53-derived peptide complex is different from that observed for apo-MDM4-N versus the MDM4-N:Nutin-3 complex. All helices are affected in terms of $^{13}\text{C}$ perturbations with $^{13}\text{CO}$ in $\alpha_2$, $\beta_3$ and $\alpha_2'$ being the most strongly perturbed (Fig. 5.26). The p53-derived peptide is a larger molecule than Nutlin-3 and its binding is likely to disrupt resonances of the binding groove more widely, and in general affect the $^{13}\text{C}$ shifts to a greater degree. Overall, combined $^{15}\text{N},^{1}\text{H}$ amide chemical shift perturbations extend to a larger region when compared with MDM4-N:Nutlin-3, but the largest chemical shift perturbation are clearly identified in $\alpha_2$ and $\alpha_2'$, the walls of the binding groove.
Fig. 5.26. Chemical shift perturbations for apo-MDM4-N vs MDM4-N p53. Plot of A. $^{13}$C chemical shift perturbations and B. $^{15}$N chemical shift perturbations, for MDM4-N in complex with p53 derived peptide compared to apo-protein.
Similar conclusions to those pertaining to the MDM4-N:Nutlin-3 complex can be drawn from the trends of $^{13}$C perturbations for residues in $\alpha_2$, $\alpha_2'$ and $\beta_3$ of MDM4-N:p53. Specifically, the $^{13}$C$\text{O}$ shifts of residues in $\alpha_2$ are most perturbed upon p53 binding. Gly$^{57}$ experiences the largest perturbation (Fig. 5.26 A). The $^{13}$C chemical shift perturbations in the $\beta_3$ region and its vicinity are to be expected from the change in dynamics when MDM4-N binds p53. Interestingly, the Gln$^{71}$ $^{13}$C$\alpha$ shift is perturbed strongly (Fig. 5.26 A) as was also observed in the MDM4-N:Nutlin-3 data; its structural neighbour Tyr$^{66}$ is also significantly perturbed (Fig. 5.27).

The perturbation of shifts for residues contributing to the Leu26 sub-pocket of MDM4-N when bound to p53 are also similar to those observed when MDM4-N binds to Nutlin-3 (Fig. 5.27), except that they are smaller in magnitude (compare Fig. 5.25 and Fig. 5.27). This could arise from a lower occupancy due to lower affinity – or less chemical bulk of the ligand being inserted into the pocket.

![Fig. 5.27. Combined chemical shift perturbations of apo-MDM4-N vs p53 complex.](image)

Chemical shift perturbation mapped on the MDM4-N surface (7FEA) colour coded from red (largest perturbations) to yellow (no perturbation). In white, residues with missing resonances in the apo-MDM4-N data set. Residues with the largest perturbation are labelled.
5.4. Analysis of the chemical shift perturbations of MDM4-N

The data were analyzed with the Chemical Shift Index program (CSI) that predicts secondary structure motifs based on $^{13}\text{C}\alpha$, $^{13}\text{C}\beta$ and $^{13}\text{CO}$ chemical shift values [168, 183]. The prediction is calculated in a conservative manner, and consensus scoring is given as follows: -1→α-helix, 1→β-strand and 0→random coil. The results show that MDM4-N β1 and β2 are well predicted for all three samples (apo-protein and the two MDM4-N complexes)(see Fig. 5.28).

![Fig 5.28. Secondary structure prediction.](image)

The histograms show the CSI consensus score for apo-MDM4-N (bottom), in complex with Nutlin-3 (middle) or with p53 (top). In red is shown the CSI predicted for the crystal structure of MDM4-N from pdb id 3DAB. The secondary structure elements are labelled in red.
While CSI predicts a helical conformation for residues 49-62 in *apo*-MDM4 and in the complex with Nutlin-3, the MDM4-N:peptide complex is predicted to have a shorter equivalent helix (49-56) (Fig. 5.28 α2). Note that the H-bond pattern of α2 (residues 52-62) in the X-ray derived crystal structure 3DAB, is not canonical throughout its length; at the C-terminus of the helix the H-bond pattern is \( i-i+3 \) (see Fig. 5.29), in contrast to the expected \( i-i+4 \). The structural stress imposed by MDM4-N:p53 complex could affect chemical shift values, differing from expected values hurrying predictions.

**Fig 5.29. Details of MDM4-N α-helix number 2.** On the left is a surface representation of MDM4-N where α2 is highlighted in green. On the right is a cartoon representation of α2 with residues Met53, Gly57, Ile60 and Met61 in sticks, and the H-bonds pattern are represented by dashed lines.

The CSI predicts accurately the three β-strands towards the C-terminus of MDM4-N in complex with Nutlin-3 but does not make this prediction when the protein is unbound or in complex with peptide. For reasons discussed above, the *apo*-MDM4-N assignment is missing large amount of data, providing a poor dataset (consistent with the flexibility of this region in the unbound system). The complex of MDM4-N with p53-derivde peptide was more comprehensively assigned, however, the CSI prediction does not match the crystallographic data.
Moreover, CSI predictions of $\beta_1$, $\alpha_1$ and $\beta_2$ for apo-MDM4-N and the two complexes conflict with the secondary structure observed in an experimentally derived 3D structure (3DAB). Note, however, that strand $\beta_1$ is longer in 3JZO than in 3DAB, and therefore secondary structure within this region could indeed vary according to the construct used or the ligand present within the binding site.

The crystallographically observed helix $\alpha_1$ sequence ($^{30}$KLPLLILMAA$^{40}$) has a proline at position 32. The diagnostic value, for $\alpha$-helix prediction, of proline chemical shifts is probably dubious due to the rarity of proline residues in helices. Finally, CSI predicts $\beta$-strand ($\beta_2$, see Fig. 5.28) for the sequence $^{41}$GAQGEMFT$^{48}$ in all three NMR samples whereas the x-ray crystal-derived structure shows that this region consists of a $\beta$-bulge with water inserted (Fig 5.30).

**Fig 5.30 Beta-bulge in MDM4-N (3DAB).** The region $^{41}$GAQGEMFT$^{48}$ of MDM4-N is highlighted on the surface representation in green (left-hand panel) and re-drawn in ribbon representation (right-hand panel) to show its conformation, the side-chains and the attached water molecule [65]. This region was predicted to be $\beta$-strand by CSI.
5.5. MDM4-N relaxation

5.5.1. Introduction

Nuclear spin relaxation measurements provide insights into protein dynamics and therefore can help to understand the molecular basis for protein function. For the purposes of this thesis $^{15}$N relaxation studies were undertaken with the goal of understanding backbone motion within MDM4-N. This in turn sheds light on the binding of ligands and the structural and dynamic differences between apo-MDM4-N and the MDM4-N:Nutlin-3 and MDM4-N:p53 complexes.

The relaxation rates $R_1$ and $R_2$ and heteronuclear NOE values probe different motional frequencies and were obtained as described in Section 2.11.8 of Materials and Methods. $R_1 (1/T_1)$ values are sensitive to motions in the range $10^8$–$10^{12}$ s$^{-1}$, whereas $R_2 (1/T_2)$ values, in addition to depending on motions occurring at these high frequencies, also monitor dynamics (chemical exchange) on the microsecond-millisecond time scale. The heteronuclear NOEs (het-noe) are also sensitive to motion such that low or negative values correspond to high mobility on the picosecond-nanosecond timescale.

5.5.2. Datasets

Several residues were excluded from the analysis of backbone ($^{15}$N) relaxation measurements for various reasons. Proline residues (21, 29, 32, 95 and 96) have no amide protons, and therefore cannot be included in the datasets. Furthermore, both apo-MDM4-N and the MDM4-N:Nutlin-3 complex have multiple missing assignments and so these could not be used either. In addition, several cross peaks in the HSQC spectra had to be excluded due to overlap or poor signal-to-noise, as summarized in table 5.31. Overall, totals of 50, 65 and 76 residues were suitable for relaxation analysis for apo-MDM4-N, MDM4-N:Nutlin-3 complex and MDM4-N:peptide complex, respectively.
System | Missing from assignment | Overlapping or weak |
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>apo-MDM4-N</td>
<td>56, 60, 62, 63, 65, 66, 68, 69, 70, 72, 73, 74, 75, 77, 78, 80, 82, 87, 88, 89, 90, 91, 93, 94, 107, 108</td>
<td>25, 30, 33, 42, 43, 67, 81, 98, 102, 104, 105, 106, 111</td>
</tr>
<tr>
<td>MDM4-N: Nutlin-3</td>
<td>54, 57, 58, 61, 65, 66, 68, 69, 70, 72, 77, 92, 98</td>
<td>16, 26, 40, 52, 63, 67, 73, 80, 88, 89, 90, 96, 99, 100, 103, 105, 108, 110, 111</td>
</tr>
<tr>
<td>MDM4-N:p53 peptide</td>
<td>14, 16, 34, 37, 40, 45, 49, 54, 58, 69, 70, 74, 79, 102, 104, 106, 107, 108, 110, 111</td>
<td></td>
</tr>
</tbody>
</table>

Table 5.31. Residues missing or excluded from analysis of $^{15}$N relaxation measurements. The lists in the table indicate residues for which backbone dynamics could not be analysed due to lack of assignment or on the grounds of signal-to-noise or overlap issues.

### 5.5.3. *apo* MDM4-N relaxation rates

The large number of residues that had to be excluded from analysis for *apo*-MDM4-N is evidence that this is a highly dynamic protein with extensive mobile or disordered regions, especially in the C-terminal half of this domain. For those resonance that could be analysed, the average $R_1$ and $R_2$ values obtained for *apo*-MDM4-N are $1.16 \pm 0.13$ s$^{-1}$ and $14.4 \pm 6.2$ s$^{-1}$, respectively, and the average heteronuclear NOE is $0.62 \pm 0.29$. The relaxation data show that two regions of the protein with defined secondary structure are relatively rigid, namely $\alpha_1$, $\alpha_2$ and $\alpha_1'$. Residues in the putative “lid” region $^{14}$ESASRISPGQINQ$^{26}$ on the other hand exhibit low $R_2$ (long $T_2$) values (Fig. 5.32 A) with heteronuclear NOEs < 0.6 (Fig 5.32 B) consistent with high mobility on the fast timescale. This is expected for this region, since it is predicted to be highly flexible. Note that relaxation analysis of MDM2-N [49] suggests similar dynamic properties for the equivalent regions.
Fig 5.32 Relaxation data for apo-MDM4-N. Histogram A and B represent $R_1$ and $R_2$ respectively plotted versus residue number for apo-MDM4-N and associated errors (in red). The horizontal lines depict the average + or - one standard deviation for $R_1$ and $R_2$. C. Heteronuclear NOE values for apo-MDM4-N with associated error (red). The horizontal line marks the 0.6 threshold. Included in the picture are schematic representations of the secondary structure distribution of MDM4-N. In red are highlighted the missing resonances for MDM4-N apo assignment.
Interestingly, the high $R_2$ values (with $R_1<$Ave-SD) for residues 57-59 indicates microsecond-millisecond motion in this region that corresponds to the C-terminus of $\alpha_2$ in the crystal structure [although with ill-defined secondary structure according to the CSI (Fig 5.28)] and forms part of the walls of the Phe19 sub-pocket (Fig 5.33). Such conformational flexibility is presumably linked to the high mobility of residues surrounding the Phe19 sub-pocket including Gln$^{71}$, Cys$^{76}$ and Asp$^{79}$, which combine $hetNOEs < 0.6$ with high $R_2$ values, and many other residues that are missing from the analysis.

![Figure 5.33](image)

**Figure 5.33. Relaxation rates mapped onto MDM4-N.** A surface representation of MDM4-N (3FEA) is shown in green. Residues missing in the assignment are coloured in red, while residues inferred as being dynamically flexible are coloured in black (see text). The Phe19 sub-pocket is indicated.

### 5.5.4. Relaxation rates for MDM4-N: Nutlin-3 complex

Many residues in this complex could not be analysed as a result of broadening or overlap, consistent with a dynamic backbone, although importantly fewer residues had to be excluded than for apo-MDM4-N. The average $R_1$ and $R_2$ values obtained for the analysed residues of MDM4-N in complex with Nutlin-3 are $1.28 \pm 0.11$ s$^{-1}$ and $11.4 \pm 3.69$ s$^{-1}$, respectively, and the average $hetNOE$ is $0.7 \pm 0.2$. Overall, relaxation times are shorter than those of apo-MDM4-N, consistent with the greater number of observable resonances in the TROSY/HSQC spectrum, and the expected rigidification of MDM4-N upon binding to Nutlin-3.

As with the apo-MDM4-N data, the region $^{14}$ESASRISPGQINQ$^{26}$ exhibits all the hallmarks of a highly flexible portion of the protein. Otherwise, the MDM4-
N:Nutlin-3 relaxation data indicate that $\alpha_1$ and $\alpha_1'$ are the least flexible structural elements of the MDM4-N:Nutlin-3 complex (Fig 5.34).

Residues in the helix $\alpha_2$ (Fig 5.34, res. 49-63) display short values for $R_1$ but increased $R_2$ and hetNOEs $> 0.6$ indicate that, while not exceptionally flexible on the fast timescale, $\alpha_2$ undergoes weak chemical exchange; several nearby residues (His$^{54}$, Gly$^{57}$, Gln$^{58}$ and Met$^{61}$) were not assigned and therefore also considered in a disordered state.
Fig 5.34 Relaxation data for MDM4-N in complex with Nutlin-3. Histograms A and B represent the $R_1$ and $R_2$, respectively, plotted versus residue number for MDM4-N:Nutlin-3 and associated errors (in red). The horizontal lines depict the average ± one standard deviation for $R_1$ and $R_2$. C. Heteronuclear NOE values for apo-MDM4-N with associated error (red). The horizontal line marks the 0.6 threshold. Included in the picture are schematic representations of the secondary structure of MDM4-N. In blue are highlighted the missing resonances for MDM4-N:Nutlin-3 assignment.
Overall, MDM4-N in complex with Nutlin-3 is dynamically similar to apo-MDM4-N in that though some rigidity is gained when MDM4-N is in complex with Nutlin-3, the lack of signal and the relaxation data suggest a flexible α2 helix and Phe19 sub-pocket.

![Fig 5.35. Surface representation of MDM4-N.](image)

**Fig 5.35. Surface representation of MDM4-N.** Missing residues in the assignment of MDM4-N:Nutlin-3 complex are highlighted in blue. Residues with dynamically flexible are coloured in black (see text). Also, Phe19 sub-pocket and α2 helix are indicated.

### 5.5.5. Relaxation rates MDM4-N p53

The reappearance of multiple cross-peaks in the TROSY of MDM4-N when peptide is bound is clear evidence for rigidification. For analysable resonances, the average $R_1$ and $R_2$ obtained for MDM4-N in complex with p53 are $1.12 \pm 0.1$ s$^{-1}$ and $11.5 \pm 4.14$ s$^{-1}$ respectively and hetNOE $0.7 \pm 0.3$. Thus relaxation rates are shorter, again implying an increase in rigidity from *apo*-MDM4-N to *holo*-MDM4-N.

Nonetheless, the data clearly show that the MDM4-N region $^{14}$ESASRISPGQIN$^{25}$ remains a flexible section of the protein even when peptide occupies the binding groove. This region apparently remains unaffected by the interaction of MDM4-N with any of the ligands used in the current study (see Fig. 5.36; and compare Fig. 5.34 with Fig. 5.32). Otherwise, the relaxation data for the MDM4-N:p53-derived peptide complex reflect relatively rigid secondary structure elements. For example,
the helices each have similar average relaxation rates, close to the overall average values. Notably, relaxation rates and hetNOE values for residues in $\alpha_2$ contrast with those of their equivalents for apo-MDM4-N MDM4-N:Nutlin-3. Signs of flexibility or conformational exchange are confined to loop regions of MDM4-N (res. 63 to 73) that, interestingly, contribute to the Phe19 sub-pocket.
Fig 5.36 Relaxation data for MDM4-N in complex with p53 derived peptide. Fig 5.34 Relaxation data for MDM4-N in complex with Nutlin-3. Histogram A and B represent the $R_1$ and $R_2$ values, respectively plotted versus residue number for MDM4-N:p53 and associated errors (in red). The horizontal lines depict the average $\pm$ one standard deviation for $R_1$ and $R_2$. C. Heteronuclear NOE values for apo-MDM4-N with associated error (red). The horizontal line marks the 0.6-threshold. Representation of secondary structure is also included in the picture.
5.6. Conclusions

The NMR spectra of apo-MDM4-N confirmed its expected, very high level of backbone flexibility. This property might have contributed to the low yield of recombinant protein production (i.e. susceptibility to proteases) and tendency to self-associate or aggregate and precipitate. The protein is clearly less flexible (and more stable) when in complex with a ligand, accounting for the numerous reports of MDM4-N:ligand complex crystal structures, and the absence to date of crystals of the apo form. A model of the apo-MDM4-N in solution based on the findings of the current work is shown in Fig 5.38, with the Phe19 sub-pocket completely disordered.

Fig 5.38. Cartoon representation of a model of apo-MDM4-N. A. Surface representation of MDM4-N PDB id = 3DAB highlighting in red residues with missing resonances in the TROSY spectrum. B. Proposed structural representation of apo-MDM4-N where, in green, are represented the stable/organized portions of the protein, while in red are represented flexible/disordered regions. A prominent feature is the disordered Phe19 sub-pocket.

It is notable that Nutlin-3 binding to MDM4-N does not stabilize the protein to the same extent, as does binding of the p53-derived peptide. This difference is particularly marked in the C-terminal portion of α2, and could therefore be interpreted as a failure of Nutlin-3 to rigidify the Phe19 sub-pocket (Fig 5.39).
**Fig 5.39. Cartoon representation of MDM4-N in complex with Nutlin-3.** A. Surface representation of MDM4-N (PDB id = 3DAB) highlighting, in blue, residues with resonances missing in the TROSY spectrum. B. Proposed structural representation of MDM4-N:Nutlin-3 complex. In green are represented the stable/organized portions of the protein, while in blue are represented flexible/disordered regions. This shows the α2 helix is partially unfolded, affecting the stability of the MDM4-N Phe19 sub-pocket.

Comparing chemical shifts of *apo*-MDM4-N with those of its complexes, Tyr⁹⁹ experiences a large perturbation of amide shifts (Figs. 5.24 and 5.25), consistent with the idea that it switches from a ‘closed’ conformation in the *apo* form to an ‘open’ conformation in the *holo* form. Moreover, the ^1^C resonances for Tyr⁹⁹ (Fig 5.26) does not change in frequency between peptide and Nutlin-3 complexes. This suggests an “open” position of Tyr⁹⁹ within the Nutlin-3 complex, contradicting previous suggestions [65].

In the ‘closed’ conformation of Tyr⁹⁹ there could be an H-bond between its hydroxyl group and the backbone carbonyl of MDM4-N Met⁵³, equivalent to the H-bond between the hydroxyl of Tyr¹⁰⁰ and the backbone carbonyl of Leu⁵⁴ in *apo*-MDM2-N (Fig 5.40). Conversely, the crystal structure of Nutlin-2 in complex with MDM2-N shows that this H-bond has been broken (due to migration of the Tyr¹⁰⁰ side-chain).
Fig 5.40. Detail of the H-Bonding of Tyr$^{100}$ in MDM2-N. A. The closed conformation of Tyr$^{100}$ in apo-MDM2-N [49] (purple) is sustained by a H-bond between the hydroxyl of Tyr$^{100}$ and the backbone carbonyl of MDM2-N Leu$^{54}$ later replaced by a H-bond between Ne of p53 Trp$^{23}$ (yellow) and the carbonyl of Leu$^{54}$ (dashed lines). B. Conversely, in the Nutlin-2 complex (blue) there is no H-bond from Nutlin-2 to Leu$^{54}$ (as reported in PDB 1RV1 [78]).
Similarly, the crystal structure of MDM4-N in complex with p53 shows a replacement H-bond from $\text{Ne}$ of p53 Trp23 to MDM4-N Met$^{53}$ is formed (Fig 5.41). So an equivalent H-bonds in the \textit{apo}-MDM4-N, \textit{i.e.} Tyr$^{99}$ - Met$^{53}$ seems likely. Indeed, this H-bond could be the key to stability of MDM4-N, for the evidences discussed above shows that while the \textit{apo}-MDM4-N and the complex with p53-peptide conserve the folding of helix $\alpha_2$, this folding is lost when bound to Nutlin-3. According to this hypothesis there is some commonality between \textit{apo}-MDM4-N and the p53-peptide complex \textit{i.e.} the backbone carbonyl of MDM4-N Met$^{53}$ is engaged in an H-bond with either the $\text{Ne}$ of p53 Trp23 or the hydroxyl of Tyr$^{99}$ thereby helping to stabilise the $\alpha_2$ helix. This model suggests that Nutlin-3 cannot form this H-bond and hence binding is accompanied by the helix being destabilised (or not properly anchored to the rest of the structure). This idea is supported by the $^{13}$C chemical shift perturbations when comparing Nutlin-3 complex and \textit{apo}-forms of MDM4-N, since these data suggested a loss of H-bond register within $\alpha_2$. It is interesting to note that in the crystal structure of MDM4-N complexed with small molecule WK2988 \textsuperscript{[80]} this Met$^{53}$-H-bond is conserved and this compound has a reasonable $K_D$ of 11 $\mu\text{M}$ (Fig 5.41) for MDM4-N unlike Nutlin-3 that has an estimated $K_D$ of 40 $\mu\text{M}$. 

Chapter 5: NMR of MDM2/4 – assignment, structure, dynamics
Fig 5.41. Details of H-bonds and WK298 binding to MDM4-N. A. Cartoon representation of the H-bond between backbone carbonyl of MDM4-N Met$^{53}$ (green) and Nε of p53 Trp$^{23}$ (yellow) (dashed lines) [65]. B. MDM4-N in complex with WK298 [80] also shows a H-bond between carbonyl of Leu$^{54}$ and Nε of chloro-tryptophan within WK298.

Not only an H-bond from Met$^{53}$ is involved in the successful interaction with MDM4-N. An H-bond with Gln$^{71}$ positions Phe$^{19}$ of the p53 peptide in the MDM4 binding groove. Given that the Tyr$^{66}$ amide (spatially close to Gln$^{71}$) as shown in Fig.
5.42) resonance is neither present in the TROSY spectrum of apo-MDM4-N nor in the TROSY spectrum of the MDM4-N:Nutlin-3 complex suggests this residue is in multiple conformations. The Phe19 of p53-peptide that stacks with Tyr\(^{66}\) as a consequence of the abovementioned H-bond between Phe\(^{19}\) (p53) and Gln\(^{71}\) (MDM4-N) provides stability to the surrounding area (Fig 5.51). Nutlin-3 could theoretically H-bond with MDM4 Gln\(^{71}\) via the keto group in the piperazidine ring but no stacking with Tyr\(^{66}\) would be possible in this complex.

Fig. 5.42. Role of MDM4 Tyr\(^{66}\) in p53 recognition. Detail of the Phe19 sub-pocket of MDM4-N (green) in complex with p53 (yellow). Phe19 of the peptide plays an important role in the p53-MDM4 complex by stabilization of the region through its H-bond with Gln\(^{71}\) (dashed line) and stacking with Tyr\(^{66}\).

Overall, the data here presented here supports the evidence presented by Kallen et al. [81] that the Leu26 sub-pocket of MDM4-N has an intrinsic plasticity, modulating the size of the binding site. The interaction with Nutlin-3 displaces Tyr\(^{99}\) (just as Nutlin-2 binding displaces Tyr\(^{100}\) of MDM2-N), contradicting previous speculation [64]. This work proposes that the failure of Nutlin-3 to bind to MDM4-N as effectively as it does to MDM2-N mainly arises from the lack of the H-bond between Nutlin-3 and Met\(^{53}\) of MDM4-N, and consequent loss of stabilisation of the \(\alpha2\) helix.
Chapter 6: Conclusions and final remarks

MDM2 and MDM4 are proteins involved in regulating the tumor suppressor p53, interacting through their respective N-terminal domains with p53 trans-activation domain. The crystal structure of MDM2-N in complex with p53 derived peptide solved in 1996 by Kussie et al, showed two globular regions divided by an axis of symmetry, where 3 key residues of p53 (Phe19, Trp23 and Leu26) are buried [46]. The particularity of the interaction was regarded as a protein-protein interface amenable for drug targeting which small organic compounds could easily disrupt. Indeed in 2004 Vasilev et al described the first high affinity inhibitor of the p53-MDM2 interaction. Known as the Nutlins, these cis-imidazolines derivates were the product of HTS and bind to MDM2-N with affinities in the nM region [78]. Ever since, many organic compounds results of different drug discovery projects have yielded MDM2-p53 disruption [184].

Although MDM4 bears striking similarity with MDM2 and is directly implicated in controlling p53 [33], less is known about the interaction MDM4-p53. The first crystal structure featured a humanized zebrafish variant of MDM4-N in complex with p53 [64], for human construct MDM4-N have proven very unstable and prone to aggregation. Efforts finally yielded a MDM4-N:p53 crystal structure of the human protein, showing a very similar fold to that of MDM2. However, MDM4 had a more shallow pocket induced by the different orientation of MDM4’s helix α2’ contributing Met53 towards the binding pocket. Many other crystal structures have been produced in recent years of MDM4 in complex with peptidomimetics [81, 84, 85], and only one of MDM4 in complex with an organic compound [80].

The current work describes the steps taken to characterize the interaction between the p53 tumour suppressor and its antagonists, MDM2 and MDM4 by means of NMR.

The first step taken was to address the expression of isotopically labelled protein, aiming to express and purify enough protein for NMR and other biophysical studies. The MDM2-N construct (11-118) was successfully purified yielding ~ 17 mg/L of expression. However, MDM4 constructs proved more elusive. Construct MDM4 1-134 instability can be ascribed to large unstructured regions in both its C-and N-
termini. Also 2 cysteins closely together at the N-termini of the construct \( \text{Cys}^{10} \) and \( \text{Cys}^{17} \) proved detrimental. Buffer screening proved insufficient, for MDM4 1-134 stability did not increase in the limited conditions explored. On the other hand, MDM4-N construct (14-111) [81] was successfully purified and proved stable enough to undergo NMR experiments and subsequent screening for expression conditions yielded \( \sim 2 \text{ mg of protein per litre of culture} \).

As part of the project, MDM2-N was screened against an organic compound library obtained by the \textit{in silico} screening tool LIDAEUS [101]. Using the NMR toolkit several compounds were identified with affinities in the high \( \mu \text{M} \) region. A subsequent docking exercise assisted by NMR data was undertaken providing insight into the binding mode of the compounds. The compounds share a core structural scaffold in the form of triazole, which remains solvent exposed, and a phenyl group in position 4 that is buried in the Trp23 sub-pocket of MDM2-N interface. The occupancy of the Leu26 and Phe19 sub-pockets depends on the substituents at position 3 and 5 of the triazole. Therefore, it appears that \textit{para}-chloro substituent to the phenyl attached to the 4-position of the triazole will increase the affinity of the compounds for MDM2-N while other substitutions should be explored to achieve better occupancy of the Leu26 and Phe19 sub-pockets.

A different screened approach was also undertaken and a fragment-based library was screened against MDM2-N with regard to the \textquotedblleft\textit{druggability}\textquotedblright\ of the MDM2-N binding pocket. The similarity of MDM2-N sub-pockets –in terms of shape and physico-chemical properties, is detrimental in fragment-based screening. The fragments are preferentially buried in the deepest sub-pocket (Trp23) while the more superficial (Phe19 and Leu26 sub-pockets) are neglected. Therefore, fragment screening should be approached in tandem, and while a fragment should be added to the mixture to occupy Trp23 sub-pocket (ie. A3 fragment), the occupancy of Phe19 and Leu26 sub-pockets should be explored with different moieties.

To better understand the p53-binding groove of MDM4-N, multidimensional NMR was used to investigate the structure and backbone dynamics of double-isotopically labelled samples of MDM4-N, both free (\textit{i.e.} \textit{apo}-MDM4-N) and in complexes with a p53-derived peptide or Nutlin-3. The \textit{apo}-MDM4-N is more dynamic than MDM2,
since it contains unstructured regions. These regions appear to become structured upon binding of a ligand, and MDM4 appears to bind its ligand through conformational selection and/or an induced fit mechanism involving reorganization of key sub-sites within the binding groove. Crucially, the Phe19 sub-pocket appears to undergo the largest reorganization, as NMR data shows how this region of MDM4-N goes from conformationally flexible, judging by the missing resonances in the TROSY experiment, to structured when bound to p53 peptide. Analysis of the carbon shifts (Cα, Cβ and CO) and relaxation data support this conclusion. On the other hand, MDM4-N in complex with Nutlin-3 does not rigidify the protein fully. The data presented in this work suggests that Nutlin-3 does not successfully anchors α helix 2 and this is translated in a loosed conformation of MDM4-N in the vicinity of Phe19 sub-pocket. The evidence suggests that the lack of H-bond between Nutlin-3 and Met53 of MDM4-N are responsible for the poor affinity of Nutlin-3 to MDM4-N. Therefore, the rational design of specific inhibitors of the MDM4:p53 interaction should conserve this H-bond with Met53 of MDM4-N. Moreover, fragment screening was hampered by the poor resolution of MDM4-N TROSY experiment, suggesting that fragment screening could benefit from the addition of a tryptophan-like moiety to provide more stability to MDM4-N in solution, easing the screening exercise in Phe19 and Leu-26 sub-pocket with other fragments.
Bibliography


163. Hajduk, P.J., *SAR by NMR: Putting the pieces together*. Molecular Interventions, 2006. 6(5): p. 266-+


Appendix A.1. ProtParam details

A.1.1. hexaHis-Tag MDM2-N

```
10  20  30  40  50  60
MGSSHHHHHH SSGLVPRGSH MDGAVTTSQI PASEQETLVR PKPLLKLKLK SVGAQKDTYT
70  80  90 100 110 120
MKEVLFLYLGQ YIMTKRLYDE KQQHVYCSN DLLGDLFGVP SFSVKEHRKI YTMIYRLVV
130
VNQQESSDSG TSVSENLE
```

Number of amino acids: 138
Molecular weight: 15601.7
Theoretical pI: 7.14
Extinction coefficients:
Extinction coefficients are in units of M⁻¹ cm⁻¹, at 280 nm measured in water.
Ext. coefficient: 10430
Abs 0.1% (=1 g/l) 0.669, assuming ALL Cys residues appear as half cystines.

A.1.2. MDM2-N

```
10  20  30  40  50  60
GSHMDGAVTT SQIPASEQET LVRPKLLLLK LLKSVGAQKD TYTMKEVLFY LGQYIMTKRL
70  80  90 100 110 120
YDEKQQHVY CSNDLLGDLF GVPSSVKEH RKIYTMIYRN LVVVNQQESS DSGTSVSENLE
```

Number of amino acids: 121
Molecular weight: 13719.6
Theoretical pI: 6.34
Extinction coefficients:
Extinction coefficients are in units of M⁻¹ cm⁻¹, at 280 nm measured in water.
Ext. coefficient: 10430
Abs 0.1% (=1 g/l) 0.760, assuming ALL Cys residues appear as half cystines.
A.1.3. hexaHis-Tag MDM4 1-134

10  20  30  40  50  60
MGSSHHHHHH SSGLVPRGSH MTSFSTSQAQC STSACRIS PGQINQVRP KLPLKILHAA
70  80  90  100  110 120
GAQGEMFTVK EVMHYLGQYI MVKQLYDQQE QHMVYCGGDL LGELLGRQSF SVKNPSPLYD
130 140 150
MLRKNLVTLA TATTDAAQTAL ALAQDHSMDI PSQD

Number of amino acids: 154

Molecular weight: 16900.2

Theoretical pI: 6.82

Extinction coefficients:

Extinction coefficients are in units of M⁻¹ cm⁻¹, at 280 nm measured in water.

Ext. coefficient    7575

Abs 0.1% (=1 g/l)  0.448, assuming ALL Cys residues appear as half cystines.

A.1.4. MDM4 1-134

10  20  30  40  50  60
GSHMTSFSTS AQCTSDACRIS PGQINQVRP KLPLKIL HAAGAQGEMFT VEVMHYLG
70  80  90  100  110 120
QYIMVKQLYD QQEHHMV YCG GDLLGELLGR QSFSVKNPS PLYDMLRKNLV TLATATTDAA
130
QTLALACQDSL MDIPSQD

Number of amino acids: 137

Molecular weight: 15018.1

Theoretical pI: 6.19

Extinction coefficients:

Extinction coefficients are in units of M⁻¹ cm⁻¹, at 280 nm measured in water.

Ext. coefficient    7575

Abs 0.1% (=1 g/l)  0.504, assuming ALL Cys residues appear as half cystines.

A.1.5 hexaHis-Tag MDM4-N

10  20  30  40  50  60
MGSSHHHHHH SSGLVPRGSH MDSASRISPG QINQVRPKLP LLKILHAAGA QGEMFTKEV
70  80  90  100  110
MHYLQGYIMV KQLYDQQE QH MVCGD LLLGGRQ SFV KNPSPLYDML RKNLVT LAT
Number of amino acids: 119
Molecular weight: 13313.3
Theoretical pI: 9.17

Extinction coefficients:
Extinction coefficients are in units of $M^{-1} \text{cm}^{-1}$, at 280 nm measured in water.
Ext. coefficient 7450
Abs 0.1% (=1 g/l) 0.560, assuming all pairs of Cys residues form cystines.

A1.5 MDM4-N

<table>
<thead>
<tr>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
<th>60</th>
</tr>
</thead>
<tbody>
<tr>
<td>GSHMDSASRI SPGQINQVRP KLPLLKILHA AGAQGEMFTV KEVMHYLGQY IMVKQLYDQQ</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td>EQHMVYCGGD LLGELLGRQS FSVKNPSPLY DMLRKNLVTL AT</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Number of amino acids: 102
Molecular weight: 11431.3
Theoretical pI: 8.81

Extinction coefficients:
Extinction coefficients are in units of $M^{-1} \text{cm}^{-1}$, at 280 nm measured in water.
Ext. coefficient 7450
Abs 0.1% (=1 g/l) 0.652, assuming all pairs of Cys residues form cystines.

A1.7 p53 derived peptide

ETFSDLWKLL P

Number of amino acids: 11
Molecular weight: 1348.5
Theoretical pI: 4.37

Extinction coefficients:
Extinction coefficients are in units of $M^{-1} \text{cm}^{-1}$, at 280 nm measured in water.
Ext. coefficient 5500
Abs 0.1% (=1 g/l) 4.079, assuming ALL Cys residues appear as half cystines.
## Appendix A.2. Sequencing results

<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>20</th>
<th>30</th>
<th>40</th>
<th>50</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>. .</td>
<td>. .</td>
<td>. .</td>
<td>. .</td>
<td>. .</td>
</tr>
</tbody>
</table>

**Fw**

```
---------TCNTTTTNACCTCTGCTCAGTGTTCAACATCTGAC
```

**Rev**

```
AGGCCCATATGACATACATTTTTCCACCTCTGCTCAGTGTTCAACATCTGAC
```

### 134

```
----------ATGACATCATTCTTCACCTCTGCTCAGTGTTCAACATCTGAC
```

<table>
<thead>
<tr>
<th></th>
<th>60</th>
<th>70</th>
<th>80</th>
<th>90</th>
<th>100</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>. .</td>
<td>. .</td>
<td>. .</td>
<td>. .</td>
<td>. .</td>
</tr>
</tbody>
</table>

**Fw**

```
AGTGCTTTGCAAGATTTTGCATGCAGCAGGTGCGCAAGGTGAAATGTTCA
```

**Rev**

```
AGTGCTTTGCAAGATTTTGCATGCAGCAGGTGCGCAAGGTGAAATGTTCA
```

### 134

```
AGTGCTTTGCAAGATTTTGCATGCAGCAGGTGCGCAAGGTGAAATGTTCA
```

<table>
<thead>
<tr>
<th></th>
<th>110</th>
<th>120</th>
<th>130</th>
<th>140</th>
<th>150</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>. .</td>
<td>. .</td>
<td>. .</td>
<td>. .</td>
<td>. .</td>
</tr>
</tbody>
</table>

**Fw**

```
GCCGCTTTTGAAGATTTTGCATGCAGCAGGTGCGCAAGGTGAAATGTTCA
```

**Rev**

```
GCCGCTTTTGAAGATTTTGCATGCAGCAGGTGCGCAAGGTGAAATGTTCA
```

### 134

```
GCCGCTTTTGAAGATTTTGCATGCAGCAGGTGCGCAAGGTGAAATGTTCA
```

<table>
<thead>
<tr>
<th></th>
<th>160</th>
<th>170</th>
<th>180</th>
<th>190</th>
<th>200</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>. .</td>
<td>. .</td>
<td>. .</td>
<td>. .</td>
<td>. .</td>
</tr>
</tbody>
</table>

**Fw**

```
CTGTTAAGAGGTCATGCACATATTGTAGGTCAGTACAATGTTGGAAGCA
```

**Rev**

```
CTGTTAAGAGGTCATGCACATATTGTAGGTCAGTACAATGTTGGAAGCA
```

### 134

```
CTGTTAAGAGGTCATGCACATATTGTAGGTCAGTACAATGTTGGAAGCA
```

<table>
<thead>
<tr>
<th></th>
<th>210</th>
<th>220</th>
<th>230</th>
<th>240</th>
<th>250</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>. .</td>
<td>. .</td>
<td>. .</td>
<td>. .</td>
<td>. .</td>
</tr>
</tbody>
</table>
Appendix
Appendix A.3. Vitamins and Micronutrients recipe

Recipe for 1 LVitamins Supplements (x1000):

0.4 g Choline chloride
0.5 g Folic acid
0.5 g Pantothenic acid
0.5 g Nicotinamide
1.0 g Myo-inositol
0.5 g Pyridoxal HCl
0.5 g Thiamine HCl
0.05 g Riboflavin
1.0 g Biotin

Micronutrients added to minimal media:
3 X 10^{-3} M (NH_4)_6(MO_7)
4 X 10^{-1} M H_3BO_3
3 X 10^{-2} M CoCl_2
1 X 10^{-2} M CuSO_4
8 X 10^{-2} M MnCl_2
1 X 10^{-2} M ZnSO_4

Appendix A.4. Peptide synthesis methodology

A.4.1. Activation of resin

Reagents for washing steps
20 ml Dimethylformamide (DMF)
20 ml Methanol (MeOH)
20 ml Dichloromethane (DCM)

Deprotection solution
20% Piperidine (v/v) in DMF

0.83 g of Rink amide Fmoc (loading 1.6 mmol/g) was mixed with 10 ml deprotection solution, and incubated for 10 min at room temperature with constant stirring. The mixture was then washed two times with 20 ml DMF using the Manifold equipment.
This sequence of steps was repeated two times. Finally, the resin was washed consecutively with the washing solutions using the Manifold equipment and let to dry.

The activation of the resin was confirmed with the ninhydrin test.

**A.4.2. Coupling of Fmoc protected aminoacid**

Reagents:

- 0.83 g activated resin (R-NH$_2$)
- 0.2 M of Fmoc protected amino acid $i$ (AA$_i$-Fmoc) in DMF
- 0.12 g 1-Hydroxybenzotriazole (HOBT)
- 0.12 ml diisopropyl-carbodiimide (DIC)

To activate the amino acid, a mixture of AA$_i$-Fmoc, HOBT and DIC were incubated for 10 min at room temperature with constant stirring. Subsequently, R-NH$_2$ was added the solution of Fmoc-AA$_i$, HOBT and DIC and incubated in a Monowave 300 (Anton Paar) for 20 min at 60°C. This was followed by a wash with 5 ml DMF, MeOH, and DCM consecutively using the Manifold equipment and the final product (R-Fmoc-AA$_i$) was dried under vacuum. The presence of unreacted primary amines was checked with the ninhydrin test.

Finally, the protected of R-AAm-Fmoc was mixed with 5 ml of piperidine 20% (v/v) in DMF and incubated under mechanical stirring for 10 minutes. Followed by a wash step with 5 mL of DMF with the Manifold equipment. These steps were repeated two times.

The R-AAm-NH$_2$ was washed consecutively with the washing solutions using the Manifold equipment and let to dry. The presence of primary amino terminal group was confirmed with the ninhydrin test.
Appendix B.1. VsL details

<table>
<thead>
<tr>
<th>Structure</th>
<th>ID</th>
<th>MW (Da)</th>
<th>MLOGP</th>
<th>HBD</th>
<th>HBA</th>
</tr>
</thead>
<tbody>
<tr>
<td><img src="" alt="Chemical Structure 1" /></td>
<td>1</td>
<td>348.53</td>
<td>2.93</td>
<td>2</td>
<td>4</td>
</tr>
<tr>
<td><img src="" alt="Chemical Structure 2" /></td>
<td>2</td>
<td>395.41</td>
<td>4.15</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td><img src="" alt="Chemical Structure 3" /></td>
<td>3</td>
<td>396.52</td>
<td>4.06</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td><img src="" alt="Chemical Structure 4" /></td>
<td>4</td>
<td>422.43</td>
<td>3.95</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td><img src="" alt="Chemical Structure 5" /></td>
<td>5</td>
<td>388.88</td>
<td>3.60</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td>Structure</td>
<td>ID</td>
<td>MW (Da)</td>
<td>MLOGP</td>
<td>HBD</td>
<td>HBA</td>
</tr>
<tr>
<td>-----------</td>
<td>----</td>
<td>---------</td>
<td>-------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td><img src="image1.png" alt="Structure 1" /></td>
<td>6</td>
<td>375.47</td>
<td>4.85</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td><img src="image2.png" alt="Structure 2" /></td>
<td>7</td>
<td>332.37</td>
<td>1.61</td>
<td>2</td>
<td>7</td>
</tr>
<tr>
<td><img src="image3.png" alt="Structure 3" /></td>
<td>8</td>
<td>294.40</td>
<td>1.31</td>
<td>0</td>
<td>4</td>
</tr>
<tr>
<td><img src="image4.png" alt="Structure 4" /></td>
<td>9</td>
<td>283.35</td>
<td>2.15</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td><img src="image5.png" alt="Structure 5" /></td>
<td>10</td>
<td>326.46</td>
<td>2.65</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>Structure</td>
<td>ID</td>
<td>MW (Da)</td>
<td>MLOGP</td>
<td>HBD</td>
<td>HBA</td>
</tr>
<tr>
<td>-----------</td>
<td>----</td>
<td>---------</td>
<td>-------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td><img src="image1.png" alt="Structure 1" /></td>
<td>11</td>
<td>367.41</td>
<td>4.20</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td><img src="image2.png" alt="Structure 2" /></td>
<td>12</td>
<td>472.52</td>
<td>2.32</td>
<td>1</td>
<td>7</td>
</tr>
<tr>
<td><img src="image3.png" alt="Structure 3" /></td>
<td>13</td>
<td>428.46</td>
<td>4.63</td>
<td>4</td>
<td>13</td>
</tr>
<tr>
<td><img src="image4.png" alt="Structure 4" /></td>
<td>14</td>
<td>297.38</td>
<td>3.22</td>
<td>1</td>
<td>4</td>
</tr>
<tr>
<td><img src="image5.png" alt="Structure 5" /></td>
<td>15</td>
<td>340.45</td>
<td>2.37</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td>Structure</td>
<td>ID</td>
<td>MW (Da)</td>
<td>MLOGP</td>
<td>HBD</td>
<td>HBA</td>
</tr>
<tr>
<td>-----------</td>
<td>----</td>
<td>---------</td>
<td>-------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td><img src="image1.png" alt="Structure 1" /></td>
<td>16</td>
<td>407.42</td>
<td>3.90</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td><img src="image2.png" alt="Structure 2" /></td>
<td>17</td>
<td>371.44</td>
<td>3.40</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td><img src="image3.png" alt="Structure 3" /></td>
<td>18</td>
<td>354.43</td>
<td>3.69</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td><img src="image4.png" alt="Structure 4" /></td>
<td>19</td>
<td>382.53</td>
<td>3.13</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td><img src="image5.png" alt="Structure 5" /></td>
<td>20</td>
<td>395.53</td>
<td>3.63</td>
<td>0</td>
<td>6</td>
</tr>
<tr>
<td>Structure</td>
<td>ID</td>
<td>MW (Da)</td>
<td>MLOGP</td>
<td>HBD</td>
<td>HBA</td>
</tr>
<tr>
<td>-----------</td>
<td>----</td>
<td>---------</td>
<td>-------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td><img src="image1" alt="Structure 21" /></td>
<td>21</td>
<td>373.87</td>
<td>3.55</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td><img src="image2" alt="Structure 22" /></td>
<td>22</td>
<td>432.35</td>
<td>3.90</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td><img src="image3" alt="Structure 23" /></td>
<td>23</td>
<td>442.52</td>
<td>3.14</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td><img src="image4" alt="Structure 24" /></td>
<td>24</td>
<td>272.37</td>
<td>1.88</td>
<td>4</td>
<td>5</td>
</tr>
<tr>
<td><img src="image5" alt="Structure 25" /></td>
<td>25</td>
<td>387.52</td>
<td>2.57</td>
<td>2</td>
<td>5</td>
</tr>
<tr>
<td>Structure</td>
<td>ID</td>
<td>MW (Da)</td>
<td>MLOGP</td>
<td>HBD</td>
<td>HBA</td>
</tr>
<tr>
<td>-----------</td>
<td>----</td>
<td>---------</td>
<td>-------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td><img src="image1.png" alt="Structure 1" /></td>
<td>26</td>
<td>371.39</td>
<td>1.51</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td><img src="image2.png" alt="Structure 2" /></td>
<td>27</td>
<td>326.42</td>
<td>2.14</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td><img src="image3.png" alt="Structure 3" /></td>
<td>28</td>
<td>387.52</td>
<td>2.14</td>
<td>0</td>
<td>5</td>
</tr>
<tr>
<td><img src="image4.png" alt="Structure 4" /></td>
<td>29</td>
<td>489.55</td>
<td>0.20</td>
<td>2</td>
<td>8</td>
</tr>
<tr>
<td><img src="image5.png" alt="Structure 5" /></td>
<td>30</td>
<td>340.40</td>
<td>2.10</td>
<td>2</td>
<td>6</td>
</tr>
<tr>
<td>Structure</td>
<td>ID</td>
<td>MW (Da)</td>
<td>MLOGP</td>
<td>HBD</td>
<td>HBA</td>
</tr>
<tr>
<td>-----------</td>
<td>----</td>
<td>---------</td>
<td>-------</td>
<td>-----</td>
<td>-----</td>
</tr>
<tr>
<td><img src="image" alt="Structure 31" /></td>
<td>31</td>
<td>300.38</td>
<td>2.52</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td><img src="image" alt="Structure 32" /></td>
<td>32</td>
<td>410.50</td>
<td>4.15</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td><img src="image" alt="Structure 33" /></td>
<td>33</td>
<td>286.36</td>
<td>2.27</td>
<td>1</td>
<td>5</td>
</tr>
<tr>
<td><img src="image" alt="Structure 34" /></td>
<td>34</td>
<td>351.45</td>
<td>3.45</td>
<td>0</td>
<td>6</td>
</tr>
</tbody>
</table>
Appendix B.2. Autodock

B.2.1. Grid details

npts 54 38 50  # num.grid points in xyz
gridfld lycr_rigid.maps.fld  # grid_data_file
spacing 0.375  # spacing(A)
receptor_types A C HD N OA SA  # receptor atom types
ligand_types A C Cl NA OA N SA HD  # ligand atom types
receptor lycr_rigid.pdbqt  # macromolecule
gridcenter 25.242 -20.777 -6.302  # xyz-coordinates or auto
smooth 0.5  # store minimum energy w/in rad(A)
map lycr_rigid.A.map  # atom-specific affinity map
map lycr_rigid.C.map  # atom-specific affinity map
map lycr_rigid.Cl.map  # atom-specific affinity map
map lycr_rigid.NA.map  # atom-specific affinity map
map lycr_rigid.OA.map  # atom-specific affinity map
map lycr_rigid.N.map  # atom-specific affinity map
map lycr_rigid.SA.map  # atom-specific affinity map
map lycr_rigid.HD.map  # atom-specific affinity map
elecmap lycr_rigid.e.map  # electrostatic potential map
dsolvemap lycr_rigid.d.map  # desolation potential
map
dielectric -0.1465  # <0, AD4 distance-
dep.diel;>0, constant

B.2.2. Docking details

autodock_parameter_version 4.2  # used by autodock to validate parameter set
outlev 1  # diagnostic output level
intelec  # calculate internal electrostatics
seed pid time  # seeds for random generator
ligand_types A C F NA OA N SA HD  # atoms types in ligand
fld lycr_rigid.maps.fld  # grid_data_file
map lycr_rigid.A.map  # atom-specific affinity map
map lycr_rigid.C.map  # atom-specific affinity map
map lycr_rigid.F.map  # atom-specific affinity map

Appendix 269
map lycr_rigid.NA.map                # atom-specific affinity map
map lycr_rigid.OA.map                # atom-specific affinity map
map lycr_rigid.N.map                 # atom-specific affinity map
map lycr_rigid.SA.map                # atom-specific affinity map
map lycr_rigid.HD.map                # atom-specific affinity map
elecmap lycr_rigid.e.map             # electrostatics map
desolvmap lycr_rigid.d.map           # desolavation map
move c17.pdbqt                        # small molecule
flexres lycr_flex.pdbqt               # file containing flexible residues
about -0.1665 -0.0998 0.1156          # small molecule center
tran0 random                          # initial coordinates/A or random
axisangle0 random                     # initial orientation
dihe0 random                          # initial dihedrals
(relative) or random
(tstep 2.0                              # translation step/A
qstep 50.0                             # quaternion step/deg
dstep 50.0                             # torsion step/deg
torsdof 8                              # torsional degrees of freedom
rmstol 2.0                             # cluster_tolerance/A
extrnrg 1000.0                          # external grid energy
e0max 0.0 10000                        # max initial energy; max number of retries
ga_pop_size 150                       # number of individuals in population
ga_num_evals 25000000                  # maximum number of energy evaluations
ga_num_generations 27000              # maximum number of generations
ga_elitism 1                           # number of top individuals to survive to next generation
ga_mutation_rate 0.02                 # rate of gene mutation
ga_crossover_rate 0.8                 # rate of crossover
ga_window_size 10                     #
ga_cauchy_alpha 0.0                    # Alpha parameter of Cauchy distribution
ga_cauchy_beta 1.0                     # Beta parameter Cauchy distribution
set_ga                                 # set the above parameters
for GA or LGA                          #
sw_max_its 300                          # iterations of Solis & Wets
local search
sw_max_succ 4  # consecutive successes
before changing rho
sw_max_fail 4  # consecutive failures
before changing rho
sw_rho 1.0  # size of local search space
to sample
sw_lb_rho 0.01  # lower bound on rho
ls_search_freq 0.06  # probability of performing
local search on individual
set_pswl  # set the above pseudo-Solis
& Wets parameters
unbound_model bound  # state of unbound ligand
ga_run 100  # do this many hybrid GA-LS
runs
analysis  # perform a ranked cluster
analysis

B.2.3. Autodock results

A file with the Autodock results for c16, c17, c21 and c22 can be found at
/Autodock_results, and can be visualized with Pymol.

Appendix C. NMR details

Appendix C.1

Details of natural abundance and values for gyromagnetic ratio of isotopes with relevance in protein NMR adapted from [185]:

<table>
<thead>
<tr>
<th>Isotope</th>
<th>Natural abundance (%)</th>
<th>Gyromagnetic ratio (10^7\text{T}^1\text{s}^{-1})</th>
</tr>
</thead>
<tbody>
<tr>
<td>(^1\text{H})</td>
<td>99.98</td>
<td>26.75</td>
</tr>
<tr>
<td>(^15\text{N})</td>
<td>0.37</td>
<td>-2.71</td>
</tr>
<tr>
<td>(^13\text{C})</td>
<td>1.11</td>
<td>6.73</td>
</tr>
</tbody>
</table>
## Appendix C.2 MDM2-N re-assignment

Chemical shift list of residues assigned for MDM2-N.

<table>
<thead>
<tr>
<th>Residue</th>
<th>Seq code</th>
<th>Residue Atom</th>
<th>Atom</th>
<th>Shift/ Error/ Ambiguity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Asp</td>
<td>N</td>
<td>H</td>
<td>H</td>
<td>8.11 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Gly</td>
<td>N</td>
<td>H</td>
<td>H</td>
<td>8.09 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Gly</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>109.11 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Ala</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>7.84 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Ala</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>122.78 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Gln</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>8.29 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Gln</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>121.14 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Ile</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>7.71 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Ile</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>121.74 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Ser</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>8.37 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Ser</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>111.58 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Glu</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>7.48 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Glu</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>122.76 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Gln</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>7.75 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Gln</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>118.9 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Thr</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>7.39 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Thr</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>117.42 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Val</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>9.62 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Val</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>116.65 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Lys</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>8.14 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Lys</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>123.21 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Leu</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>8.75 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Leu</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>118.21 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Leu</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>7.19 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Leu</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>118.08 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Leu</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>8.44 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Leu</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>120.1 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Lys</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>7.79 ppm 0 ppm 1</td>
</tr>
<tr>
<td>Position</td>
<td>Residue</td>
<td>Charge</td>
<td>pKa</td>
<td>N-Protonation</td>
</tr>
<tr>
<td>----------</td>
<td>---------</td>
<td>--------</td>
<td>-----</td>
<td>---------------</td>
</tr>
<tr>
<td>36</td>
<td>Lys</td>
<td>N</td>
<td>118.13</td>
<td>0</td>
</tr>
<tr>
<td>37</td>
<td>Leu</td>
<td>H</td>
<td>7.19</td>
<td>0</td>
</tr>
<tr>
<td>37</td>
<td>Leu</td>
<td>N</td>
<td>120.96</td>
<td>0</td>
</tr>
<tr>
<td>39</td>
<td>Lys</td>
<td>H</td>
<td>8.7</td>
<td>0</td>
</tr>
<tr>
<td>39</td>
<td>Lys</td>
<td>N</td>
<td>118.21</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>Ser</td>
<td>H</td>
<td>7.75</td>
<td>0</td>
</tr>
<tr>
<td>40</td>
<td>Ser</td>
<td>N</td>
<td>116.72</td>
<td>0</td>
</tr>
<tr>
<td>41</td>
<td>Val</td>
<td>H</td>
<td>7.14</td>
<td>0</td>
</tr>
<tr>
<td>41</td>
<td>Val</td>
<td>N</td>
<td>112.66</td>
<td>0</td>
</tr>
<tr>
<td>42</td>
<td>Gly</td>
<td>H</td>
<td>7.39</td>
<td>0</td>
</tr>
<tr>
<td>42</td>
<td>Gly</td>
<td>N</td>
<td>106.19</td>
<td>0</td>
</tr>
<tr>
<td>43</td>
<td>Ala</td>
<td>H</td>
<td>7.22</td>
<td>0</td>
</tr>
<tr>
<td>43</td>
<td>Ala</td>
<td>N</td>
<td>124.82</td>
<td>0</td>
</tr>
<tr>
<td>44</td>
<td>Gln</td>
<td>H</td>
<td>8.57</td>
<td>0</td>
</tr>
<tr>
<td>44</td>
<td>Gln</td>
<td>N</td>
<td>118.09</td>
<td>0</td>
</tr>
<tr>
<td>45</td>
<td>Lys</td>
<td>H</td>
<td>7.23</td>
<td>0</td>
</tr>
<tr>
<td>45</td>
<td>Lys</td>
<td>N</td>
<td>118.12</td>
<td>0</td>
</tr>
<tr>
<td>46</td>
<td>Asp</td>
<td>H</td>
<td>8.29</td>
<td>0</td>
</tr>
<tr>
<td>46</td>
<td>Asp</td>
<td>N</td>
<td>119.33</td>
<td>0</td>
</tr>
<tr>
<td>47</td>
<td>Thr</td>
<td>H</td>
<td>6.86</td>
<td>0</td>
</tr>
<tr>
<td>47</td>
<td>Thr</td>
<td>N</td>
<td>111.24</td>
<td>0</td>
</tr>
<tr>
<td>48</td>
<td>Tyr</td>
<td>H</td>
<td>8.62</td>
<td>0</td>
</tr>
<tr>
<td>48</td>
<td>Tyr</td>
<td>N</td>
<td>120.02</td>
<td>0</td>
</tr>
<tr>
<td>49</td>
<td>Thr</td>
<td>H</td>
<td>8.83</td>
<td>0</td>
</tr>
<tr>
<td>49</td>
<td>Thr</td>
<td>N</td>
<td>110.62</td>
<td>0</td>
</tr>
<tr>
<td>50</td>
<td>Met</td>
<td>H</td>
<td>8.06</td>
<td>0</td>
</tr>
<tr>
<td>50</td>
<td>Met</td>
<td>N</td>
<td>121.53</td>
<td>0</td>
</tr>
<tr>
<td>51</td>
<td>Lys</td>
<td>H</td>
<td>8.41</td>
<td>0</td>
</tr>
<tr>
<td>51</td>
<td>Lys</td>
<td>N</td>
<td>117.61</td>
<td>0</td>
</tr>
<tr>
<td>52</td>
<td>Glu</td>
<td>H</td>
<td>7.74</td>
<td>0</td>
</tr>
<tr>
<td>52</td>
<td>Glu</td>
<td>N</td>
<td>120.26</td>
<td>0</td>
</tr>
<tr>
<td>53</td>
<td>Val</td>
<td>H</td>
<td>8.09</td>
<td>0</td>
</tr>
<tr>
<td>53</td>
<td>Val</td>
<td>N</td>
<td>119.58</td>
<td>0</td>
</tr>
<tr>
<td>54</td>
<td>Leu</td>
<td>H</td>
<td>8.19</td>
<td>0</td>
</tr>
<tr>
<td>54</td>
<td>Leu</td>
<td>N</td>
<td>119.17</td>
<td>0</td>
</tr>
<tr>
<td>55</td>
<td>Phe</td>
<td>H</td>
<td>7.96</td>
<td>0</td>
</tr>
<tr>
<td>Position</td>
<td>Amino Acid</td>
<td>Side Chain</td>
<td>Secondary Structure</td>
<td>Mass (Da)</td>
</tr>
<tr>
<td>----------</td>
<td>------------</td>
<td>------------</td>
<td>---------------------</td>
<td>----------</td>
</tr>
<tr>
<td>55</td>
<td>Phe</td>
<td>N</td>
<td>N</td>
<td>122.04</td>
</tr>
<tr>
<td>56</td>
<td>Tyr</td>
<td>H</td>
<td>H</td>
<td>8.23</td>
</tr>
<tr>
<td>56</td>
<td>Tyr</td>
<td>N</td>
<td>N</td>
<td>118.59</td>
</tr>
<tr>
<td>57</td>
<td>Leu</td>
<td>H</td>
<td>H</td>
<td>8.65</td>
</tr>
<tr>
<td>57</td>
<td>Leu</td>
<td>N</td>
<td>N</td>
<td>121.04</td>
</tr>
<tr>
<td>58</td>
<td>Gly</td>
<td>H</td>
<td>H</td>
<td>8.15</td>
</tr>
<tr>
<td>58</td>
<td>Gly</td>
<td>N</td>
<td>N</td>
<td>106.68</td>
</tr>
<tr>
<td>59</td>
<td>Gln</td>
<td>H</td>
<td>H</td>
<td>7.59</td>
</tr>
<tr>
<td>59</td>
<td>Gln</td>
<td>N</td>
<td>N</td>
<td>120.52</td>
</tr>
<tr>
<td>60</td>
<td>Tyr</td>
<td>H</td>
<td>H</td>
<td>8.55</td>
</tr>
<tr>
<td>60</td>
<td>Tyr</td>
<td>N</td>
<td>N</td>
<td>124.17</td>
</tr>
<tr>
<td>61</td>
<td>Ile</td>
<td>H</td>
<td>H</td>
<td>8.3</td>
</tr>
<tr>
<td>61</td>
<td>Ile</td>
<td>N</td>
<td>N</td>
<td>118.8</td>
</tr>
<tr>
<td>62</td>
<td>Met</td>
<td>H</td>
<td>H</td>
<td>7.78</td>
</tr>
<tr>
<td>62</td>
<td>Met</td>
<td>N</td>
<td>N</td>
<td>115.6</td>
</tr>
<tr>
<td>63</td>
<td>Thr</td>
<td>H</td>
<td>H</td>
<td>8.43</td>
</tr>
<tr>
<td>63</td>
<td>Thr</td>
<td>N</td>
<td>N</td>
<td>116.13</td>
</tr>
<tr>
<td>64</td>
<td>Lys</td>
<td>H</td>
<td>H</td>
<td>7.71</td>
</tr>
<tr>
<td>64</td>
<td>Lys</td>
<td>N</td>
<td>N</td>
<td>118.48</td>
</tr>
<tr>
<td>65</td>
<td>Arg</td>
<td>H</td>
<td>H</td>
<td>7.48</td>
</tr>
<tr>
<td>65</td>
<td>Arg</td>
<td>N</td>
<td>N</td>
<td>115.99</td>
</tr>
<tr>
<td>66</td>
<td>Leu</td>
<td>H</td>
<td>H</td>
<td>7.52</td>
</tr>
<tr>
<td>66</td>
<td>Leu</td>
<td>N</td>
<td>N</td>
<td>115.23</td>
</tr>
<tr>
<td>67</td>
<td>Tyr</td>
<td>H</td>
<td>H</td>
<td>8.44</td>
</tr>
<tr>
<td>67</td>
<td>Tyr</td>
<td>N</td>
<td>N</td>
<td>123.12</td>
</tr>
<tr>
<td>68</td>
<td>Asp</td>
<td>H</td>
<td>H</td>
<td>8.48</td>
</tr>
<tr>
<td>68</td>
<td>Asp</td>
<td>N</td>
<td>N</td>
<td>124.06</td>
</tr>
<tr>
<td>69</td>
<td>Glu</td>
<td>H</td>
<td>H</td>
<td>8.41</td>
</tr>
<tr>
<td>69</td>
<td>Glu</td>
<td>N</td>
<td>N</td>
<td>122.25</td>
</tr>
<tr>
<td>70</td>
<td>Lys</td>
<td>H</td>
<td>H</td>
<td>8.15</td>
</tr>
<tr>
<td>70</td>
<td>Lys</td>
<td>N</td>
<td>N</td>
<td>116.47</td>
</tr>
<tr>
<td>71</td>
<td>Gln</td>
<td>H</td>
<td>H</td>
<td>8.25</td>
</tr>
<tr>
<td>71</td>
<td>Gln</td>
<td>N</td>
<td>N</td>
<td>121.04</td>
</tr>
<tr>
<td>72</td>
<td>Gln</td>
<td>H</td>
<td>H</td>
<td>8.13</td>
</tr>
<tr>
<td>72</td>
<td>Gln</td>
<td>N</td>
<td>N</td>
<td>118.83</td>
</tr>
<tr>
<td>73</td>
<td>His</td>
<td>H</td>
<td>H</td>
<td>7.64</td>
</tr>
<tr>
<td>73</td>
<td>His</td>
<td>N</td>
<td>N</td>
<td>112.5</td>
</tr>
<tr>
<td>74</td>
<td>Ile</td>
<td>H</td>
<td>H</td>
<td>7.54</td>
</tr>
<tr>
<td>74</td>
<td>Ile</td>
<td>N</td>
<td>N</td>
<td>122.41</td>
</tr>
<tr>
<td>75</td>
<td>Val</td>
<td>H</td>
<td>H</td>
<td>8.22</td>
</tr>
<tr>
<td>75</td>
<td>Val</td>
<td>N</td>
<td>N</td>
<td>126.17</td>
</tr>
<tr>
<td>76</td>
<td>Tyr</td>
<td>H</td>
<td>H</td>
<td>8.2</td>
</tr>
<tr>
<td>76</td>
<td>Tyr</td>
<td>N</td>
<td>N</td>
<td>125.62</td>
</tr>
<tr>
<td>77</td>
<td>Cys</td>
<td>H</td>
<td>H</td>
<td>8.28</td>
</tr>
<tr>
<td>77</td>
<td>Cys</td>
<td>N</td>
<td>N</td>
<td>117.1</td>
</tr>
<tr>
<td>78</td>
<td>Ser</td>
<td>H</td>
<td>H</td>
<td>7.61</td>
</tr>
<tr>
<td>78</td>
<td>Ser</td>
<td>N</td>
<td>N</td>
<td>116.47</td>
</tr>
<tr>
<td>79</td>
<td>Asn</td>
<td>H</td>
<td>H</td>
<td>8.59</td>
</tr>
<tr>
<td>79</td>
<td>Asn</td>
<td>N</td>
<td>N</td>
<td>118.06</td>
</tr>
<tr>
<td>80</td>
<td>Asp</td>
<td>H</td>
<td>H</td>
<td>7.39</td>
</tr>
<tr>
<td>80</td>
<td>Asp</td>
<td>N</td>
<td>N</td>
<td>118.25</td>
</tr>
<tr>
<td>82</td>
<td>Leu</td>
<td>H</td>
<td>H</td>
<td>9.69</td>
</tr>
<tr>
<td>82</td>
<td>Leu</td>
<td>N</td>
<td>N</td>
<td>118.5</td>
</tr>
<tr>
<td>83</td>
<td>Gly</td>
<td>H</td>
<td>H</td>
<td>7.65</td>
</tr>
<tr>
<td>83</td>
<td>Gly</td>
<td>N</td>
<td>N</td>
<td>106.57</td>
</tr>
<tr>
<td>84</td>
<td>Asp</td>
<td>H</td>
<td>H</td>
<td>7.39</td>
</tr>
<tr>
<td>84</td>
<td>Asp</td>
<td>N</td>
<td>N</td>
<td>122.94</td>
</tr>
<tr>
<td>85</td>
<td>Leu</td>
<td>H</td>
<td>H</td>
<td>7.92</td>
</tr>
<tr>
<td>85</td>
<td>Leu</td>
<td>N</td>
<td>N</td>
<td>118.55</td>
</tr>
<tr>
<td>87</td>
<td>Gly</td>
<td>H</td>
<td>H</td>
<td>8.17</td>
</tr>
<tr>
<td>87</td>
<td>Gly</td>
<td>N</td>
<td>N</td>
<td>108.21</td>
</tr>
<tr>
<td>88</td>
<td>Val</td>
<td>H</td>
<td>H</td>
<td>6.96</td>
</tr>
<tr>
<td>88</td>
<td>Val</td>
<td>N</td>
<td>N</td>
<td>108.45</td>
</tr>
<tr>
<td>90</td>
<td>Ser</td>
<td>H</td>
<td>H</td>
<td>7.12</td>
</tr>
<tr>
<td>90</td>
<td>Ser</td>
<td>N</td>
<td>N</td>
<td>111.23</td>
</tr>
<tr>
<td>91</td>
<td>Phe</td>
<td>H</td>
<td>H</td>
<td>8.11</td>
</tr>
<tr>
<td>91</td>
<td>Phe</td>
<td>N</td>
<td>N</td>
<td>114.38</td>
</tr>
<tr>
<td>92</td>
<td>Ser</td>
<td>H</td>
<td>H</td>
<td>8.62</td>
</tr>
<tr>
<td>92</td>
<td>Ser</td>
<td>N</td>
<td>N</td>
<td>112.67</td>
</tr>
<tr>
<td>93</td>
<td>Val</td>
<td>H</td>
<td>H</td>
<td>9.31</td>
</tr>
<tr>
<td>93</td>
<td>Val</td>
<td>N</td>
<td>N</td>
<td>123.22</td>
</tr>
<tr>
<td>94</td>
<td>Lys</td>
<td>H</td>
<td>H</td>
<td>7.86</td>
</tr>
<tr>
<td></td>
<td>Residue</td>
<td>Charge</td>
<td>pKᵢ</td>
<td>pKᵢ+1</td>
</tr>
<tr>
<td>---</td>
<td>---------</td>
<td>--------</td>
<td>------</td>
<td>--------</td>
</tr>
<tr>
<td>94</td>
<td>Lys</td>
<td>N</td>
<td>N</td>
<td>116.45</td>
</tr>
<tr>
<td>95</td>
<td>Glu</td>
<td>H</td>
<td>H</td>
<td>7.24</td>
</tr>
<tr>
<td>95</td>
<td>Glu</td>
<td>N</td>
<td>N</td>
<td>119.79</td>
</tr>
<tr>
<td>96</td>
<td>His</td>
<td>H</td>
<td>H</td>
<td>7.43</td>
</tr>
<tr>
<td>96</td>
<td>His</td>
<td>N</td>
<td>N</td>
<td>120.67</td>
</tr>
<tr>
<td>97</td>
<td>Arg</td>
<td>H</td>
<td>H</td>
<td>8.67</td>
</tr>
<tr>
<td>97</td>
<td>Arg</td>
<td>N</td>
<td>N</td>
<td>118.17</td>
</tr>
<tr>
<td>98</td>
<td>Lys</td>
<td>H</td>
<td>H</td>
<td>7.61</td>
</tr>
<tr>
<td>98</td>
<td>Lys</td>
<td>N</td>
<td>N</td>
<td>120.79</td>
</tr>
<tr>
<td>99</td>
<td>Ile</td>
<td>H</td>
<td>H</td>
<td>7.82</td>
</tr>
<tr>
<td>99</td>
<td>Ile</td>
<td>N</td>
<td>N</td>
<td>118.88</td>
</tr>
<tr>
<td>100</td>
<td>Tyr</td>
<td>H</td>
<td>H</td>
<td>8.37</td>
</tr>
<tr>
<td>100</td>
<td>Tyr</td>
<td>N</td>
<td>N</td>
<td>120.63</td>
</tr>
<tr>
<td>101</td>
<td>Thr</td>
<td>H</td>
<td>H</td>
<td>7.99</td>
</tr>
<tr>
<td>101</td>
<td>Thr</td>
<td>N</td>
<td>N</td>
<td>113.88</td>
</tr>
<tr>
<td>102</td>
<td>Met</td>
<td>H</td>
<td>H</td>
<td>7.52</td>
</tr>
<tr>
<td>102</td>
<td>Met</td>
<td>N</td>
<td>N</td>
<td>119.83</td>
</tr>
<tr>
<td>103</td>
<td>Ile</td>
<td>H</td>
<td>H</td>
<td>8.07</td>
</tr>
<tr>
<td>103</td>
<td>Ile</td>
<td>N</td>
<td>N</td>
<td>120.42</td>
</tr>
<tr>
<td>104</td>
<td>Tyr</td>
<td>H</td>
<td>H</td>
<td>8.7</td>
</tr>
<tr>
<td>104</td>
<td>Tyr</td>
<td>N</td>
<td>N</td>
<td>119.82</td>
</tr>
<tr>
<td>110</td>
<td>Val</td>
<td>H</td>
<td>H</td>
<td>8.27</td>
</tr>
<tr>
<td>110</td>
<td>Val</td>
<td>N</td>
<td>N</td>
<td>126.64</td>
</tr>
<tr>
<td>115</td>
<td>Ser</td>
<td>H</td>
<td>H</td>
<td>8.22</td>
</tr>
<tr>
<td>115</td>
<td>Ser</td>
<td>N</td>
<td>N</td>
<td>116.71</td>
</tr>
</tbody>
</table>
… y colorín Colorado este cuento se ha acabado.