

## Computer Lab – Chimera 1

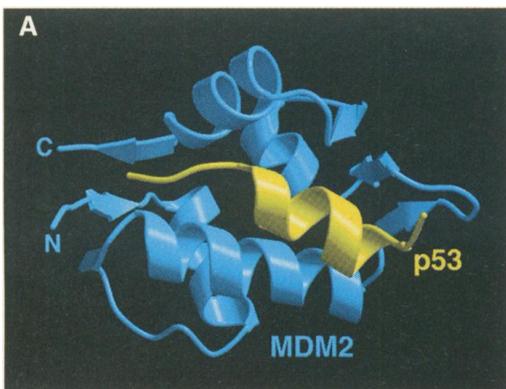
---

**Goal:** The purpose of this lab is to get familiar with protein structures determined by X-ray crystallography vs. nuclear magnetic resonance methods.

### Introduction

In this exercise, we will use a protein called MDM2 as an example. MDM2 is an important protein in cancer—it has the ability to bind the transactivation domain of tumor suppressor protein p53, downregulating its ability to activate transcription. Many researchers have sought to develop anti-cancer drugs that target the binding site of MDM2, blocking p53 from binding, and thus upregulating the tumor suppressor activity of p53. Several of these drugs are currently in clinical trials.<sup>1</sup>

The first high-resolution atomic structure of MDM2 bound to p53 came from an X-ray crystal structure by Kussie et al. (1996).<sup>2</sup> This structure shows a small helical region of p53 that binds a hydrophobic cleft in MDM2. To get these proteins to co-crystallize well, constructs with only residues 17-125 of MDM2, and residues 15-29 of p53, were used. Only the atomic structures of residues 25-109 could be resolved, however.



The X-ray crystal structure of MDM2-p53 (PDB ID: 1YCR)

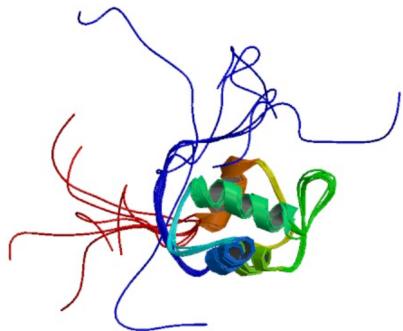
In 2005, Uhrinova et al. used 2D NMR spectroscopy to determine the *apo* structure of MDM2 (“*apo*” = unbound by ligand).<sup>3</sup> This structure contains more residues (1-119) but most of the terminal regions are unstructured.

---

<sup>1</sup> Hoe, K. K., Verma, C. S., & Lane, D. P. (2014). Drugging the p53 pathway: understanding the route to clinical efficacy. *Nature Reviews Drug Discovery*, 13(3), 217–236. doi:10.1038/nrd4236

<sup>2</sup> Kussie, P. H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A. J., & Pavletich, N. P. (1996). Structure of the MDM2 oncprotein bound to the p53 tumor suppressor transactivation domain. *Science*, 274(5289), 948–953.

<sup>3</sup> Uhrinova, S., Uhrin, D., Powers, H., Watt, K., Zheleva, D., Fischer, P., et al. (2005). Structure of Free MDM2 N-terminal Domain Reveals Conformational Adjustments that Accompany p53-binding. *Journal of Molecular Biology*, 350(3), 587–598. doi:10.1016/j.jmb.2005.05.010



The NMR structure of *apo*-MDM2 (PDB ID: 1Z1M)

### Viewing and visualizing PDB files

To find a given PDB structure, open a browser and go to <http://pdb.org>. In the search window, type in the PDB ID, which will bring you to the entry page. This page has a great deal of useful information about the protein sequence and structure.

To view the contents of the deposited PDB file containing the atomic coordinates, click on the **Display Files** drop-down in the upper right and choose **PDB File**. The full contents (text) of the PDB file should be viewable in your browser.

To view the molecular structure in 3D, open the **Chimera** application and choose **File > Fetch by ID ...**. Type the PDB ID into the text field and click **Fetch**. The ribbon structure of the protein should appear, which you can rotate in 3D by clicking and dragging.

#### Chimera Tips:

1. To view side chain atoms: Actions > Atoms/Bonds > show
2. A number of “preset” renderings are found under the Presets menu
3. To select atoms, Control-click
4. To select multiple atoms Shift-Control click
5. Option-click and mouse drag to translate the view
6. If your view gets off-center for any reason, try Action > Focus to recenter.

### Procedure

Use (1) the information in the Protein Data Bank (visual inspection of the \*.pdb file, and molecular visualization in Chimera), and (2) the original publications (posted on Blackboard) to answer the questions below.

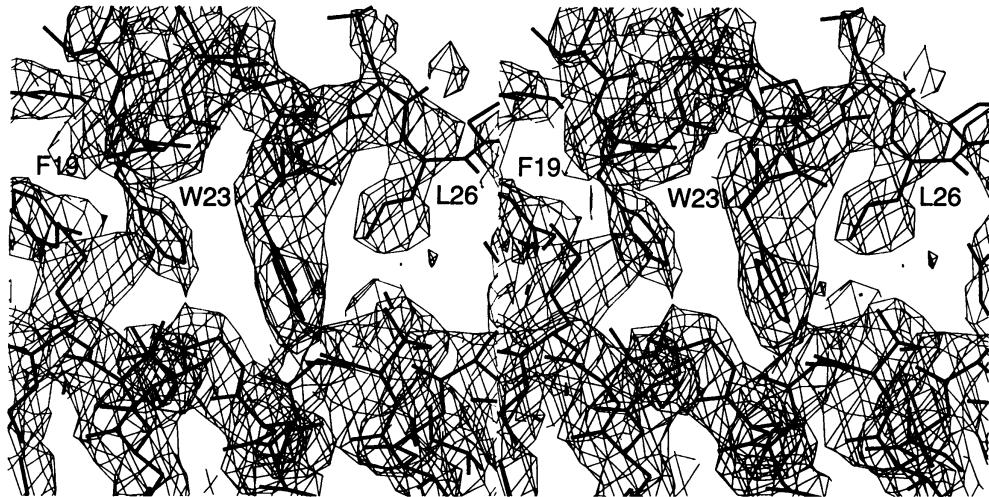
### X-ray crystal structure of MDM2-p53

**PDB:** The atomic coordinates for this structure are deposited in the Protein Data Bank under PDB ID: **1YCR**.

**Publication:** Kussie, P. H., Gorina, S., Marechal, V., Elenbaas, B., Moreau, J., Levine, A. J., & Pavletich, N. P. (1996). Structure of the MDM2 oncoprotein bound to the p53 tumor suppressor transactivation domain. *Science*, 274(5289), 948–953.

Last week we learned about X-ray crystallography, and how it can be used to determine the atomic structure of proteins. Recall that this involves: (1) collecting diffraction data (i.e. the intensities of a large number of diffraction spots and their identification with Miller indices and a unit cell), and (2) using a model of protein electron density to calculate (by Fourier Transform) computed structure factors  $F_c$  that can be compared to the observed structure factors  $|F_o|$ .<sup>4</sup>

The resolution of the computed electron density depends on the number of diffraction spots (i.e. the number of “reflections”). To give you some sense of the resolution of the final model, consider Figure 1 from Kussie et al (1996) below:



**Fig. 1.** MIR electron density map of the *X. laevis* MDM2-p53 interface at 3.0 Å resolution, contoured at 1.0  $\sigma$ , with the refined 2.3 Å resolution atomic model in a stick representation. Stereo view focuses on the interactions of Phe<sup>19</sup>, Trp<sup>23</sup>, and Leu<sup>26</sup> of p53 (labeled) with the  $\alpha$ 2 helix of MDM2.

As you can see, the final model of the electron density really is “blobby”, but it is high-resolution enough that backbone and side-chain atoms (solid black lines, hard to see, I know) can be fit unambiguously. Some regions of this density may be too diffuse to fit (often the thermally mobile termini). For these regions, the protein structure is typically left undefined (hence, “missing residues” are common in X-ray crystal structures.)

---

<sup>4</sup> Note that we are careful here to denote *absolute values* of  $F_o$  because of the so-called “phase problem” in which only *intensities* can be measured, not phase shifts.

**Table 1.** Statistics from the crystallographic analysis.

Data set	Resolution (Å)	Reflections		Data coverage (%)	$R_{\text{sym}}^*$	MFID†	Phasing power‡		
		Measured	Unique						
Native ( <i>X. laevis</i> )	2.3	15953	6257	97.0	0.055				
Native (human)	2.6	16069	3559	90.0	0.052				
Thimerosal	3.0	6707	2886	94.8	0.056	0.25	1.9		
$\text{UO}_2(\text{OAc})_2$	3.1	6039	2545	93.2	0.078	0.12	1.4		
$\text{UO}_2(\text{OAc})_2 + \text{K}_2\text{Pt}(\text{CN})_4$	3.2	14350	2181	86.9	0.088	0.16	1.8		
$\text{K}_2\text{Au}(\text{CN})_4$	3.0	4316	2592	86.4	0.024	0.08	0.8		
Refinement statistics									
Resolution (Å)	Reflections§	Protein atoms	Water atoms	$R_{\parallel}$	$R_{\text{free}}\ddagger$	rmsd#			
						Bonds (Å)	Angles (°)		
<i>X. laevis</i>	7.0–2.3	5423	801	40	0.188	0.253	0.010	1.39	3.1
Human	8.0–2.6	3293	807		0.200	0.276	0.013	1.55	2.7

\* $R_{\text{sym}} = \sum_i \sum_j |I_{h,i} - I_{j,i}| / \sum_i \sum_j I_{h,i}$  for the intensity ( $I$ ) of  $i$  observations of reflection  $h$ . †MFID, mean isomorphous difference =  $\sum |F_{\text{PH}} - F_{\text{P}}| / \sum F_{\text{PH}}$ , where  $F_{\text{PH}}$  and  $F_{\text{P}}$  are the derivative and native structure factors, respectively.

‡Phasing power =  $\{[F_{\text{HgO}}]^2 / [F_{\text{PH(O)}} - F_{\text{PH(C)}}]\}^{1/2}$ . §Reflections with  $|F| > 2\sigma$  and  $|F| > 1\sigma$  for the data sets from *X. laevis* and humans, respectively. ¶R factor =  $\sum |F_{\text{o}} - F_{\text{c}}| / \sum |F_{\text{o}}|$ , where  $F_{\text{o}}$  and  $F_{\text{c}}$  are the observed and calculated structure factors, respectively. ¶Free R factor calculated from 10 percent of the data chosen randomly and omitted from simulated annealing refinement from 3000 K. #Root-mean-square deviations from ideal geometry and root-mean-square variation of the  $B$  factors of bonded atoms.

## Questions

1. Consider **Table 1** from Kussie et al. (1996), in which statistics are given for a number of similar crystal structures the authors obtained.
  - a. Why is the number of unique reflections smaller than the total number?
  - b. Explain the correlation between the number of unique reflections and the resolution of the crystal structure.
2. The  $R$ -factor is a measure of the quality of the (electron density) model, defined as  $R = \sum |F_{\text{o}} - F_{\text{c}}| / \sum |F_{\text{o}}|$ , where  $F_{\text{o}}$  and  $F_{\text{c}}$  are the observed and calculated structure factors, respectively. Because of the iterative refinement of computed structure factors from the observed data, there is a danger of overfitting. A more useful measure is called the  $R_{\text{free}}$  value, where 10% of the observed structure factors are left out of the data set, so they can be used for testing the predictions.
  - a. Which should have the larger value,  $R$  or  $R_{\text{free}}$ ?
  - b. Based on the reported values for (human) MDM2-p53, do you think overfitting is a problem for this structure?

3. Why doesn't this structure have any hydrogens?

In Chimera, go to **Tools > Higher-Order Structure > Unit Cell**, and click **Outline**. Click on **Make Copies** to see the symmetrical copies of MDM2 inside the unit cell.

4. What are the unit cell dimensions  $a$ ,  $b$  and  $c$ ? (in Å, see also <http://pdb.org>)
5. What *shape* is the unit cell?
6. Visualize the side chains using wire-frame (**Actions > Atoms/Bonds > Wire**).
  - a. Based on what you can see, does it look like there are strong interactions between unit cell copies?
  - b. How careful do you think we need to be about such crystal packing artifacts when interpreting X-ray structures of proteins?

Recall that the crystallographic B-factor reports how diffuse the model of electronic density is for each atom. It is highly correlated to thermal motion. In Chimera, color the protein by B-factor, using **Actions > Color > all options... > Tools > Render by Attribute** (choose bfactor in the pull-down menu for Attribute:, then click Apply)

7. What parts of the protein have the most thermal motion?

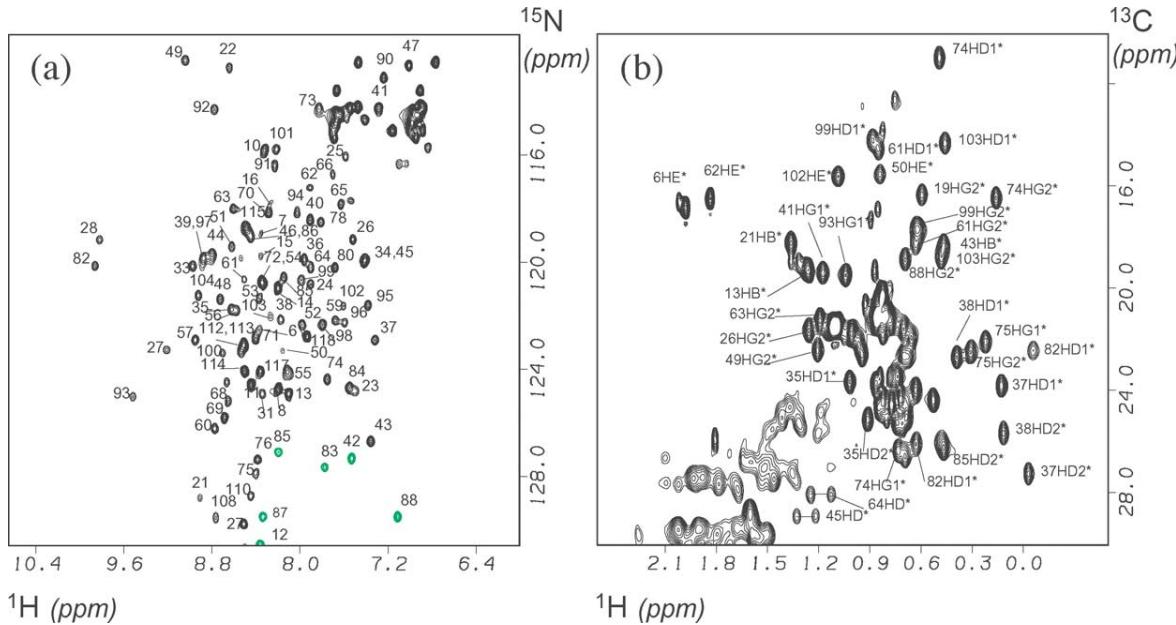
## NMR structure of *apo*-MDM2

**PDB:** The atomic coordinates for this structure are deposited in the Protein Data Bank under PDB ID: **1Z1M**.

**Publication:** Uhrinova, S., Uhrin, D., Powers, H., Watt, K., Zheleva, D., Fischer, P., et al. (2005). Structure of Free MDM2 N-terminal Domain Reveals Conformational Adjustments that Accompany p53-binding. *Journal of Molecular Biology*, 350 (3), 587–598. doi:10.1016/j.jmb.2005.05.010

Solution structures of proteins are usually determined using 2D NMR spectroscopy. Two-dimensional techniques measure how the magnetization of one kind of nuclei (say,  $^1\text{H}$ ) is affected by another kind of nuclei (say,  $^{15}\text{N}$ ). The first step in 2D NMR is to *assign* the various magnetic resonances to particular protons. As long as the spectroscopic peaks are disperse (i.e. well-defined and far enough apart to be distinguished from one another), assignment is possible because resonance can only occur for nearby (through-bond) nuclei. For example, Figure 1 in

Uhrinova et al. (2005) shows HSQC (heteronuclear single quantum coherence) spectra disperse enough to assign (a) backbone NH protons and (b) side-chain protons.



**Figure 1.** Assigned HSQC spectra of MDM2<sup>N</sup>. (a) <sup>1</sup>H,<sup>15</sup>N HSQC spectrum. Folded peaks are shown in grey. <sup>HN-NH</sup> assignments are indicated by residue number. (b) A portion of the <sup>1</sup>H,<sup>13</sup>C HSQC spectrum showing methyl groups, with assignments indicated by residue and atom identifier (based on CNS notation).

Once the peaks are assigned, studies of NMR relaxation times can determine the presence of Nuclear Overhauser Effects, or **NOEs**. An NOE is a short-range transfer of nuclear spin polarization from one nuclei to another that occurs through *space* (rather than through-bond). Like its optical cousin FRET (Förster Resonance Energy Transfer), it is an induced-dipole-dipole interaction that falls off as  $\sim 1/r^6$ . If an NOE is observed between two protons, it means that they must be close in three-dimensional space, with the strength of the NOE related to proximity, for example:

1. Strong NOEs: 1.7 Å to 2.5 Å
2. Medium strong NOEs: 2.5 Å to 3.5 Å
3. Medium weak NOEs: 3.5 Å to 4.5 Å
4. Weak NOEs: 4.5 Å to 5.5 Å

A set of measured NOE distances can put enough restraints on the set of possible protein structures, so as to completely determine the solution conformation. A computer is used to explore possible structures that best satisfy the NOE distance restraints, generating an *ensemble* of different conformations, each equally compatible with the experimental restraints.

## Questions

In Chimera, load structure **1Z1M**.

1. Why does this structure include protons, whereas the X-ray crystal structure does not?
2. Why are there so many structures in this PDB?
3. Obviously, there are very few NOE distance restraints for the N-terminal region. The modeled ensemble has many possible conformations for this region. How well do you think this represents reality? In other words: how much is the conformational heterogeneity an artifact of the computer models used for structural refinement, versus a real feature of the MDM2 N-terminal region? Can you think of any other experimental biophysical methods that you could use to answer this question?

Bonus: Load both the X-ray structure (1YCR) and the NMR structure (1Z1M) into Chimera and align them using **Tools > Structure Comparison > MatchMaker** (select 1YCR as the reference, and shift-click to select all NMR structures to match).

4. Can you see any changes in the binding cleft as described by Uhrinova et al.?