



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 31, 2022 – 02:08 PM JST

PDB ID : 7V97  
Title : Arsenic-bound p53 DNA-binding domain mutant V272M  
Authors : Lu, M.; Xing, Y.F.; Wang, Z.Y.; Ni, Y.; Song, H.X.  
Deposited on : 2021-08-24  
Resolution : 2.02 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at <http://www.wwpdb.org/validation/2017/FAQs#types>.

---

The following versions of software and data (see [references ⓘ](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.30
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.30

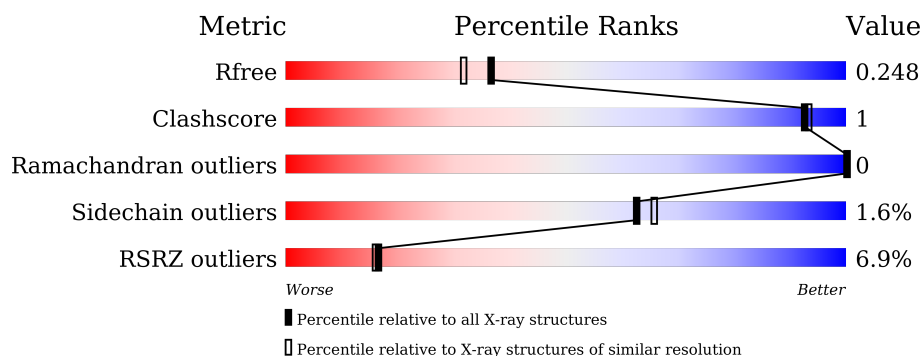
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.02 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	10434 (2.04-2.00)
Clashscore	141614	11643 (2.04-2.00)
Ramachandran outliers	138981	11493 (2.04-2.00)
Sidechain outliers	138945	11492 (2.04-2.00)
RSRZ outliers	127900	10220 (2.04-2.00)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions  $\leq 5\%$ . The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	200	<div> <div>3%</div> <div>90%</div> <div>6%</div> </div>
1	B	200	<div> <div>5%</div> <div>90%</div> <div>6%</div> </div>
1	C	200	<div> <div>6%</div> <div>92%</div> <div>.</div> </div>
1	D	200	<div> <div>12%</div> <div>93%</div> <div>.</div> </div>

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6196 atoms, of which 0 are hydrogens and 0 are deuteriums.

In the tables below, the ZeroOcc column contains the number of atoms modelled with zero occupancy, the AltConf column contains the number of residues with at least one atom in alternate conformation and the Trace column contains the number of residues modelled with at most 2 atoms.

- Molecule 1 is a protein called Cellular tumor antigen p53.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	194	Total	C	N	O	S	0	2	0
			1533	944	282	289	18			
1	B	193	Total	C	N	O	S	0	4	0
			1547	953	287	289	18			
1	C	194	Total	C	N	O	S	0	2	0
			1533	944	282	289	18			
1	D	192	Total	C	N	O	S	0	1	0
			1514	933	279	285	17			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	272	MET	VAL	variant	UNP P04637
B	272	MET	VAL	variant	UNP P04637
C	272	MET	VAL	variant	UNP P04637
D	272	MET	VAL	variant	UNP P04637

- Molecule 2 is ARSENIC (three-letter code: ARS) (formula: As) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total	As	0	0
			1	1		
2	B	1	Total	As	0	0
			1	1		
2	C	1	Total	As	0	0
			1	1		
2	D	1	Total	As	0	0
			1	1		

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	Zn 1	0	0
3	B	1	Total 1	Zn 1	0	0
3	C	1	Total 1	Zn 1	0	0
3	D	1	Total 1	Zn 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	18	Total 18	O 18	0	0
4	B	11	Total 11	O 11	0	0
4	C	16	Total 16	O 16	0	0
4	D	16	Total 16	O 16	0	0

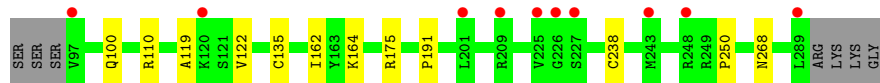
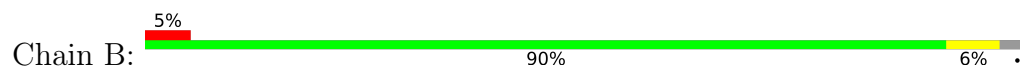
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density ( $RSRZ > 2$ ). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

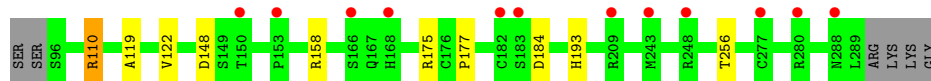
- Molecule 1: Cellular tumor antigen p53



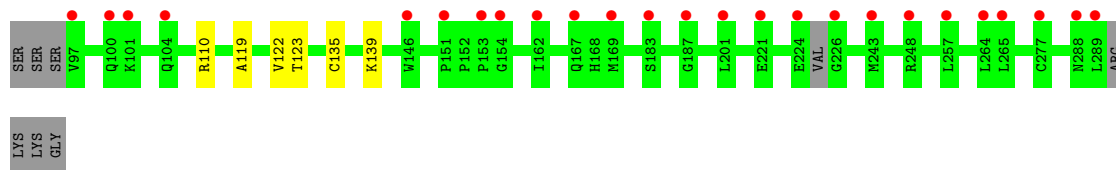
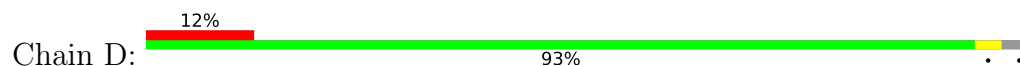
- Molecule 1: Cellular tumor antigen p53



- Molecule 1: Cellular tumor antigen p53



- Molecule 1: Cellular tumor antigen p53



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	68.84Å 71.52Å 84.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.80 – 2.02 42.80 – 2.02	Depositor EDS
% Data completeness (in resolution range)	98.3 (42.80-2.02) 99.8 (42.80-2.02)	Depositor EDS
$R_{merge}$	0.18	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.20 (at 2.01Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, $R_{free}$	0.234 , 0.251 0.234 , 0.248	Depositor DCC
$R_{free}$ test set	2593 reflections (4.80%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	24.4	Xtriage
Anisotropy	0.062	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 39.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtriage
Estimated twinning fraction	0.023 for -k,-h,-l 0.018 for k,h,-l 0.110 for h,-k,-l	Xtriage
Reported twinning fraction	0.568 for H, K, L 0.432 for -h,-k,l	Depositor
Outliers	0 of 54021 reflections	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	6196	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	27.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.89% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup> Intensities estimated from amplitudes.

<sup>2</sup> Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

## 5 Model quality [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: ZN, ARS

The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 5$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Chain	Bond lengths		Bond angles	
		RMSZ	$\# Z  > 5$	RMSZ	$\# Z  > 5$
1	A	0.67	0/1571	0.72	0/2130
1	B	0.67	0/1585	0.70	0/2147
1	C	0.67	0/1571	0.71	0/2129
1	D	0.67	0/1551	0.71	0/2100
All	All	0.67	0/6278	0.71	0/8506

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

In the following table, the Non-H and H(model) columns list the number of non-hydrogen atoms and hydrogen atoms in the chain respectively. The H(added) column lists the number of hydrogen atoms added and optimized by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, whereas Symm-Clashes lists symmetry-related clashes.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1533	0	1487	7	0
1	B	1547	0	1506	4	0
1	C	1533	0	1486	7	0
1	D	1514	0	1467	2	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
2	D	1	0	0	0	0
3	A	1	0	0	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	18	0	0	0	0
4	B	11	0	0	0	0
4	C	16	0	0	0	0
4	D	16	0	0	0	0
All	All	6196	0	5946	18	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 1.

All (18) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:263:ASN:HD21	1:C:177:PRO:HB2	1.11	1.08
1:A:263:ASN:ND2	1:C:177:PRO:HB2	1.94	0.82
1:B:119:ALA:O	1:B:122:VAL:HG22	1.94	0.68
1:A:110:ARG:NH1	1:A:148:ASP:OD1	2.42	0.52
1:C:175:ARG:NH2	1:C:184:ASP:OD2	2.47	0.47
1:C:110:ARG:NH1	1:C:148:ASP:OD1	2.48	0.47
1:D:123:THR:CG2	1:D:139:LYS:HB3	2.45	0.47
1:B:100:GLN:OE1	1:B:162:ILE:HD11	2.15	0.46
1:A:119:ALA:O	1:A:122:VAL:HG22	2.17	0.45
1:B:175:ARG:HD3	1:B:191:PRO:O	2.17	0.45
1:A:158:ARG:HB3	1:A:256:THR:OG1	2.17	0.44
1:D:119:ALA:O	1:D:122:VAL:HG22	2.18	0.44
1:C:158:ARG:HB3	1:C:256:THR:OG1	2.20	0.42
1:A:259:ASP:OD1	1:A:259:ASP:C	2.59	0.41
1:C:119:ALA:O	1:C:122:VAL:HG22	2.21	0.41
1:A:133:MET:HB3	1:A:272:MET:SD	2.61	0.41
1:B:164[B]:LYS:HE3	1:B:250:PRO:HB2	2.01	0.41
1:C:175:ARG:HD2	1:C:193:HIS:O	2.21	0.40

There are no symmetry-related clashes.



## 5.3 Torsion angles

### 5.3.1 Protein backbone

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	194/200 (97%)	191 (98%)	3 (2%)	0	100	100
1	B	195/200 (98%)	192 (98%)	3 (2%)	0	100	100
1	C	194/200 (97%)	192 (99%)	2 (1%)	0	100	100
1	D	189/200 (94%)	183 (97%)	6 (3%)	0	100	100
All	All	772/800 (96%)	758 (98%)	14 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 5.3.2 Protein sidechains

In the following table, the Percentiles column shows the percent sidechain outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the sidechain conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	176/179 (98%)	172 (98%)	4 (2%)	50	51
1	B	177/179 (99%)	173 (98%)	4 (2%)	50	51
1	C	176/179 (98%)	175 (99%)	1 (1%)	86	89
1	D	173/179 (97%)	171 (99%)	2 (1%)	71	75
All	All	702/716 (98%)	691 (98%)	11 (2%)	62	66

All (11) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	110	ARG
1	A	123	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	135	CYS
1	A	238	CYS
1	B	110	ARG
1	B	135	CYS
1	B	238	CYS
1	B	268	ASN
1	C	110	ARG
1	D	110	ARG
1	D	135	CYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	210	ASN
1	A	263	ASN
1	C	210	ASN
1	C	235	ASN
1	C	239	ASN
1	C	288	ASN
1	D	214	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 8 ligands modelled in this entry, 8 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers

There are no such residues in this entry.

## 5.8 Polymer linkage issues

There are no chain breaks in this entry.

## 6 Fit of model and data

### 6.1 Protein, DNA and RNA chains

In the following table, the column labelled ‘#RSRZ > 2’ contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95<sup>th</sup> percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled ‘Q < 0.9’ lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2	OWAB(Å <sup>2</sup> )	Q<0.9
1	A	194/200 (97%)	0.35	6 (3%) 49 48	12, 22, 32, 47	0
1	B	193/200 (96%)	0.55	10 (5%) 27 26	16, 26, 40, 48	0
1	C	194/200 (97%)	0.71	12 (6%) 20 19	14, 27, 46, 57	0
1	D	192/200 (96%)	0.91	25 (13%) 3 3	16, 29, 49, 56	0
All	All	773/800 (96%)	0.63	53 (6%) 16 16	12, 26, 45, 57	0

All (53) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	289	LEU	4.7
1	C	183	SER	4.7
1	B	226	GLY	4.4
1	C	166	SER	4.3
1	C	280	ARG	4.3
1	D	153	PRO	4.3
1	C	248	ARG	4.1
1	B	289	LEU	4.1
1	D	243	MET	4.0
1	C	153	PRO	3.9
1	D	288	ASN	3.7
1	B	225	VAL	3.7
1	D	97	VAL	3.6
1	D	201	LEU	3.3
1	D	265	LEU	3.2
1	A	225	VAL	3.2
1	D	226	GLY	3.2
1	C	150	THR	3.1
1	C	288	ASN	3.0
1	C	277	CYS	3.0
1	C	243	MET	2.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	146	TRP	2.7
1	A	153	PRO	2.7
1	D	224	GLU	2.7
1	D	289	LEU	2.7
1	D	248	ARG	2.6
1	D	221	GLU	2.6
1	B	227	SER	2.6
1	C	182[A]	CYS	2.6
1	A	288	ASN	2.6
1	D	187	GLY	2.5
1	D	183	SER	2.5
1	B	209	ARG	2.5
1	B	243	MET	2.4
1	A	248	ARG	2.4
1	D	104	GLN	2.4
1	C	209	ARG	2.3
1	D	151	PRO	2.3
1	D	169	MET	2.3
1	B	120	LYS	2.3
1	D	101	LYS	2.3
1	B	97	VAL	2.3
1	C	168	HIS	2.3
1	B	248	ARG	2.2
1	D	277	CYS	2.2
1	D	162	ILE	2.2
1	D	154	GLY	2.2
1	D	167	GLN	2.1
1	B	201	LEU	2.1
1	A	96	SER	2.1
1	D	257	LEU	2.1
1	D	264	LEU	2.1
1	D	100	GLN	2.1

## 6.2 Non-standard residues in protein, DNA, RNA chains [i](#)

There are no non-standard protein/DNA/RNA residues in this entry.

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands

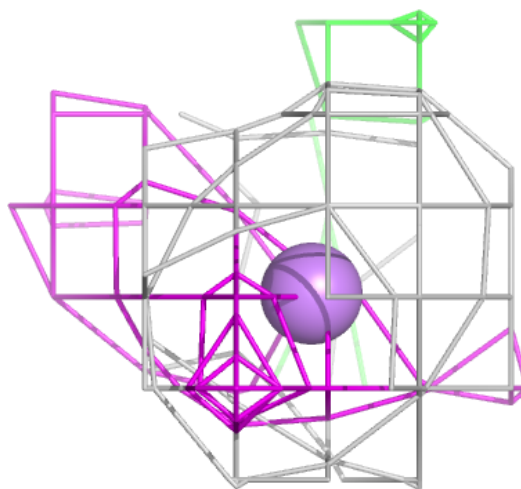
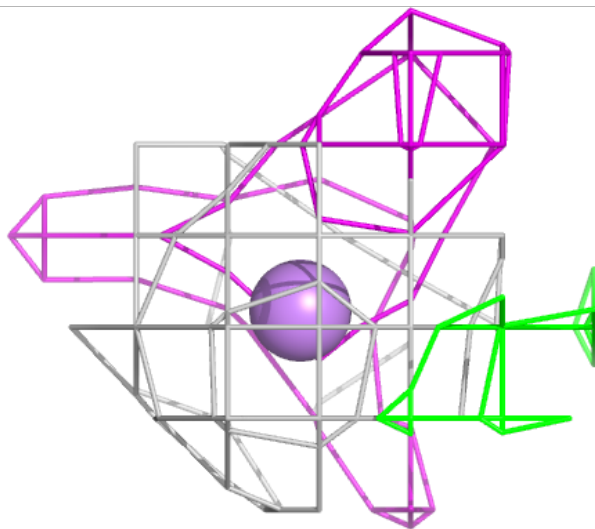
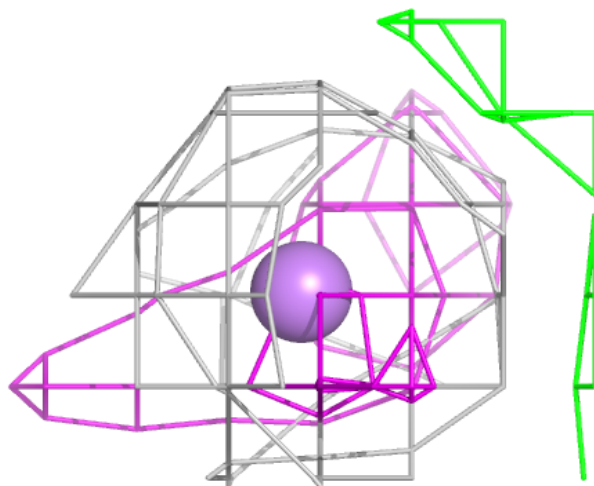
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95<sup>th</sup> percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' lists the number of atoms with occupancy less than 0.9.

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors( $\text{\AA}^2$ )	Q<0.9
2	ARS	C	301	1/1	0.85	0.11	29,29,29,29	1
2	ARS	D	301	1/1	0.91	0.09	23,23,23,23	1
2	ARS	A	301	1/1	0.96	0.07	21,21,21,21	1
2	ARS	B	301	1/1	0.96	0.06	29,29,29,29	1
3	ZN	C	302	1/1	0.97	0.06	22,22,22,22	1
3	ZN	B	302	1/1	0.98	0.12	22,22,22,22	1
3	ZN	D	302	1/1	0.98	0.08	21,21,21,21	1
3	ZN	A	302	1/1	0.99	0.12	20,20,20,20	1

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

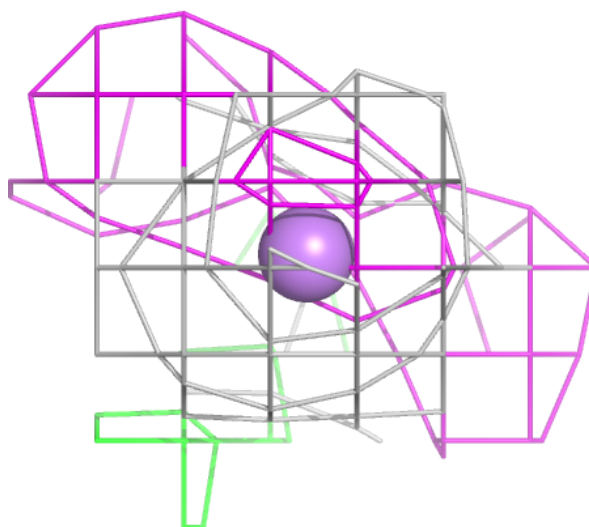
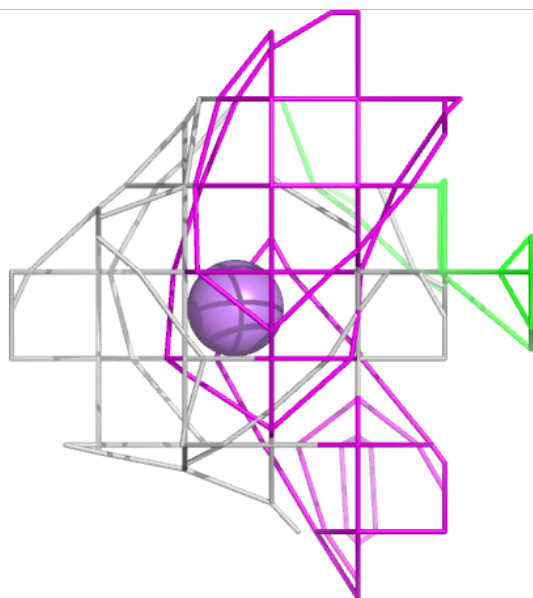
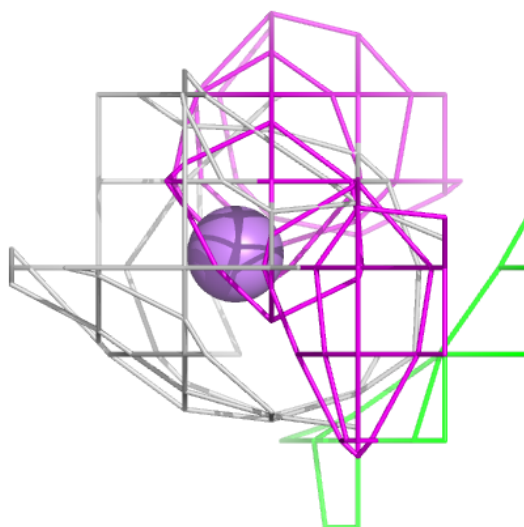
**Electron density around ARS C 301:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)



**Electron density around ARS D 301:**

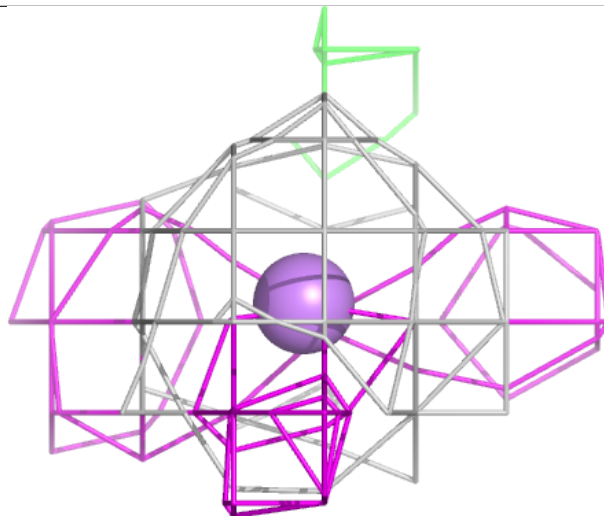
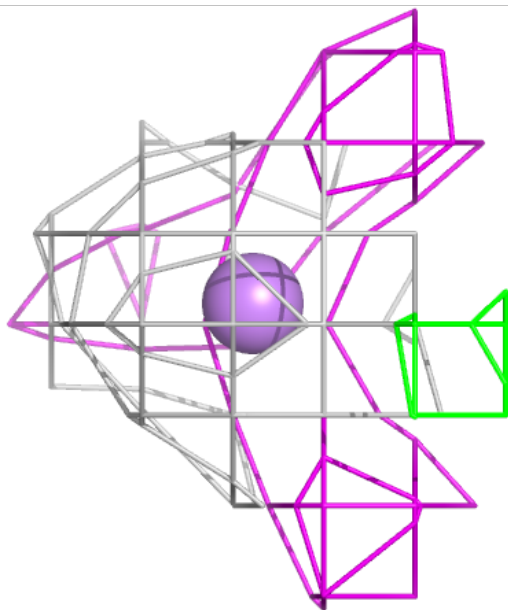
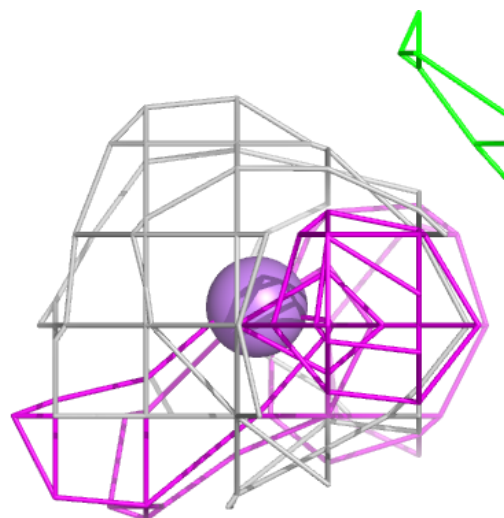
$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)

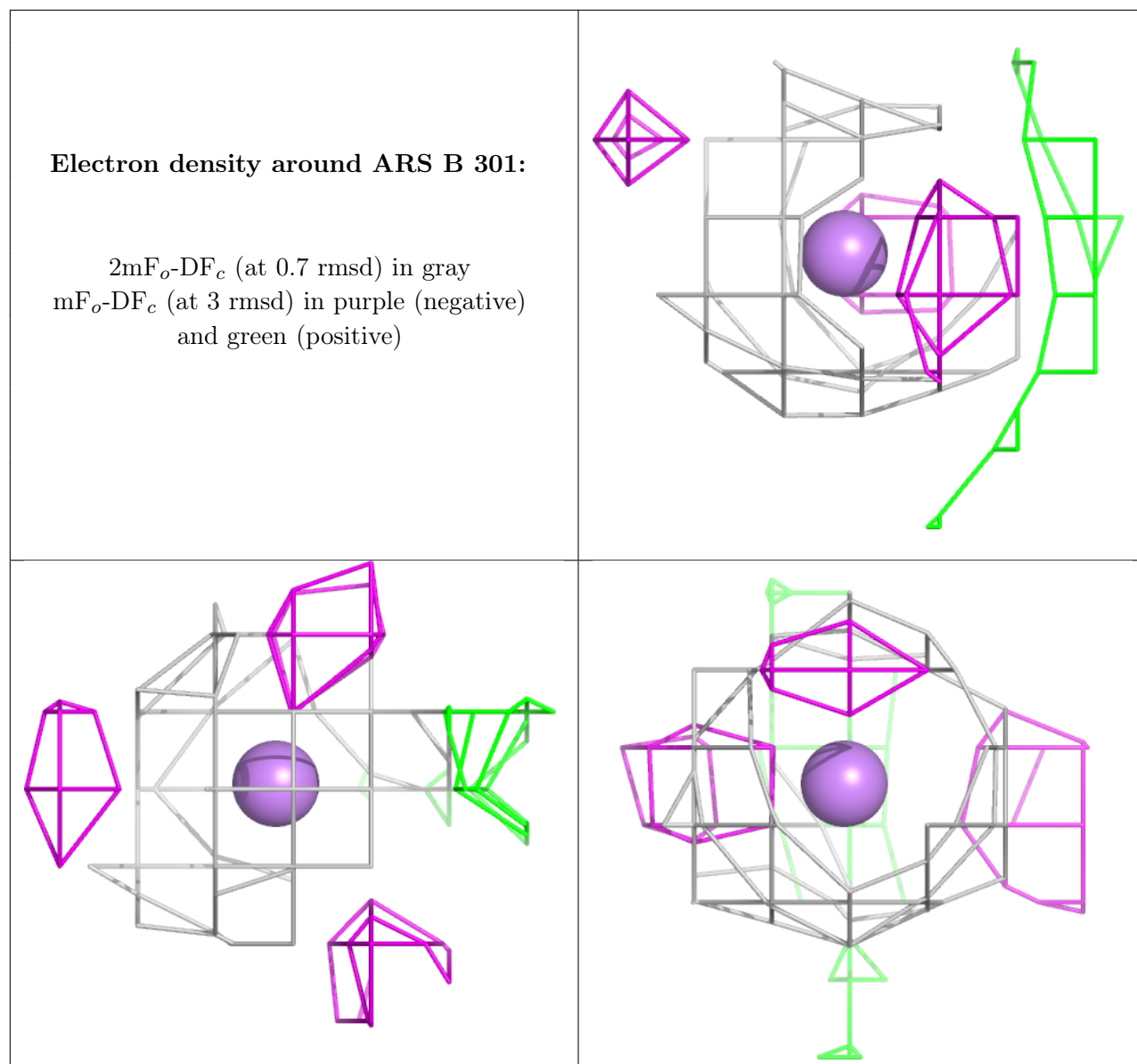




**Electron density around ARS A 301:**

$2mF_o-DF_c$  (at 0.7 rmsd) in gray  
 $mF_o-DF_c$  (at 3 rmsd) in purple (negative)  
and green (positive)





## 6.5 Other polymers ⓘ

There are no such residues in this entry.