



Full wwPDB X-ray Structure Validation Report ⓘ

May 29, 2020 – 11:38 am BST

PDB ID : 6VYC
Title : Crystal structure of WD-repeat domain of human WDR91
Authors : Halabelian, L.; Hutchinson, A.; Li, Y.; Seitova, A.; Bountra, C.; Edwards, A.M.; Arrowsmith, C.H.; Structural Genomics Consortium (SGC)
Deposited on : 2020-02-26
Resolution : 2.10 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

A user guide is available at

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

with specific help available everywhere you see the ⓘ symbol.

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 : 2.11
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) : 2.11

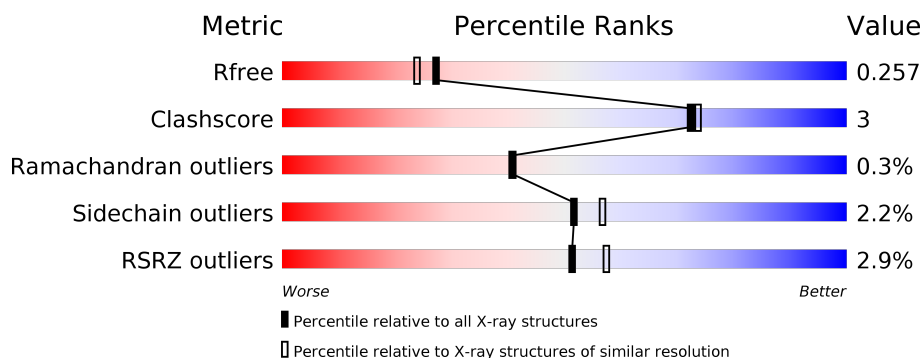
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.10 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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 on the lower bar indicate the fraction of residues that contain outliers for ≥ 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	374	<div> <div style="width: 100%; height: 10px; position: relative;"> <div style="position: absolute; top: -10px; left: 0;">%</div> <div style="position: absolute; top: 0; left: 0; width: 100%; height: 100%; background: linear-gradient(to right, red 0%, orange 1%, yellow 2, green 82%, yellow 89%, grey 100%);"></div> <div style="position: absolute; bottom: -10px; left: 0;">82% 7% 11%</div> </div> </div>
1	B	374	<div> <div style="width: 100%; height: 10px; position: relative;"> <div style="position: absolute; top: -10px; left: 0;">4%</div> <div style="position: absolute; top: 0; left: 0; width: 100%; height: 100%; background: linear-gradient(to right, red 0%, orange 4%, yellow 8%, green 79%, yellow 87%, grey 100%);"></div> <div style="position: absolute; bottom: -10px; left: 0;">79% 8% 12%</div> </div> </div>

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	UNX	A	801	-	-	-	X
2	UNX	A	803	-	-	-	X

2 Entry composition [i](#)

There are 3 unique types of molecules in this entry. The entry contains 5037 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 WD repeat-containing protein 91.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	333	Total	C	N	O	S	0	3	0
			2493	1577	418	477	21			
1	B	328	Total	C	N	O	S	0	2	0
			2413	1522	406	465	20			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	374	MET	-	expression tag	UNP A4D1P6
A	375	HIS	-	expression tag	UNP A4D1P6
A	376	HIS	-	expression tag	UNP A4D1P6
A	377	HIS	-	expression tag	UNP A4D1P6
A	378	HIS	-	expression tag	UNP A4D1P6
A	379	HIS	-	expression tag	UNP A4D1P6
A	380	HIS	-	expression tag	UNP A4D1P6
A	381	SER	-	expression tag	UNP A4D1P6
A	382	SER	-	expression tag	UNP A4D1P6
A	383	GLY	-	expression tag	UNP A4D1P6
A	384	ARG	-	expression tag	UNP A4D1P6
A	385	GLU	-	expression tag	UNP A4D1P6
A	386	ASN	-	expression tag	UNP A4D1P6
A	387	LEU	-	expression tag	UNP A4D1P6
A	388	TYR	-	expression tag	UNP A4D1P6
A	389	PHE	-	expression tag	UNP A4D1P6
A	390	GLN	-	expression tag	UNP A4D1P6
A	391	GLY	-	expression tag	UNP A4D1P6
B	374	MET	-	expression tag	UNP A4D1P6
B	375	HIS	-	expression tag	UNP A4D1P6
B	376	HIS	-	expression tag	UNP A4D1P6
B	377	HIS	-	expression tag	UNP A4D1P6
B	378	HIS	-	expression tag	UNP A4D1P6
B	379	HIS	-	expression tag	UNP A4D1P6
B	380	HIS	-	expression tag	UNP A4D1P6

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	381	SER	-	expression tag	UNP A4D1P6
B	382	SER	-	expression tag	UNP A4D1P6
B	383	GLY	-	expression tag	UNP A4D1P6
B	384	ARG	-	expression tag	UNP A4D1P6
B	385	GLU	-	expression tag	UNP A4D1P6
B	386	ASN	-	expression tag	UNP A4D1P6
B	387	LEU	-	expression tag	UNP A4D1P6
B	388	TYR	-	expression tag	UNP A4D1P6
B	389	PHE	-	expression tag	UNP A4D1P6
B	390	GLN	-	expression tag	UNP A4D1P6
B	391	GLY	-	expression tag	UNP A4D1P6

- Molecule 2 is UNKNOWN ATOM OR ION (three-letter code: UNX) (formula: X) (labeled as "Ligand of Interest" by author).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total X 1 1	0	0
2	A	3	Total X 3 3	0	0

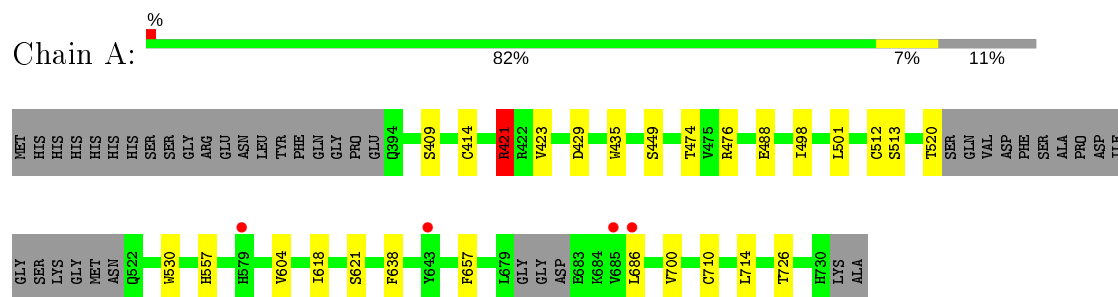
- Molecule 3 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	74	Total O 74 74	0	0
3	B	53	Total O 53 53	0	0

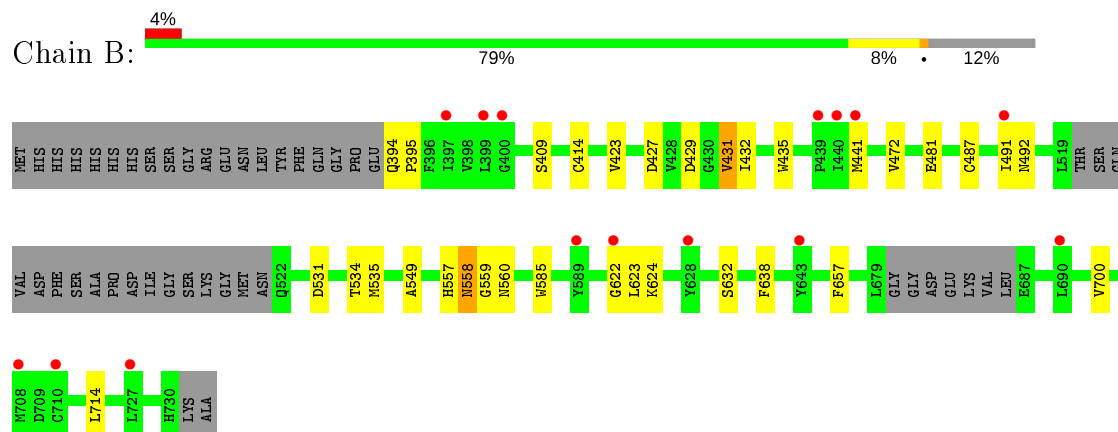
3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA 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: WD repeat-containing protein 91



- Molecule 1: WD repeat-containing protein 91



4 Data and refinement statistics

Property	Value	Source
Space group	P 2 ₁ 2 ₁ 2 ₁	Depositor
Cell constants a, b, c, α , β , γ	77.00 Å 84.05 Å 110.17 Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.07 – 2.10 46.07 – 2.10	Depositor EDS
% Data completeness (in resolution range)	97.3 (46.07-2.10) 97.3 (46.07-2.10)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.87 (at 2.10 Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.221 , 0.251 0.228 , 0.257	Depositor DCC
R_{free} test set	2085 reflections (5.05%)	wwPDB-VP
Wilson B-factor (Å ²)	34.2	Xtriage
Anisotropy	0.596	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 38.5	EDS
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o, F_c correlation	0.94	EDS
Total number of atoms	5037	wwPDB-VP
Average B, all atoms (Å ²)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.32% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical 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: UNX

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	0.66	0/2548	0.78	1/3465 (0.0%)
1	B	0.66	0/2464	0.78	0/3358
All	All	0.66	0/5012	0.78	1/6823 (0.0%)

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	421	ARG	CG-CD-NE	6.20	124.82	111.80

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	2493	0	2388	16	0
1	B	2413	0	2249	20	0
2	A	3	0	0	0	0
2	B	1	0	0	0	0
3	A	74	0	0	0	0
3	B	53	0	0	0	0
All	All	5037	0	4637	32	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 3.

All (32) 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:557[A]:HIS:NE2	1:B:557:HIS:CD2	1.82	1.47
1:A:557[A]:HIS:NE2	1:B:557:HIS:HD2	0.91	1.40
1:A:557[A]:HIS:CD2	1:B:557:HIS:HD2	1.62	1.16
1:A:557[A]:HIS:CD2	1:B:557:HIS:CD2	2.39	1.09
1:B:558:ASN:ND2	1:B:560:ASN:OD1	2.10	0.83
1:A:414[A]:CYS:SG	1:A:714:LEU:HD13	2.34	0.66
1:B:472:VAL:HG12	1:B:472:VAL:O	1.96	0.66
1:A:421:ARG:HG2	1:A:421:ARG:HH11	1.64	0.61
1:A:449:SER:OG	1:A:476:ARG:NH2	2.35	0.60
1:B:394:GLN:CB	1:B:395:PRO:HD3	2.33	0.59
1:B:414[A]:CYS:SG	1:B:714:LEU:HD23	2.46	0.56
1:B:409:SER:OG	1:B:429:ASP:OD2	2.22	0.56
1:B:487:CYS:SG	1:B:535:MET:HG2	2.45	0.56
1:A:409:SER:OG	1:A:429:ASP:OD2	2.23	0.55
1:B:491:ILE:HD12	1:B:491:ILE:N	2.24	0.51
1:B:558:ASN:HD22	1:B:559:GLY:N	2.11	0.49
1:B:394:GLN:CB	1:B:395:PRO:CD	2.91	0.48
1:B:585:TRP:CE3	1:B:622:GLY:HA3	2.50	0.47
1:B:491:ILE:HG22	1:B:492:ASN:N	2.29	0.47
1:A:604:VAL:HG13	1:A:618:ILE:HD13	1.98	0.46
1:A:638:PHE:CG	1:A:700:VAL:HG12	2.52	0.45
1:A:686:LEU:HD23	1:A:686:LEU:HA	1.86	0.44
1:A:474:THR:HB	1:A:488:GLU:HG3	2.00	0.44
1:B:638:PHE:CG	1:B:700:VAL:HG12	2.53	0.43
1:B:531:ASP:OD1	1:B:534:THR:HG22	2.19	0.43
1:A:501:LEU:HD23	1:A:512:CYS:HB3	2.01	0.42
1:A:710:CYS:SG	1:A:726:THR:HG23	2.59	0.42
1:B:423:VAL:HG11	1:B:435:TRP:CZ2	2.55	0.42
1:A:498:ILE:HA	1:A:513:SER:O	2.20	0.41
1:B:623:LEU:HD23	1:B:624:LYS:N	2.36	0.41
1:A:423:VAL:HG11	1:A:435:TRP:CZ2	2.56	0.41
1:B:427:ASP:OD2	1:B:431:VAL:HG22	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	330/374 (88%)	321 (97%)	8 (2%)	1 (0%)	41	41
1	B	324/374 (87%)	316 (98%)	7 (2%)	1 (0%)	41	41
All	All	654/748 (87%)	637 (97%)	15 (2%)	2 (0%)	41	41

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	621	SER
1	B	549	ALA

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	266/317 (84%)	262 (98%)	4 (2%)	65	71
1	B	249/317 (78%)	242 (97%)	7 (3%)	43	47
All	All	515/634 (81%)	504 (98%)	11 (2%)	52	59

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

Mol	Chain	Res	Type
1	A	421	ARG
1	A	520	THR
1	A	530	TRP
1	A	657	PHE

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	431	VAL
1	B	432	ILE
1	B	441	MET
1	B	481	GLU
1	B	558	ASN
1	B	632	SER
1	B	657	PHE

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	557	HIS
1	B	558	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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

5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 4 are unknown - 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 [i](#)

There are no such residues in this entry.

5.8 Polymer linkage issues [i](#)

There are no chain breaks in this entry.

6 Fit of model and data [i](#)

6.1 Protein, DNA and RNA chains [i](#)

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, 95th 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 ²)	Q<0.9
1	A	333/374 (89%)	0.02	4 (1%) 79 82	22, 38, 64, 85	0
1	B	328/374 (87%)	0.25	15 (4%) 32 38	27, 46, 75, 89	2 (0%)
All	All	661/748 (88%)	0.13	19 (2%) 51 57	22, 42, 72, 89	2 (0%)

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	439	PRO	4.7
1	B	397	ILE	4.6
1	B	628	TYR	4.4
1	B	440	ILE	4.0
1	B	400	GLY	3.4
1	A	579	HIS	3.3
1	B	710	CYS	3.3
1	B	491	ILE	3.0
1	A	685	VAL	2.7
1	B	690	LEU	2.7
1	A	686	LEU	2.4
1	B	727	LEU	2.3
1	B	708	MET	2.3
1	B	643	TYR	2.3
1	B	399	LEU	2.2
1	A	643	TYR	2.2
1	B	441	MET	2.2
1	B	589	TYR	2.1
1	B	622	GLY	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

There are no carbohydrates in this entry.

6.4 Ligands

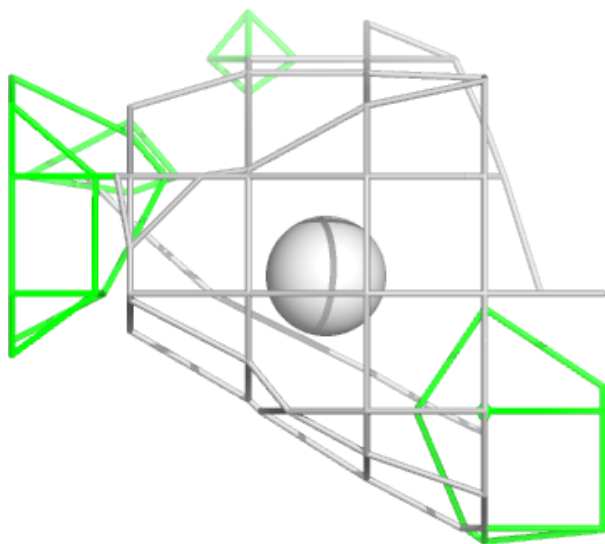
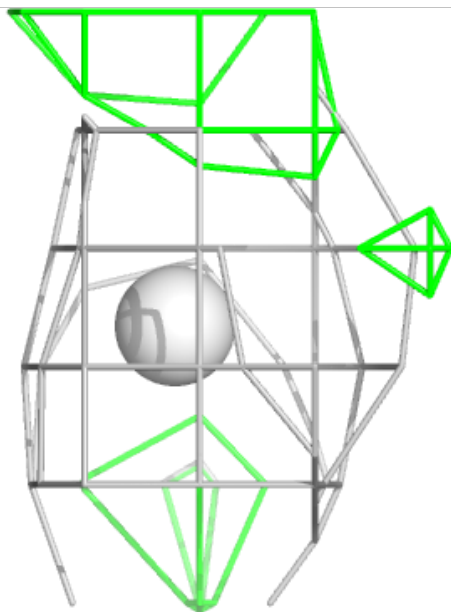
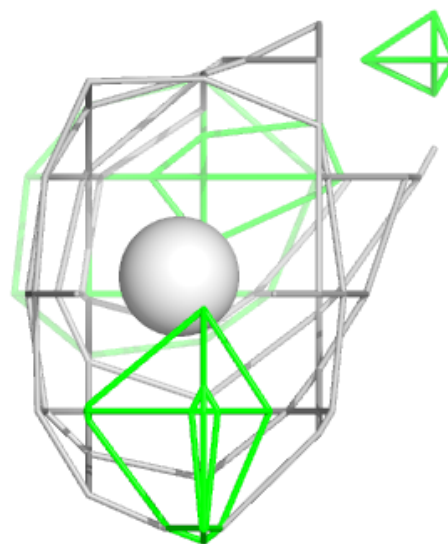
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95th percentile and maximum values of B factors of atoms in the group. The column labelled 'Q<0.9' lists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

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
2	UNX	A	801	1/1	0.39	0.54	45,45,45,45	0
2	UNX	A	802	1/1	0.63	0.22	49,49,49,49	0
2	UNX	A	803	1/1	0.72	0.64	47,47,47,47	0
2	UNX	B	801	1/1	0.86	0.49	47,47,47,47	0

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.

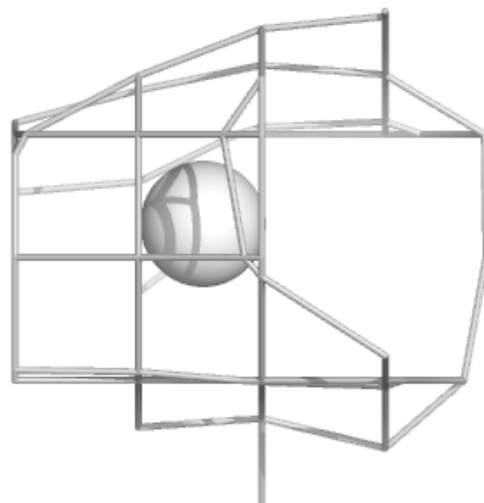
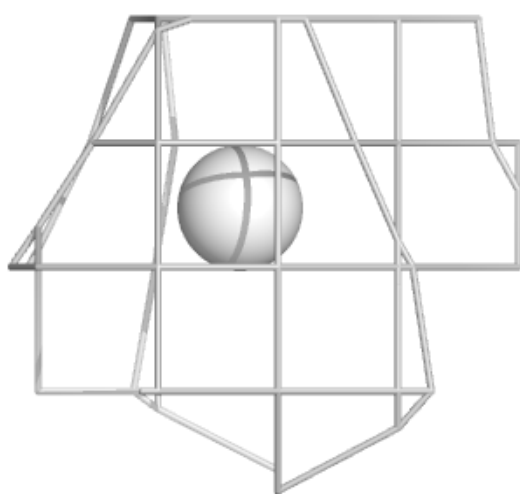
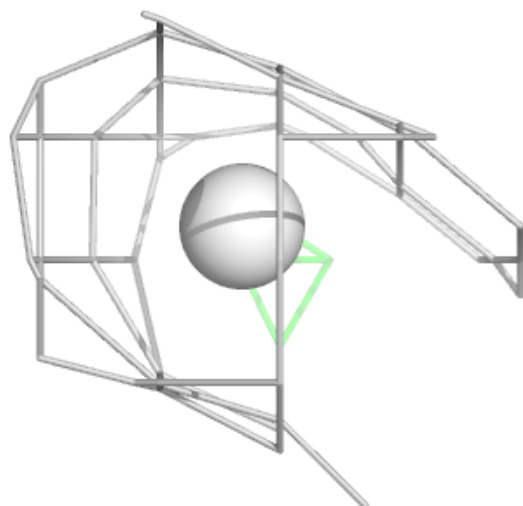
Electron density around UNX A 801:

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



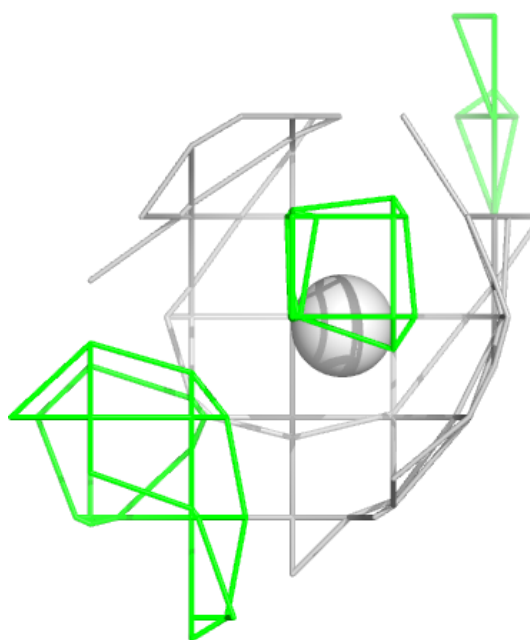
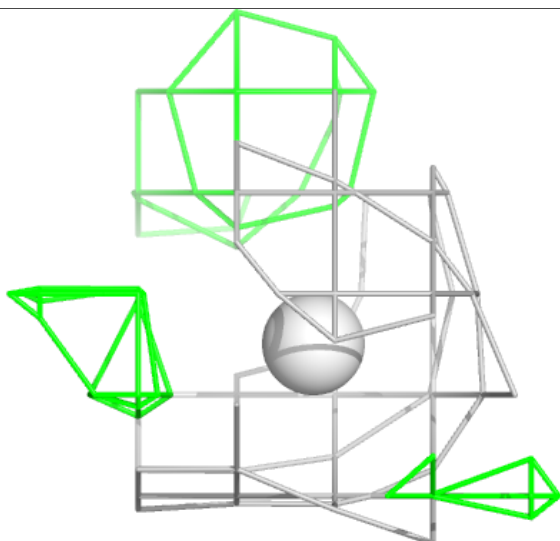
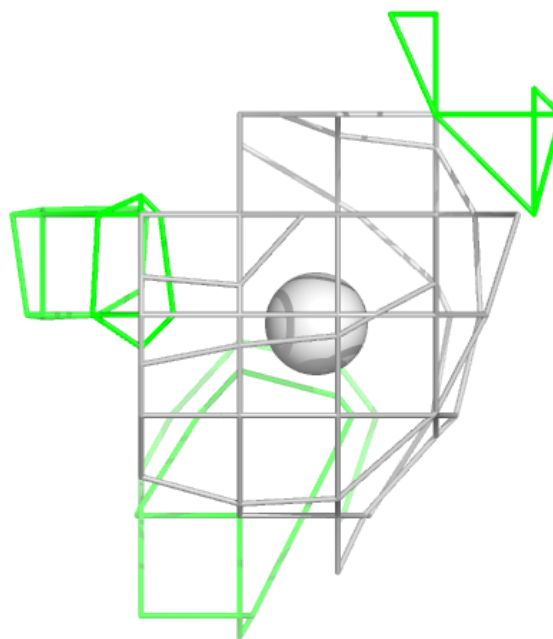
Electron density around UNX A 802:

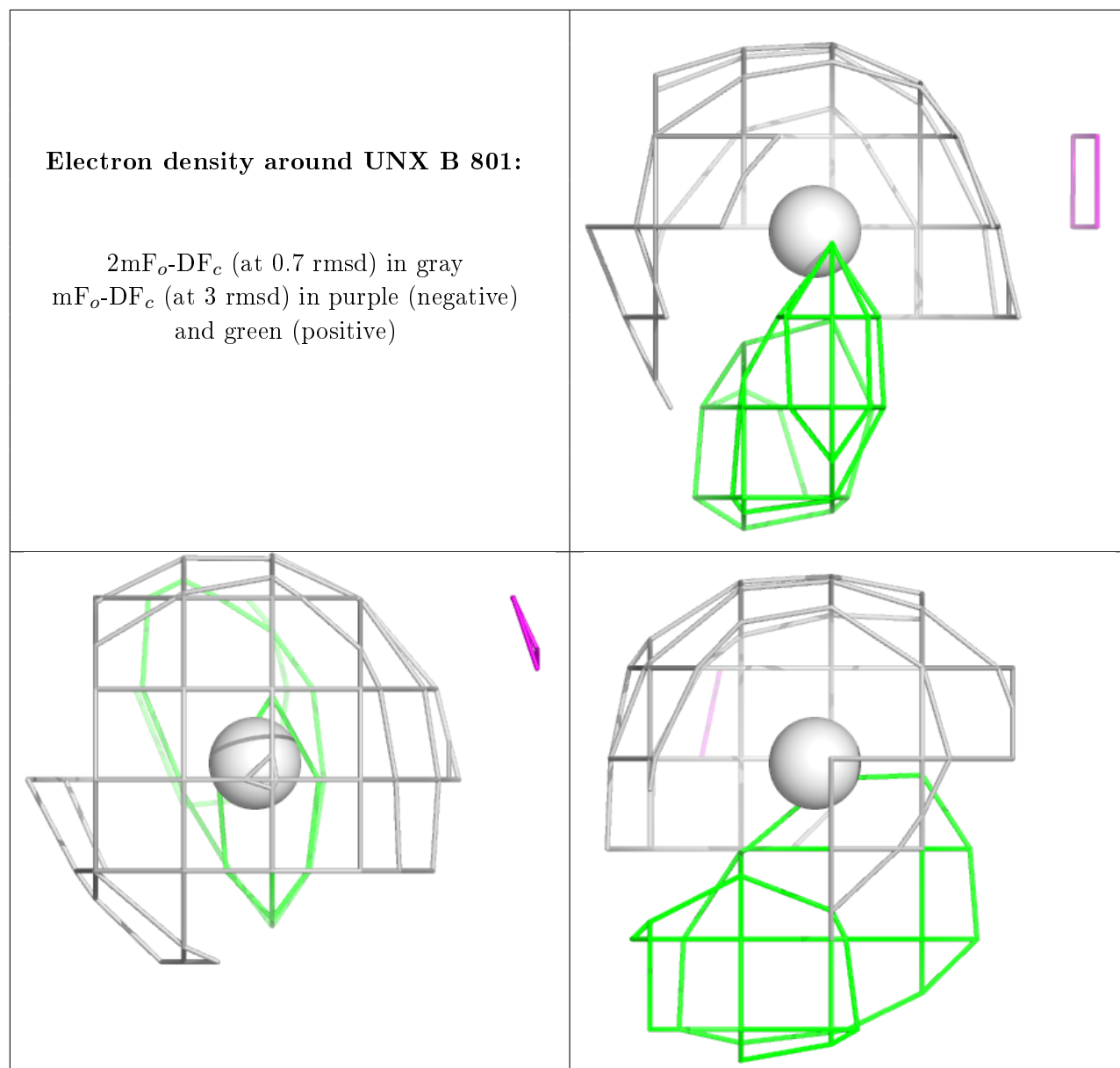
$2mF_o-DF_c$ (at 0.7 rmsd) in gray
 mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)



Electron density around UNX A 803:

$2mF_o-DF_c$ (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.