WORLDWIDE wwPDB PROTEIN DATA BANK

wwPDB X-ray Structure Validation Summary Report (i)

Oct 10, 2023 – 12:35 AM EDT

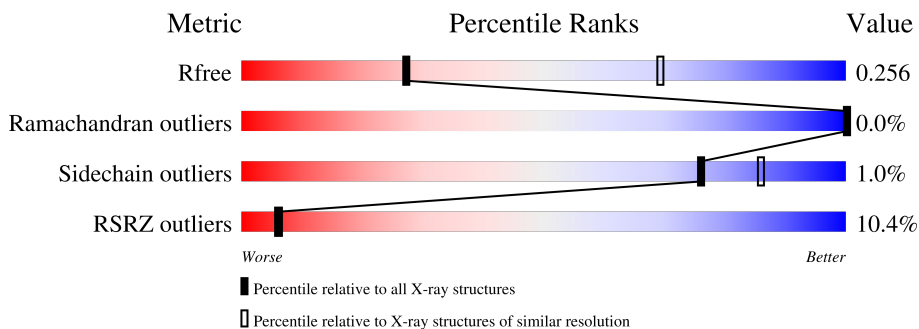|		
PDB ID	:	7S00
Title	:	X-ray structure of the phage AR9 non-virion RNA polymerase core
Authors	:	Leiman, P.G.; Sokolova, M.L.; Fraser, A.
Deposited on	:	2021-08-28
Resolution	:	3.30 Å(reported)

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 3.30 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.
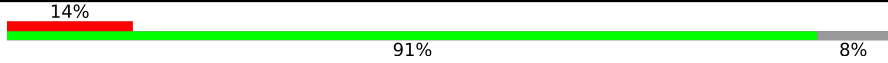
Metric	Percentile Ranks	Value
Rfree		0.256
Ramachandran outliers		0.0%
Sidechain outliers		1.0%
RSRZ outliers		10.4%

Worse   *Better*

▊ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1149 (3.34-3.26)
Ramachandran outliers	138981	1183 (3.34-3.26)
Sidechain outliers	138945	1182 (3.34-3.26)
RSRZ outliers	127900	1115 (3.34-3.26)

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 >=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 <=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	c	496	9% / 96% ··
1	e	496	11% / 97% ··
2	C	665	5% / 95% · ·
2	E	665	8% / 96% · ·
3	D	631	4% / 76% · 24%
3	F	631	14% / 94% · 6%
4	d	448	17% / 89% · 8%

Continued from previous page...

Mol	Chain	Length	Quality of chain
4	f	448	

For the chain above: 14%, 91%, 8%

2 Entry composition ⓘ

There are 5 unique types of molecules in this entry. The entry contains 33894 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 DNA-directed RNA polymerase beta subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	c	484	Total	C	N	O	S	0	0	0
			4003	2580	658	754	11			
1	e	484	Total	C	N	O	S	0	0	0
			4003	2580	658	754	11			

- Molecule 2 is a protein called DNA-directed RNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	640	Total	C	N	O	S	0	1	0
			5208	3296	857	1031	24			
2	E	640	Total	C	N	O	S	0	1	0
			5208	3296	857	1031	24			

- Molecule 3 is a protein called DNA-directed RNA polymerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	D	481	Total	C	N	O	S	0	0	0
			3893	2481	640	757	15			
3	F	596	Total	C	N	O	S	0	0	0
			4853	3092	791	952	18			

- Molecule 4 is a protein called DNA-directed RNA polymerase beta' subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	d	410	Total	C	N	O	S	0	0	0
			3362	2182	541	631	8			
4	f	410	Total	C	N	O	S	0	0	0
			3362	2182	541	631	8			

There are 44 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
d	-21	MET	-	expression tag	UNP A0A172JIH0
d	-20	GLY	-	expression tag	UNP A0A172JIH0
d	-19	SER	-	expression tag	UNP A0A172JIH0
d	-18	SER	-	expression tag	UNP A0A172JIH0
d	-17	HIS	-	expression tag	UNP A0A172JIH0
d	-16	HIS	-	expression tag	UNP A0A172JIH0
d	-15	HIS	-	expression tag	UNP A0A172JIH0
d	-14	HIS	-	expression tag	UNP A0A172JIH0
d	-13	HIS	-	expression tag	UNP A0A172JIH0
d	-12	HIS	-	expression tag	UNP A0A172JIH0
d	-11	SER	-	expression tag	UNP A0A172JIH0
d	-10	SER	-	expression tag	UNP A0A172JIH0
d	-9	GLY	-	expression tag	UNP A0A172JIH0
d	-8	GLU	-	expression tag	UNP A0A172JIH0
d	-7	ASN	-	expression tag	UNP A0A172JIH0
d	-6	LEU	-	expression tag	UNP A0A172JIH0
d	-5	TYR	-	expression tag	UNP A0A172JIH0
d	-4	PHE	-	expression tag	UNP A0A172JIH0
d	-3	GLN	-	expression tag	UNP A0A172JIH0
d	-2	GLY	-	expression tag	UNP A0A172JIH0
d	-1	HIS	-	expression tag	UNP A0A172JIH0
d	0	HIS	-	expression tag	UNP A0A172JIH0
f	-21	MET	-	expression tag	UNP A0A172JIH0
f	-20	GLY	-	expression tag	UNP A0A172JIH0
f	-19	SER	-	expression tag	UNP A0A172JIH0
f	-18	SER	-	expression tag	UNP A0A172JIH0
f	-17	HIS	-	expression tag	UNP A0A172JIH0
f	-16	HIS	-	expression tag	UNP A0A172JIH0
f	-15	HIS	-	expression tag	UNP A0A172JIH0
f	-14	HIS	-	expression tag	UNP A0A172JIH0
f	-13	HIS	-	expression tag	UNP A0A172JIH0
f	-12	HIS	-	expression tag	UNP A0A172JIH0
f	-11	SER	-	expression tag	UNP A0A172JIH0
f	-10	SER	-	expression tag	UNP A0A172JIH0
f	-9	GLY	-	expression tag	UNP A0A172JIH0
f	-8	GLU	-	expression tag	UNP A0A172JIH0
f	-7	ASN	-	expression tag	UNP A0A172JIH0
f	-6	LEU	-	expression tag	UNP A0A172JIH0
f	-5	TYR	-	expression tag	UNP A0A172JIH0
f	-4	PHE	-	expression tag	UNP A0A172JIH0
f	-3	GLN	-	expression tag	UNP A0A172JIH0
f	-2	GLY	-	expression tag	UNP A0A172JIH0
f	-1	HIS	-	expression tag	UNP A0A172JIH0

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
f	0	HIS	-	expression tag	UNP A0A172JIH0

- Molecule 5 is ZINC ION (three-letter code: ZN) (formula: Zn) (labeled as "Ligand of Interest" by depositor).

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	D	1	Total 1	Zn 1	0	0
5	F	1	Total 1	Zn 1	0	0

3 Residue-property plots ⓘ

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.
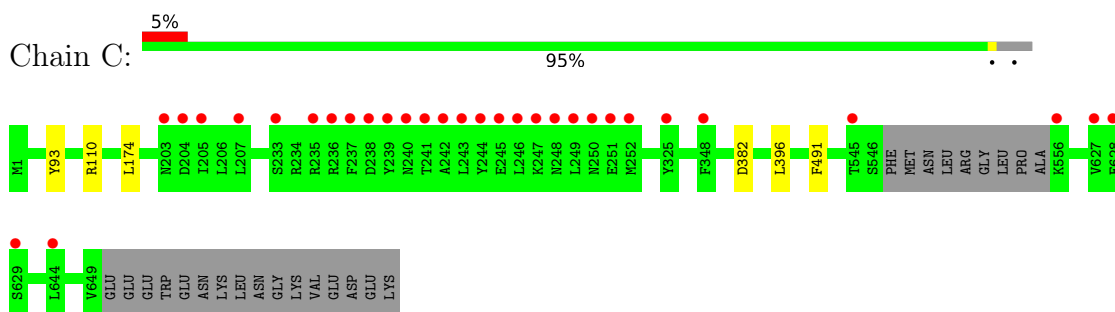
- Molecule 1: DNA-directed RNA polymerase beta subunit
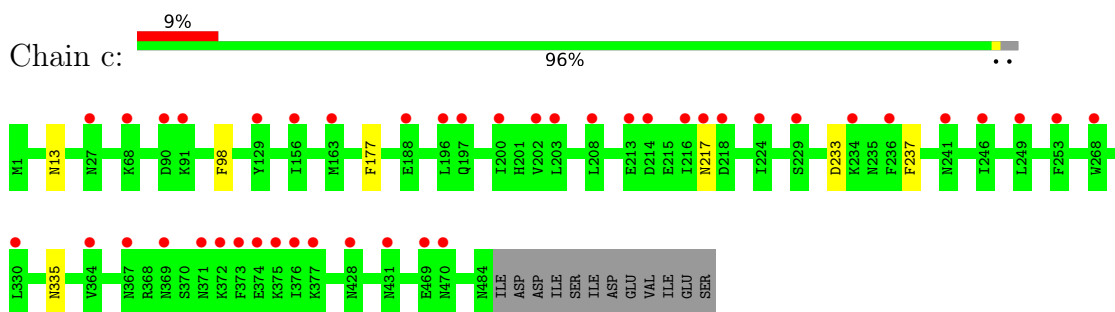
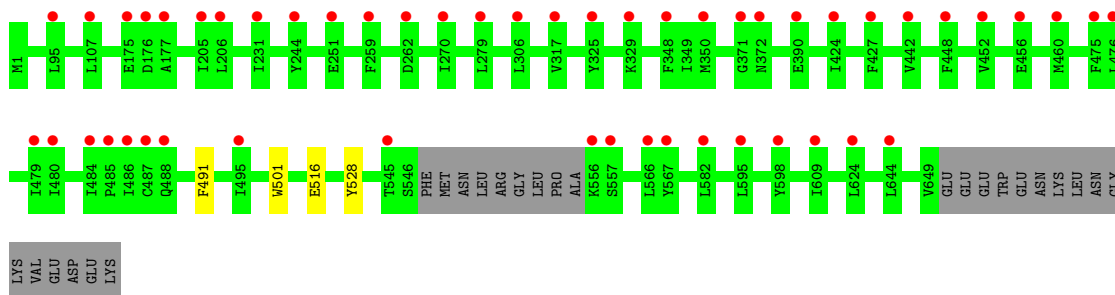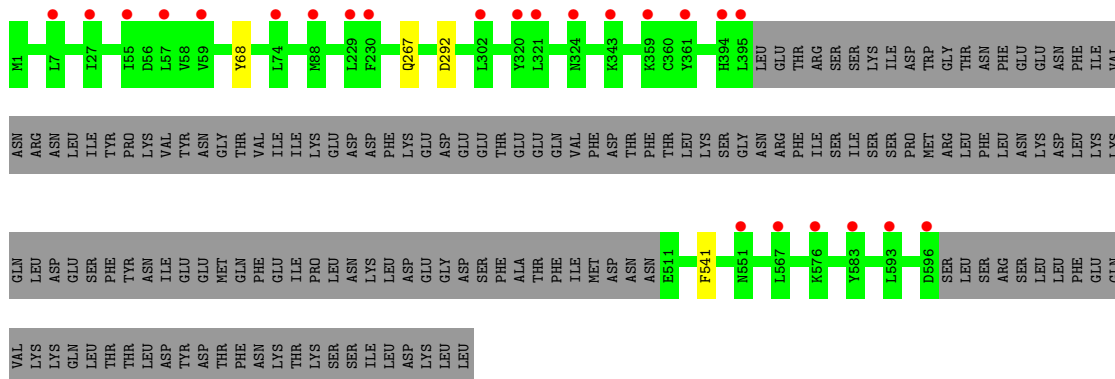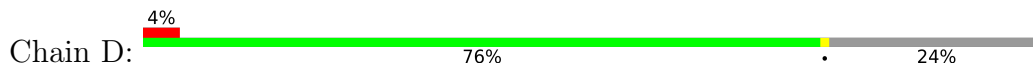- Molecule 1: DNA-directed RNA polymerase beta subunit
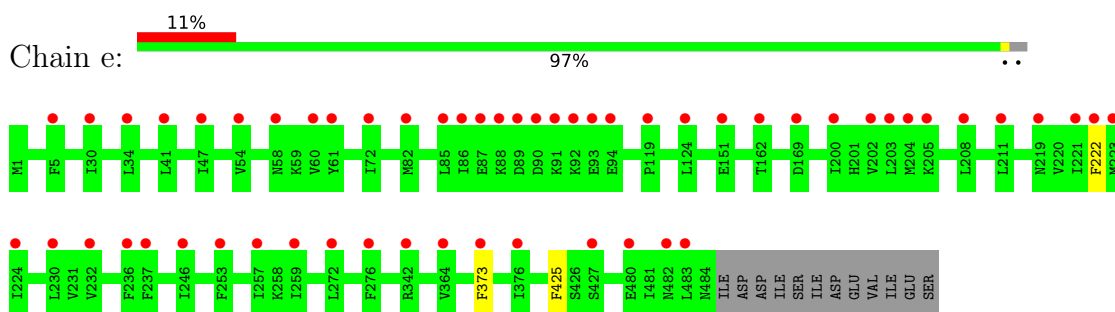
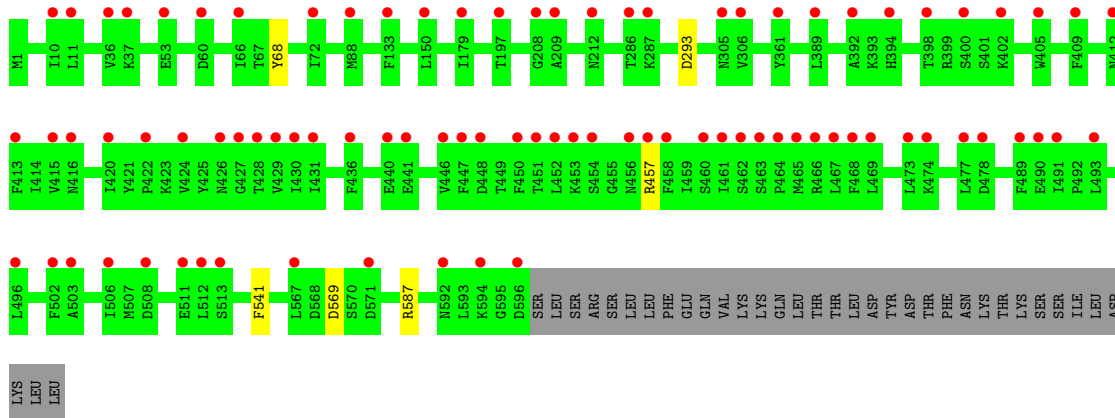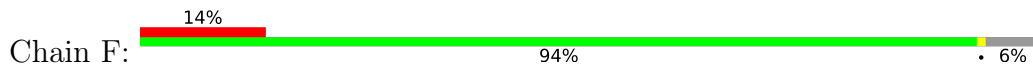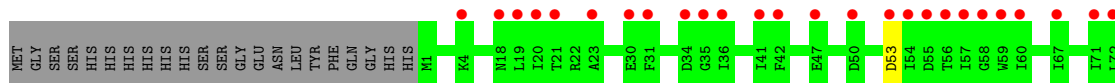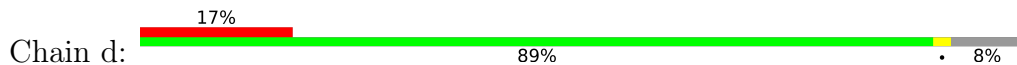- Molecule 2: DNA-directed RNA polymerase

- Molecule 2: DNA-directed RNA polymerase

• Molecule 3: DNA-directed RNA polymerase

Chain D:

• Molecule 3: DNA-directed RNA polymerase

Chain F:

• Molecule 4: DNA-directed RNA polymerase beta' subunit

Chain d:

F73 I76 K77 K78 C79 I80 K85 Y89 Q90 GLN SER ILE ASP GLN ASN GLY GLU ASN ILE ASP LEU THR GLU GLU ILE G107 E108 D109 I112 G113 L114 V115 K116 F117 F121 D122 L125 E137 Y138 D139 F140 L141 N144 H145 I148 F149 I150 P154 V155 F156

S157 H158 K159 L160 R161 P162 A163 T164 L165 L166 T167 K172 V173 L174 Y183 N184 N194 L208 F216 Y229 K233 K234 G235 F236 L237 N240 I241 F248 D374 K385 D416 N426

● Molecule 4: DNA-directed RNA polymerase beta' subunit

Chain f:

14%

91%

8%

MET GLY SER SER HIS HIS HIS HIS HIS HIS SER SER GLY GLU ASN LEU TYR PHE GLN GLY HIS HIS M1 F10 I13 Y14 N15 E16 I20 T21 N28 H29 E30 F31 I41 D48 D49 I57 G58 W59 I62 I67 I68 L72 Q90 GLN SER ILE ASP

GLN ASN GLY GLU ASN ILE ASP LEU THR GLU GLU ILE G107 L114 F117 L124 L125 K131 K132 F140 N144 K147 I148 L160 R161 P162 A163 T164 S169 V173 Y183 D201 S202 I203 D204 L207 Y229 K233 K234 N240 I241 F248

L275 Y276 K277 F278 Q279 L280 L283 I284 L296 W299 M314 S365 G369 D370 Y371 D372 G373 D374 P380 V381 F382 L424 N425 N426

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	112.86Å 166.27Å 307.21Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	49.42 – 3.30	Depositor
	49.42 – 3.30	EDS
% Data completeness (in resolution range)	98.6 (49.42-3.30)	Depositor
	98.6 (49.42-3.30)	EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.59 (at 3.33Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.219 , 0.258	Depositor
	0.220 , 0.256	DCC
R_{free} test set	4308 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	107.3	Xtriage
Anisotropy	0.364	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 96.5	EDS
L-test for twinning[2]	$< |L| > = 0.44, < L^2 > = 0.27$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	33894	wwPDB-VP
Average B, all atoms (Å^2)	155.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.46% of the height of the origin peak. No significant pseudotranslation is detected.*

---

[1]Intensities estimated from amplitudes.
[2]Theoretical values of $< |L| >$, $< L^2 >$ 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

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	c	0.26	0/4082	0.48	0/5502
1	e	0.26	0/4082	0.47	0/5502
2	C	0.27	0/5308	0.48	0/7156
2	E	0.26	0/5308	0.47	0/7156
3	D	0.25	0/3952	0.44	0/5328
3	F	0.25	0/4932	0.44	0/6650
4	d	0.25	0/3431	0.44	0/4630
4	f	0.25	0/3431	0.44	0/4630
All	All	0.26	0/34526	0.46	0/46554

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)

Due to software issues we are unable to calculate clashes - this section is therefore empty.

5.3   Torsion angles (i)

5.3.1   Protein backbone (i)

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	c	482/496 (97%)	473 (98%)	9 (2%)	0	100	100
1	e	482/496 (97%)	474 (98%)	8 (2%)	0	100	100
2	C	637/665 (96%)	625 (98%)	12 (2%)	0	100	100
2	E	637/665 (96%)	627 (98%)	10 (2%)	0	100	100
3	D	477/631 (76%)	471 (99%)	6 (1%)	0	100	100
3	F	594/631 (94%)	584 (98%)	10 (2%)	0	100	100
4	d	406/448 (91%)	398 (98%)	8 (2%)	0	100	100
4	f	406/448 (91%)	399 (98%)	6 (2%)	1 (0%)	47	77
All	All	4121/4480 (92%)	4051 (98%)	69 (2%)	1 (0%)	100	100

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	f	201	ASP

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	c	457/469 (97%)	450 (98%)	7 (2%)	65	81
1	e	457/469 (97%)	454 (99%)	3 (1%)	84	90
2	C	587/608 (96%)	581 (99%)	6 (1%)	76	86
2	E	587/608 (96%)	583 (99%)	4 (1%)	84	90
3	D	445/590 (75%)	441 (99%)	4 (1%)	78	87
3	F	555/590 (94%)	549 (99%)	6 (1%)	73	85
4	d	375/409 (92%)	365 (97%)	10 (3%)	44	71
4	f	375/409 (92%)	375 (100%)	0	100	100
All	All	3838/4152 (92%)	3798 (99%)	40 (1%)	76	86

5 of 40 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	e	373	PHE
3	F	293	ASP
1	e	425	PHE
2	E	516	GLU
3	F	541	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
4	d	15	ASN
4	d	180	ASN
4	d	377	ASN
4	f	247	ASN
4	f	377	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 monosaccharides in this entry.

5.6 Ligand geometry ⓘ

Of 2 ligands modelled in this entry, 2 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, 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	c	484/496 (97%)	0.57	43 (8%)	9	10	57, 129, 227, 365	0
1	e	484/496 (97%)	0.56	56 (11%)	4	4	109, 173, 251, 319	0
2	C	640/665 (96%)	0.37	31 (4%)	30	28	49, 92, 202, 317	0
2	E	640/665 (96%)	0.51	51 (7%)	12	11	98, 151, 214, 279	0
3	D	481/631 (76%)	0.33	24 (4%)	28	27	54, 112, 194, 261	0
3	F	596/631 (94%)	0.80	87 (14%)	2	2	121, 175, 272, 329	0
4	d	410/448 (91%)	0.94	78 (19%)	1	1	71, 179, 259, 367	0
4	f	410/448 (91%)	0.83	63 (15%)	2	2	123, 196, 278, 427	0
All	All	4145/4480 (92%)	0.60	433 (10%)	6	6	49, 152, 250, 427	0

The worst 5 of 433 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	F	450	PHE	9.6
4	f	49	ASP	9.1
2	E	557	SER	8.7
3	F	429	VAL	8.5
3	F	491	ILE	7.9

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

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

6.3 Carbohydrates ⓘ

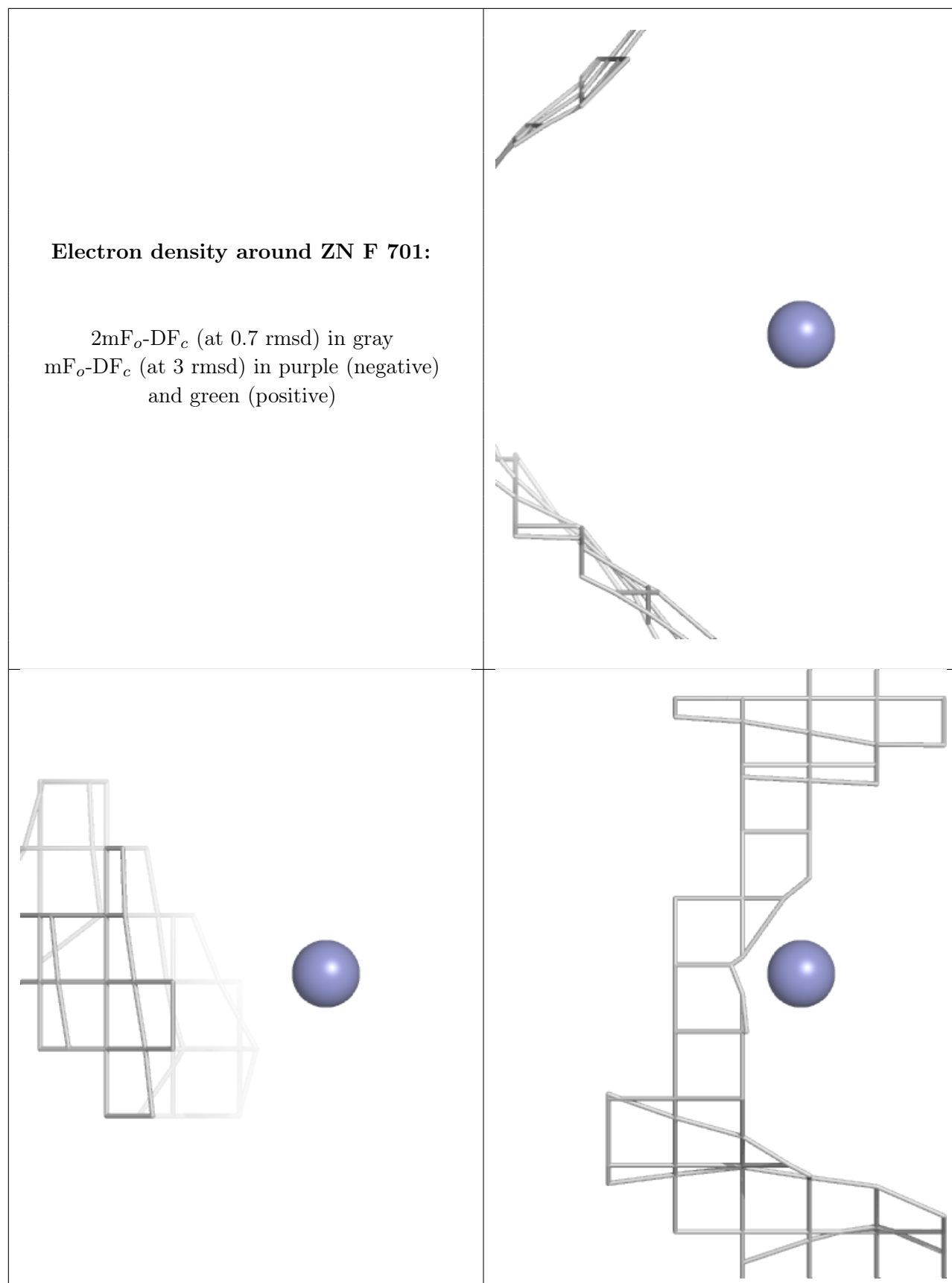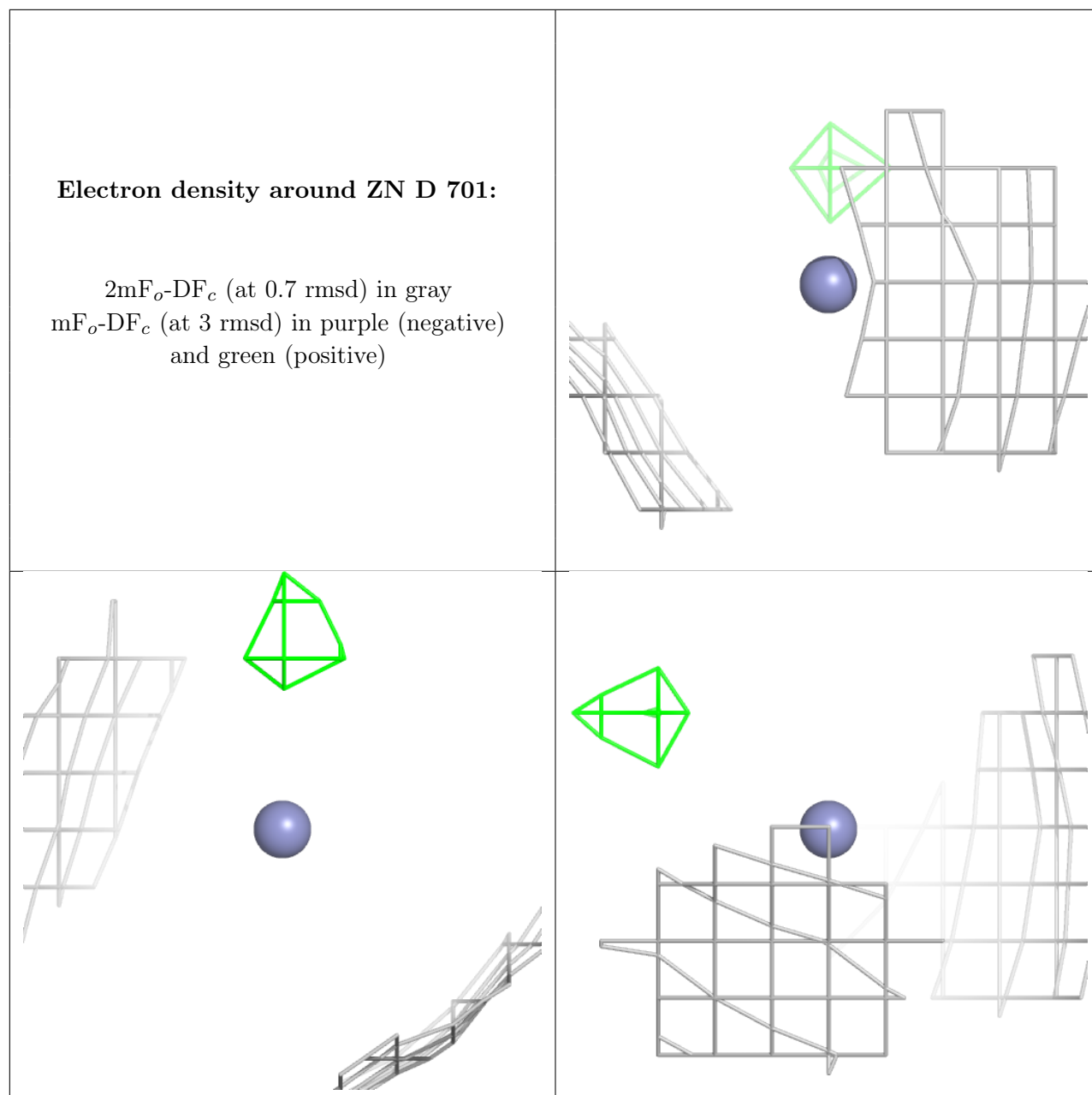There are no monosaccharides in this entry.

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, 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($Å^2$)	Q<0.9
5	ZN	F	701	1/1	0.71	0.15	188,188,188,188	0
5	ZN	D	701	1/1	0.91	0.17	135,135,135,135	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 ZN F 701:

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

Electron density around ZN D 701:

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

6.5 Other polymers ⓘ

There are no such residues in this entry.