Full wwPDB X-ray Structure Validation Report ⓘ

Apr 22, 2025 – 12:08 PM JST

PDB ID	:	8ZBF / pdb_00008zbf
Title	:	Crystal structure of the A58-T10 DNA aptamer in complex with SARS-CoV-2 N-NTD
Authors	:	Chen, X.; Huang, L.
Deposited on	:	2024-04-26
Resolution	:	2.80 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 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:

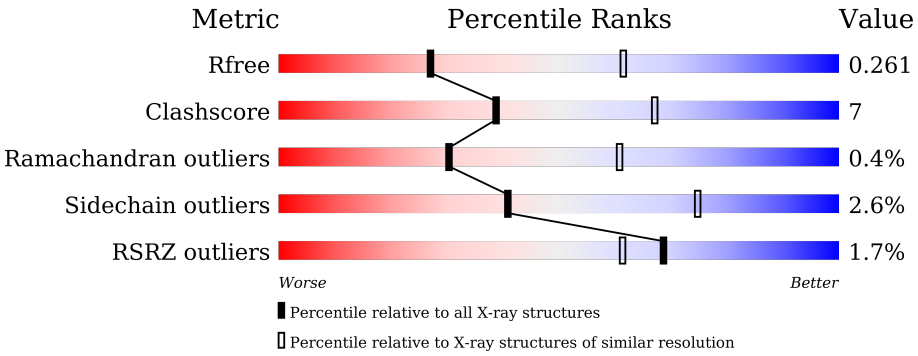
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.006 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.42

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.80 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}	164625	3657 (2.80-2.80)
Clashscore	180529	4123 (2.80-2.80)
Ramachandran outliers	177936	4071 (2.80-2.80)
Sidechain outliers	177891	4073 (2.80-2.80)
RSRZ outliers	164620	3659 (2.80-2.80)

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	A	135	91% · 7%
1	B	135	73% 16% · 10%
1	C	135	4% 79% 10% · 10%
1	D	135	65% 23% · 11%
2	E	40	58% 40% ·
2	F	40	62% 35% ·

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	40	
2	H	40	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 7034 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 Nucleoprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	126	Total	C	N	O	S	0	0	0
			980	621	177	181	1			
1	B	121	Total	C	N	O	S	0	0	0
			939	599	165	174	1			
1	C	121	Total	C	N	O	S	0	0	0
			945	602	168	174	1			
1	D	120	Total	C	N	O	S	0	0	0
			933	596	166	170	1			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	40	SER	-	expression tag	UNP P0DTC9
B	40	SER	-	expression tag	UNP P0DTC9
C	40	SER	-	expression tag	UNP P0DTC9
D	40	SER	-	expression tag	UNP P0DTC9

- Molecule 2 is a DNA chain called DNA (40-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	40	Total	C	N	O	P	0	0	0
			819	391	155	234	39			
2	F	39	Total	C	N	O	P	0	0	0
			799	381	150	230	38			
2	G	40	Total	C	N	O	P	0	0	0
			819	391	155	234	39			
2	H	39	Total	C	N	O	P	0	0	0
			798	380	150	230	38			

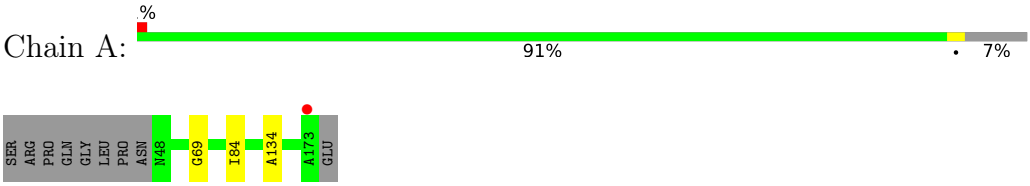
- Molecule 3 is SODIUM ION (CCD ID: NA) (formula: Na) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total Na 1 1	0	0
3	C	1	Total Na 1 1	0	0

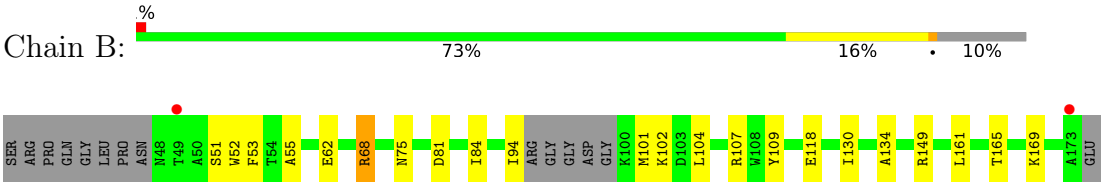
3 Residue-property plots (i)

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.

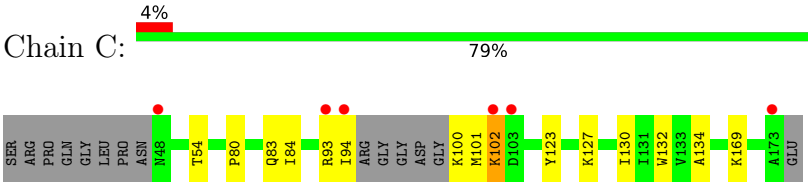
• Molecule 1: Nucleoprotein



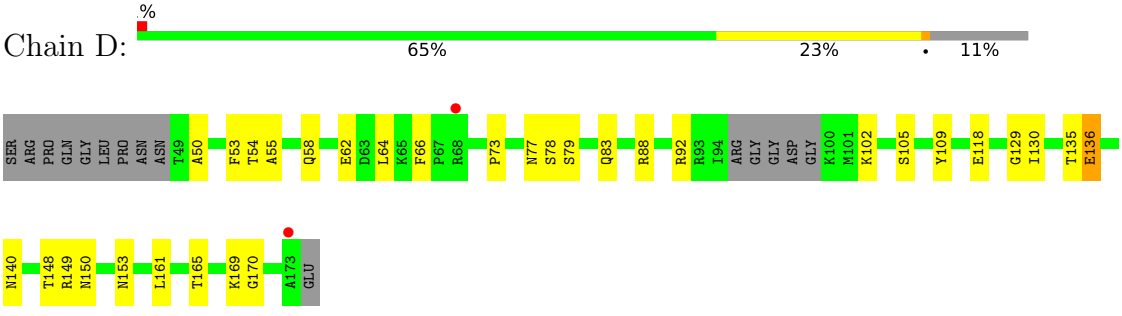
• Molecule 1: Nucleoprotein



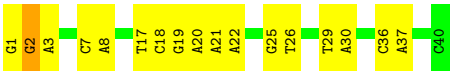
• Molecule 1: Nucleoprotein



• Molecule 1: Nucleoprotein

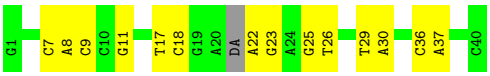


• Molecule 2: DNA (40-MER)



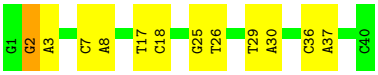
• Molecule 2: DNA (40-MER)

Chain F:



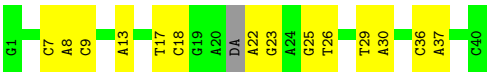
• Molecule 2: DNA (40-MER)

Chain G:



• Molecule 2: DNA (40-MER)

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	70.97Å 92.48Å 77.92Å 90.00° 90.54° 90.00°	Depositor
Resolution (Å)	35.48 – 2.80 35.48 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.5 (35.48-2.80) 99.7 (35.48-2.80)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.00 (at 2.81Å)	Xtriage
Refinement program	PHENIX (1.21.1_5286: ???)	Depositor
R, R_{free}	0.217 , 0.261 0.218 , 0.261	Depositor DCC
R_{free} test set	1142 reflections (4.59%)	wwPDB-VP
Wilson B-factor (Å^2)	65.1	Xtriage
Anisotropy	0.685	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 55.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.108 for h,-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	7034	wwPDB-VP
Average B, all atoms (Å^2)	81.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 18.69% 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

5.1 Standard geometry

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

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	0.33	0/1008	0.60	0/1371
1	B	0.32	0/966	0.62	0/1317
1	C	0.37	0/972	0.62	0/1324
1	D	0.33	0/960	0.62	0/1308
2	E	0.60	0/920	0.96	3/1419 (0.2%)
2	F	0.62	0/896	0.93	0/1379
2	G	0.64	0/920	0.96	2/1419 (0.1%)
2	H	0.59	0/894	0.94	0/1374
All	All	0.49	0/7536	0.80	5/10911 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand.A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	B	0	1
1	C	0	1
All	All	0	2

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	G	2	DG	P-O3'-C3'	6.00	126.90	119.70
2	E	2	DG	P-O3'-C3'	5.62	126.44	119.70
2	E	2	DG	OP2-P-O3'	5.18	116.60	105.20
2	E	21	DA	O4'-C1'-N9	5.10	111.57	108.00
2	G	2	DG	OP2-P-O3'	5.03	116.26	105.20

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	68	ARG	Sidechain
1	C	101	MET	Peptide

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	980	0	953	2	0
1	B	939	0	904	12	0
1	C	945	0	915	8	1
1	D	933	0	905	20	0
2	E	819	0	448	14	1
2	F	799	0	441	9	0
2	G	819	0	448	10	0
2	H	798	0	437	10	0
3	A	1	0	0	0	0
3	C	1	0	0	0	0
All	All	7034	0	5451	83	1

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 7.

All (83) 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:C:94:ILE:HG13	1:C:102:LYS:HG3	1.56	0.87
1:B:53:PHE:HZ	1:B:75:ASN:HB2	1.49	0.78
1:D:118:GLU:HG3	1:D:130:ILE:HD11	1.73	0.70
1:C:94:ILE:CG1	1:C:102:LYS:HG3	2.23	0.69
1:D:62:GLU:OE2	1:D:170:GLY:N	2.26	0.68
1:B:53:PHE:CZ	1:B:75:ASN:HB2	2.30	0.64
2:E:19:DG:H2"	2:E:20:DA:C8	2.34	0.62
2:G:2:DG:H2"	2:G:3:DA:OP2	1.99	0.62
1:D:150:ASN:HB3	1:D:153:ASN:OD1	2.01	0.61

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:E:2:DG:H2"	2:E:3:DA:OP2	2.02	0.60
2:E:19:DG:H8	2:E:22:DA:H62	1.49	0.60
2:F:29:DT:H2"	2:F:30:DA:H5'	1.85	0.58
2:H:29:DT:H2"	2:H:30:DA:H5'	1.87	0.57
1:B:107:ARG:NH2	2:F:11:DG:OP2	2.38	0.57
2:G:29:DT:H2"	2:G:30:DA:H5'	1.88	0.56
2:E:18:DC:H2"	2:E:19:DG:H5"	1.87	0.55
2:E:3:DA:OP2	2:E:3:DA:H8	1.89	0.55
1:B:52:TRP:CD2	1:B:149:ARG:HB2	2.42	0.54
1:C:94:ILE:HG13	1:C:102:LYS:CG	2.35	0.54
2:G:3:DA:OP2	2:G:3:DA:H8	1.92	0.53
1:D:66:PHE:HE1	1:D:165:THR:HG23	1.73	0.52
2:E:29:DT:H2"	2:E:30:DA:H5'	1.91	0.52
2:H:7:DC:H2'	2:H:8:DA:C8	2.44	0.51
1:C:127:LYS:HG2	1:C:130:ILE:HD13	1.92	0.50
2:E:7:DC:H2'	2:E:8:DA:C8	2.47	0.50
1:B:118:GLU:HG3	1:B:130:ILE:HD11	1.94	0.49
2:F:36:DC:H2"	2:F:37:DA:C8	2.48	0.48
2:G:7:DC:H2'	2:G:8:DA:C8	2.49	0.48
2:F:7:DC:H2'	2:F:8:DA:C8	2.49	0.48
2:E:36:DC:H2"	2:E:37:DA:C8	2.49	0.47
1:D:78:SER:OG	1:D:83:GLN:NE2	2.48	0.47
1:D:55:ALA:HB2	1:D:109:TYR:CD2	2.51	0.46
2:H:36:DC:H2"	2:H:37:DA:C8	2.50	0.46
2:H:25:DG:H2'	2:H:26:DT:C6	2.51	0.46
2:G:17:DT:H2"	2:G:18:DC:C6	2.50	0.46
2:F:25:DG:H2"	2:F:26:DT:H5'	1.98	0.46
1:D:62:GLU:OE1	1:D:62:GLU:N	2.31	0.45
2:G:36:DC:H2"	2:G:37:DA:C8	2.51	0.45
2:H:29:DT:H2"	2:H:30:DA:H8	1.80	0.45
1:A:84:ILE:HG13	1:A:134:ALA:HB2	1.99	0.45
1:C:123:TYR:CE1	1:C:132:TRP:HB3	2.52	0.45
2:E:1:DG:H2"	2:E:2:DG:C8	2.52	0.45
2:H:17:DT:H2"	2:H:18:DC:C6	2.52	0.45
2:F:22:DA:H2"	2:F:23:DG:O5'	2.17	0.44
2:G:25:DG:H2'	2:G:26:DT:C6	2.53	0.44
1:B:84:ILE:HG13	1:B:134:ALA:HB2	2.00	0.44
2:E:17:DT:H2"	2:E:18:DC:C6	2.53	0.44
2:E:29:DT:H2"	2:E:30:DA:H8	1.82	0.44
1:D:64:LEU:HD11	1:D:66:PHE:HE2	1.83	0.44
1:D:153:ASN:OD1	1:D:153:ASN:N	2.49	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:E:36:DC:H2"	2:E:37:DA:H8	1.82	0.44
1:D:92:ARG:NH1	2:H:13:DA:H2"	2.32	0.43
2:F:29:DT:H2"	2:F:30:DA:H8	1.83	0.43
1:B:94:ILE:HG12	1:B:104:LEU:HD11	2.01	0.43
1:D:135:THR:HG22	1:D:136:GLU:N	2.34	0.43
2:G:25:DG:H2"	2:G:26:DT:H5'	2.00	0.43
2:F:17:DT:H2"	2:F:18:DC:C6	2.54	0.43
2:F:9:DC:O2	2:F:9:DC:H2'	2.18	0.43
2:H:22:DA:H2"	2:H:23:DG:O5'	2.19	0.43
1:C:100:LYS:O	1:C:102:LYS:HD3	2.19	0.43
1:D:102:LYS:HB3	1:D:102:LYS:HE3	1.81	0.42
1:C:84:ILE:HG13	1:C:134:ALA:HB2	2.01	0.42
1:D:58:GLN:O	1:D:105:SER:HB3	2.19	0.42
2:E:19:DG:H1'	2:E:22:DA:N6	2.33	0.42
1:D:50:ALA:N	1:D:149:ARG:O	2.51	0.42
1:B:52:TRP:CE3	1:B:149:ARG:HD3	2.54	0.42
1:B:81:ASP:N	1:B:81:ASP:OD2	2.53	0.42
1:C:80:PRO:HA	1:C:83:GLN:HG3	2.01	0.42
1:B:161:LEU:HD13	1:B:165:THR:HG21	2.01	0.42
2:E:25:DG:H2"	2:E:26:DT:H5'	2.01	0.42
1:A:69:GLY:HA2	1:A:134:ALA:O	2.20	0.41
1:D:161:LEU:HD13	1:D:165:THR:HG21	2.01	0.41
2:G:36:DC:H2"	2:G:37:DA:H8	1.85	0.41
1:B:55:ALA:HB2	1:B:109:TYR:CE2	2.56	0.41
1:B:101:MET:C	1:B:102:LYS:HG3	2.41	0.41
1:D:54:THR:O	1:D:55:ALA:HB3	2.20	0.41
1:D:66:PHE:HE1	1:D:165:THR:CG2	2.33	0.41
1:D:50:ALA:O	1:D:148:THR:HB	2.20	0.41
1:D:53:PHE:CE2	1:D:73:PRO:HB2	2.56	0.41
2:G:29:DT:H2"	2:G:30:DA:H8	1.86	0.41
2:H:9:DC:O2	2:H:9:DC:H2'	2.20	0.40
2:H:29:DT:H2"	2:H:30:DA:C8	2.55	0.40
1:D:88:ARG:HA	1:D:129:GLY:O	2.22	0.40

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:169:LYS:NZ	2:E:8:DA:OP1[2_554]	2.11	0.09

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	A	124/135 (92%)	122 (98%)	2 (2%)	0	100	100
1	B	117/135 (87%)	112 (96%)	5 (4%)	0	100	100
1	C	117/135 (87%)	112 (96%)	4 (3%)	1 (1%)	14	42
1	D	116/135 (86%)	112 (97%)	3 (3%)	1 (1%)	14	42
All	All	474/540 (88%)	458 (97%)	14 (3%)	2 (0%)	30	61

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	136	GLU
1	C	54	THR

5.3.2 Protein sidechains (i)

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	99/107 (92%)	99 (100%)	0	100	100
1	B	95/107 (89%)	91 (96%)	4 (4%)	25	58
1	C	96/107 (90%)	94 (98%)	2 (2%)	48	80
1	D	94/107 (88%)	90 (96%)	4 (4%)	25	57
All	All	384/428 (90%)	374 (97%)	10 (3%)	41	75

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	51	SER
1	B	62	GLU
1	B	68	ARG
1	B	169	LYS
1	C	93	ARG
1	C	102	LYS
1	D	77	ASN
1	D	79	SER
1	D	140	ASN
1	D	169	LYS

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

Mol	Chain	Res	Type
1	A	140	ASN
1	B	163	GLN
1	D	77	ASN
1	D	83	GLN
1	D	140	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 oligosaccharides 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 (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

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^{th} 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	126/135 (93%)	-0.12	1 (0%) 82 77	37, 56, 89, 126	0
1	B	121/135 (89%)	0.13	2 (1%) 69 61	48, 73, 103, 148	0
1	C	121/135 (89%)	-0.01	6 (4%) 35 28	45, 68, 105, 130	0
1	D	120/135 (88%)	0.27	2 (1%) 69 61	55, 81, 118, 161	0
2	E	40/40 (100%)	0.05	0 100 100	45, 102, 140, 219	0
2	F	39/40 (97%)	-0.13	0 100 100	49, 93, 118, 121	0
2	G	40/40 (100%)	-0.17	0 100 100	48, 85, 128, 137	0
2	H	39/40 (97%)	-0.13	0 100 100	59, 90, 122, 131	0
All	All	646/700 (92%)	0.03	11 (1%) 69 61	37, 74, 118, 219	0

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	173	ALA	3.8
1	C	94	ILE	3.6
1	C	93	ARG	3.4
1	C	102	LYS	3.3
1	D	68	ARG	2.8
1	B	173	ALA	2.4
1	D	173	ALA	2.2
1	C	48	ASN	2.2
1	B	49	THR	2.1
1	A	173	ALA	2.1
1	C	103	ASP	2.0

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 (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

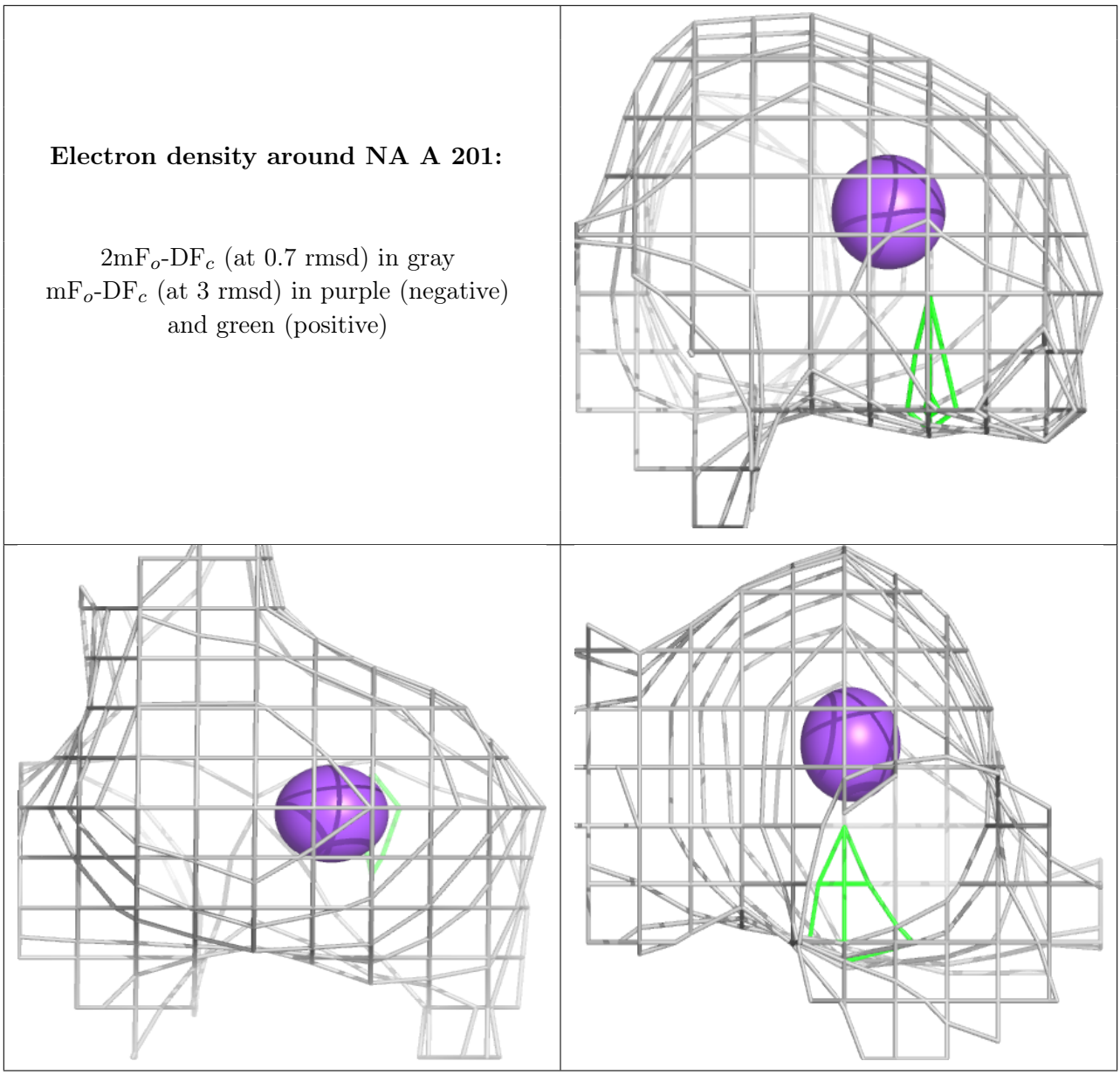
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
3	NA	A	201	1/1	0.75	0.21	67,67,67,67	0
3	NA	C	201	1/1	0.78	0.35	118,118,118,118	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.

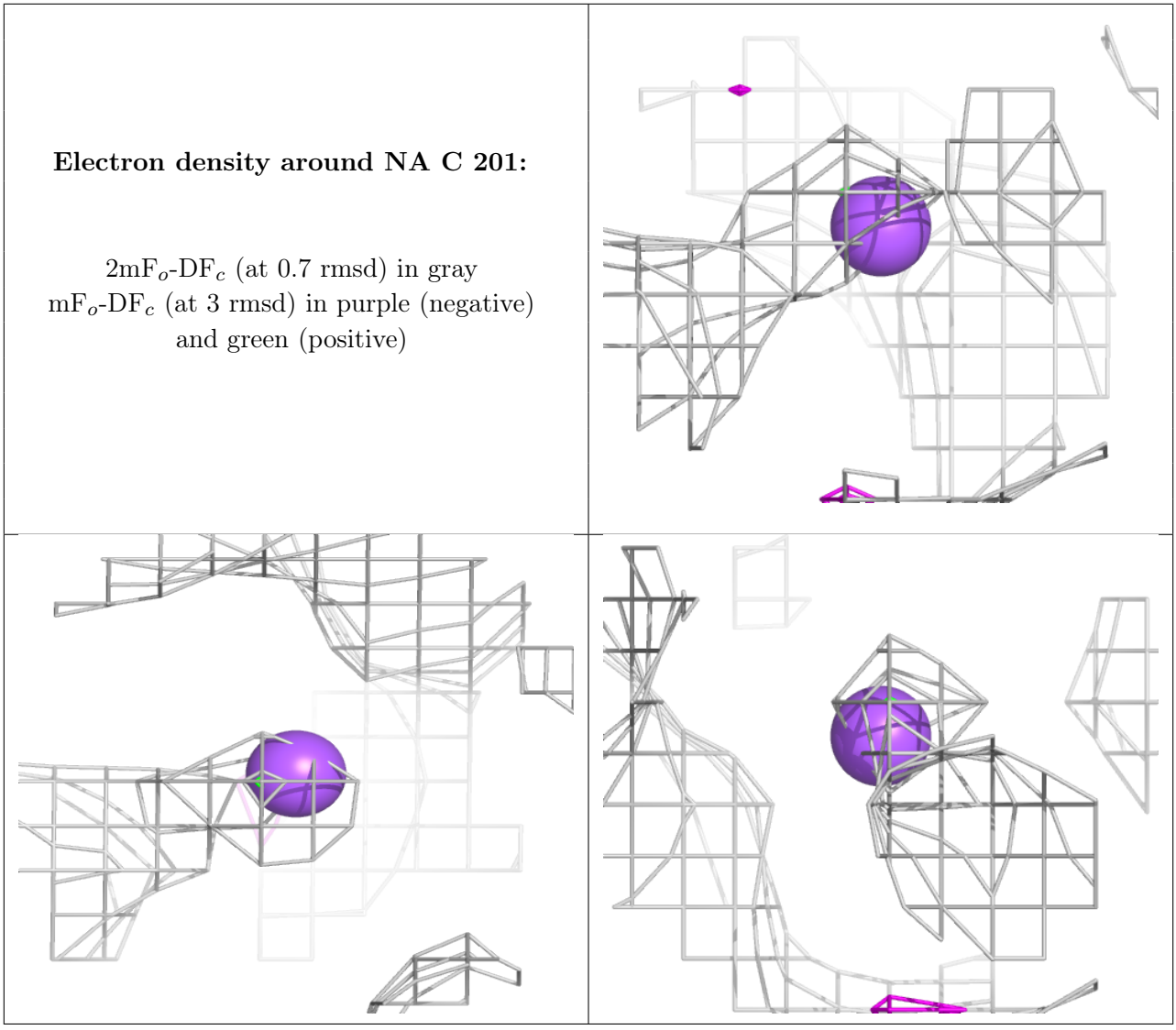
Electron density around NA A 201:

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



Electron density around NA C 201:

$2mF_o$-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 (i)

There are no such residues in this entry.