Full wwPDB X-ray Structure Validation Report ⓘ

Oct 14, 2023 – 11:26 PM EDT

PDB ID	:	8FN9
Title	:	Crystal structure of the C-terminal Fg domain of TNR with the single FN domain of PTPRZ
Authors	:	Bouyain, S.
Deposited on	:	2022-12-27
Resolution	:	1.80 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:

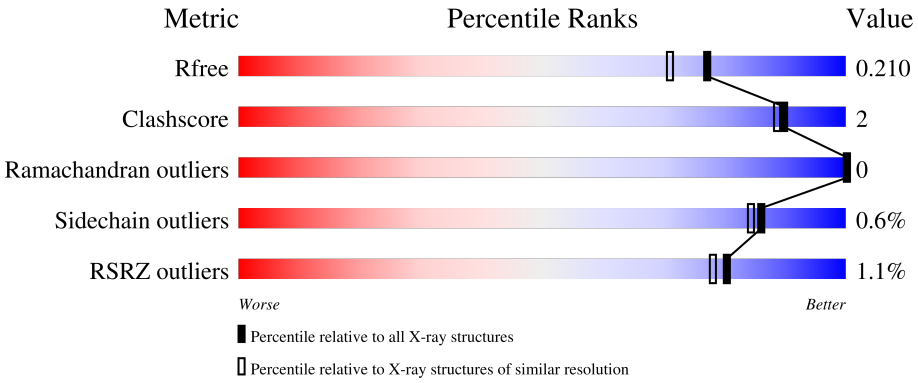
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.36
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.80 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}	130704	5950 (1.80-1.80)
Clashscore	141614	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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	234	87% 7% 6%
1	C	234	91% · 6%
1	E	234	87% 5% 8%
1	G	234	89% · 8%
2	B	106	5% 85% 8% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	D	106	3% 92% 6%
2	F	106	% 92% 6%
2	H	106	2% 87% 8% 6%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 11035 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 Tenascin-R.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	219	Total	C	N	O	S	0	8	0
			1845	1162	332	339	12			
1	C	219	Total	C	N	O	S	0	5	0
			1836	1156	335	333	12			
1	E	215	Total	C	N	O	S	0	3	0
			1781	1121	321	328	11			
1	G	215	Total	C	N	O	S	0	4	0
			1785	1125	321	328	11			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1125	GLY	-	expression tag	UNP Q92752
A	1126	THR	-	expression tag	UNP Q92752
A	1127	GLY	-	expression tag	UNP Q92752
A	1128	SER	-	expression tag	UNP Q92752
C	1125	GLY	-	expression tag	UNP Q92752
C	1126	THR	-	expression tag	UNP Q92752
C	1127	GLY	-	expression tag	UNP Q92752
C	1128	SER	-	expression tag	UNP Q92752
E	1125	GLY	-	expression tag	UNP Q92752
E	1126	THR	-	expression tag	UNP Q92752
E	1127	GLY	-	expression tag	UNP Q92752
E	1128	SER	-	expression tag	UNP Q92752
G	1125	GLY	-	expression tag	UNP Q92752
G	1126	THR	-	expression tag	UNP Q92752
G	1127	GLY	-	expression tag	UNP Q92752
G	1128	SER	-	expression tag	UNP Q92752

- Molecule 2 is a protein called Receptor-type tyrosine-protein phosphatase zeta.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	99	Total	C	N	O	S	0	0	0
			790	502	124	159	5			
2	D	100	Total	C	N	O	S	0	2	0
			808	516	125	161	6			
2	F	103	Total	C	N	O	S	0	0	0
			816	517	128	166	5			
2	H	100	Total	C	N	O	S	0	0	0
			798	508	125	160	5			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	306	GLY	-	expression tag	UNP P23471
B	307	THR	-	expression tag	UNP P23471
B	308	GLY	-	expression tag	UNP P23471
B	309	SER	-	expression tag	UNP P23471
D	306	GLY	-	expression tag	UNP P23471
D	307	THR	-	expression tag	UNP P23471
D	308	GLY	-	expression tag	UNP P23471
D	309	SER	-	expression tag	UNP P23471
F	306	GLY	-	expression tag	UNP P23471
F	307	THR	-	expression tag	UNP P23471
F	308	GLY	-	expression tag	UNP P23471
F	309	SER	-	expression tag	UNP P23471
H	306	GLY	-	expression tag	UNP P23471
H	307	THR	-	expression tag	UNP P23471
H	308	GLY	-	expression tag	UNP P23471
H	309	SER	-	expression tag	UNP P23471

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca) (labeled as "Ligand of Interest" by depositor).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Ca	0	0
			1	1		
3	C	1	Total	Ca	0	0
			1	1		
3	E	1	Total	Ca	0	0
			1	1		
3	G	1	Total	Ca	0	0
			1	1		

- Molecule 4 is water.

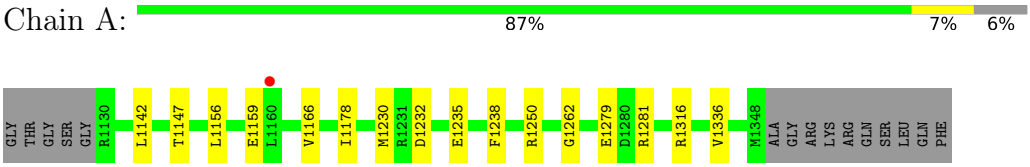
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	134	Total O 134 134	0	0
4	B	23	Total O 23 23	0	0
4	C	100	Total O 100 100	0	0
4	D	25	Total O 25 25	0	0
4	E	127	Total O 127 127	0	0
4	F	37	Total O 37 37	0	0
4	G	94	Total O 94 94	0	0
4	H	32	Total O 32 32	0	0

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.

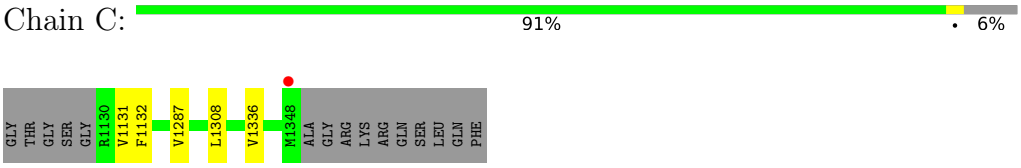
• Molecule 1: Tenascin-R

Chain A:



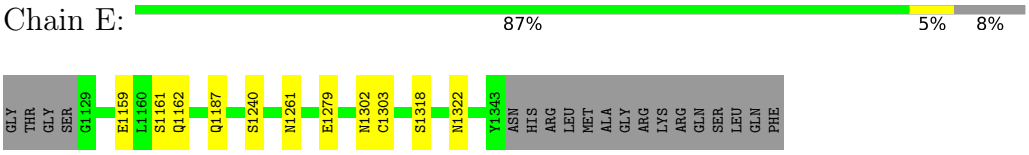
• Molecule 1: Tenascin-R

Chain C:



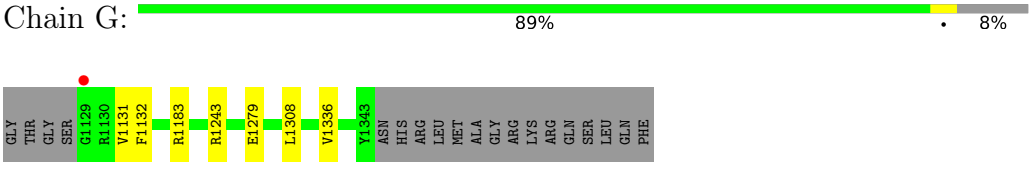
• Molecule 1: Tenascin-R

Chain E:



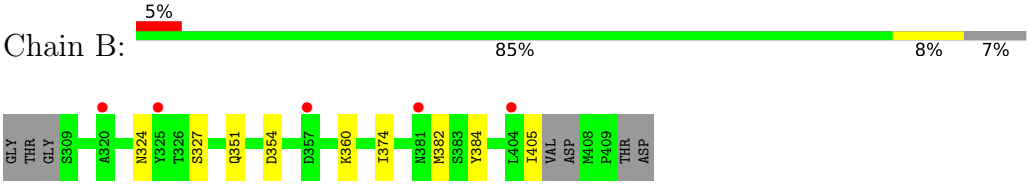
• Molecule 1: Tenascin-R

Chain G:



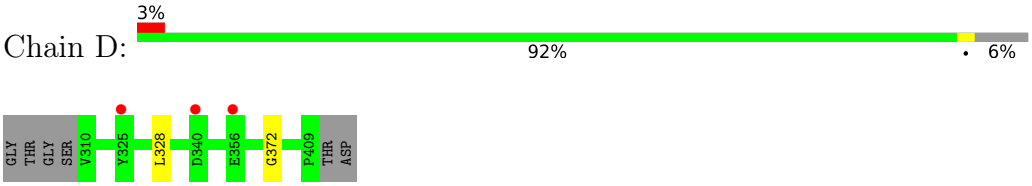
• Molecule 2: Receptor-type tyrosine-protein phosphatase zeta

Chain B:



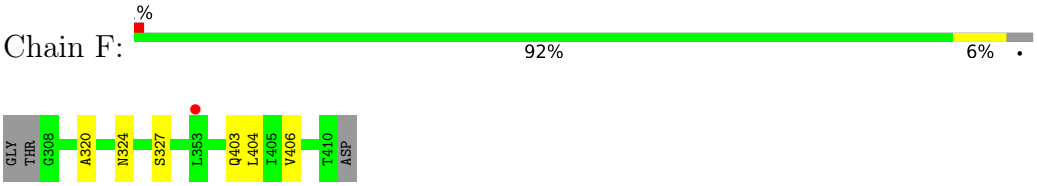
• Molecule 2: Receptor-type tyrosine-protein phosphatase zeta

Chain D:



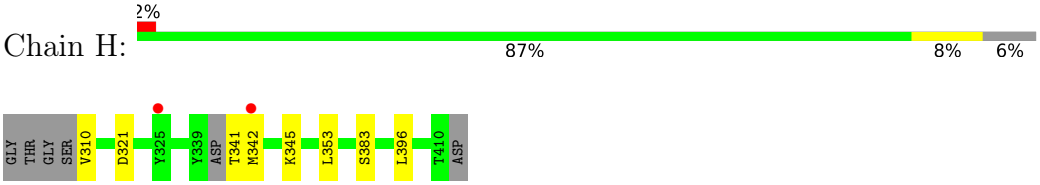
• Molecule 2: Receptor-type tyrosine-protein phosphatase zeta

Chain F:



• Molecule 2: Receptor-type tyrosine-protein phosphatase zeta

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α, β, γ	48.33Å 75.66Å 88.32Å 104.41° 92.40° 91.81°	Depositor
Resolution (Å)	42.96 – 1.80 49.77 – 1.80	Depositor EDS
% Data completeness (in resolution range)	91.3 (42.96-1.80) 87.2 (49.77-1.80)	Depositor EDS
R_{merge}	0.16	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.37 (at 1.81Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.174 , 0.210 0.175 , 0.210	Depositor DCC
R_{free} test set	2178 reflections (2.13%)	wwPDB-VP
Wilson B-factor (Å^2)	23.4	Xtriage
Anisotropy	0.215	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 43.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	11035	wwPDB-VP
Average B, all atoms (Å^2)	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 30.53 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.2877e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: CA

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	0/1922	0.72	0/2595
1	C	0.56	0/1904	0.68	0/2571
1	E	0.62	0/1842	0.69	0/2487
1	G	0.60	0/1849	0.68	0/2497
2	B	0.41	0/805	0.56	0/1096
2	D	0.41	0/830	0.54	0/1132
2	F	0.46	0/832	0.55	0/1135
2	H	0.47	0/813	0.57	0/1108
All	All	0.57	0/10797	0.66	0/14621

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

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	1845	0	1725	9	0
1	C	1836	0	1718	3	0
1	E	1781	0	1652	5	0
1	G	1785	0	1661	4	0
2	B	790	0	759	5	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	D	808	0	786	2	0
2	F	816	0	783	6	0
2	H	798	0	770	4	0
3	A	1	0	0	0	0
3	C	1	0	0	0	0
3	E	1	0	0	0	0
3	G	1	0	0	0	0
4	A	134	0	0	1	0
4	B	23	0	0	0	0
4	C	100	0	0	0	0
4	D	25	0	0	0	0
4	E	127	0	0	0	0
4	F	37	0	0	1	0
4	G	94	0	0	1	0
4	H	32	0	0	2	0
All	All	11035	0	9854	37	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 2.

All (37) 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)
2:F:320:ALA:HB3	2:F:406:VAL:HG11	1.52	0.91
1:A:1166[B]:VAL:HG11	1:A:1178:ILE:HG23	1.75	0.67
1:A:1166[B]:VAL:CG1	1:A:1178:ILE:HG23	2.27	0.65
2:H:353:LEU:HD12	2:H:383:SER:HB3	1.84	0.59
2:F:320:ALA:CB	2:F:406:VAL:HG11	2.31	0.59
1:A:1230[A]:MET:HG2	1:A:1336[A]:VAL:HG23	1.86	0.58
2:H:341:THR:OG1	2:H:342:MET:N	2.36	0.57
2:F:403:GLN:NE2	4:F:502:HOH:O	2.29	0.56
2:F:404:LEU:HD23	2:F:406:VAL:HG13	1.89	0.54
1:G:1131:VAL:HG13	1:G:1132:PHE:CD2	2.43	0.54
1:E:1240[B]:SER:OG	1:E:1261:ASN:OD1	2.27	0.53
2:F:404:LEU:CD2	2:F:406:VAL:HG13	2.40	0.52
2:H:310:VAL:HG13	2:H:396:LEU:HD13	1.93	0.51
1:A:1238:PHE:CZ	1:A:1262:GLY:HA2	2.47	0.50
1:E:1159:GLU:OE1	1:E:1162:GLN:NE2	2.44	0.50
1:C:1131:VAL:HG23	1:C:1132:PHE:CD2	2.48	0.49
2:B:354:ASP:HB2	2:B:382:MET:HE1	1.95	0.48
2:B:351:GLN:HB3	2:B:360:LYS:HA	1.96	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:G:1243:ARG:NE	4:G:1505:HOH:O	2.44	0.47
1:G:1308:LEU:HD13	1:G:1336[B]:VAL:HG11	1.95	0.47
1:A:1281:ARG:NH1	4:A:1511:HOH:O	2.46	0.47
2:H:345:LYS:NZ	4:H:502:HOH:O	2.38	0.47
1:A:1232:ASP:O	1:A:1235[B]:GLU:HG2	2.15	0.47
1:A:1142:LEU:HD12	1:A:1147:THR:HG22	1.97	0.46
1:E:1187:GLN:NE2	4:H:502:HOH:O	2.48	0.46
2:B:384:TYR:O	2:B:405:ILE:HA	2.17	0.45
2:B:374:ILE:O	2:B:374:ILE:HG13	2.17	0.45
1:G:1183:ARG:HB2	1:G:1336[A]:VAL:HG22	1.99	0.45
1:A:1156:LEU:O	1:A:1159:GLU:HG2	2.17	0.44
2:B:324:ASN:CG	2:B:327:SER:HG	2.18	0.44
1:C:1308:LEU:HD13	1:C:1336[B]:VAL:HG11	2.01	0.43
2:F:324:ASN:OD1	2:F:327:SER:OG	2.29	0.42
2:D:328:LEU:HD12	2:D:328:LEU:O	2.20	0.41
1:E:1302:ASN:N	1:E:1303:CYS:HA	2.34	0.41
1:C:1287[B]:VAL:HG21	2:D:372:GLY:C	2.41	0.41
1:E:1318:SER:HA	1:E:1322:ASN:OD1	2.21	0.41
1:A:1166[B]:VAL:HG13	1:A:1178:ILE:HG23	2.03	0.40

There are no symmetry-related clashes.

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	225/234 (96%)	215 (96%)	10 (4%)	0	100	100
1	C	222/234 (95%)	215 (97%)	7 (3%)	0	100	100
1	E	216/234 (92%)	209 (97%)	7 (3%)	0	100	100
1	G	217/234 (93%)	206 (95%)	11 (5%)	0	100	100
2	B	95/106 (90%)	91 (96%)	4 (4%)	0	100	100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	D	100/106 (94%)	96 (96%)	4 (4%)	0	100	100
2	F	101/106 (95%)	97 (96%)	4 (4%)	0	100	100
2	H	96/106 (91%)	90 (94%)	6 (6%)	0	100	100
All	All	1272/1360 (94%)	1219 (96%)	53 (4%)	0	100	100

There are no Ramachandran outliers to report.

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	197/199 (99%)	194 (98%)	3 (2%)	65	56
1	C	194/199 (98%)	194 (100%)	0	100	100
1	E	187/199 (94%)	185 (99%)	2 (1%)	73	68
1	G	188/199 (94%)	187 (100%)	1 (0%)	88	87
2	B	90/95 (95%)	90 (100%)	0	100	100
2	D	93/95 (98%)	93 (100%)	0	100	100
2	F	93/95 (98%)	93 (100%)	0	100	100
2	H	91/95 (96%)	90 (99%)	1 (1%)	73	68
All	All	1133/1176 (96%)	1126 (99%)	7 (1%)	86	84

All (7) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1250	ARG
1	A	1279	GLU
1	A	1316	ARG
1	E	1161	SER
1	E	1279	GLU
1	G	1279	GLU
2	H	321	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

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 4 ligands modelled in this entry, 4 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^{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	219/234 (93%)	-0.53	1 (0%) 91 89	16, 25, 49, 69	0
1	C	219/234 (93%)	-0.55	1 (0%) 91 89	19, 31, 49, 70	0
1	E	215/234 (91%)	-0.59	0 100 100	18, 28, 42, 57	0
1	G	215/234 (91%)	-0.57	1 (0%) 91 89	20, 29, 45, 75	0
2	B	99/106 (93%)	0.39	5 (5%) 28 22	28, 50, 77, 90	0
2	D	100/106 (94%)	0.06	3 (3%) 50 44	32, 47, 67, 90	0
2	F	103/106 (97%)	-0.06	1 (0%) 82 80	29, 40, 63, 83	0
2	H	100/106 (94%)	-0.23	2 (2%) 65 61	28, 39, 59, 70	0
All	All	1270/1360 (93%)	-0.37	14 (1%) 80 78	16, 33, 60, 90	0

All (14) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	325	TYR	4.4
2	H	325	TYR	3.4
1	G	1129	GLY	3.4
2	B	320	ALA	3.0
2	D	340	ASP	2.8
2	B	325	TYR	2.7
2	D	356	GLU	2.7
1	A	1160	LEU	2.6
2	B	381	ASN	2.5
2	B	357	ASP	2.5
2	B	404	LEU	2.1
1	C	1348	MET	2.1
2	F	353	LEU	2.1
2	H	342	MET	2.0

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 (i)

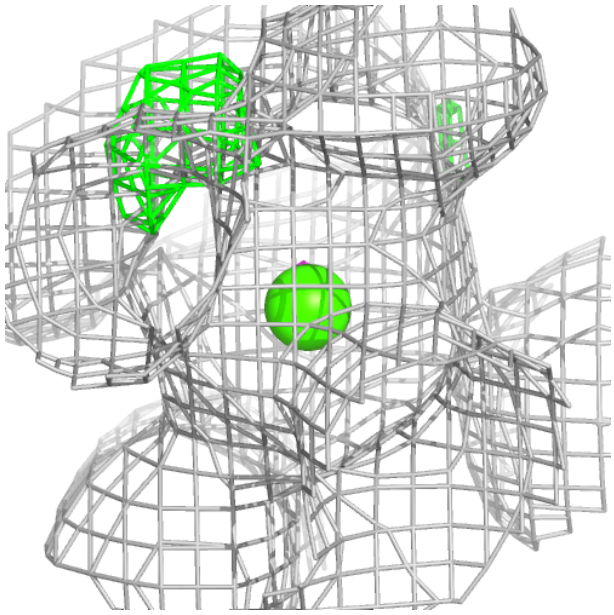
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9'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

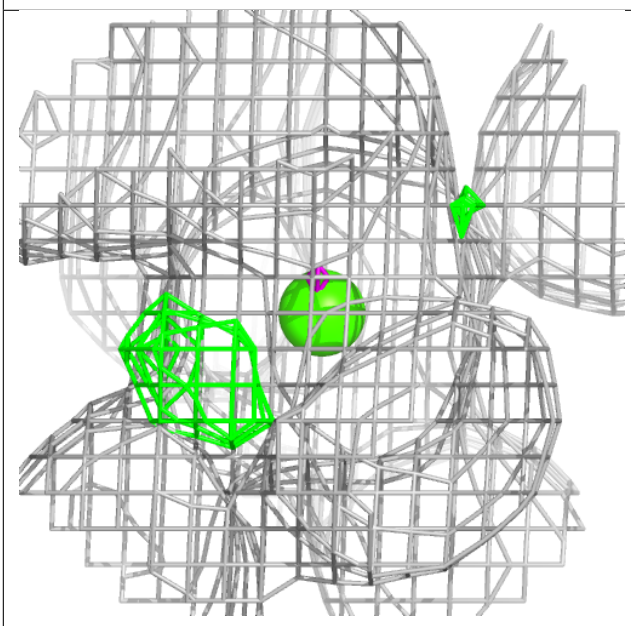
Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
3	CA	A	1401	1/1	0.99	0.03	24,24,24,24	0
3	CA	C	1401	1/1	0.99	0.05	34,34,34,34	0
3	CA	E	1401	1/1	0.99	0.03	33,33,33,33	0
3	CA	G	1401	1/1	0.99	0.04	34,34,34,34	0

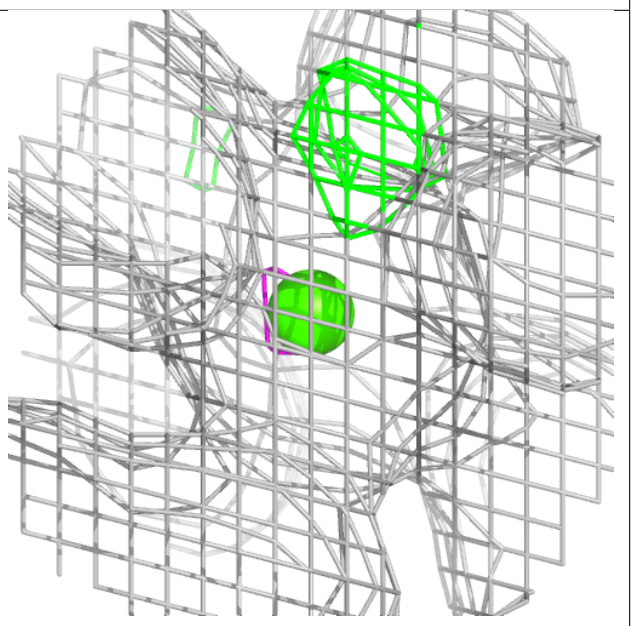
The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation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

Electron density around CA A 1401:

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

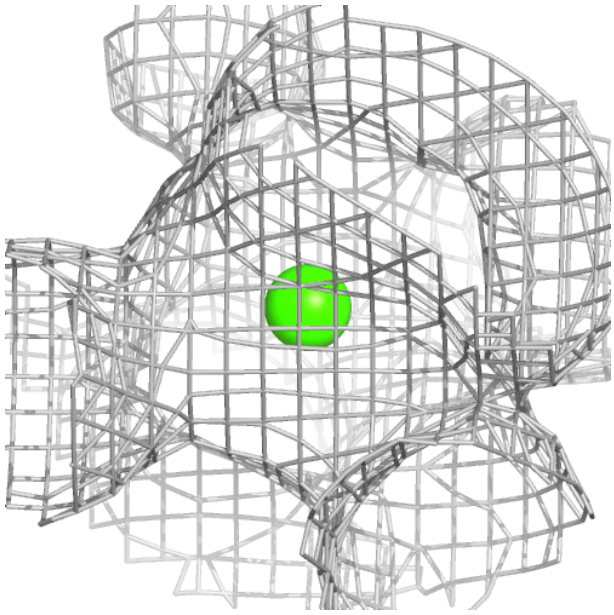


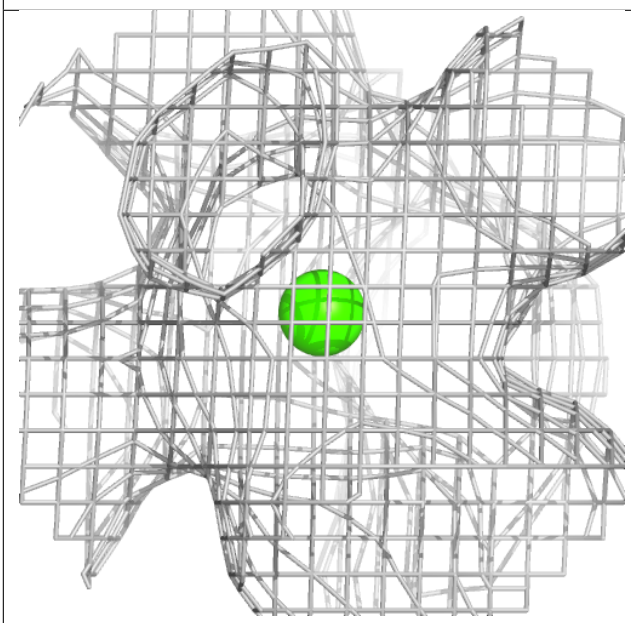


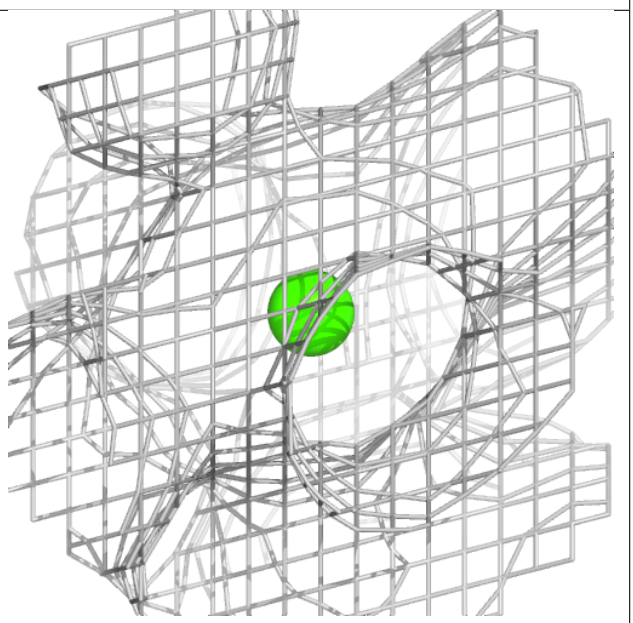


Electron density around CA C 1401:

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

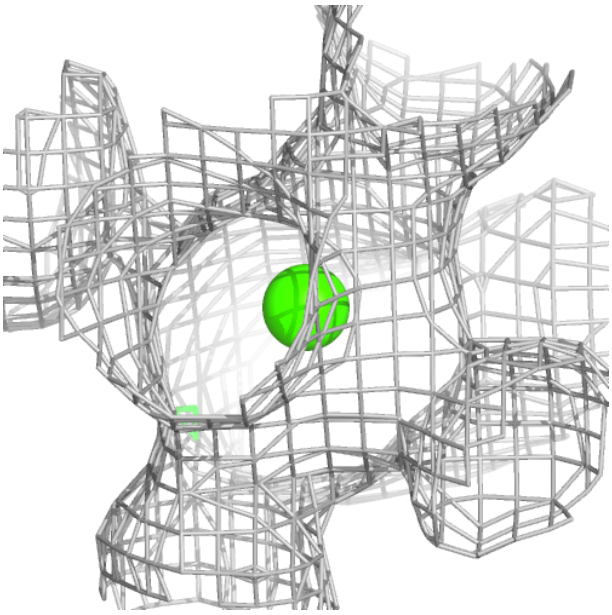


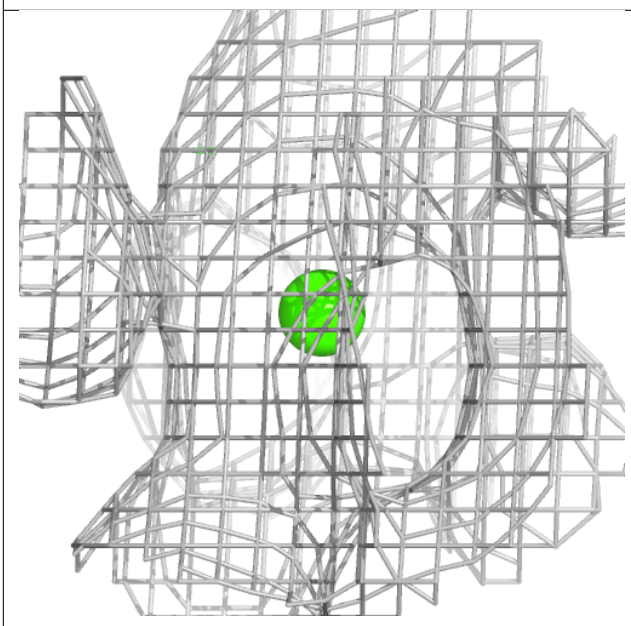


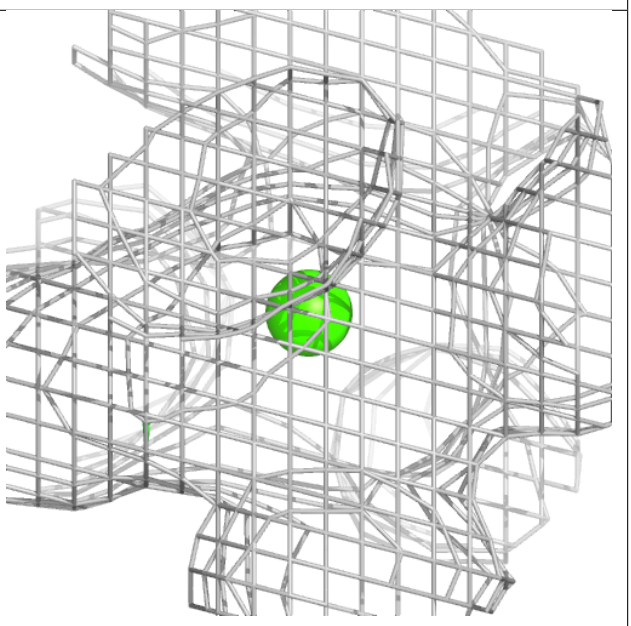


Electron density around CA E 1401:

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

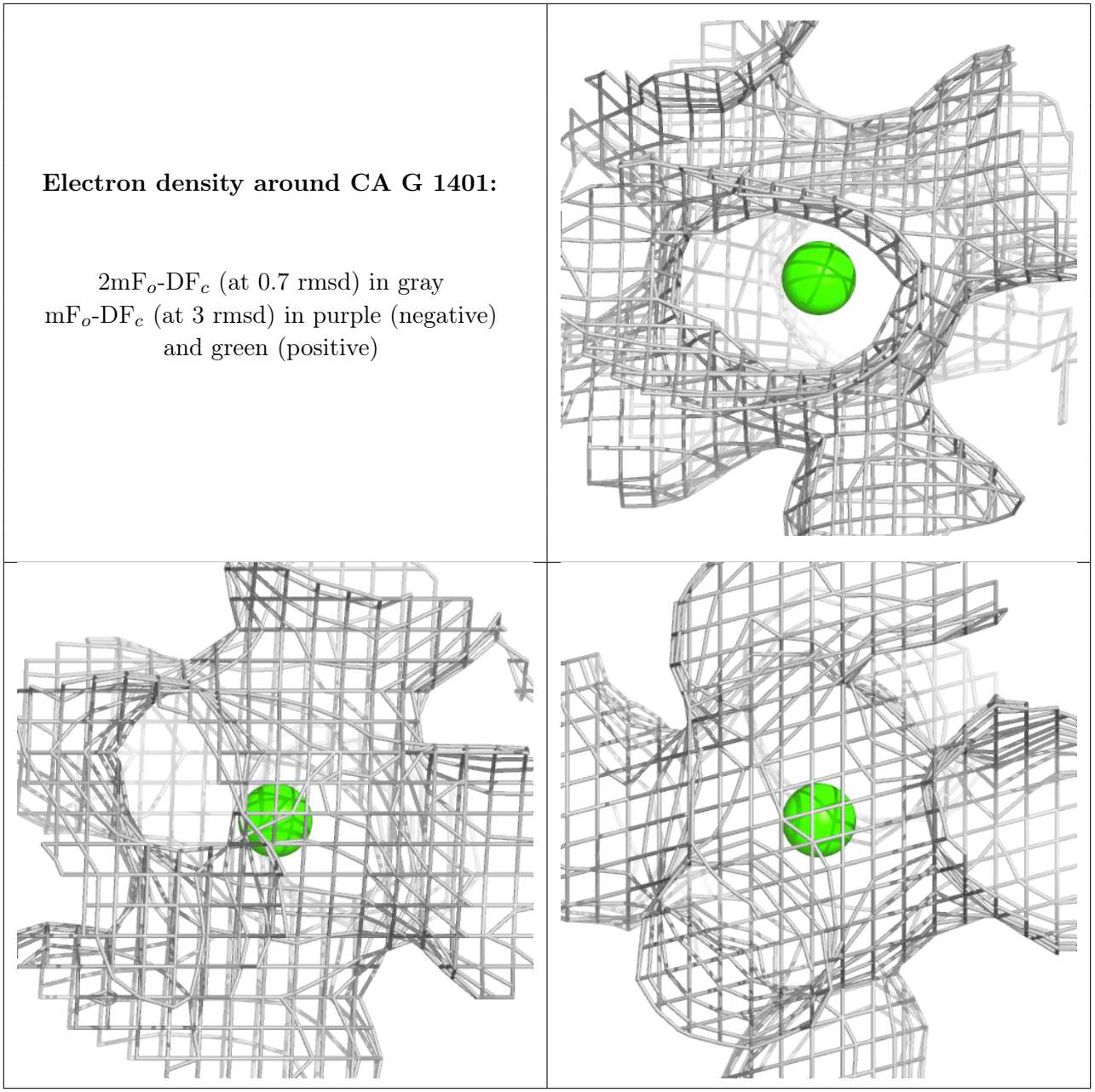






Electron density around CA G 1401:

$2mF_o$-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.