wwPDB X-ray Structure Validation Summary Report (i)

Sep 28, 2024 – 08:20 AM EDT

PDB ID	:	6V55
Title	:	Full extracellular region of zebrafish Gpr126/Adgrg6
Authors	:	Leon, K.; Arac, D.
Deposited on	:	2019-12-03
Resolution	:	2.38 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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 (i) 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 (1)) were used in the production of this report:

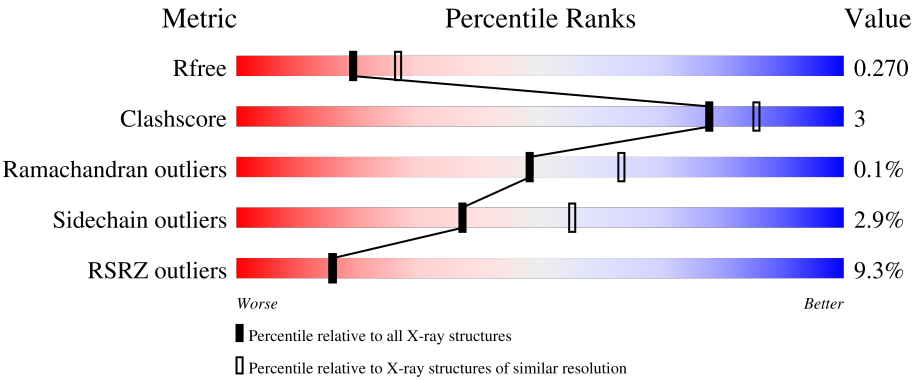
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.39

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 2.38 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	6699 (2.40-2.36)
Clashscore	180529	7414 (2.40-2.36)
Ramachandran outliers	177936	7337 (2.40-2.36)
Sidechain outliers	177891	7338 (2.40-2.36)
RSRZ outliers	164620	6699 (2.40-2.36)

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	788	9%, 85%, 9%, • 5%
2	Q	12	50%, 8%, 42%

2 Entry composition (i)

There are 5 unique types of molecules in this entry. The entry contains 6114 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 Adhesion G-protein coupled receptor G6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	746	Total	C	N	O	S	0	0	0
			5706	3590	941	1143	32			

There are 36 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	37	ALA	SER	conflict	UNP C6KFA3
A	234	PHE	LEU	conflict	UNP C6KFA3
A	263	GLU	GLY	conflict	UNP C6KFA3
A	?	-	ALA	deletion	UNP C6KFA3
A	?	-	THR	deletion	UNP C6KFA3
A	?	-	LEU	deletion	UNP C6KFA3
A	?	-	THR	deletion	UNP C6KFA3
A	?	-	VAL	deletion	UNP C6KFA3
A	?	-	THR	deletion	UNP C6KFA3
A	?	-	ILE	deletion	UNP C6KFA3
A	?	-	THR	deletion	UNP C6KFA3
A	?	-	SER	deletion	UNP C6KFA3
A	?	-	ILE	deletion	UNP C6KFA3
A	?	-	ALA	deletion	UNP C6KFA3
A	?	-	THR	deletion	UNP C6KFA3
A	?	-	THR	deletion	UNP C6KFA3
A	?	-	ASN	deletion	UNP C6KFA3
A	?	-	ILE	deletion	UNP C6KFA3
A	?	-	ILE	deletion	UNP C6KFA3
A	?	-	PRO	deletion	UNP C6KFA3
A	?	-	THR	deletion	UNP C6KFA3
A	?	-	ASN	deletion	UNP C6KFA3
A	?	-	ALA	deletion	UNP C6KFA3
A	?	-	THR	deletion	UNP C6KFA3
A	?	-	THR	deletion	UNP C6KFA3
A	?	-	HIS	deletion	UNP C6KFA3
A	461	ILE	THR	conflict	UNP C6KFA3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
A	804	TRP	CYS	conflict	UNP C6KFA3
A	840	HIS	-	expression tag	UNP C6KFA3
A	841	HIS	-	expression tag	UNP C6KFA3
A	842	HIS	-	expression tag	UNP C6KFA3
A	843	HIS	-	expression tag	UNP C6KFA3
A	844	HIS	-	expression tag	UNP C6KFA3
A	845	HIS	-	expression tag	UNP C6KFA3
A	846	HIS	-	expression tag	UNP C6KFA3
A	847	HIS	-	expression tag	UNP C6KFA3

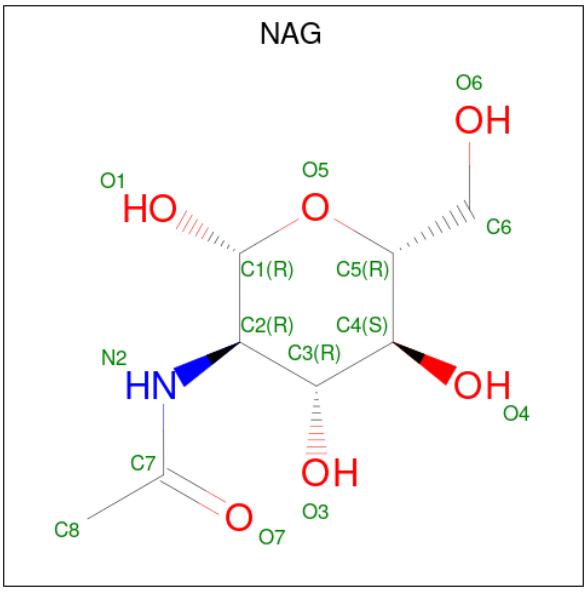
- Molecule 2 is a protein called Adhesion G-protein coupled receptor G6.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	Q	7	Total	C	N	O	S	0	0	0
			56	38	9	8	1			

- 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		

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C N O 14 8 1 5	0	0
4	A	1	Total C N O 14 8 1 5	0	0
4	A	1	Total C N O 14 8 1 5	0	0
4	A	1	Total C N O 14 8 1 5	0	0
4	A	1	Total C N O 14 8 1 5	0	0
4	A	1	Total C N O 14 8 1 5	0	0
4	A	1	Total C N O 14 8 1 5	0	0
4	A	1	Total C N O 14 8 1 5	0	0
4	A	1	Total C N O 14 8 1 5	0	0
4	A	1	Total C N O 14 8 1 5	0	0
4	A	1	Total C N O 14 8 1 5	0	0

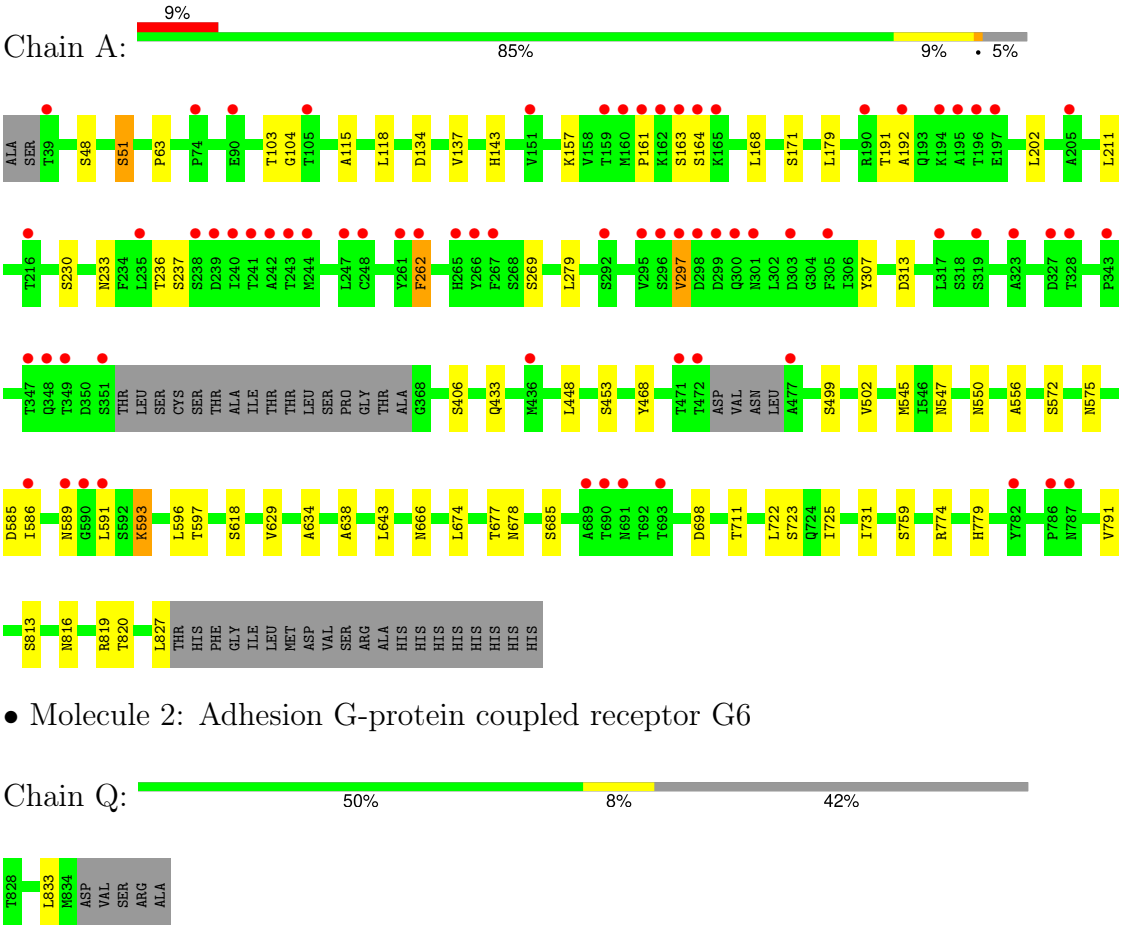
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	197	Total O 197 197	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: Adhesion G-protein coupled receptor G6



• Molecule 2: Adhesion G-protein coupled receptor G6

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	144.96Å 59.45Å 168.38Å 90.00° 107.82° 90.00°	Depositor
Resolution (Å)	47.90 – 2.38 47.90 – 2.38	Depositor EDS
% Data completeness (in resolution range)	62.3 (47.90-2.38) 62.3 (47.90-2.38)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.50 (at 2.37Å)	Xtriage
Refinement program	REFMAC 1.17.1_3660, PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.212 , 0.272 0.223 , 0.270	Depositor DCC
R_{free} test set	2739 reflections (5.09%)	wwPDB-VP
Wilson B-factor (Å^2)	24.2	Xtriage
Anisotropy	1.166	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.31 , 33.9	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	0.91	EDS
Total number of atoms	6114	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.86% 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: NAG, 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	0.49	0/5816	0.71	4/7934 (0.1%)
2	Q	0.53	0/57	0.59	0/75
All	All	0.49	0/5873	0.71	4/8009 (0.0%)

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	448	LEU	CA-CB-CG	7.69	133.00	115.30
1	A	297	VAL	CA-CB-CG2	6.46	120.59	110.90
1	A	134	ASP	CB-CG-OD2	5.31	123.08	118.30
1	A	262	PHE	CB-CG-CD2	5.12	124.39	120.80

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	5706	0	5492	31	0
2	Q	56	0	56	1	0
3	A	1	0	0	0	0
4	A	154	0	143	0	0
5	A	197	0	0	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	6114	0	5691	31	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.

The worst 5 of 31 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:164:SER:HB3	1:A:297:VAL:HG13	1.57	0.86
1:A:586:ILE:HG22	1:A:591:LEU:HD21	1.69	0.73
1:A:722:LEU:HA	1:A:725:ILE:HG22	1.76	0.68
1:A:629:VAL:HG13	1:A:666:ASN:ND2	2.17	0.60
1:A:685:SER:HB3	1:A:698:ASP:HB3	1.87	0.57

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	740/788 (94%)	703 (95%)	36 (5%)	1 (0%)	48 63
2	Q	5/12 (42%)	5 (100%)	0	0	100 100
All	All	745/800 (93%)	708 (95%)	36 (5%)	1 (0%)	48 63

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	137	VAL

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	647/695 (93%)	628 (97%)	19 (3%)	37 55
2	Q	6/10 (60%)	6 (100%)	0	100 100
All	All	653/705 (93%)	634 (97%)	19 (3%)	37 55

5 of 19 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	593	LYS
1	A	723	SER
1	A	813	SER
1	A	618	SER
1	A	313	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	72	GLN
1	A	348	GLN
1	A	493	ASN

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

Of 12 ligands modelled in this entry, 1 is monoatomic - leaving 11 for Mogul analysis.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the Chemical Component Dictionary. The Link column lists molecule types, if any, to which the group is linked. 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 $|Z| > 2$ is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\| > 2$	Counts	RMSZ	$\#\|Z\| > 2$
4	NAG	A	912	1	14,14,15	0.90	1 (7%)	17,19,21	0.82	1 (5%)
4	NAG	A	908	1	14,14,15	1.21	2 (14%)	17,19,21	0.72	1 (5%)
4	NAG	A	907	1	14,14,15	0.87	2 (14%)	17,19,21	0.55	0
4	NAG	A	905	1	14,14,15	0.99	2 (14%)	17,19,21	0.90	1 (5%)
4	NAG	A	902	1	14,14,15	1.08	1 (7%)	17,19,21	0.66	1 (5%)
4	NAG	A	909	1	14,14,15	0.91	1 (7%)	17,19,21	0.65	1 (5%)
4	NAG	A	911	1	14,14,15	0.79	1 (7%)	17,19,21	0.59	0
4	NAG	A	904	1	14,14,15	1.45	2 (14%)	17,19,21	1.05	1 (5%)
4	NAG	A	906	1	14,14,15	1.15	2 (14%)	17,19,21	0.94	1 (5%)
4	NAG	A	903	1	14,14,15	0.97	2 (14%)	17,19,21	1.03	1 (5%)
4	NAG	A	910	1	14,14,15	0.78	1 (7%)	17,19,21	0.65	1 (5%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	A	912	1	-	0/6/23/26	0/1/1/1
4	NAG	A	908	1	-	2/6/23/26	0/1/1/1
4	NAG	A	907	1	-	2/6/23/26	0/1/1/1
4	NAG	A	905	1	-	2/6/23/26	0/1/1/1
4	NAG	A	902	1	-	2/6/23/26	0/1/1/1
4	NAG	A	909	1	-	2/6/23/26	0/1/1/1
4	NAG	A	911	1	-	1/6/23/26	0/1/1/1
4	NAG	A	904	1	-	1/6/23/26	0/1/1/1
4	NAG	A	906	1	-	1/6/23/26	0/1/1/1
4	NAG	A	903	1	-	2/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
4	NAG	A	910	1	-	2/6/23/26	0/1/1/1

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	904	NAG	C1-C2	3.96	1.57	1.52
4	A	902	NAG	O5-C1	3.45	1.49	1.43
4	A	908	NAG	O5-C1	3.43	1.49	1.43
4	A	904	NAG	O5-C1	3.36	1.49	1.43
4	A	906	NAG	O5-C1	3.07	1.48	1.43

The worst 5 of 9 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
4	A	903	NAG	C1-O5-C5	3.58	116.98	112.19
4	A	904	NAG	C1-O5-C5	3.35	116.67	112.19
4	A	906	NAG	C1-O5-C5	3.33	116.64	112.19
4	A	905	NAG	C1-O5-C5	3.22	116.50	112.19
4	A	912	NAG	C1-O5-C5	2.62	115.69	112.19

There are no chirality outliers.

5 of 17 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	907	NAG	O5-C5-C6-O6
4	A	908	NAG	C4-C5-C6-O6
4	A	908	NAG	O5-C5-C6-O6
4	A	902	NAG	O5-C5-C6-O6
4	A	907	NAG	C4-C5-C6-O6

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	746/788 (94%)	0.37	70 (9%) 15 15	13, 37, 73, 102	0
2	Q	7/12 (58%)	-0.10	0 100 100	20, 24, 36, 50	0
All	All	753/800 (94%)	0.37	70 (9%) 16 16	13, 37, 73, 102	0

The worst 5 of 70 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	192	ALA	6.5
1	A	296	SER	5.7
1	A	690	THR	5.4
1	A	297	VAL	5.4
1	A	241	THR	5.3

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

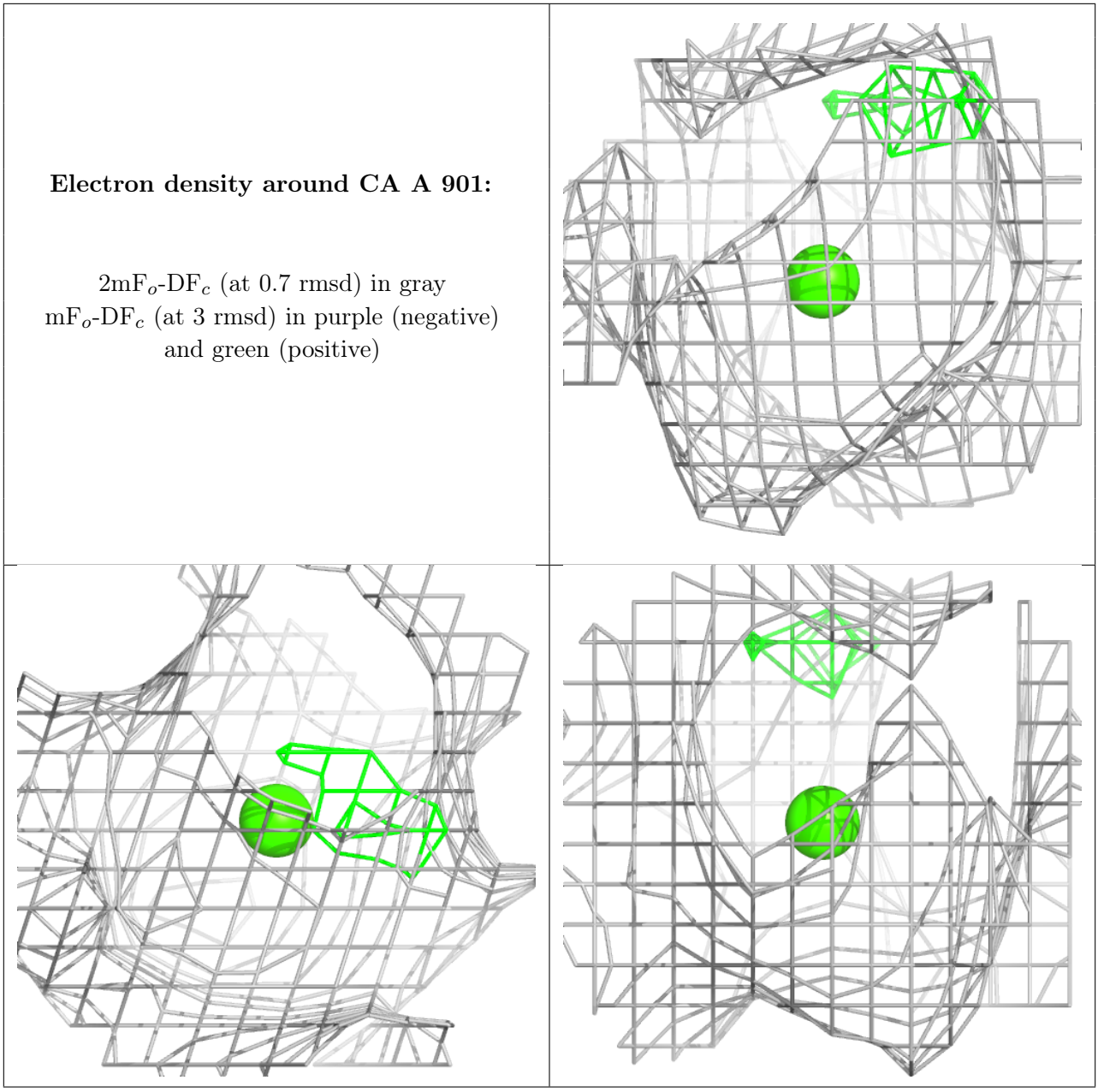
In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} percentile and maximum values of B factors of atoms in the group. The column labelled 'Q< 0.9' 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
4	NAG	A	904	14/15	0.41	0.21	59,70,75,78	0
4	NAG	A	911	14/15	0.47	0.20	68,74,79,79	0
4	NAG	A	912	14/15	0.70	0.20	59,68,81,85	0
4	NAG	A	905	14/15	0.71	0.18	60,67,80,83	0
4	NAG	A	907	14/15	0.74	0.18	49,59,74,79	0
4	NAG	A	910	14/15	0.77	0.14	62,68,73,73	0
4	NAG	A	906	14/15	0.78	0.13	56,62,65,66	0
4	NAG	A	909	14/15	0.84	0.12	46,51,56,56	0
4	NAG	A	902	14/15	0.84	0.13	41,46,52,53	0
4	NAG	A	908	14/15	0.86	0.12	52,60,70,70	0
3	CA	A	901	1/1	0.94	0.03	19,19,19,19	0
4	NAG	A	903	14/15	0.95	0.07	30,38,43,44	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 901:

$2mF_o$-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.