



wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 19, 2016 – 09:18 PM GMT

PDB ID : 4ZG7
Title : Structural basis for inhibition of human autotaxin by four novel compounds
Authors : Stein, A.J.; Bain, G.; Hutchinson, J.H.; Evans, J.F.
Deposited on : 2015-04-22
Resolution : 1.75 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
<http://wwpdb.org/validation/2016/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
Mogul : 1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix) : 1.9-1692
EDS : rb-20026982
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac : 5.8.0135
CCP4 : 6.5.0
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP) : rb-20026982

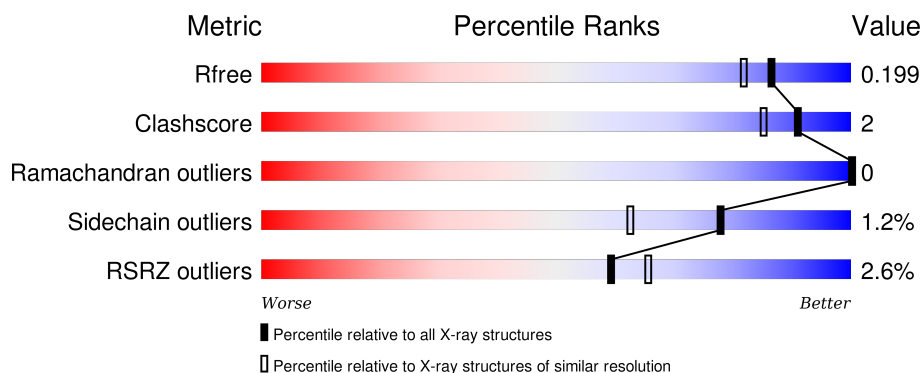
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION


The reported resolution of this entry is 1.75 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}	91344	1609 (1.76-1.76)
Clashscore	102246	1730 (1.76-1.76)
Ramachandran outliers	100387	1711 (1.76-1.76)
Sidechain outliers	100360	1711 (1.76-1.76)
RSRZ outliers	91569	1610 (1.76-1.76)

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. 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	806	

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
7	NKN	A	909	-	-	-	X

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
8	GOL	A	910	-	-	-	X
8	GOL	A	912	-	-	-	X
8	GOL	A	915	-	-	-	X
8	GOL	A	916	-	-	-	X

2 Entry composition [i](#)

There are 10 unique types of molecules in this entry. The entry contains 6888 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.

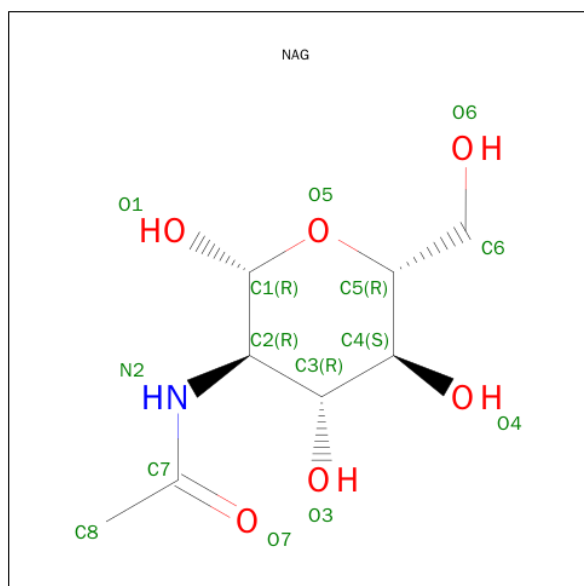
- Molecule 1 is a protein called Ectonucleotide pyrophosphatase/phosphodiesterase family member 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	776	6255	3982	1062	1163	48	0	12	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	411	ALA	ASN	conflict	UNP Q13822
A	493	PRO	SER	variant	UNP Q13822

- Molecule 2 is N-ACETYL-D-GLUCOSAMINE (three-letter code: NAG) (formula: C₈H₁₅NO₆).



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

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	2	Total	Zn	0	0
			2	2		

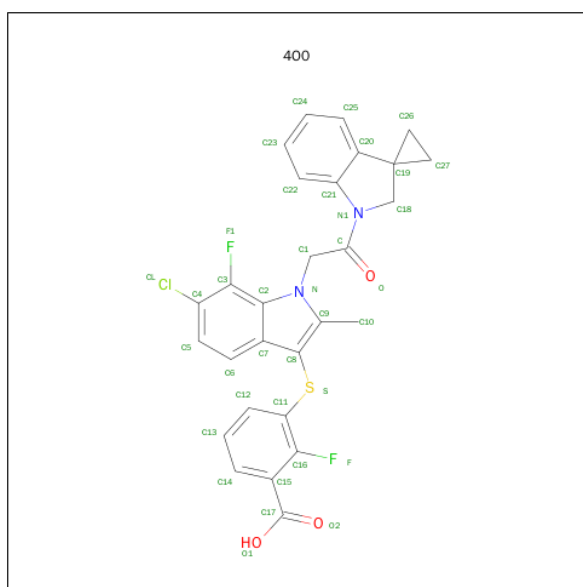
- Molecule 4 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	1	Total	Ca	0	0
			1	1		

- Molecule 5 is SODIUM ION (three-letter code: NA) (formula: Na).

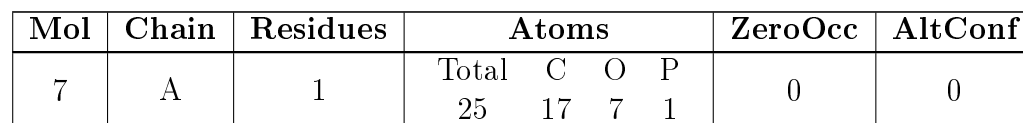
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	2	Total	Na	0	0
			2	2		

- Molecule 6 is 3-({6-chloro-7-fluoro-2-methyl-1-[2-oxo-2-(spiro[cyclopropane-1,3'-indol]-1'-(2'-H)-yl)ethyl]-1H-indol-3-yl}sulfanyl)-2-fluorobenzoic acid (three-letter code: 4O0) (formula: C₂₈H₂₁ClF₂N₂O₃S).



Mol	Chain	Residues	Atoms							ZeroOcc	AltConf
6	A	1	Total	C	Cl	F	N	O	S	0	0
			37	28	1	2	2	3	1		

- Molecule 7 is (2R)-2-hydroxy-3-(phosphonoxy)propyl tetradecanoate (three-letter code: NKN) (formula: C₁₇H₃₅O₇P).



- GOL
-
- The diagram shows a skeletal structure of 1,2,3-propanetriol (glycerol). The carbon chain is represented by three vertices labeled C1, C2, and C3 in green. C1 is on the left, C2 is in the middle, and C3 is on the right. Each carbon is bonded to an oxygen atom, labeled O1, O2, and O3 respectively, in green. The bonds to the oxygen atoms are shown in red. The hydrogens are represented by the letters H in red, which are part of the HO and OH groups. The structure is drawn in a zig-zag conformation.

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total 6	C 3	O 3	0	0
8	A	1	Total 6	C 3	O 3	0	0
8	A	1	Total 6	C 3	O 3	0	0



Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	A	1	Total	C	O	0	0
			6	3	3		
8	A	1	Total	C	O	0	0
			6	3	3		
8	A	1	Total	C	O	0	0
			6	3	3		
8	A	1	Total	C	O	0	0
			6	3	3		

- Molecule 9 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	1	Total	Cl	0	0
			1	1		

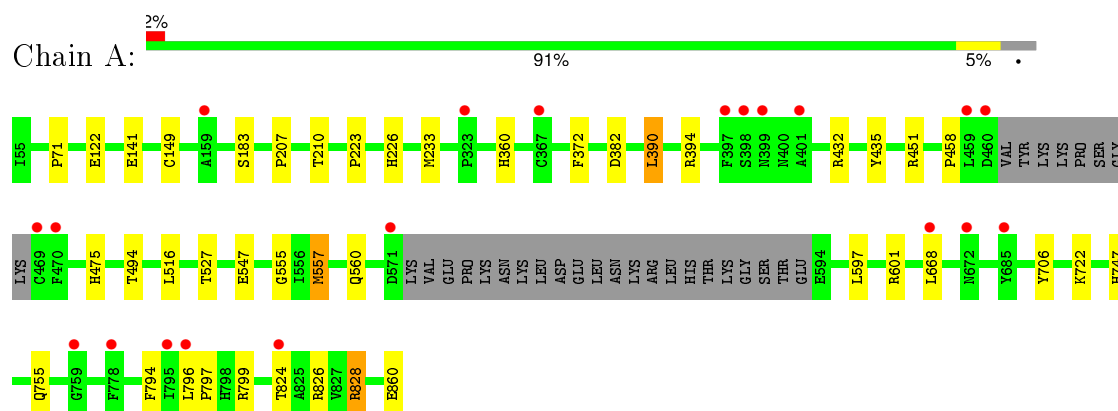
- Molecule 10 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	A	495	Total	O	0	0
			495	495		

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 errors 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: Ectonucleotide pyrophosphatase/phosphodiesterase family member 2



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	61.77Å 85.82Å 83.86Å 90.00° 111.34° 90.00°	Depositor
Resolution (Å)	37.61 – 1.75 37.61 – 1.75	Depositor EDS
% Data completeness (in resolution range)	98.5 (37.61-1.75) 98.5 (37.61-1.75)	Depositor EDS
R_{merge}	0.06	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.03 (at 1.75Å)	Xtriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.164 , 0.195 0.173 , 0.199	Depositor DCC
R_{free} test set	4053 reflections (5.29%)	DCC
Wilson B-factor (Å ²)	19.9	Xtriage
Anisotropy	0.079	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.39 , 46.0	EDS
Estimated twinning fraction	0.025 for h,-k,-h-l	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.50$, $\langle L^2 \rangle = 0.33$	Xtriage
Outliers	0 of 80640 reflections	Xtriage
F_o, F_c correlation	0.96	EDS
Total number of atoms	6888	wwPDB-VP
Average B, all atoms (Å ²)	23.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.74% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: GOL, ZN, NAG, CL, NA, 4O0, NKN, 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	0.52	0/6471	0.70	7/8795 (0.1%)

There are no bond length outliers.

The worst 5 of 7 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	451	ARG	NE-CZ-NH1	5.66	123.13	120.30
1	A	799	ARG	NE-CZ-NH2	5.61	123.11	120.30
1	A	149	CYS	CA-CB-SG	5.51	123.92	114.00
1	A	557	MET	CG-SD-CE	-5.45	91.48	100.20
1	A	799	ARG	NE-CZ-NH1	-5.38	117.61	120.30

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	6255	0	5966	27	0
2	A	28	0	25	0	0
3	A	2	0	0	0	0
4	A	1	0	0	0	0
5	A	2	0	0	0	0
6	A	37	0	20	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
7	A	25	0	33	1	0
8	A	42	0	56	0	0
9	A	1	0	0	0	0
10	A	495	0	0	7	1
All	All	6888	0	6100	27	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 2.

The worst 5 of 27 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:547:GLU:OE1	10:A:1001:HOH:O	1.85	0.93
1:A:122:GLU:CD	10:A:1008:HOH:O	2.15	0.83
1:A:794:PHE:CD1	1:A:796[B]:LEU:HD13	2.15	0.81
1:A:601[B]:ARG:HH21	1:A:601[B]:ARG:HG3	1.53	0.74
1:A:71:PRO:O	10:A:1002:HOH:O	2.06	0.72

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)
10:A:1002:HOH:O	10:A:1010:HOH:O[2_454]	1.35	0.85

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	782/806 (97%)	760 (97%)	22 (3%)	0	100 100

There are no Ramachandran outliers to report.

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	697/730 (96%)	688 (99%)	9 (1%)	76 60

5 of 9 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	560	GLN
1	A	828[B]	ARG
1	A	722	LYS
1	A	382	ASP
1	A	668	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 7 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	472	GLN
1	A	857	HIS
1	A	747	HIS
1	A	291	GLN
1	A	755	GLN

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 17 ligands modelled in this entry, 6 are 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$
2	NAG	A	901	1,2	14,14,15	0.65	0	15,19,21	1.19	1 (6%)
2	NAG	A	902	2	14,14,15	0.40	0	15,19,21	1.47	2 (13%)
6	4O0	A	908	-	36,42,42	2.37	11 (30%)	42,65,65	2.21	16 (38%)
7	NKN	A	909	3	24,24,24	1.60	3 (12%)	26,28,28	1.51	4 (15%)
8	GOL	A	910	-	5,5,5	0.25	0	5,5,5	0.69	0
8	GOL	A	911	-	5,5,5	0.23	0	5,5,5	0.35	0
8	GOL	A	912	-	5,5,5	0.37	0	5,5,5	0.43	0
8	GOL	A	913	-	5,5,5	0.25	0	5,5,5	0.29	0
8	GOL	A	914	-	5,5,5	0.37	0	5,5,5	0.32	0
8	GOL	A	915	-	5,5,5	0.23	0	5,5,5	0.54	0
8	GOL	A	916	-	5,5,5	0.32	0	5,5,5	0.56	0

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
2	NAG	A	901	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	902	2	-	0/6/23/26	0/1/1/1
6	4O0	A	908	-	-	0/10/35/35	0/5/6/6
7	NKN	A	909	3	-	0/24/24/24	0/0/0/0
8	GOL	A	910	-	-	0/4/4/4	0/0/0/0
8	GOL	A	911	-	-	0/4/4/4	0/0/0/0
8	GOL	A	912	-	-	0/4/4/4	0/0/0/0
8	GOL	A	913	-	-	0/4/4/4	0/0/0/0
8	GOL	A	914	-	-	0/4/4/4	0/0/0/0
8	GOL	A	915	-	-	0/4/4/4	0/0/0/0
8	GOL	A	916	-	-	0/4/4/4	0/0/0/0

The worst 5 of 14 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
6	A	908	4O0	C11-S	-5.34	1.70	1.78
6	A	908	4O0	C19-C20	-4.19	1.49	1.52
6	A	908	4O0	C8-S	-3.56	1.73	1.78
6	A	908	4O0	C1-C	-2.84	1.49	1.52
7	A	909	NKN	PAC-OAA	-2.77	1.45	1.54

The worst 5 of 23 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	A	908	4O0	C4-C3-C2	-6.56	115.45	121.00
6	A	908	4O0	C21-N1-C	-3.89	121.66	125.19
6	A	908	4O0	C16-C11-S	-3.71	116.19	121.36
6	A	908	4O0	C8-C7-C2	-3.33	100.48	112.77
7	A	909	NKN	OAB-PAC-OAD	-3.06	100.64	110.63

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
7	A	909	NKN	1	0

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	776/806 (96%)	-0.09	20 (2%) 59 65	13, 21, 38, 59	0

The worst 5 of 20 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	469	CYS	5.2
1	A	401	ALA	4.7
1	A	685	TYR	4.2
1	A	159	ALA	3.4
1	A	459	LEU	3.2

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 carbohydrates 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. 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	LLDF	B-factors(Å ²)	Q<0.9
8	GOL	A	912	6/6	0.78	0.20	14.95	27,29,31,35	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
8	GOL	A	916	6/6	0.84	0.14	4.91	29,39,41,42	0
8	GOL	A	915	6/6	0.77	0.18	3.75	39,44,47,49	0
7	NKN	A	909	25/25	0.89	0.18	3.67	23,36,44,48	0
8	GOL	A	910	6/6	0.85	0.11	2.38	28,33,38,42	0
8	GOL	A	911	6/6	0.93	0.12	1.52	21,26,27,27	0
8	GOL	A	913	6/6	0.87	0.10	0.39	30,34,36,38	0
8	GOL	A	914	6/6	0.87	0.11	-0.16	27,31,32,36	0
5	NA	A	906	1/1	0.91	0.09	-0.72	26,26,26,26	0
2	NAG	A	901	14/15	0.96	0.06	-1.14	16,17,20,21	0
6	4O0	A	908	37/37	0.98	0.06	-1.17	15,18,20,21	0
3	ZN	A	903	1/1	1.00	0.05	-1.63	16,16,16,16	0
4	CA	A	905	1/1	0.99	0.04	-1.92	16,16,16,16	0
9	CL	A	917	1/1	1.00	0.02	-3.10	22,22,22,22	0
3	ZN	A	904	1/1	1.00	0.04	-3.28	20,20,20,20	0
5	NA	A	907	1/1	0.99	0.02	-3.58	21,21,21,21	0
2	NAG	A	902	14/15	0.93	0.11	-	23,28,32,33	0

6.5 Other polymers ⓘ

There are no such residues in this entry.