wwPDB X-ray Structure Validation Summary Report (i)

May 26, 2020 – 03:01 am BST

PDB ID	:	2WJ3
Title	:	CRYSTAL STRUCTURE OF THE COFACTOR-DEVOID 1-H-3-HYDROXY-4- OXOQUINALDINE 2,4-DIOXYGENASE (HOD) FROM ARTHROBACTER NITROGUAJACOLICUS RU61A
Authors	:	Steiner, R.A.
Deposited on	:	2009-05-20
Resolution	:	2.09 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 following versions of software and data (see references (1)) were used in the production of this report:

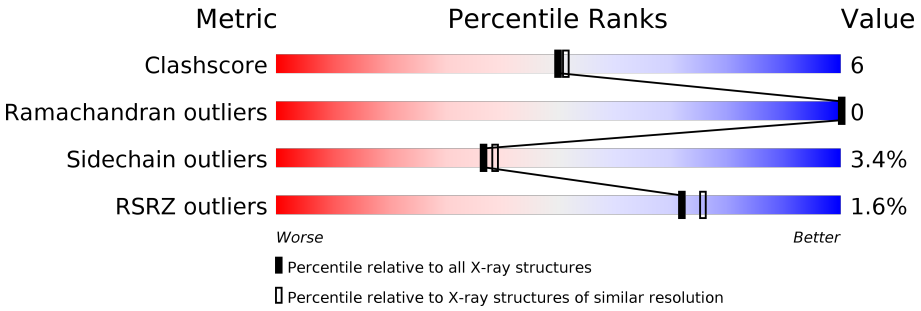
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.11
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

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 2.09 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))
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 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	276	3% 88% 11%
1	B	276	% 83% 15%
1	C	276	84% 14%
1	D	276	3% 84% 14%

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
2	GOL	B	1276	-	-	X	-

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 9322 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.

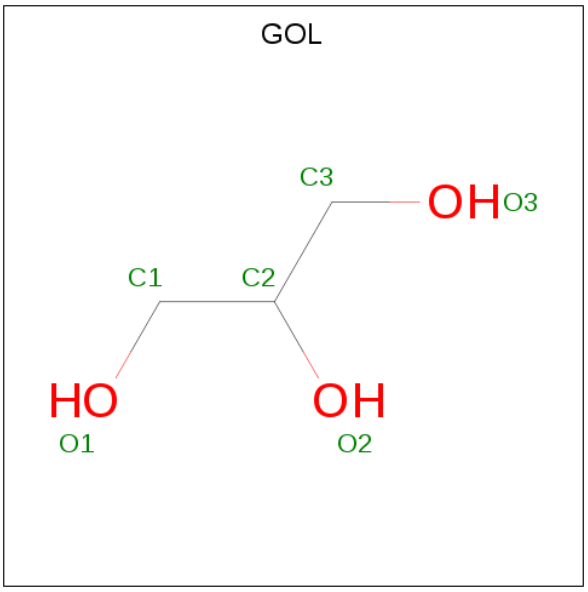
- Molecule 1 is a protein called 1-H-3-HYDROXY-4-OXOQUINALDINE 2,4-DIOXYGENASE.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	274	Total	C	N	O	S	0	0	0
			2234	1426	391	409	8			
1	B	273	Total	C	N	O	S	0	1	0
			2233	1426	391	408	8			
1	C	273	Total	C	N	O	S	0	1	0
			2236	1428	391	409	8			
1	D	273	Total	C	N	O	S	0	1	0
			2238	1429	393	408	8			

There are 4 discrepancies between the modelled and reference sequences:

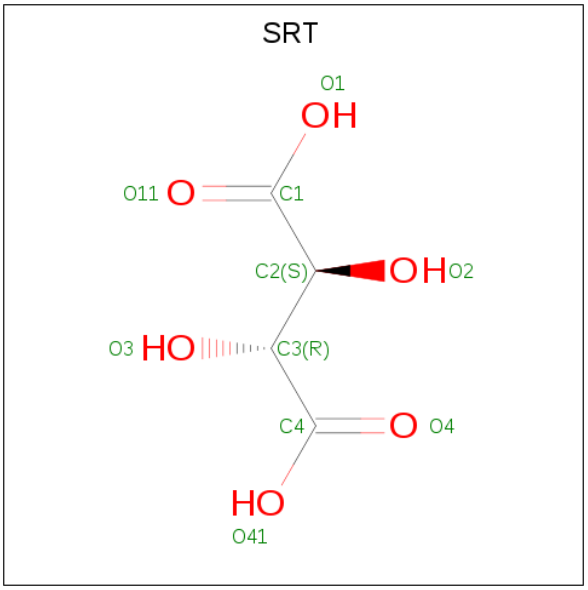
Chain	Residue	Modelled	Actual	Comment	Reference
A	69	SER	CYS	engineered mutation	UNP A4V8M9
B	69	SER	CYS	engineered mutation	UNP A4V8M9
C	69	SER	CYS	engineered mutation	UNP A4V8M9
D	69	SER	CYS	engineered mutation	UNP A4V8M9

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



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

- Molecule 3 is S,R MESO-TARTARIC ACID (three-letter code: SRT) (formula: $C_4H_6O_6$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 10 4 6	0	0
3	A	1	Total C O 10 4 6	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total C O 10 4 6	0	0
3	B	1	Total C O 10 4 6	0	0
3	C	1	Total C O 10 4 6	0	0
3	C	1	Total C O 10 4 6	0	0
3	D	1	Total C O 10 4 6	0	0
3	D	1	Total C O 20 8 12	0	1

- Molecule 4 is water.

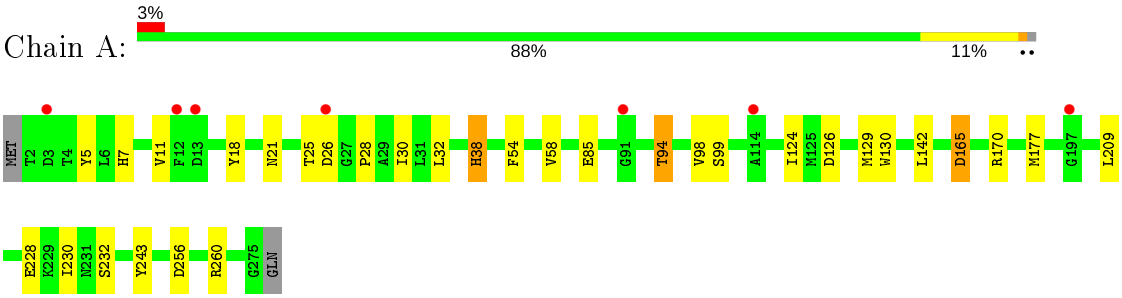
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	64	Total O 64 64	0	0
4	B	71	Total O 71 71	0	0
4	C	88	Total O 88 88	0	0
4	D	56	Total O 56 56	0	0

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 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.

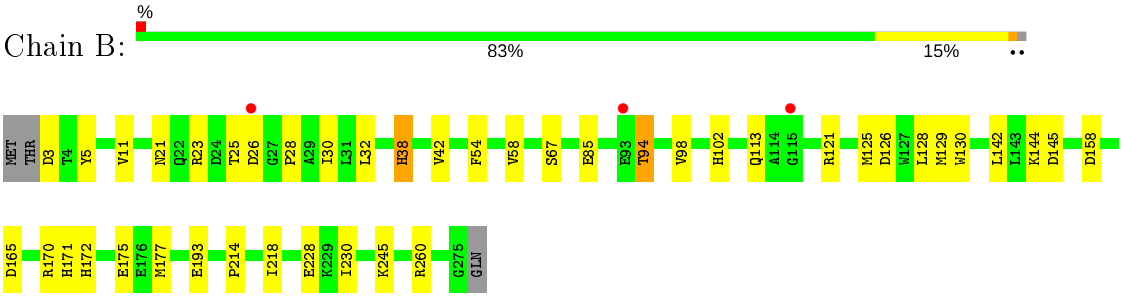
- Molecule 1: 1-H-3-HYDROXY-4-OXOQUINALDINE 2,4-DIOXYGENASE

Chain A:



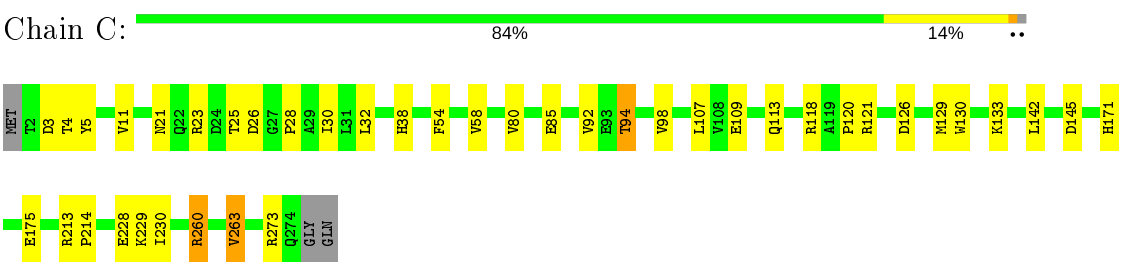
- Molecule 1: 1-H-3-HYDROXY-4-OXOQUINALDINE 2,4-DIOXYGENASE

Chain B:



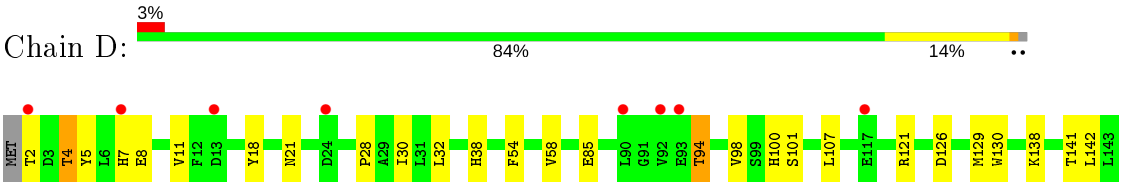
- Molecule 1: 1-H-3-HYDROXY-4-OXOQUINALDINE 2,4-DIOXYGENASE

Chain C:



- Molecule 1: 1-H-3-HYDROXY-4-OXOQUINALDINE 2,4-DIOXYGENASE

Chain D:




K144
D145
E193
P202
M203
Q204
H205
M206
P214
E228
K229
I230
N231
S232
Y243
R260
V263
H264
I265
Q274
GLY
GLN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	45.17Å 167.86Å 168.03Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	53.16 – 2.09 53.13 – 2.10	Depositor EDS
% Data completeness (in resolution range)	97.5 (53.16-2.09) 92.5 (53.13-2.10)	Depositor EDS
R_{merge}	0.07	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.07 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.5.0070	Depositor
R, R_{free}	0.196 , 0.232 0.208 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	40.5	Xtriage
Anisotropy	0.048	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 56.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.136 for -h,l,k	Xtriage
Reported twinning fraction	0.510 for H, K, L 0.490 for -H, L, K	Depositor
Outliers	0 of 70649 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	9322	wwPDB-VP
Average B, all atoms (Å^2)	38.0	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 26.92 % 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 2.4133e-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: GOL, SRT

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	0.77	0/2305	0.79	4/3137 (0.1%)
1	B	0.77	0/2307	0.80	2/3139 (0.1%)
1	C	0.79	0/2310	0.78	3/3144 (0.1%)
1	D	0.77	0/2312	0.79	2/3146 (0.1%)
All	All	0.78	0/9234	0.79	11/12566 (0.1%)

There are no bond length outliers.

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	260	ARG	NE-CZ-NH2	8.72	124.66	120.30
1	B	260	ARG	NE-CZ-NH1	-7.93	116.34	120.30
1	A	260	ARG	NE-CZ-NH1	-7.21	116.69	120.30
1	C	260	ARG	NE-CZ-NH1	-7.18	116.71	120.30
1	A	260	ARG	NE-CZ-NH2	6.96	123.78	120.30

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	2234	0	2125	20	0
1	B	2233	0	2126	35	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	2236	0	2130	25	0
1	D	2238	0	2135	25	0
2	A	6	0	8	0	0
2	B	6	0	8	6	0
3	A	20	0	8	3	0
3	B	20	0	8	1	0
3	C	20	0	8	0	0
3	D	30	0	12	2	0
4	A	64	0	0	1	0
4	B	71	0	0	3	0
4	C	88	0	0	4	0
4	D	56	0	0	1	0
All	All	9322	0	8568	102	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 6.

The worst 5 of 102 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:B:113[B]:GLN:HA	1:B:113[B]:GLN:HE21	1.12	1.11
1:B:113[B]:GLN:CA	1:B:113[B]:GLN:HE21	1.66	1.08
1:B:128:LEU:HG	2:B:1276:GOL:H11	1.38	1.01
1:B:113[B]:GLN:NE2	1:B:113[B]:GLN:HA	1.90	0.86
3:D:1276[A]:SRT:O3	3:D:1276[A]:SRT:O11	1.67	0.85

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	272/276 (99%)	261 (96%)	11 (4%)	0	100	100
1	B	272/276 (99%)	263 (97%)	9 (3%)	0	100	100
1	C	272/276 (99%)	263 (97%)	9 (3%)	0	100	100
1	D	272/276 (99%)	261 (96%)	11 (4%)	0	100	100
All	All	1088/1104 (99%)	1048 (96%)	40 (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	236/238 (99%)	230 (98%)	6 (2%)	47	52
1	B	236/238 (99%)	228 (97%)	8 (3%)	37	39
1	C	237/238 (100%)	227 (96%)	10 (4%)	30	30
1	D	237/238 (100%)	229 (97%)	8 (3%)	37	39
All	All	946/952 (99%)	914 (97%)	32 (3%)	37	39

5 of 32 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	3	ASP
1	C	38	HIS
1	D	142	LEU
1	C	23	ARG
1	C	94	THR

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	171	HIS

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

11 ligands are modelled in this entry.

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	GOL	A	1276	-	5,5,5	0.47	0	5,5,5	0.40	0
3	SRT	D	1276[A]	-	3,9,9	0.44	0	6,12,12	2.66	3 (50%)
2	GOL	B	1276	-	5,5,5	0.68	0	5,5,5	1.11	0
3	SRT	B	1278	-	3,9,9	1.11	0	6,12,12	1.67	3 (50%)
3	SRT	B	1277	-	3,9,9	0.38	0	6,12,12	1.19	1 (16%)
3	SRT	A	1277	-	3,9,9	1.17	0	6,12,12	1.01	0
3	SRT	D	1275	-	3,9,9	1.52	1 (33%)	6,12,12	1.34	1 (16%)
3	SRT	C	1275	-	3,9,9	0.91	0	6,12,12	1.52	1 (16%)
3	SRT	C	1276	-	3,9,9	0.74	0	6,12,12	1.78	1 (16%)
3	SRT	D	1276[B]	-	3,9,9	0.16	0	6,12,12	1.81	1 (16%)
3	SRT	A	1278	-	3,9,9	0.72	0	6,12,12	3.07	2 (33%)

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	GOL	A	1276	-	-	4/4/4/4	-
3	SRT	D	1276[A]	-	-	4/4/12/12	-
2	GOL	B	1276	-	-	4/4/4/4	-
3	SRT	B	1278	-	-	0/4/12/12	-
3	SRT	B	1277	-	-	0/4/12/12	-
3	SRT	A	1277	-	-	3/4/12/12	-
3	SRT	D	1275	-	-	1/4/12/12	-
3	SRT	C	1275	-	-	0/4/12/12	-
3	SRT	C	1276	-	-	2/4/12/12	-
3	SRT	D	1276[B]	-	-	4/4/12/12	-
3	SRT	A	1278	-	-	4/4/12/12	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	1275	SRT	O2-C2	2.36	1.47	1.42

The worst 5 of 13 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	A	1278	SRT	C4-C3-C2	6.09	126.21	113.11
3	D	1276[A]	SRT	C1-C2-C3	-4.37	103.70	113.11
3	A	1278	SRT	C1-C2-C3	4.08	121.89	113.11
3	D	1276[A]	SRT	C4-C3-C2	-3.83	104.86	113.11
3	C	1276	SRT	O2-C2-C1	-3.74	102.11	111.10

There are no chirality outliers.

5 of 26 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1276	GOL	O1-C1-C2-C3
3	D	1276[A]	SRT	C1-C2-C3-O3
3	D	1276[A]	SRT	O2-C2-C3-O3
3	D	1276[A]	SRT	O2-C2-C3-C4
3	A	1277	SRT	O2-C2-C3-O3

There are no ring outliers.

5 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	1276[A]	SRT	1	0
2	B	1276	GOL	6	0
3	B	1278	SRT	1	0
3	A	1277	SRT	3	0
3	D	1276[B]	SRT	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

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	274/276 (99%)	0.30	7 (2%) 56 61	27, 38, 47, 60	0
1	B	273/276 (98%)	0.26	3 (1%) 80 84	27, 38, 47, 54	0
1	C	273/276 (98%)	0.23	0 100 100	27, 38, 47, 52	0
1	D	273/276 (98%)	0.31	8 (2%) 51 57	27, 38, 47, 62	0
All	All	1093/1104 (99%)	0.27	18 (1%) 72 75	27, 38, 47, 62	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	24	ASP	4.4
1	D	92	VAL	3.3
1	A	13	ASP	3.2
1	D	2	THR	3.1
1	D	90	LEU	2.9

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 carbohydrates 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
3	SRT	A	1278	10/10	0.82	0.15	21,40,43,43	0
3	SRT	B	1277	10/10	0.90	0.14	42,44,45,46	0
2	GOL	B	1276	6/6	0.90	0.17	32,35,36,40	0
3	SRT	A	1277	10/10	0.91	0.13	38,39,43,47	0
3	SRT	B	1278	10/10	0.91	0.19	23,41,49,53	0
3	SRT	D	1276[B]	10/10	0.92	0.20	16,21,21,25	10
3	SRT	D	1276[A]	10/10	0.92	0.20	14,19,23,27	10
3	SRT	C	1275	10/10	0.93	0.13	27,40,46,47	0
3	SRT	C	1276	10/10	0.93	0.17	33,39,40,40	0
2	GOL	A	1276	6/6	0.94	0.15	28,32,34,38	0
3	SRT	D	1275	10/10	0.94	0.11	28,44,46,50	0

6.5 Other polymers (i)

There are no such residues in this entry.