Full wwPDB X-ray Structure Validation Report ⓘ

Oct 12, 2022 – 07:17 pm BST

PDB ID	:	7QNF
Title	:	CRYSTAL STRUCTURE OF E.coli ALCOHOL DEHYDROGENASE - FucO MUTANT N151G, L259V COMPLEXED WITH FE, NAD+, AND ETHYLENE GLYCOL
Authors	:	Sridhar, S.; Kiema, T.R.; Wierenga, R.K.; Widersten, M.
Deposited on	:	2021-12-20
Resolution	:	2.14 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:

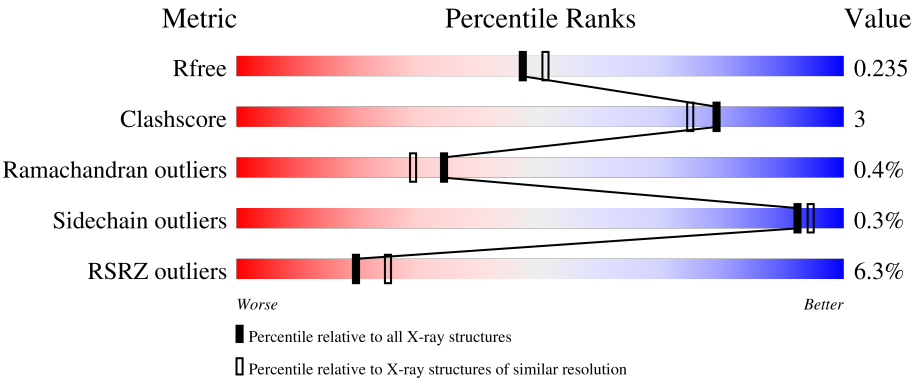
MolProbity	:	4.02b-467
Mogul	:	FAILED
Xtriage (Phenix)	:	1.13
EDS	:	2.31.2
buster-report	:	FAILED
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0267
CCP4	:	7.1.010 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.31.2

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 2.14 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	2523 (2.16-2.12)
Clashscore	141614	2653 (2.16-2.12)
Ramachandran outliers	138981	2618 (2.16-2.12)
Sidechain outliers	138945	2617 (2.16-2.12)
RSRZ outliers	127900	2485 (2.16-2.12)

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	AAA	390	3% 92% 6% ·
1	BBB	390	9% 90% 8% ··

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
3	EDO	AAA	403	-	-	-	X
3	EDO	BBB	404	-	-	-	X

2 Entry composition (i)

There are 6 unique types of molecules in this entry. The entry contains 11873 atoms, of which 5798 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 Lactaldehyde reductase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	383	Total	C	H	N	O	S	65	2	0
			5723	1803	2864	497	544	15			
1	BBB	383	Total	C	H	N	O	S	65	0	0
			5688	1793	2846	493	541	15			

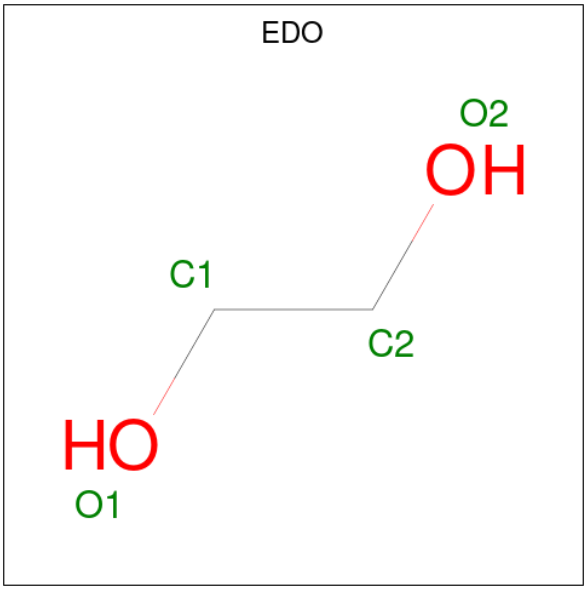
There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
AAA	1	MET	-	initiating methionine	UNP P0A9S2
AAA	151	GLY	ASN	engineered mutation	UNP P0A9S2
AAA	259	VAL	LEU	engineered mutation	UNP P0A9S2
AAA	315	GLY	SER	engineered mutation	UNP P0A9S2
AAA	384	THR	-	expression tag	UNP P0A9S2
AAA	385	SER	-	expression tag	UNP P0A9S2
AAA	386	HIS	-	expression tag	UNP P0A9S2
AAA	387	HIS	-	expression tag	UNP P0A9S2
AAA	388	HIS	-	expression tag	UNP P0A9S2
AAA	389	HIS	-	expression tag	UNP P0A9S2
AAA	390	HIS	-	expression tag	UNP P0A9S2
BBB	1	MET	-	initiating methionine	UNP P0A9S2
BBB	151	GLY	ASN	engineered mutation	UNP P0A9S2
BBB	259	VAL	LEU	engineered mutation	UNP P0A9S2
BBB	315	GLY	SER	engineered mutation	UNP P0A9S2
BBB	384	THR	-	expression tag	UNP P0A9S2
BBB	385	SER	-	expression tag	UNP P0A9S2
BBB	386	HIS	-	expression tag	UNP P0A9S2
BBB	387	HIS	-	expression tag	UNP P0A9S2
BBB	388	HIS	-	expression tag	UNP P0A9S2
BBB	389	HIS	-	expression tag	UNP P0A9S2
BBB	390	HIS	-	expression tag	UNP P0A9S2

- Molecule 2 is FE (III) ION (three-letter code: FE) (formula: Fe) (labeled as "Ligand of Interest" by depositor).

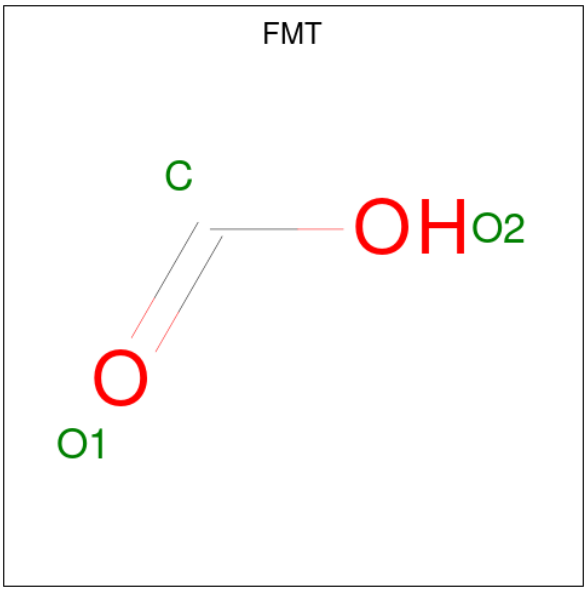
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	AAA	1	Total Fe 1 1	0	0
2	BBB	1	Total Fe 1 1	0	0

- Molecule 3 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$) (labeled as "Ligand of Interest" by depositor).



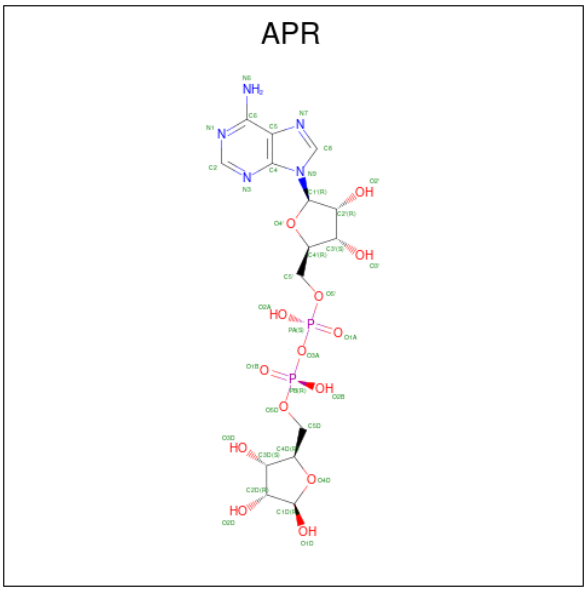
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	AAA	1	Total C H O 10 2 6 2	1	0
3	AAA	1	Total C H O 10 2 6 2	1	0
3	AAA	1	Total C H O 10 2 6 2	1	0
3	AAA	1	Total C H O 10 2 6 2	1	0
3	BBB	1	Total C H O 10 2 6 2	1	0
3	BBB	1	Total C H O 10 2 6 2	1	0
3	BBB	1	Total C H O 10 2 6 2	1	0

- Molecule 4 is FORMIC ACID (three-letter code: FMT) (formula: CH_2O_2) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	AAA	1	Total	C	H	O	1	0
			5	1	2	2		
4	AAA	1	Total	C	H	O	1	0
			5	1	2	2		

- Molecule 5 is ADENOSINE-5-DIPHOSPHORIBOSE (three-letter code: APR) (formula: $C_{15}H_{23}N_5O_{14}P_2$) (labeled as "Ligand of Interest" by depositor).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
5	AAA	1	Total	C	H	N	O	P	5	0
			57	15	21	5	14	2		

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
5	BBB	1	Total	C	H	N	O	P	5	0
			57	15	21	5	14	2		

- Molecule 6 is water.

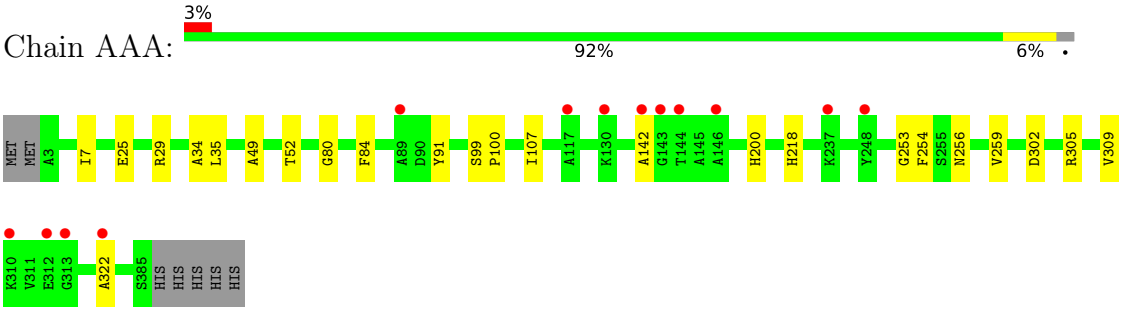
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	AAA	153	Total	O	0	0
			153	153		
6	BBB	113	Total	O	0	0
			113	113		

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.

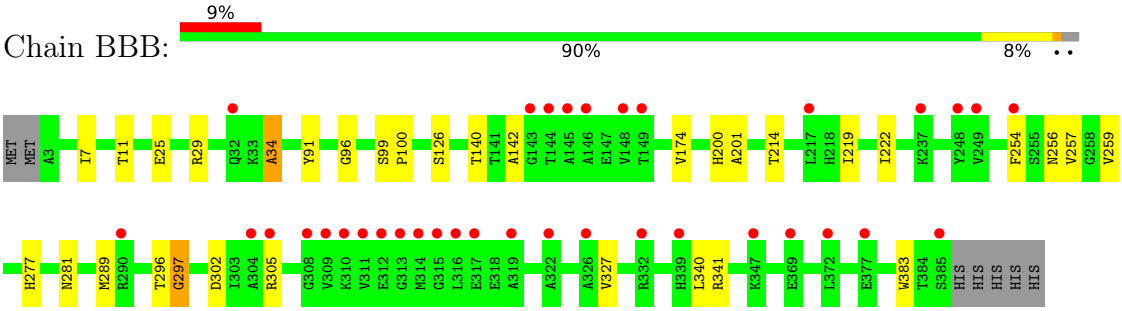
• Molecule 1: Lactaldehyde reductase

Chain AAA:



• Molecule 1: Lactaldehyde reductase

Chain BBB:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	61.84Å 85.77Å 136.68Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	36.71 – 2.14 36.68 – 2.14	Depositor EDS
% Data completeness (in resolution range)	99.9 (36.71-2.14) 100.0 (36.68-2.14)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.76 (at 2.14Å)	Xtriage
Refinement program	REFMAC 5.8.0267	Depositor
R, R_{free}	0.179 , 0.226 0.188 , 0.235	Depositor DCC
R_{free} test set	2050 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	37.7	Xtriage
Anisotropy	0.182	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	(Not available) , (Not available)	EDS
L-test for twinning[2]	$< \|L\| > = 0.51$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	11873	wwPDB-VP
Average B, all atoms (Å^2)	44.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.00% 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: APR, EDO, FE, FMT

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	AAA	0.68	0/2916	0.77	0/3968
1	BBB	0.68	0/2896	0.77	0/3942
All	All	0.68	0/5812	0.77	0/7910

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	AAA	0	1
1	BBB	0	2
All	All	0	3

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	AAA	142	ALA	Peptide
1	BBB	142	ALA	Peptide
1	BBB	34	ALA	Peptide

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	AAA	2859	2864	2854	13	0
1	BBB	2842	2846	2836	17	0
2	AAA	1	0	0	0	0
2	BBB	1	0	0	0	0
3	AAA	16	24	24	0	0
3	BBB	12	18	17	2	0
4	AAA	6	4	2	0	0
5	AAA	36	21	21	0	0
5	BBB	36	21	21	0	0
6	AAA	153	0	0	0	0
6	BBB	113	0	0	2	0
All	All	6075	5798	5775	32	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.

All (32) 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:AAA:7:ILE:H	1:AAA:256:ASN:ND2	1.92	0.67
1:BBB:7:ILE:H	1:BBB:256:ASN:HD21	1.44	0.64
1:AAA:7:ILE:H	1:AAA:256:ASN:HD21	1.46	0.64
1:AAA:34:ALA:HB2	1:AAA:91:TYR:CZ	2.35	0.62
1:BBB:340:LEU:HB2	1:BBB:383:TRP:HA	1.82	0.62
1:AAA:99:SER:HB2	1:AAA:100:PRO:HD3	1.83	0.59
1:BBB:7:ILE:H	1:BBB:256:ASN:ND2	2.03	0.56
1:AAA:35:LEU:HD22	1:AAA:84:PHE:HA	1.90	0.54
1:AAA:309:VAL:HG11	1:AAA:322:ALA:HB1	1.93	0.51
1:BBB:25:GLU:OE1	1:BBB:29:ARG:NH1	2.39	0.49
1:BBB:289:MET:SD	1:BBB:327:VAL:HG21	2.53	0.48
1:AAA:200:HIS:HB3	1:AAA:254:PHE:CD2	2.49	0.48
1:BBB:34:ALA:HB2	1:BBB:91:TYR:CZ	2.49	0.48
1:BBB:99:SER:HB2	1:BBB:100:PRO:HD3	1.96	0.46
1:BBB:201:ALA:HB1	1:BBB:222:ILE:HG13	1.97	0.45
1:BBB:11:THR:O	1:BBB:174:VAL:HA	2.16	0.45
1:BBB:96:GLY:HA2	1:BBB:140:THR:OG1	2.17	0.44
1:AAA:25:GLU:OE1	1:AAA:29:ARG:NH1	2.45	0.43
1:BBB:277:HIS:CE1	1:BBB:281:ASN:HD21	2.36	0.43
1:BBB:214:THR:HG21	1:BBB:257:VAL:HG13	1.99	0.43
3:BBB:402:EDO:H21	6:BBB:511:HOH:O	2.18	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:BBB:296:THR:O	1:BBB:297:GLY:C	2.57	0.43
1:AAA:302:ASP:OD1	1:AAA:305:ARG:NH2	2.52	0.42
1:AAA:80:GLY:HA3	1:AAA:107:ILE:HD11	2.01	0.42
1:AAA:49:ALA:HA	1:AAA:52:THR:OG1	2.20	0.42
3:BBB:402:EDO:C2	6:BBB:511:HOH:O	2.66	0.42
1:BBB:219:ILE:HD12	1:BBB:219:ILE:HA	1.88	0.42
1:BBB:200:HIS:HB3	1:BBB:254:PHE:CD2	2.55	0.42
1:BBB:214:THR:HG21	1:BBB:257:VAL:CG1	2.49	0.42
1:AAA:218:HIS:CE1	1:AAA:253:GLY:HA3	2.55	0.41
1:BBB:302:ASP:HA	1:BBB:305:ARG:NH1	2.34	0.41
1:AAA:218:HIS:HE1	1:AAA:254:PHE:N	2.20	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	AAA	383/390 (98%)	372 (97%)	10 (3%)	1 (0%)	41	36
1	BBB	381/390 (98%)	367 (96%)	12 (3%)	2 (0%)	29	22
All	All	764/780 (98%)	739 (97%)	22 (3%)	3 (0%)	34	29

All (3) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	259	VAL
1	BBB	297	GLY
1	BBB	259	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	AAA	292/297 (98%)	292 (100%)	0	100	100
1	BBB	290/297 (98%)	288 (99%)	2 (1%)	84	87
All	All	582/594 (98%)	580 (100%)	2 (0%)	92	94

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	BBB	126	SER
1	BBB	341	ARG

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

5.3.3 RNA (i)

There are no RNA molecules in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains (i)

Mogul failed to run properly - this section is therefore empty.

5.5 Carbohydrates (i)

Mogul failed to run properly - this section is therefore empty.

5.6 Ligand geometry (i)

Mogul failed to run properly - this section is therefore empty.

5.7 Other polymers (i)

Mogul failed to run properly - this section is therefore empty.

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	AAA	383/390 (98%)	0.39	13 (3%) 45 52	28, 40, 59, 76	0
1	BBB	383/390 (98%)	0.65	35 (9%) 9 12	28, 43, 73, 106	0
All	All	766/780 (98%)	0.52	48 (6%) 20 25	28, 42, 67, 106	0

All (48) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	BBB	319	ALA	8.7
1	BBB	308	GLY	8.6
1	BBB	312	GLU	5.7
1	BBB	311	VAL	5.6
1	BBB	316	LEU	4.5
1	BBB	314	MET	4.3
1	BBB	254	PHE	4.2
1	BBB	313	GLY	4.0
1	BBB	310	LYS	3.8
1	BBB	143	GLY	3.8
1	BBB	145	ALA	3.7
1	BBB	144	THR	3.6
1	BBB	309	VAL	3.5
1	AAA	89	ALA	3.4
1	AAA	312	GLU	3.3
1	BBB	317	GLU	3.1
1	BBB	305	ARG	2.9
1	BBB	149	THR	2.8
1	AAA	313	GLY	2.8
1	BBB	315	GLY	2.7
1	BBB	248	TYR	2.7
1	BBB	385	SER	2.7
1	AAA	117	ALA	2.6
1	AAA	144	THR	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	BBB	148	VAL	2.5
1	BBB	146	ALA	2.5
1	AAA	237	LYS	2.5
1	BBB	322	ALA	2.4
1	BBB	332	ARG	2.4
1	BBB	377	GLU	2.4
1	BBB	347	LYS	2.3
1	BBB	249	VAL	2.3
1	BBB	304	ALA	2.3
1	BBB	326	ALA	2.3
1	AAA	248	TYR	2.3
1	AAA	143	GLY	2.3
1	AAA	142	ALA	2.3
1	BBB	32	GLN	2.3
1	BBB	369	GLU	2.3
1	AAA	310	LYS	2.2
1	BBB	217	LEU	2.2
1	BBB	237	LYS	2.1
1	AAA	146	ALA	2.1
1	BBB	339	HIS	2.1
1	BBB	372	LEU	2.1
1	AAA	322	ALA	2.1
1	BBB	290	ARG	2.0
1	AAA	130	LYS	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	EDO	BBB	404	4/4	0.65	0.55	63,70,74,74	1
4	FMT	AAA	405	3/3	0.67	0.28	68,68,70,72	1
3	EDO	AAA	406	4/4	0.73	0.31	61,63,63,63	1
3	EDO	AAA	403	4/4	0.76	0.46	59,60,63,63	1
3	EDO	BBB	403	4/4	0.78	0.18	61,63,64,64	1
3	EDO	AAA	402	4/4	0.88	0.08	48,52,55,55	1
3	EDO	BBB	402	4/4	0.89	0.48	46,52,54,54	1
3	EDO	AAA	404	4/4	0.91	0.39	60,61,62,62	1
4	FMT	AAA	408	3/3	0.91	0.41	63,63,63,65	1
5	APR	AAA	407	36/36	0.94	0.18	32,36,41,42	5
5	APR	BBB	405	36/36	0.94	0.16	33,41,47,48	5
2	FE	AAA	401	1/1	0.99	0.17	30,30,30,30	0
2	FE	BBB	401	1/1	1.00	0.16	33,33,33,33	0

6.5 Other polymers (i)

There are no such residues in this entry.