Full wwPDB X-ray Structure Validation Report (i)

Aug 22, 2023 – 04:51 AM EDT

PDB ID : 2Q01

Title : Crystal structure of glucuronate isomerase from Caulobacter crescentus

Authors : Patskovsky, Y.; Bonanno, J.; Sridhar, V.; Sauder, J.M.; Freeman, J.; Powell, A.; Koss, J.; Groshong, C.; Gheyi, T.; Wasserman, S.R.; Raushel, F.; Burley, S.K.; Almo, S.C.; New York SGX Research Center for Structural Genomics (NYSGXRC)

Deposited on : 2007-05-18

Resolution : 2.34 Å(reported)

The following versions of software and data (see references (i)) were used in the production of this report:

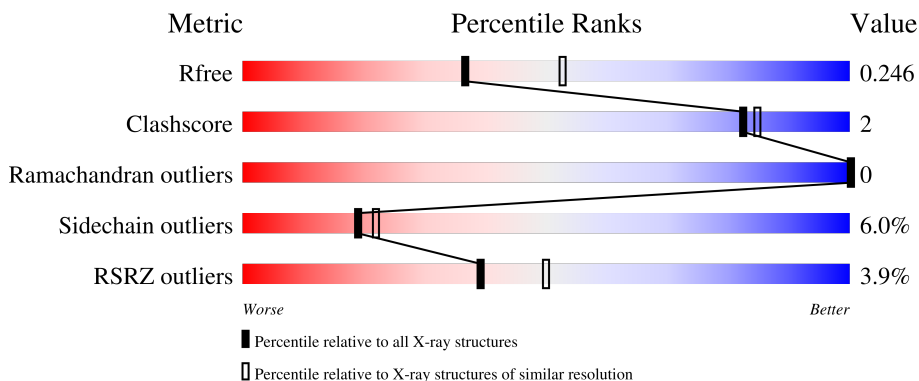|		
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.34 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	Percentile Ranks	Value
Rfree		0.246
Clashscore		2
Ramachandran outliers		0
Sidechain outliers		6.0%
RSRZ outliers		3.9%

▮ Percentile relative to all X-ray structures
▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R$_{free}$	130704	2096 (2.36-2.32)
Clashscore	141614	2193 (2.36-2.32)
Ramachandran outliers	138981	2159 (2.36-2.32)
Sidechain outliers	138945	2160 (2.36-2.32)
RSRZ outliers	127900	2067 (2.36-2.32)

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	497	4% 87% 9% · ·
1	B	497	4% 89% 7% · ·
1	C	497	3% 85% 10% · ·

2 Entry composition ⓘ

There are 3 unique types of molecules in this entry. The entry contains 12130 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 Uronate isomerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	479	Total 3887	C 2468	N 685	O 722	S 12	0	10	0
1	B	480	Total 3871	C 2459	N 680	O 721	S 11	0	7	0
1	C	480	Total 3881	C 2465	N 686	O 719	S 11	0	7	0

There are 33 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	cloning artifact	UNP Q9A874
A	0	SER	-	cloning artifact	UNP Q9A874
A	1	LEU	-	cloning artifact	UNP Q9A874
A	488	GLU	-	cloning artifact	UNP Q9A874
A	489	GLY	-	cloning artifact	UNP Q9A874
A	490	HIS	-	cloning artifact	UNP Q9A874
A	491	HIS	-	cloning artifact	UNP Q9A874
A	492	HIS	-	cloning artifact	UNP Q9A874
A	493	HIS	-	cloning artifact	UNP Q9A874
A	494	HIS	-	cloning artifact	UNP Q9A874
A	495	HIS	-	cloning artifact	UNP Q9A874
B	-1	MET	-	cloning artifact	UNP Q9A874
B	0	SER	-	cloning artifact	UNP Q9A874
B	1	LEU	-	cloning artifact	UNP Q9A874
B	488	GLU	-	cloning artifact	UNP Q9A874
B	489	GLY	-	cloning artifact	UNP Q9A874
B	490	HIS	-	cloning artifact	UNP Q9A874
B	491	HIS	-	cloning artifact	UNP Q9A874
B	492	HIS	-	cloning artifact	UNP Q9A874
B	493	HIS	-	cloning artifact	UNP Q9A874
B	494	HIS	-	cloning artifact	UNP Q9A874
B	495	HIS	-	cloning artifact	UNP Q9A874
C	-1	MET	-	cloning artifact	UNP Q9A874

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	SER	-	cloning artifact	UNP Q9A874
C	1	LEU	-	cloning artifact	UNP Q9A874
C	488	GLU	-	cloning artifact	UNP Q9A874
C	489	GLY	-	cloning artifact	UNP Q9A874
C	490	HIS	-	cloning artifact	UNP Q9A874
C	491	HIS	-	cloning artifact	UNP Q9A874
C	492	HIS	-	cloning artifact	UNP Q9A874
C	493	HIS	-	cloning artifact	UNP Q9A874
C	494	HIS	-	cloning artifact	UNP Q9A874
C	495	HIS	-	cloning artifact	UNP Q9A874

• Molecule 2 is POTASSIUM ION (three-letter code: K) (formula: K).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	A	1	Total 1	K 1	0	0
2	B	1	Total 1	K 1	0	0
2	C	1	Total 1	K 1	0	0

• Molecule 3 is water.

| Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	151	Total 151	O 151	0	0
3	B	188	Total 188	O 188	0	0
3	C	149	Total 149	O 149	0	0

3 Residue-property plots ⓘ

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.
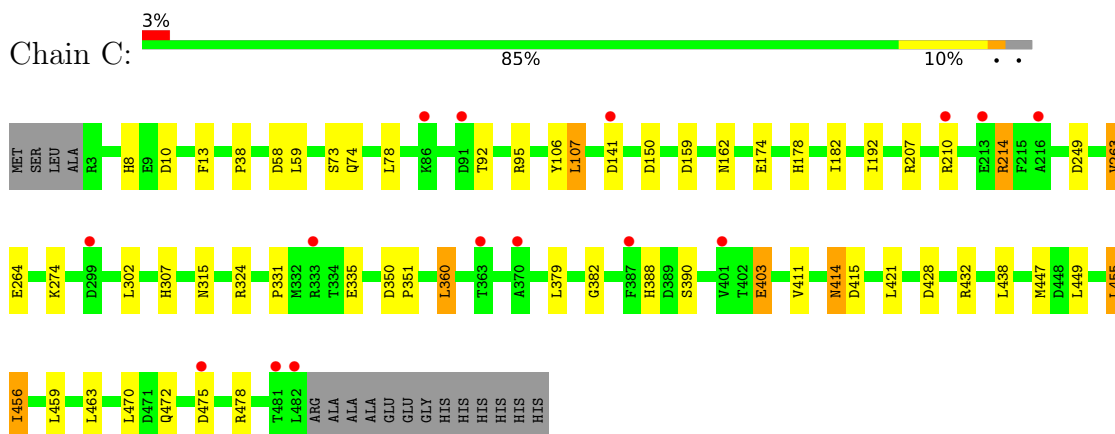
- Molecule 1: Uronate isomerase
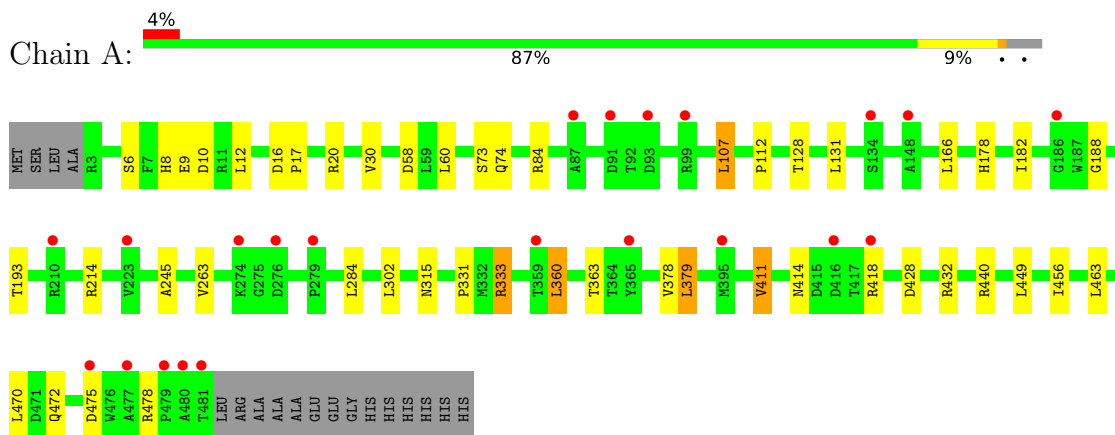
- Molecule 1: Uronate isomerase
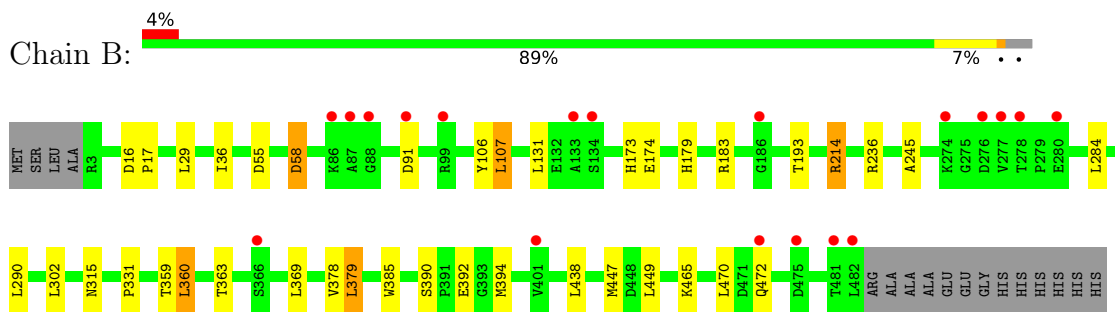
- Molecule 1: Uronate isomerase

4 Data and refinement statistics ⓘ

Property	Value	Source
Space group	F 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	177.33Å 190.18Å 319.29Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.34 46.44 – 2.34	Depositor EDS
% Data completeness (in resolution range)	91.7 (20.00-2.34) 91.7 (46.44-2.34)	Depositor EDS
R_{merge}	0.19	Depositor
R_{sym}	0.18	Depositor
$< I/\sigma(I) >$ [1]	1.24 (at 2.34Å)	Xtriage
Refinement program	REFMAC 5.3.0034	Depositor
R, R_{free}	0.198 , 0.251 0.197 , 0.246	Depositor DCC
R_{free} test set	3040 reflections (2.95%)	wwPDB-VP
Wilson B-factor (Å^2)	46.0	Xtriage
Anisotropy	0.231	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 51.2	EDS
L-test for twinning[2]	$< |L| > = 0.48, < L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	12130	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.67% 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: K

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	0.36	0/4029	0.58	0/5478
1	B	0.36	0/4003	0.56	0/5445
1	C	0.36	0/4013	0.58	0/5457
All	All	0.36	0/12045	0.58	0/16380

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	3887	0	3744	20	0
1	B	3871	0	3724	15	0
1	C	3881	0	3740	24	0
2	A	1	0	0	0	0
2	B	1	0	0	0	0
2	C	1	0	0	0	0
3	A	151	0	0	0	0
3	B	188	0	0	0	0
3	C	149	0	0	0	0
All	All	12130	0	11208	56	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 (56) 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:8:HIS:HD2	1:A:10[B]:ASP:H	1.36	0.72
1:A:8:HIS:HD2	1:A:10[A]:ASP:H	1.37	0.72
1:C:8:HIS:HD2	1:C:10[B]:ASP:H	1.48	0.61
1:C:8:HIS:HD2	1:C:10[A]:ASP:H	1.49	0.60
1:A:411:VAL:O	1:A:411:VAL:HG13	2.01	0.60
1:C:331:PRO:HB3	1:C:360:LEU:HB3	1.85	0.59
1:B:179:HIS:O	1:B:183:ARG:HG3	2.02	0.58
1:A:9:GLU:O	1:A:432:ARG:NH2	2.38	0.56
1:B:173:HIS:HB2	1:B:214:ARG:HG3	1.88	0.56
1:A:74:GLN:HG3	1:A:107:LEU:HD21	1.89	0.55
1:A:12:LEU:HD12	1:A:432:ARG:HD3	1.92	0.51
1:C:207:ARG:HE	1:C:324:ARG:HD2	1.76	0.50
1:C:159:ASP:OD2	1:C:478:ARG:NH2	2.45	0.50
1:B:331:PRO:HB3	1:B:360:LEU:HB3	1.95	0.49
1:A:331:PRO:HB3	1:A:360:LEU:HB3	1.95	0.47
1:A:363:THR:HB	1:B:363:THR:HG21	1.96	0.47
1:B:378:VAL:HG23	1:B:379:LEU:HD13	1.96	0.47
1:C:106:TYR:CE2	1:C:107:LEU:CD1	2.97	0.47
1:C:38:PRO:HD2	1:C:414:ASN:HB3	1.97	0.47
1:B:390:SER:O	1:B:394:MET:HG3	2.15	0.47
1:C:249:ASP:OD1	1:C:307:HIS:ND1	2.45	0.47
1:C:428:ASP:O	1:C:432[A]:ARG:HG3	2.16	0.46
1:A:178:HIS:O	1:A:182:ILE:HG12	2.16	0.46
1:B:16:ASP:HA	1:B:17:PRO:HD3	1.79	0.46
1:B:36:ILE:HD13	1:B:465:LYS:HG3	1.98	0.46
1:C:95:ARG:NH1	1:C:141:ASP:OD1	2.48	0.46
1:A:16:ASP:HA	1:A:17:PRO:HD3	1.76	0.45
1:A:74:GLN:HG3	1:A:107:LEU:CD2	2.46	0.45
1:C:59:LEU:HD11	1:C:421:LEU:HD13	1.99	0.45
1:C:263:VAL:CG1	1:C:264:GLU:N	2.79	0.45
1:B:236:ARG:HA	1:B:236:ARG:HD2	1.78	0.45
1:A:378:VAL:HG23	1:A:379:LEU:HD13	1.99	0.44
1:B:360:LEU:HD13	1:B:385:TRP:CE3	2.53	0.44
1:C:106:TYR:CE2	1:C:107:LEU:HD13	2.52	0.43
1:A:188:GLY:HA3	1:A:478:ARG:HG3	2.00	0.43
1:C:382:GLY:HA2	1:C:411:VAL:HG22	2.01	0.43
1:A:112:PRO:HD3	1:C:403:GLU:HG3	1.99	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:447:MET:HE1	1:C:455:LEU:HD23	2.00	0.43
1:A:428:ASP:O	1:A:432:ARG:HG3	2.19	0.42
1:B:290:LEU:HD23	1:B:290:LEU:HA	1.95	0.42
1:C:74:GLN:HG3	1:C:107:LEU:HD21	2.01	0.42
1:C:178:HIS:O	1:C:182:ILE:HG12	2.19	0.42
1:C:350:ASP:HA	1:C:351:PRO:HD3	1.91	0.41
1:B:359:THR:HB	1:B:369:LEU:HD11	2.03	0.41
1:A:193:THR:HG23	1:A:245:ALA:HB2	2.02	0.41
1:A:333[A]:ARG:HH11	1:C:335:GLU:HG3	1.86	0.41
1:B:55:ASP:HB2	1:B:58:ASP:H	1.86	0.41
1:C:210:ARG:O	1:C:214:ARG:HG2	2.21	0.41
1:B:193:THR:HG23	1:B:245:ALA:HB2	2.03	0.41
1:A:10[B]:ASP:OD2	1:A:20:ARG:NH1	2.54	0.41
1:C:456:ILE:H	1:C:456:ILE:HG13	1.61	0.41
1:C:388:HIS:C	1:C:390:SER:H	2.24	0.40
1:B:106:TYR:CE2	1:B:107:LEU:HD13	2.56	0.40
1:C:10[B]:ASP:HB2	1:C:13:PHE:HD2	1.87	0.40

There are no symmetry-related clashes.

5.3 Torsion angles ⓘ

5.3.1 Protein backbone ⓘ

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	487/497 (98%)	471 (97%)	16 (3%)	0	100	100
1	B	485/497 (98%)	472 (97%)	13 (3%)	0	100	100
1	C	485/497 (98%)	468 (96%)	17 (4%)	0	100	100
All	All	1457/1491 (98%)	1411 (97%)	46 (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	412/415 (99%)	383 (93%)	29 (7%)	15	16
1	B	410/415 (99%)	392 (96%)	18 (4%)	28	35
1	C	410/415 (99%)	382 (93%)	28 (7%)	16	17
All	All	1232/1245 (99%)	1157 (94%)	75 (6%)	19	21

All (75) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	30	VAL
1	A	58	ASP
1	A	60	LEU
1	A	73	SER
1	A	84	ARG
1	A	107	LEU
1	A	128	THR
1	A	131	LEU
1	A	166	LEU
1	A	214	ARG
1	A	263	VAL
1	A	284	LEU
1	A	302	LEU
1	A	315	ASN
1	A	333[A]	ARG
1	A	333[B]	ARG
1	A	360	LEU
1	A	379	LEU
1	A	411	VAL
1	A	414	ASN
1	A	418[A]	ARG
1	A	418[B]	ARG
1	A	440	ARG
1	A	449	LEU
1	A	456	ILE
1	A	463	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	470	LEU
1	A	472	GLN
1	A	475	ASP
1	B	29	LEU
1	B	58	ASP
1	B	91	ASP
1	B	107	LEU
1	B	131	LEU
1	B	174	GLU
1	B	214	ARG
1	B	284	LEU
1	B	302	LEU
1	B	315	ASN
1	B	360	LEU
1	B	379	LEU
1	B	392	GLU
1	B	438	LEU
1	B	447	MET
1	B	449	LEU
1	B	470	LEU
1	B	472	GLN
1	C	58	ASP
1	C	73	SER
1	C	78	LEU
1	C	92	THR
1	C	107	LEU
1	C	150	ASP
1	C	162	ASN
1	C	174	GLU
1	C	192	ILE
1	C	214	ARG
1	C	263	VAL
1	C	274	LYS
1	C	302	LEU
1	C	315	ASN
1	C	360	LEU
1	C	379	LEU
1	C	403	GLU
1	C	414	ASN
1	C	415	ASP
1	C	438	LEU
1	C	449	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	C	455	LEU
1	C	456	ILE
1	C	459	LEU
1	C	463	LEU
1	C	470	LEU
1	C	472	GLN
1	C	475	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (27) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	8	HIS
1	A	162	ASN
1	A	190	HIS
1	A	270	ASN
1	A	305	GLN
1	A	311	HIS
1	A	315	ASN
1	A	349	ASN
1	A	400	GLN
1	A	414	ASN
1	B	162	ASN
1	B	177	GLN
1	B	270	ASN
1	B	311	HIS
1	B	315	ASN
1	B	349	ASN
1	B	400	GLN
1	B	414	ASN
1	C	8	HIS
1	C	162	ASN
1	C	270	ASN
1	C	305	GLN
1	C	311	HIS
1	C	315	ASN
1	C	349	ASN
1	C	400	GLN
1	C	414	ASN

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 3 ligands modelled in this entry, 3 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, 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(Å²)	Q<0.9
1	A	479/497 (96%)	0.27	22 (4%)	32	43	29, 52, 82, 108	0
1	B	480/497 (96%)	0.11	19 (3%)	38	49	28, 46, 74, 109	0
1	C	480/497 (96%)	0.15	15 (3%)	49	59	31, 47, 75, 119	0
All	All	1439/1491 (96%)	0.18	56 (3%)	39	50	28, 48, 77, 119	0

All (56) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	91	ASP	5.7
1	B	482	LEU	4.7
1	A	186	GLY	4.3
1	A	87	ALA	4.2
1	C	91	ASP	3.9
1	A	276[A]	ASP	3.8
1	B	481	THR	3.6
1	A	481	THR	3.6
1	B	276[A]	ASP	3.5
1	A	480	ALA	3.3
1	C	481	THR	3.1
1	A	479	PRO	3.1
1	A	475	ASP	3.1
1	C	482	LEU	3.0
1	B	91	ASP	2.9
1	A	134	SER	2.8
1	B	475	ASP	2.6
1	C	216	ALA	2.5
1	C	333[A]	ARG	2.5
1	B	274	LYS	2.5
1	B	277	VAL	2.5
1	A	210	ARG	2.5
1	C	363	THR	2.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	133	ALA	2.4
1	A	274	LYS	2.4
1	A	148	ALA	2.4
1	B	278	THR	2.4
1	C	213	GLU	2.3
1	C	86	LYS	2.3
1	A	99	ARG	2.3
1	A	418[A]	ARG	2.3
1	B	401	VAL	2.3
1	C	370	ALA	2.3
1	A	223	VAL	2.2
1	A	359	THR	2.2
1	B	472	GLN	2.2
1	A	365	TYR	2.2
1	B	87	ALA	2.2
1	B	134	SER	2.2
1	B	186	GLY	2.2
1	A	279	PRO	2.2
1	A	93	ASP	2.2
1	C	299	ASP	2.1
1	B	88	GLY	2.1
1	C	401	VAL	2.1
1	B	86	LYS	2.1
1	A	395	MET	2.1
1	C	387	PHE	2.1
1	B	99	ARG	2.1
1	B	366	SER	2.1
1	B	280	GLU	2.1
1	A	416	ASP	2.1
1	A	477	ALA	2.1
1	C	141	ASP	2.0
1	C	475	ASP	2.0
1	C	210	ARG	2.0

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, 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	B-factors(Å^2)	Q<0.9
2	K	B	502	1/1	0.98	0.10	37,37,37,37	0
2	K	C	503	1/1	0.98	0.09	36,36,36,36	0
2	K	A	501	1/1	0.99	0.08	35,35,35,35	0

6.5 Other polymers ⓘ

There are no such residues in this entry.