Full wwPDB X-ray Structure Validation Report (i)

Feb 10, 2024 – 08:29 AM EST

PDB ID	:	2RTH
Title	:	STREPTAVIDIN-GLYCOLURIL COMPLEX, PH 2.50, SPACE GROUP I222
Authors	:	Katz, B.A.
Deposited on	:	1997-09-11
Resolution	:	1.56 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 (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 (i)) were used in the production of this report:

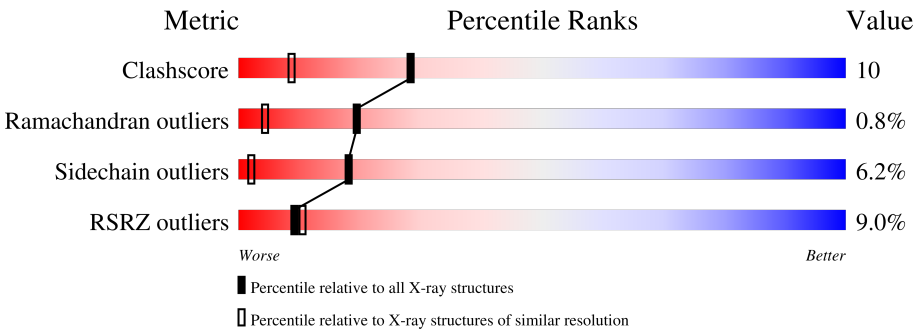
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 1.56 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	1529 (1.56-1.56)
Ramachandran outliers	138981	1498 (1.56-1.56)
Sidechain outliers	138945	1495 (1.56-1.56)
RSRZ outliers	127900	1465 (1.56-1.56)

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	B	135	7%; 60%; 26%; 5%; 9%
1	D	135	10%; 67%; 18%; •; 10%

2 Entry composition (i)

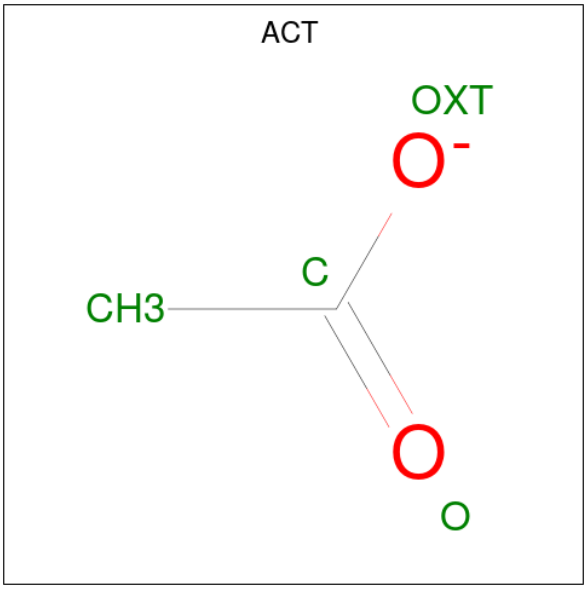
There are 4 unique types of molecules in this entry. The entry contains 4246 atoms, of which 2109 are hydrogens and 0 are deuteriums.

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 STREPTAVIDIN.

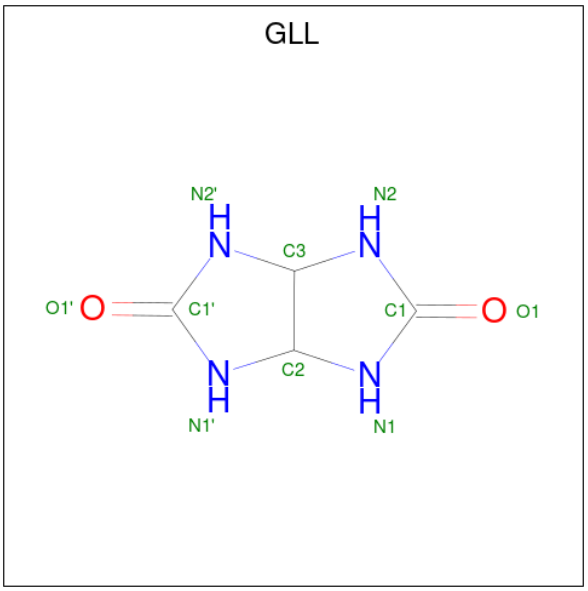
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	H	N	O			
1	B	123	1954	631	949	169	205	66	14	0
1	D	121	1899	612	916	167	204	97	14	0

- Molecule 2 is ACETATE ION (three-letter code: ACT) (formula: $C_2H_3O_2$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	H	O		
2	B	1	14	4	6	4	0	1
2	D	1	14	4	6	4	0	1

- Molecule 3 is GLYCOLURIL (three-letter code: GLL) (formula: $C_4H_6N_4O_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	B	1	Total	C	H	N	O	0	0
			16	4	6	4	2		
3	D	1	Total	C	H	N	O	0	0
			16	4	6	4	2		

- Molecule 4 is water.

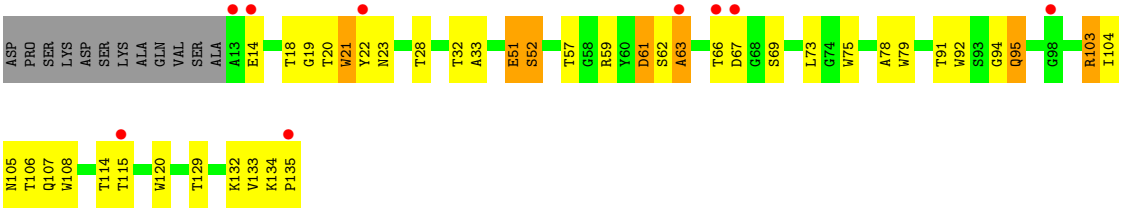
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	B	54	Total	H	O	0	2
			158	104	54		
4	D	59	Total	H	O	0	2
			175	116	59		

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.

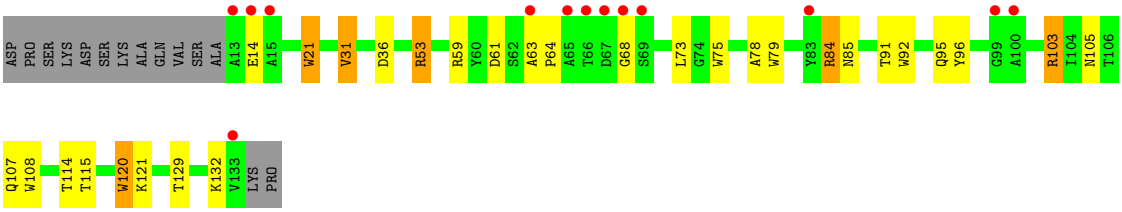
- Molecule 1: STREPTAVIDIN

Chain B:



- Molecule 1: STREPTAVIDIN

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, α, β, γ	94.70Å 104.77Å 47.77Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	7.50 – 1.56 28.19 – 1.36	Depositor EDS
% Data completeness (in resolution range)	64.0 (7.50-1.56) 47.1 (28.19-1.36)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.26 (at 1.36Å)	Xtriage
Refinement program	X-PLOR	Depositor
R, R_{free}	0.203 , 0.269 0.210 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å^2)	13.0	Xtriage
Anisotropy	0.201	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.41 , 121.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.31$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	4246	wwPDB-VP
Average B, all atoms (Å^2)	30.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 10.27% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: GLL, ACT

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	B	1.39	1/1047 (0.1%)	1.77	31/1434 (2.2%)
1	D	1.36	2/1019 (0.2%)	1.81	31/1396 (2.2%)
All	All	1.37	3/2066 (0.1%)	1.79	62/2830 (2.2%)

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	B	0	3
1	D	0	4
All	All	0	7

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	75	TRP	CG-CD2	-5.60	1.34	1.43
1	B	75	TRP	CG-CD2	-5.49	1.34	1.43
1	D	21	TRP	CG-CD2	-5.00	1.35	1.43

All (62) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	92	TRP	CD1-NE1-CE2	10.56	118.51	109.00
1	B	92	TRP	CD1-NE1-CE2	10.17	118.15	109.00
1	B	75	TRP	CD1-NE1-CE2	9.70	117.73	109.00
1	B	21	TRP	CD1-NE1-CE2	9.64	117.68	109.00
1	B	108	TRP	CD1-NE1-CE2	9.60	117.64	109.00
1	D	21	TRP	CD1-NE1-CE2	9.46	117.52	109.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	75	TRP	CD1-NE1-CE2	9.46	117.52	109.00
1	D	79	TRP	CD1-NE1-CE2	9.31	117.38	109.00
1	D	120	TRP	CD1-NE1-CE2	9.31	117.38	109.00
1	D	108	TRP	CD1-NE1-CE2	9.20	117.28	109.00
1	D	79	TRP	NE1-CE2-CZ2	8.97	140.27	130.40
1	B	120	TRP	CD1-NE1-CE2	8.95	117.06	109.00
1	D	96	TYR	CB-CG-CD1	-8.78	115.73	121.00
1	B	79	TRP	CD1-NE1-CE2	8.73	116.86	109.00
1	D	92	TRP	CG-CD1-NE1	-8.23	101.87	110.10
1	B	75	TRP	CG-CD1-NE1	-7.96	102.14	110.10
1	D	75	TRP	NE1-CE2-CZ2	7.94	139.13	130.40
1	B	120	TRP	NE1-CE2-CZ2	7.92	139.11	130.40
1	B	92	TRP	CG-CD1-NE1	-7.88	102.22	110.10
1	D	75	TRP	CG-CD1-NE1	-7.68	102.42	110.10
1	B	108	TRP	CG-CD1-NE1	-7.64	102.46	110.10
1	D	21	TRP	CG-CD1-NE1	-7.48	102.62	110.10
1	D	21	TRP	NE1-CE2-CZ2	7.46	138.60	130.40
1	B	120	TRP	CG-CD1-NE1	-7.34	102.76	110.10
1	B	92	TRP	NE1-CE2-CZ2	7.30	138.43	130.40
1	D	108	TRP	CG-CD1-NE1	-7.24	102.86	110.10
1	B	21	TRP	CG-CD1-NE1	-7.21	102.89	110.10
1	D	78	ALA	N-CA-C	-7.15	91.70	111.00
1	D	92	TRP	NE1-CE2-CZ2	7.00	138.09	130.40
1	D	120	TRP	CG-CD1-NE1	-6.99	103.11	110.10
1	B	21	TRP	NE1-CE2-CZ2	6.94	138.04	130.40
1	B	78	ALA	N-CA-C	-6.87	92.46	111.00
1	B	75	TRP	NE1-CE2-CZ2	6.82	137.90	130.40
1	D	79	TRP	NE1-CE2-CD2	-6.71	100.59	107.30
1	B	79	TRP	CG-CD1-NE1	-6.69	103.41	110.10
1	D	79	TRP	CG-CD1-NE1	-6.67	103.44	110.10
1	D	75	TRP	NE1-CE2-CD2	-6.50	100.81	107.30
1	B	79	TRP	NE1-CE2-CZ2	6.34	137.37	130.40
1	B	92	TRP	NE1-CE2-CD2	-6.29	101.01	107.30
1	B	108	TRP	NE1-CE2-CZ2	6.19	137.21	130.40
1	B	21	TRP	NE1-CE2-CD2	-6.17	101.14	107.30
1	D	31	VAL	CB-CA-C	-6.12	99.78	111.40
1	D	92	TRP	NE1-CE2-CD2	-6.03	101.28	107.30
1	D	108	TRP	NE1-CE2-CZ2	5.92	136.91	130.40
1	D	21	TRP	NE1-CE2-CD2	-5.90	101.40	107.30
1	B	75	TRP	NE1-CE2-CD2	-5.86	101.44	107.30
1	B	106	THR	CA-CB-CG2	5.77	120.47	112.40
1	D	108	TRP	NE1-CE2-CD2	-5.70	101.60	107.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	120	TRP	NE1-CE2-CD2	-5.69	101.61	107.30
1	D	96	TYR	CB-CG-CD2	5.61	124.37	121.00
1	D	79	TRP	CG-CD2-CE3	-5.57	128.89	133.90
1	B	61[A]	ASP	CB-CG-OD2	5.43	123.19	118.30
1	B	61[B]	ASP	CB-CG-OD2	5.43	123.19	118.30
1	B	104	ILE	N-CA-C	-5.42	96.37	111.00
1	D	120	TRP	NE1-CE2-CD2	-5.40	101.90	107.30
1	B	103	ARG	NE-CZ-NH2	-5.38	117.61	120.30
1	B	107	GLN	N-CA-C	-5.37	96.50	111.00
1	D	84	ARG	NE-CZ-NH2	-5.32	117.64	120.30
1	B	108	TRP	NE1-CE2-CD2	-5.28	102.02	107.30
1	D	120	TRP	NE1-CE2-CZ2	5.18	136.09	130.40
1	D	115	THR	N-CA-C	-5.01	97.46	111.00
1	B	79	TRP	NE1-CE2-CD2	-5.01	102.29	107.30

There are no chirality outliers.

All (7) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	103	ARG	Sidechain
1	B	59	ARG	Sidechain
1	B	63[A]	ALA	Peptide
1	D	103	ARG	Sidechain
1	D	53	ARG	Sidechain
1	D	63[A]	ALA	Peptide
1	D	84	ARG	Sidechain

CLOSE-CONTACTS INFOmissingINFO

5.2 Torsion angles (i)

5.2.1 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	B	135/135 (100%)	127 (94%)	7 (5%)	1 (1%)	22 5
1	D	132/135 (98%)	127 (96%)	4 (3%)	1 (1%)	19 4
All	All	267/270 (99%)	254 (95%)	11 (4%)	2 (1%)	19 5

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	14[A]	GLU
1	B	52	SER

5.2.2 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	B	101/100 (101%)	93 (92%)	8 (8%)	12 1
1	D	98/100 (98%)	91 (93%)	7 (7%)	14 1
All	All	199/200 (100%)	184 (92%)	15 (8%)	18 1

All (15) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	51	GLU
1	B	62[A]	SER
1	B	62[B]	SER
1	B	66[A]	THR
1	B	66[B]	THR
1	B	73[A]	LEU
1	B	73[B]	LEU
1	B	95	GLN
1	D	31	VAL
1	D	36	ASP
1	D	53	ARG
1	D	73	LEU
1	D	103	ARG
1	D	107[A]	GLN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	107[B]	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	95	GLN
1	D	95	GLN

5.2.3 RNA (i)

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.4 Carbohydrates (i)

There are no monosaccharides in this entry.

5.5 Ligand geometry (i)

6 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	$\#\|Z\| > 2$
3	GLL	B	300	-	11,11,11	3.61	8 (72%)	16,16,16	3.08	5 (31%)
3	GLL	D	300	-	11,11,11	3.51	8 (72%)	16,16,16	2.60	5 (31%)
2	ACT	D	500[B]	-	3,3,3	1.38	0	3,3,3	0.74	0
2	ACT	B	501[B]	-	3,3,3	1.17	0	3,3,3	1.10	0
2	ACT	D	500[A]	-	3,3,3	1.18	0	3,3,3	1.37	0
2	ACT	B	501[A]	-	3,3,3	1.23	0	3,3,3	1.41	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
3	GLL	D	300	-	-	-	0/2/2/2
3	GLL	B	300	-	-	-	0/2/2/2

All (16) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	D	300	GLL	C3-N2'	6.02	1.51	1.44
3	B	300	GLL	C2-N1	5.41	1.51	1.44
3	B	300	GLL	C3-N2'	4.82	1.50	1.44
3	B	300	GLL	C3-N2	4.45	1.49	1.44
3	D	300	GLL	C2-N1'	4.41	1.49	1.44
3	B	300	GLL	C2-N1'	4.29	1.49	1.44
3	D	300	GLL	C2-N1	4.11	1.49	1.44
3	B	300	GLL	C1-N2	-3.91	1.28	1.35
3	D	300	GLL	C3-N2	3.87	1.49	1.44
3	D	300	GLL	C1-N1	-3.70	1.29	1.35
3	B	300	GLL	C1-N1	-3.65	1.29	1.35
3	D	300	GLL	C1'-N1'	-3.34	1.29	1.35
3	D	300	GLL	C1'-N2'	-3.25	1.29	1.35
3	B	300	GLL	C1'-N2'	-3.12	1.30	1.35
3	B	300	GLL	C1'-N1'	-3.09	1.30	1.35
3	D	300	GLL	C1-N2	-2.97	1.30	1.35

All (10) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	300	GLL	N2'-C3-N2	6.38	119.29	114.38
3	D	300	GLL	C2-C3-N2	-6.08	100.05	102.75
3	B	300	GLL	N1'-C2-N1	5.94	118.95	114.38
3	B	300	GLL	C2-C3-N2	-5.42	100.35	102.75
3	D	300	GLL	N2'-C3-N2	4.86	118.12	114.38
3	D	300	GLL	C3-C2-N1'	-4.67	100.68	102.75
3	B	300	GLL	C3-C2-N1'	-4.50	100.75	102.75
3	B	300	GLL	O1-C1-N2	-2.89	121.79	125.94
3	D	300	GLL	C2-C3-N2'	-2.86	101.48	102.75
3	D	300	GLL	C2-N1'-C1'	2.49	114.57	112.89

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

5.6 Other polymers (i)

There are no such residues in this entry.

5.7 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	B	123/135 (91%)	0.13	9 (7%) 15 17	11, 24, 48, 54	19 (15%)
1	D	121/135 (89%)	0.52	13 (10%) 6 6	12, 27, 51, 66	21 (17%)
All	All	244/270 (90%)	0.32	22 (9%) 9 10	11, 25, 50, 66	40 (16%)

All (22) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	66[A]	THR	9.4
1	D	68[A]	GLY	8.7
1	B	135	PRO	8.0
1	B	14	GLU	7.4
1	D	14[A]	GLU	6.0
1	B	22[A]	TYR	5.5
1	D	63[A]	ALA	4.9
1	B	66[A]	THR	4.5
1	D	15	ALA	4.3
1	D	100	ALA	3.9
1	D	67[A]	ASP	3.9
1	D	83	TYR	3.4
1	D	65[A]	ALA	3.1
1	B	67[A]	ASP	3.0
1	D	99	GLY	2.9
1	D	133	VAL	2.7
1	B	13	ALA	2.6
1	B	115	THR	2.4
1	B	98	GLY	2.3
1	B	63[A]	ALA	2.1
1	D	69[A]	SER	2.0
1	D	13	ALA	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, 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	ACT	B	501[A]	4/4	0.69	0.28	51,55,57,57	7
2	ACT	B	501[B]	4/4	0.69	0.28	40,49,54,54	7
2	ACT	D	500[A]	4/4	0.84	0.38	55,57,58,58	7
2	ACT	D	500[B]	4/4	0.84	0.38	49,52,54,54	7
3	GLL	B	300	10/10	0.96	0.06	17,21,26,27	0
3	GLL	D	300	10/10	0.97	0.05	19,23,28,30	0

6.5 Other polymers (i)

There are no such residues in this entry.