



wwPDB X-ray Structure Validation Summary Report ⓘ

May 24, 2020 – 08:19 am BST

PDB ID : 3BN1
Title : Crystal structure of GDP-perosamine synthase
Authors : Cook, P.D.; Holden, H.M.
Deposited on : 2007-12-13
Resolution : 1.80 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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

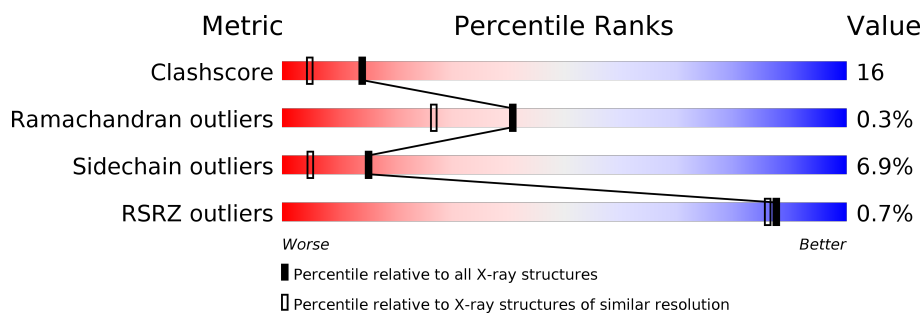
1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 1.80 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	6793 (1.80-1.80)
Ramachandran outliers	138981	6697 (1.80-1.80)
Sidechain outliers	138945	6696 (1.80-1.80)
RSRZ outliers	127900	5850 (1.80-1.80)

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 $\leq 5\%$. 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	373	<div> <div>73%</div> <div>23%</div> <div>..</div> </div>
1	B	373	<div> <div>77%</div> <div>20%</div> <div>..</div> </div>
1	C	373	<div> <div>%</div> <div>66%</div> <div>26%</div> <div>5%</div> <div>.</div> </div>
1	D	373	<div> <div>%</div> <div>65%</div> <div>26%</div> <div>6%</div> <div>.</div> </div>

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	ACT	C	372	-	-	X	-
2	ACT	D	372	-	-	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 12420 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 Perosamine synthetase.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	368	Total	C	N	O	P	S	0	1	0
			2852	1800	501	532	1	18			
1	B	368	Total	C	N	O	P	S	0	2	0
			2857	1806	501	531	1	18			
1	C	365	Total	C	N	O	P	S	0	2	0
			2830	1789	495	527	1	18			
1	D	364	Total	C	N	O	P	S	0	1	0
			2816	1777	493	526	1	19			

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	GLY	-	EXPRESSION TAG	UNP Q9A9H3
A	0	HIS	-	EXPRESSION TAG	UNP Q9A9H3
A	1	VAL	-	EXPRESSION TAG	UNP Q9A9H3
B	-1	GLY	-	EXPRESSION TAG	UNP Q9A9H3
B	0	HIS	-	EXPRESSION TAG	UNP Q9A9H3
B	1	VAL	-	EXPRESSION TAG	UNP Q9A9H3
C	-1	GLY	-	EXPRESSION TAG	UNP Q9A9H3
C	0	HIS	-	EXPRESSION TAG	UNP Q9A9H3
C	1	VAL	-	EXPRESSION TAG	UNP Q9A9H3
D	-1	GLY	-	EXPRESSION TAG	UNP Q9A9H3
D	0	HIS	-	EXPRESSION TAG	UNP Q9A9H3
D	1	VAL	-	EXPRESSION TAG	UNP Q9A9H3

- Molecule 2 is ACETATE ION (three-letter code: ACT) (formula: C₂H₃O₂).

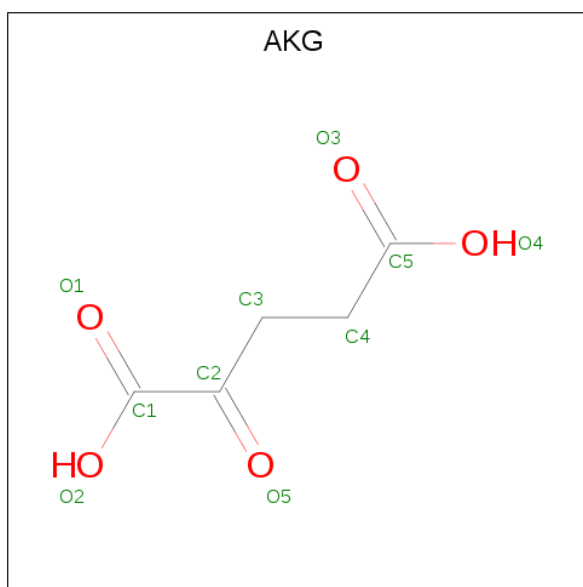


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			4	2	2		
2	C	1	Total	C	O	0	0
			4	2	2		
2	D	1	Total	C	O	0	0
			4	2	2		

- Molecule 3 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total	Na	0	0
			1	1		

- Molecule 4 is 2-OXOGLUTARIC ACID (three-letter code: AKG) (formula: C₅H₆O₅).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			10	5	5		

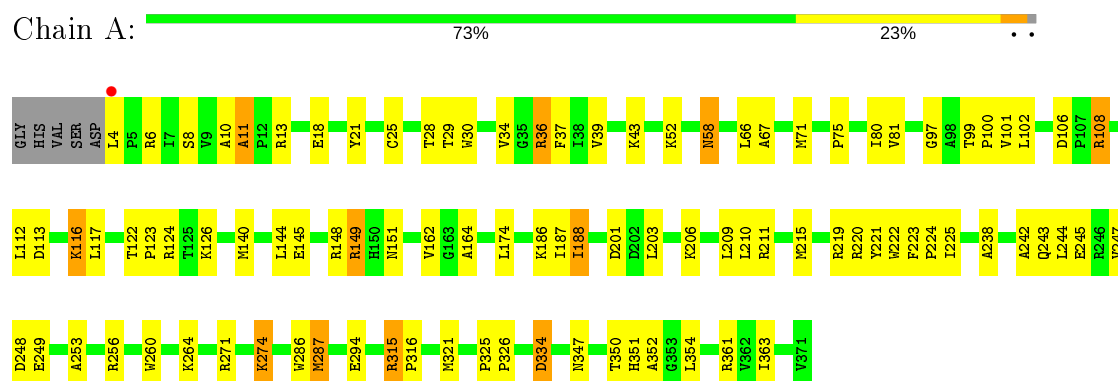
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	304	Total	O	0	0
			304	304		
5	B	284	Total	O	0	0
			284	284		
5	C	238	Total	O	0	0
			238	238		
5	D	216	Total	O	0	0
			216	216		

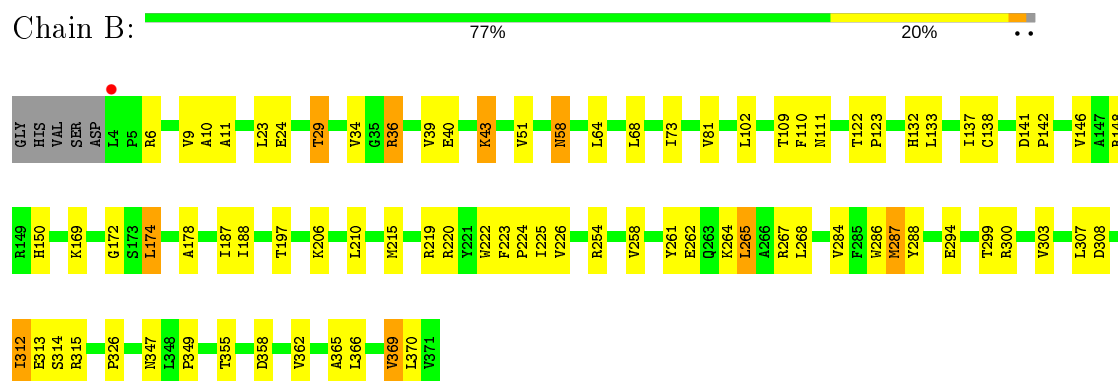
3 Residue-property plots

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.

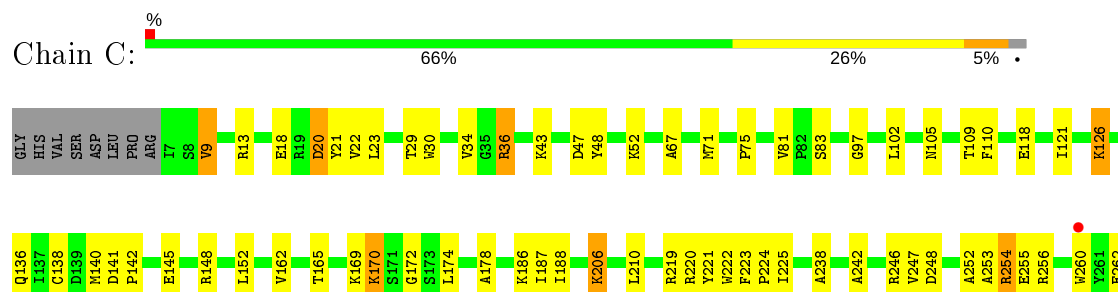
• Molecule 1: Perosamine synthetase

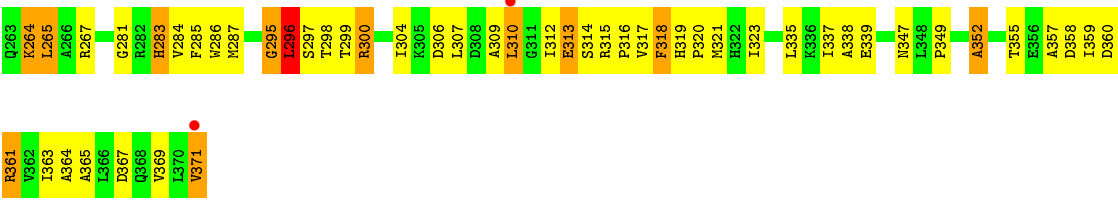


• Molecule 1: Perosamine synthetase

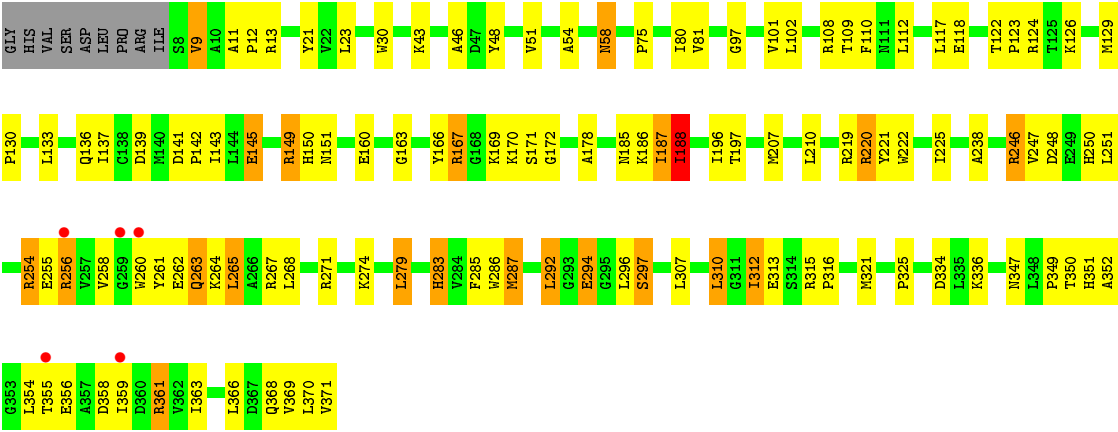


• Molecule 1: Perosamine synthetase





● Molecule 1: Perosamine synthetase



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	50.27Å 152.93Å 105.73Å 90.00° 102.09° 90.00°	Depositor
Resolution (Å)	30.00 – 1.80 35.48 – 1.80	Depositor EDS
% Data completeness (in resolution range)	98.5 (30.00-1.80) 98.5 (35.48-1.80)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.08	Depositor
$\langle I/\sigma(I) \rangle$ ¹	0.80 (at 1.79Å)	Xtriage
Refinement program	TNT	Depositor
R, R_{free}	0.176 , 0.253 0.175 , (Not available)	Depositor DCC
R_{free} test set	No test flags present.	wwPDB-VP
Wilson B-factor (Å ²)	11.5	Xtriage
Anisotropy	0.045	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.34 , 86.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.036 for h,-k,-h-l	Xtriage
F_o, F_c correlation	0.95	EDS
Total number of atoms	12420	wwPDB-VP
Average B, all atoms (Å ²)	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.65% 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, 0.333 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: NA, LLP, AKG, 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	A	0.85	1/2884 (0.0%)	1.19	11/3920 (0.3%)
1	B	0.81	0/2892	1.18	5/3931 (0.1%)
1	C	0.80	3/2867 (0.1%)	1.14	2/3895 (0.1%)
1	D	0.74	0/2850	1.14	8/3872 (0.2%)
All	All	0.80	4/11493 (0.0%)	1.16	26/15618 (0.2%)

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	371	VAL	C-OXT	8.03	1.38	1.23
1	C	371	VAL	CB-CG2	-6.85	1.38	1.52
1	C	295	GLY	C-O	-5.43	1.15	1.23
1	A	164	ALA	CA-CB	5.08	1.63	1.52

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	300	ARG	NE-CZ-NH2	9.19	124.90	120.30
1	B	148	ARG	NE-CZ-NH2	-9.08	115.76	120.30
1	D	11	ALA	C-N-CD	-8.68	101.50	120.60
1	C	300	ARG	NE-CZ-NH1	-8.55	116.02	120.30
1	B	148	ARG	NE-CZ-NH1	7.72	124.16	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	2852	0	2838	72	0
1	B	2857	0	2851	70	0
1	C	2830	0	2825	106	0
1	D	2816	0	2799	109	0
2	A	4	0	3	0	0
2	C	4	0	3	2	0
2	D	4	0	3	3	0
3	A	1	0	0	0	0
4	A	10	0	5	2	0
5	A	304	0	0	1	0
5	B	284	0	0	6	0
5	C	238	0	0	4	0
5	D	216	0	0	6	0
All	All	12420	0	11327	354	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 16.

The worst 5 of 354 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:C:206:LYS:HB2	1:C:206:LYS:NZ	1.77	0.99
1:C:359:ILE:HG22	1:C:363:ILE:HD11	1.51	0.92
1:D:219:ARG:HG2	1:D:222:TRP:CB	2.02	0.90
1:C:36:ARG:HG3	1:C:36:ARG:HH21	1.37	0.89
1:A:149:ARG:HH21	1:A:149:ARG:HG2	1.43	0.83

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	366/373 (98%)	357 (98%)	8 (2%)	1 (0%)	41	27
1	B	367/373 (98%)	352 (96%)	15 (4%)	0	100	100
1	C	364/373 (98%)	343 (94%)	18 (5%)	3 (1%)	19	7
1	D	362/373 (97%)	342 (94%)	19 (5%)	1 (0%)	41	27
All	All	1459/1492 (98%)	1394 (96%)	60 (4%)	5 (0%)	41	27

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	296	LEU
1	C	352	ALA
1	A	11	ALA
1	C	318	PHE
1	D	12	PRO

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	298/301 (99%)	282 (95%)	16 (5%)	22	9
1	B	299/301 (99%)	288 (96%)	11 (4%)	34	19
1	C	296/301 (98%)	270 (91%)	26 (9%)	10	3
1	D	294/301 (98%)	265 (90%)	29 (10%)	8	2
All	All	1187/1204 (99%)	1105 (93%)	82 (7%)	15	5

5 of 82 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	206	LYS
1	C	296	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	292	LEU
1	C	246	ARG
1	C	262	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	347	ASN
1	D	58	ASN
1	C	105	ASN
1	B	150	HIS
1	C	91	ASN

5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

4 non-standard protein/DNA/RNA residues 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	$\# Z > 2$
1	LLP	D	186	1	23,24,25	1.09	2 (8%)	25,32,34	1.57	6 (24%)
1	LLP	C	186	1	23,24,25	1.33	3 (13%)	25,32,34	2.15	8 (32%)
1	LLP	B	186	1	23,24,25	1.31	4 (17%)	25,32,34	1.56	6 (24%)
1	LLP	A	186	1	23,24,25	1.13	1 (4%)	25,32,34	1.69	8 (32%)

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
1	LLP	D	186	1	-	5/16/17/19	0/1/1/1
1	LLP	C	186	1	-	5/16/17/19	0/1/1/1
1	LLP	B	186	1	-	5/16/17/19	0/1/1/1
1	LLP	A	186	1	-	5/16/17/19	0/1/1/1

The worst 5 of 10 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	186	LLP	C4-C5	2.93	1.45	1.42
1	A	186	LLP	P-OP1	2.84	1.59	1.50
1	D	186	LLP	P-OP1	2.71	1.59	1.50
1	C	186	LLP	P-OP1	2.60	1.58	1.50
1	C	186	LLP	C3-C2	-2.47	1.38	1.40

The worst 5 of 28 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	186	LLP	C4-C3-C2	5.60	123.65	120.19
1	C	186	LLP	C3-C4-C5	-4.91	114.49	118.26
1	C	186	LLP	CE-NZ-C4'	3.58	129.89	118.90
1	B	186	LLP	CE-NZ-C4'	3.41	129.37	118.90
1	A	186	LLP	CE-NZ-C4'	3.37	129.25	118.90

There are no chirality outliers.

5 of 20 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	D	186	LLP	C5-C4-C4'-NZ
1	C	186	LLP	CG-CD-CE-NZ
1	B	186	LLP	CG-CD-CE-NZ
1	A	186	LLP	CG-CD-CE-NZ
1	D	186	LLP	CG-CD-CE-NZ

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	D	186	LLP	3	0
1	C	186	LLP	3	0
1	A	186	LLP	1	0

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

Of 5 ligands modelled in this entry, 1 is monoatomic - leaving 4 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	ACT	A	372	-	1,3,3	1.64	0	0,3,3	0.00	-
2	ACT	C	372	-	1,3,3	3.34	1 (100%)	0,3,3	0.00	-
2	ACT	D	372	-	1,3,3	2.48	1 (100%)	0,3,3	0.00	-
4	AKG	A	374	-	3,9,9	0.41	0	4,11,11	1.68	1 (25%)

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
4	AKG	A	374	-	-	0/3/9/9	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	372	ACT	CH3-C	3.34	1.53	1.48
2	D	372	ACT	CH3-C	2.48	1.51	1.48

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	374	AKG	C3-C4-C5	-2.43	108.59	112.67

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	372	ACT	2	0
2	D	372	ACT	3	0
4	A	374	AKG	2	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	367/373 (98%)	-0.72	1 (0%) 94 92	6, 14, 38, 75	0
1	B	367/373 (98%)	-0.61	1 (0%) 94 92	5, 15, 45, 86	0
1	C	364/373 (97%)	-0.43	3 (0%) 86 84	7, 18, 56, 83	0
1	D	363/373 (97%)	-0.39	5 (1%) 75 72	7, 20, 58, 82	0
All	All	1461/1492 (97%)	-0.54	10 (0%) 87 86	5, 16, 50, 86	0

The worst 5 of 10 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	4	LEU	4.2
1	C	371	VAL	3.3
1	D	355	THR	2.6
1	D	259	GLY	2.4
1	C	310	LEU	2.4

6.2 Non-standard residues in protein, DNA, RNA chains [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
1	LLP	C	186	24/25	0.97	0.13	8,15,27,99	0
1	LLP	A	186	24/25	0.97	0.11	7,12,19,33	0
1	LLP	B	186	24/25	0.98	0.13	4,10,22,30	0
1	LLP	D	186	24/25	0.98	0.13	7,15,28,34	0

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. 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(\AA^2)	Q<0.9
3	NA	A	373	1/1	0.82	0.10	42,42,42,42	0
2	ACT	C	372	4/4	0.94	0.14	14,22,22,25	0
4	AKG	A	374	10/10	0.94	0.15	16,29,99,99	0
2	ACT	D	372	4/4	0.95	0.15	13,16,22,38	0
2	ACT	A	372	4/4	0.96	0.10	11,14,15,33	0

6.5 Other polymers [i](#)

There are no such residues in this entry.