



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 16, 2024 – 07:33 PM EDT

PDB ID : 5G5O  
Title : Structure of the snake adenovirus 1 hexon-interlacing LH3 protein, native  
Authors : Nguyen, T.H.; Singh, A.K.; Albala-Perez, B.; van Raaij, M.J.  
Deposited on : 2016-05-26  
Resolution : 2.00 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](mailto: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 : 1.8.5 (274361), CSD as541be (2020)  
Xtrriage (Phenix) : 1.13  
EDS : 2.37.1  
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) : 2.37.1

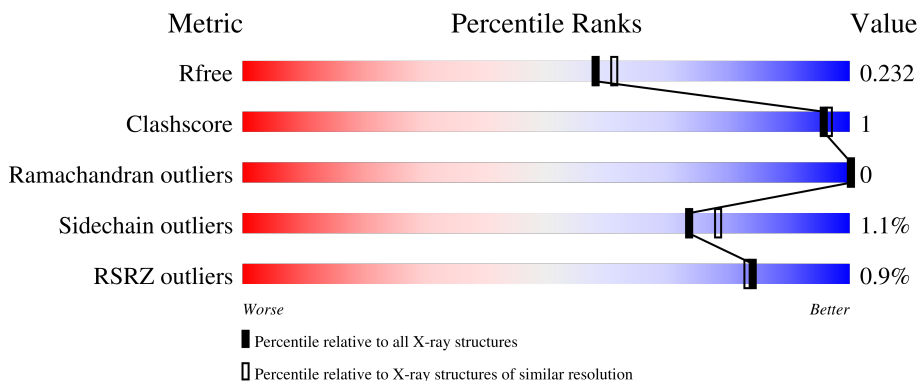
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.00 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.



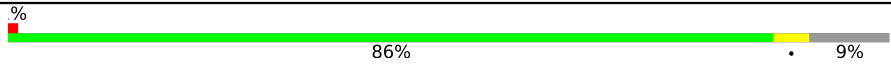
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	8085 (2.00-2.00)
Clashscore	141614	9178 (2.00-2.00)
Ramachandran outliers	138981	9054 (2.00-2.00)
Sidechain outliers	138945	9053 (2.00-2.00)
RSRZ outliers	127900	7900 (2.00-2.00)

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  $\geq 3$ , 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	 88% 9%
1	B	373	 89% 8%
1	C	373	 89% 8%
1	D	373	 86% 9%
1	E	373	 89% 7%

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	F	373	 A horizontal bar chart representing the quality of the chain. The bar is primarily green, indicating a high quality score of 86%. A small yellow segment at the end indicates a lower quality score of 9%. The percentage values are printed below the bar.

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	F	1374	-	-	X	-

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 17167 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 LH3 HEXON-INTERLACING CAPSID PROTEIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	338	2561	1617	448	479	17	0	0	0
1	B	345	2620	1655	461	487	17	0	0	0
1	C	342	2592	1636	455	484	17	0	0	0
1	D	338	2561	1617	448	479	17	0	0	0
1	E	348	2649	1673	465	494	17	0	1	0
1	F	338	2561	1617	448	479	17	0	0	0

There are 108 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	21	GLY	ALA	conflict	UNP A9CB85
A	22	LEU	PHE	conflict	UNP A9CB85
A	23	PRO	LEU	conflict	UNP A9CB85
A	24	VAL	TYR	conflict	UNP A9CB85
A	26	SER	LEU	conflict	UNP A9CB85
A	28	PRO	ARG	conflict	UNP A9CB85
A	71	ASN	GLU	conflict	UNP A9CB85
A	72	GLY	ARG	conflict	UNP A9CB85
A	73	ALA	CYS	conflict	UNP A9CB85
A	74	THR	HIS	conflict	UNP A9CB85
A	75	VAL	GLY	conflict	UNP A9CB85
A	76	LYS	GLU	conflict	UNP A9CB85
A	77	THR	ASP	conflict	UNP A9CB85
A	78	SER	PHE	conflict	UNP A9CB85
A	79	GLY	ARG	conflict	UNP A9CB85
A	80	LEU	PRO	conflict	UNP A9CB85
A	81	GLY	ARG	conflict	UNP A9CB85

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	82	PRO	ALA	conflict	UNP A9CB85
B	21	GLY	ALA	conflict	UNP A9CB85
B	22	LEU	PHE	conflict	UNP A9CB85
B	23	PRO	LEU	conflict	UNP A9CB85
B	24	VAL	TYR	conflict	UNP A9CB85
B	26	SER	LEU	conflict	UNP A9CB85
B	28	PRO	ARG	conflict	UNP A9CB85
B	71	ASN	GLU	conflict	UNP A9CB85
B	72	GLY	ARG	conflict	UNP A9CB85
B	73	ALA	CYS	conflict	UNP A9CB85
B	74	THR	HIS	conflict	UNP A9CB85
B	75	VAL	GLY	conflict	UNP A9CB85
B	76	LYS	GLU	conflict	UNP A9CB85
B	77	THR	ASP	conflict	UNP A9CB85
B	78	SER	PHE	conflict	UNP A9CB85
B	79	GLY	ARG	conflict	UNP A9CB85
B	80	LEU	PRO	conflict	UNP A9CB85
B	81	GLY	ARG	conflict	UNP A9CB85
B	82	PRO	ALA	conflict	UNP A9CB85
C	21	GLY	ALA	conflict	UNP A9CB85
C	22	LEU	PHE	conflict	UNP A9CB85
C	23	PRO	LEU	conflict	UNP A9CB85
C	24	VAL	TYR	conflict	UNP A9CB85
C	26	SER	LEU	conflict	UNP A9CB85
C	28	PRO	ARG	conflict	UNP A9CB85
C	71	ASN	GLU	conflict	UNP A9CB85
C	72	GLY	ARG	conflict	UNP A9CB85
C	73	ALA	CYS	conflict	UNP A9CB85
C	74	THR	HIS	conflict	UNP A9CB85
C	75	VAL	GLY	conflict	UNP A9CB85
C	76	LYS	GLU	conflict	UNP A9CB85
C	77	THR	ASP	conflict	UNP A9CB85
C	78	SER	PHE	conflict	UNP A9CB85
C	79	GLY	ARG	conflict	UNP A9CB85
C	80	LEU	PRO	conflict	UNP A9CB85
C	81	GLY	ARG	conflict	UNP A9CB85
C	82	PRO	ALA	conflict	UNP A9CB85
D	21	GLY	ALA	conflict	UNP A9CB85
D	22	LEU	PHE	conflict	UNP A9CB85
D	23	PRO	LEU	conflict	UNP A9CB85
D	24	VAL	TYR	conflict	UNP A9CB85
D	26	SER	LEU	conflict	UNP A9CB85

*Continued on next page...*

*Continued from previous page...*

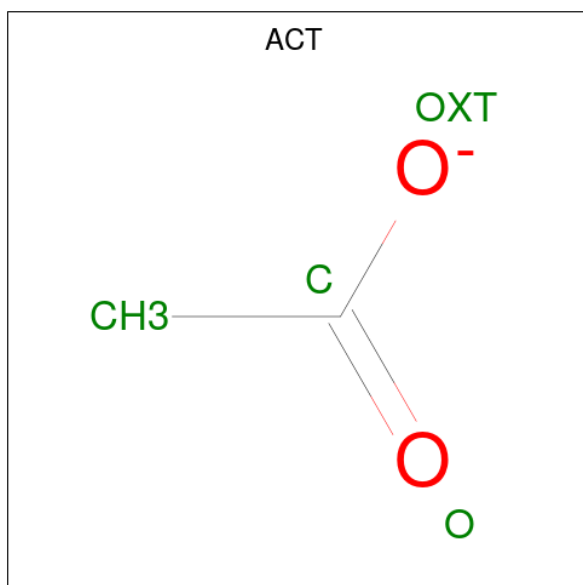
Chain	Residue	Modelled	Actual	Comment	Reference
D	28	PRO	ARG	conflict	UNP A9CB85
D	71	ASN	GLU	conflict	UNP A9CB85
D	72	GLY	ARG	conflict	UNP A9CB85
D	73	ALA	CYS	conflict	UNP A9CB85
D	74	THR	HIS	conflict	UNP A9CB85
D	75	VAL	GLY	conflict	UNP A9CB85
D	76	LYS	GLU	conflict	UNP A9CB85
D	77	THR	ASP	conflict	UNP A9CB85
D	78	SER	PHE	conflict	UNP A9CB85
D	79	GLY	ARG	conflict	UNP A9CB85
D	80	LEU	PRO	conflict	UNP A9CB85
D	81	GLY	ARG	conflict	UNP A9CB85
D	82	PRO	ALA	conflict	UNP A9CB85
E	21	GLY	ALA	conflict	UNP A9CB85
E	22	LEU	PHE	conflict	UNP A9CB85
E	23	PRO	LEU	conflict	UNP A9CB85
E	24	VAL	TYR	conflict	UNP A9CB85
E	26	SER	LEU	conflict	UNP A9CB85
E	28	PRO	ARG	conflict	UNP A9CB85
E	71	ASN	GLU	conflict	UNP A9CB85
E	72	GLY	ARG	conflict	UNP A9CB85
E	73	ALA	CYS	conflict	UNP A9CB85
E	74	THR	HIS	conflict	UNP A9CB85
E	75	VAL	GLY	conflict	UNP A9CB85
E	76	LYS	GLU	conflict	UNP A9CB85
E	77	THR	ASP	conflict	UNP A9CB85
E	78	SER	PHE	conflict	UNP A9CB85
E	79	GLY	ARG	conflict	UNP A9CB85
E	80	LEU	PRO	conflict	UNP A9CB85
E	81	GLY	ARG	conflict	UNP A9CB85
E	82	PRO	ALA	conflict	UNP A9CB85
F	21	GLY	ALA	conflict	UNP A9CB85
F	22	LEU	PHE	conflict	UNP A9CB85
F	23	PRO	LEU	conflict	UNP A9CB85
F	24	VAL	TYR	conflict	UNP A9CB85
F	26	SER	LEU	conflict	UNP A9CB85
F	28	PRO	ARG	conflict	UNP A9CB85
F	71	ASN	GLU	conflict	UNP A9CB85
F	72	GLY	ARG	conflict	UNP A9CB85
F	73	ALA	CYS	conflict	UNP A9CB85
F	74	THR	HIS	conflict	UNP A9CB85
F	75	VAL	GLY	conflict	UNP A9CB85

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
F	76	LYS	GLU	conflict	UNP A9CB85
F	77	THR	ASP	conflict	UNP A9CB85
F	78	SER	PHE	conflict	UNP A9CB85
F	79	GLY	ARG	conflict	UNP A9CB85
F	80	LEU	PRO	conflict	UNP A9CB85
F	81	GLY	ARG	conflict	UNP A9CB85
F	82	PRO	ALA	conflict	UNP A9CB85

- Molecule 2 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



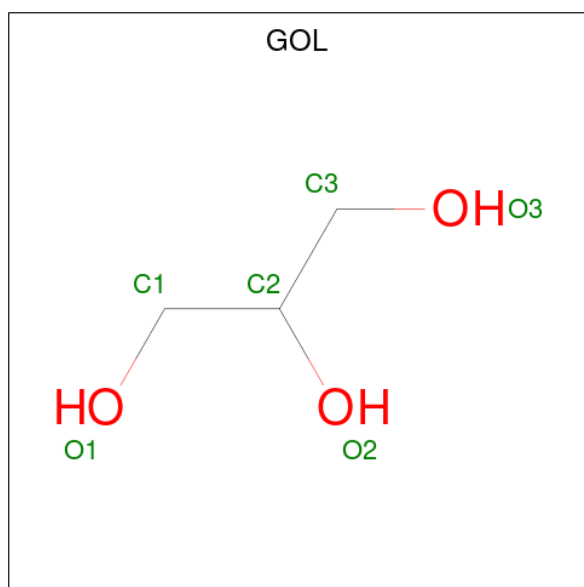
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 4 2 2	0	0
2	A	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0
2	C	1	Total C O 4 2 2	0	0
2	D	1	Total C O 4 2 2	0	0
2	D	1	Total C O 4 2 2	0	0
2	D	1	Total C O 4 2 2	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	E	1	Total	C	O	0	0
			4	2	2		
2	E	1	Total	C	O	0	0
			4	2	2		
2	F	1	Total	C	O	0	0
			4	2	2		

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		
3	D	1	Total	C	O	0	0
			6	3	3		
3	D	1	Total	C	O	0	0
			6	3	3		
3	E	1	Total	C	O	0	0
			6	3	3		
3	E	1	Total	C	O	0	0
			6	3	3		
3	F	1	Total	C	O	0	0
			6	3	3		



- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total Cl 1 1	0	0
4	B	1	Total Cl 1 1	0	0
4	C	1	Total Cl 1 1	0	0
4	D	1	Total Cl 1 1	0	0
4	E	1	Total Cl 1 1	0	0
4	F	1	Total Cl 1 1	0	0

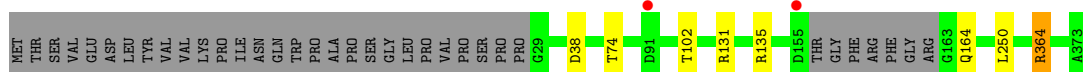
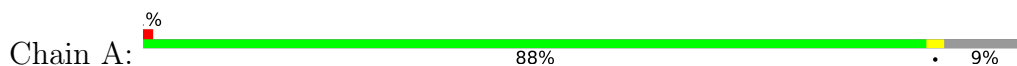
- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	277	Total O 277 277	0	0
5	B	257	Total O 257 257	0	0
5	C	225	Total O 225 225	0	0
5	D	281	Total O 281 281	0	0
5	E	230	Total O 230 230	0	0
5	F	255	Total O 255 255	0	0

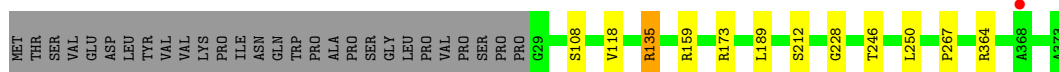
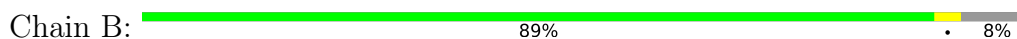
### 3 Residue-property plots [i](#)

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.

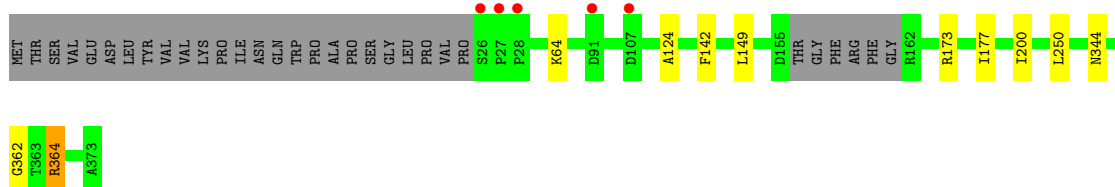
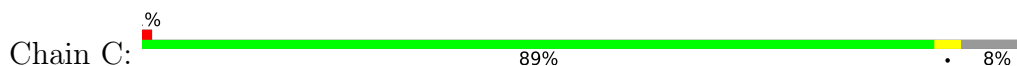
- Molecule 1: LH3 HEXON-INTERLACING CAPSID PROTEIN



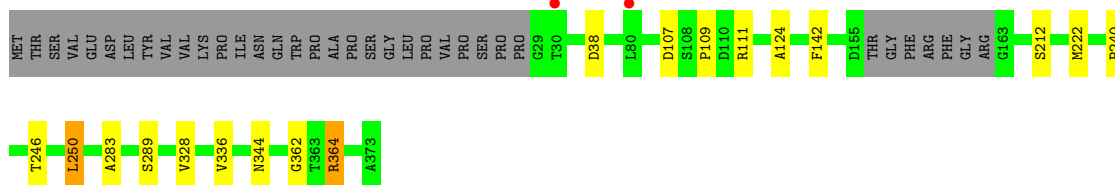
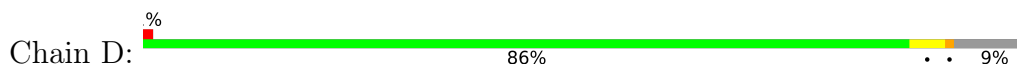
- Molecule 1: LH3 HEXON-INTERLACING CAPSID PROTEIN



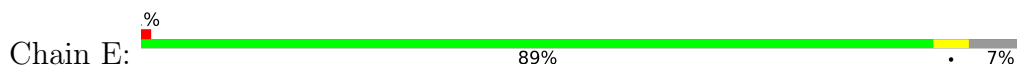
- Molecule 1: LH3 HEXON-INTERLACING CAPSID PROTEIN

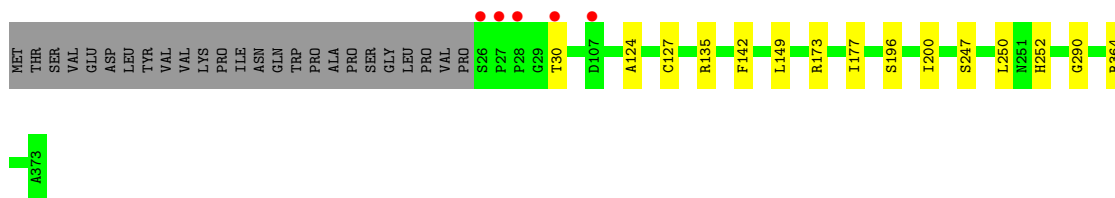


- Molecule 1: LH3 HEXON-INTERLACING CAPSID PROTEIN



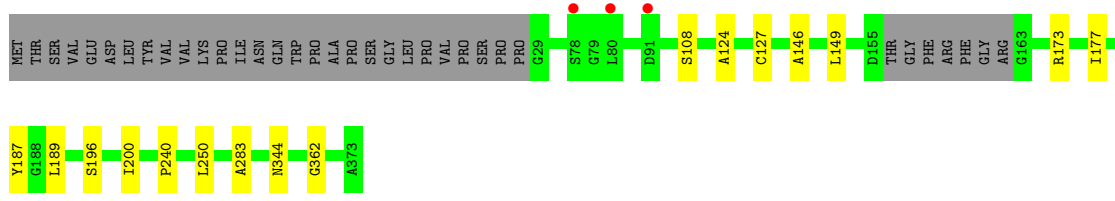
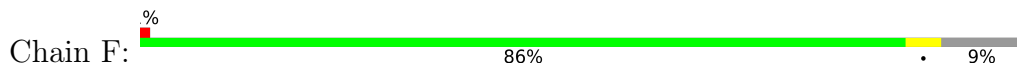
- Molecule 1: LH3 HEXON-INTERLACING CAPSID PROTEIN





A373

• Molecule 1: LH3 HEXON-INTERLACING CAPSID PROTEIN



Y187  
G188  
L189  
S196  
I200  
P240  
L250  
A283  
N344  
G362  
A373

## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	70.79Å 126.88Å 120.94Å 90.00° 106.15° 90.00°	Depositor
Resolution (Å)	30.00 – 2.00 29.97 – 2.00	Depositor EDS
% Data completeness (in resolution range)	94.2 (30.00-2.00) 94.2 (29.97-2.00)	Depositor EDS
$R_{merge}$	0.20	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.69 (at 2.00Å)	Xtrriage
Refinement program	REFMAC 5.8.0151	Depositor
R, $R_{free}$	0.172 , 0.226 0.180 , 0.232	Depositor DCC
$R_{free}$ test set	1945 reflections (1.49%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	21.3	Xtrriage
Anisotropy	0.278	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 46.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.027 for h,-k,-h-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	17167	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.80% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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 [i](#)

### 5.1 Standard geometry [i](#)

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

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	0.56	0/2624	0.76	2/3571 (0.1%)
1	B	0.54	0/2686	0.73	4/3654 (0.1%)
1	C	0.57	0/2657	0.72	1/3617 (0.0%)
1	D	0.56	0/2624	0.75	2/3571 (0.1%)
1	E	0.56	0/2717	0.73	3/3698 (0.1%)
1	F	0.56	0/2624	0.71	1/3571 (0.0%)
All	All	0.56	0/15932	0.73	13/21682 (0.1%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	364	ARG	NE-CZ-NH1	6.83	123.71	120.30
1	B	135	ARG	NE-CZ-NH2	-6.51	117.05	120.30
1	A	364	ARG	NE-CZ-NH1	6.40	123.50	120.30
1	B	364	ARG	NE-CZ-NH2	-6.29	117.15	120.30
1	A	364	ARG	NE-CZ-NH2	-6.22	117.19	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	2561	0	2455	4	0
1	B	2620	0	2513	5	0
1	C	2592	0	2487	4	0
1	D	2561	0	2455	9	0
1	E	2649	0	2537	7	0
1	F	2561	0	2455	7	0
2	A	8	0	6	0	0
2	B	8	0	6	0	0
2	C	4	0	3	0	0
2	D	12	0	9	0	0
2	E	8	0	6	0	0
2	F	4	0	3	4	0
3	A	6	0	8	0	0
3	B	6	0	8	0	0
3	C	6	0	8	0	0
3	D	12	0	16	0	0
3	E	12	0	16	0	0
3	F	6	0	8	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
4	C	1	0	0	0	0
4	D	1	0	0	0	0
4	E	1	0	0	0	0
4	F	1	0	0	0	0
5	A	277	0	0	1	0
5	B	257	0	0	1	0
5	C	225	0	0	0	0
5	D	281	0	0	0	0
5	E	230	0	0	0	0
5	F	255	0	0	1	0
All	All	17167	0	14999	34	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 1.

The worst 5 of 34 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)
2:F:1374:ACT:H1	5:F:2117:HOH:O	1.83	0.78
1:F:187:TYR:OH	2:F:1374:ACT:H3	1.82	0.77
1:D:328:VAL:HG21	1:D:336:VAL:HG21	1.71	0.71
1:E:252:HIS:NE2	2:F:1374:ACT:H2	2.12	0.64

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:74:THR:HG23	5:A:2020:HOH:O	1.96	0.64

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	334/373 (90%)	325 (97%)	9 (3%)	0	100	100
1	B	343/373 (92%)	333 (97%)	10 (3%)	0	100	100
1	C	338/373 (91%)	328 (97%)	10 (3%)	0	100	100
1	D	334/373 (90%)	326 (98%)	8 (2%)	0	100	100
1	E	347/373 (93%)	336 (97%)	11 (3%)	0	100	100
1	F	334/373 (90%)	321 (96%)	13 (4%)	0	100	100
All	All	2030/2238 (91%)	1969 (97%)	61 (3%)	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	267/298 (90%)	263 (98%)	4 (2%)	65	69
1	B	272/298 (91%)	269 (99%)	3 (1%)	73	78

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	C	271/298 (91%)	267 (98%)	4 (2%)	65	69
1	D	267/298 (90%)	264 (99%)	3 (1%)	73	78
1	E	276/298 (93%)	274 (99%)	2 (1%)	84	88
1	F	267/298 (90%)	265 (99%)	2 (1%)	84	88
All	All	1620/1788 (91%)	1602 (99%)	18 (1%)	73	78

5 of 18 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	30	THR
1	F	250	LEU
1	F	108	SER
1	C	149	LEU
1	D	364	ARG

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	227	GLN
1	F	192	GLN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

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

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

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 25 ligands modelled in this entry, 6 are monoatomic - leaving 19 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	D	1374	-	3,3,3	0.84	0	3,3,3	1.25	0
2	ACT	D	1376	-	3,3,3	0.72	0	3,3,3	0.87	0
2	ACT	C	1374	-	3,3,3	0.90	0	3,3,3	0.40	0
3	GOL	E	1377	-	5,5,5	0.40	0	5,5,5	0.57	0
2	ACT	A	1374	-	3,3,3	0.74	0	3,3,3	0.92	0
2	ACT	E	1374	-	3,3,3	0.83	0	3,3,3	0.62	0
2	ACT	A	1375	-	3,3,3	0.72	0	3,3,3	0.75	0
3	GOL	F	1375	-	5,5,5	0.45	0	5,5,5	0.37	0
3	GOL	E	1376	-	5,5,5	0.61	0	5,5,5	0.70	0
2	ACT	B	1375	-	3,3,3	0.67	0	3,3,3	0.85	0
2	ACT	E	1375	-	3,3,3	0.88	0	3,3,3	0.37	0
3	GOL	D	1378	-	5,5,5	0.23	0	5,5,5	0.53	0
3	GOL	A	1376	-	5,5,5	0.44	0	5,5,5	0.74	0
2	ACT	D	1375	-	3,3,3	0.55	0	3,3,3	0.81	0
2	ACT	B	1374	-	3,3,3	0.80	0	3,3,3	0.93	0
2	ACT	F	1374	-	3,3,3	0.95	0	3,3,3	0.55	0
3	GOL	C	1375	-	5,5,5	0.27	0	5,5,5	0.51	0
3	GOL	B	1376	-	5,5,5	0.56	0	5,5,5	0.69	0
3	GOL	D	1377	-	5,5,5	0.49	0	5,5,5	0.51	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	GOL	F	1375	-	-	2/4/4/4	-
3	GOL	E	1377	-	-	1/4/4/4	-
3	GOL	E	1376	-	-	2/4/4/4	-
3	GOL	C	1375	-	-	0/4/4/4	-
3	GOL	B	1376	-	-	2/4/4/4	-
3	GOL	D	1377	-	-	3/4/4/4	-
3	GOL	D	1378	-	-	2/4/4/4	-
3	GOL	A	1376	-	-	1/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

5 of 13 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	1376	GOL	O1-C1-C2-C3
3	E	1376	GOL	O1-C1-C2-C3
3	F	1375	GOL	O1-C1-C2-C3
3	D	1377	GOL	C1-C2-C3-O3
3	D	1378	GOL	C1-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	1374	ACT	4	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	A	338/373 (90%)	-0.36	2 (0%) 89 88	13, 20, 36, 55	0
1	B	345/373 (92%)	-0.42	1 (0%) 94 93	13, 19, 32, 50	0
1	C	342/373 (91%)	-0.41	5 (1%) 73 72	13, 19, 36, 68	0
1	D	338/373 (90%)	-0.53	2 (0%) 89 88	11, 18, 32, 51	0
1	E	348/373 (93%)	-0.32	5 (1%) 75 74	12, 20, 36, 88	0
1	F	338/373 (90%)	-0.36	3 (0%) 84 83	13, 21, 37, 52	0
All	All	2049/2238 (91%)	-0.40	18 (0%) 84 83	11, 20, 35, 88	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	E	26	SER	6.5
1	E	27	PRO	5.8
1	C	26	SER	5.4
1	C	27	PRO	5.4
1	C	28	PRO	4.1

### 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
3	GOL	F	1375	6/6	0.74	0.21	47,47,48,48	0
2	ACT	D	1374	4/4	0.81	0.18	29,32,32,32	0
3	GOL	B	1376	6/6	0.82	0.20	39,46,48,50	0
3	GOL	E	1376	6/6	0.85	0.17	32,37,38,39	0
3	GOL	E	1377	6/6	0.86	0.24	41,45,45,47	0
2	ACT	F	1374	4/4	0.87	0.16	25,29,30,32	0
3	GOL	C	1375	6/6	0.88	0.12	49,52,56,57	0
2	ACT	A	1375	4/4	0.88	0.21	47,48,50,50	0
2	ACT	D	1376	4/4	0.91	0.20	48,49,51,51	0
2	ACT	D	1375	4/4	0.92	0.13	18,19,20,24	0
3	GOL	D	1377	6/6	0.93	0.11	28,32,34,36	0
3	GOL	A	1376	6/6	0.93	0.11	24,28,30,31	0
2	ACT	B	1374	4/4	0.94	0.10	16,17,17,17	0
2	ACT	C	1374	4/4	0.94	0.12	27,27,27,30	0
3	GOL	D	1378	6/6	0.94	0.10	22,26,27,29	0
2	ACT	E	1375	4/4	0.95	0.09	17,17,17,19	0
2	ACT	E	1374	4/4	0.95	0.08	23,23,23,23	0
2	ACT	B	1375	4/4	0.97	0.08	21,21,22,23	0
2	ACT	A	1374	4/4	0.98	0.06	29,29,30,31	0
4	CL	A	1377	1/1	0.99	0.09	18,18,18,18	0
4	CL	C	1376	1/1	0.99	0.05	15,15,15,15	0
4	CL	E	1378	1/1	0.99	0.08	16,16,16,16	0
4	CL	D	1379	1/1	1.00	0.06	14,14,14,14	0
4	CL	B	1377	1/1	1.00	0.09	16,16,16,16	0
4	CL	F	1376	1/1	1.00	0.04	16,16,16,16	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.