



Full wwPDB X-ray Structure Validation Report ⓘ

Nov 1, 2023 – 02:04 AM EDT

PDB ID : 3HPH
Title : Closed tetramer of Visna virus integrase (residues 1-219) in complex with LEDGF IBD
Authors : Hare, S.; Wang, J.; Cherepanov, P.
Deposited on : 2009-06-04
Resolution : 2.64 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 : 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

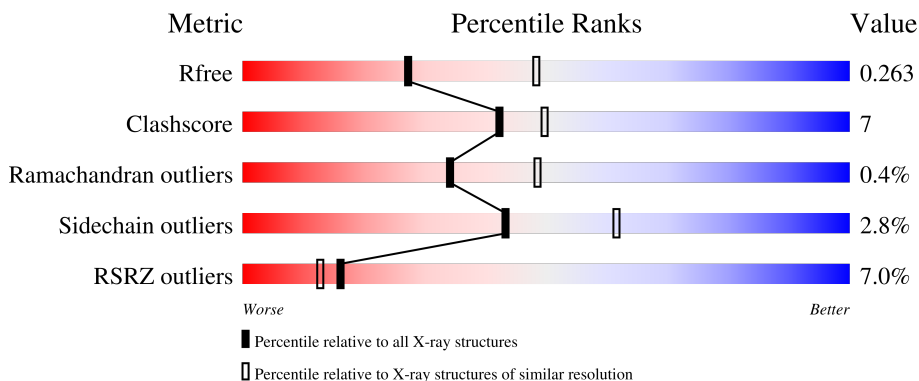
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 2.64 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.




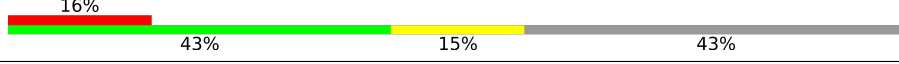

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1426 (2.66-2.62)
Clashscore	141614	1472 (2.66-2.62)
Ramachandran outliers	138981	1446 (2.66-2.62)
Sidechain outliers	138945	1446 (2.66-2.62)
RSRZ outliers	127900	1408 (2.66-2.62)

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 $\leq 5\%$. 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	219	 4% 78% 11% 11%
1	B	219	 2% 79% 12% 9%
1	C	219	 3% 75% 16% 8%
1	D	219	 8% 75% 15% 11%
2	E	94	 4% 83% 12% 5%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	94	
2	G	94	
2	H	94	

2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 8778 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 Integrase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	194	Total 1539	C 982	N 259	O 291	S 7	0	0	0
1	B	200	Total 1569	C 995	N 271	O 296	S 7	0	0	0
1	C	202	Total 1583	C 1003	N 274	O 299	S 7	0	0	0
1	D	196	Total 1554	C 990	N 262	O 295	S 7	0	0	0

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1	MET	-	expression tag	UNP P35956
A	2	VAL	-	expression tag	UNP P35956
B	1	MET	-	expression tag	UNP P35956
B	2	VAL	-	expression tag	UNP P35956
C	1	MET	-	expression tag	UNP P35956
C	2	VAL	-	expression tag	UNP P35956
D	1	MET	-	expression tag	UNP P35956
D	2	VAL	-	expression tag	UNP P35956

- Molecule 2 is a protein called PC4 and SFRS1-interacting protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	E	89	Total 730	C 463	N 128	O 132	S 7	0	0	0
2	F	60	Total 482	C 306	N 87	O 86	S 3	0	0	0
2	G	54	Total 429	C 272	N 75	O 79	S 3	0	0	0
2	H	91	Total 739	C 468	N 130	O 134	S 7	0	0	0

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	436	LEU	-	expression tag	UNP O75475
E	437	GLU	-	expression tag	UNP O75475
E	438	VAL	-	expression tag	UNP O75475
E	439	LEU	-	expression tag	UNP O75475
E	440	PHE	-	expression tag	UNP O75475
E	441	GLN	-	expression tag	UNP O75475
F	436	LEU	-	expression tag	UNP O75475
F	437	GLU	-	expression tag	UNP O75475
F	438	VAL	-	expression tag	UNP O75475
F	439	LEU	-	expression tag	UNP O75475
F	440	PHE	-	expression tag	UNP O75475
F	441	GLN	-	expression tag	UNP O75475
G	436	LEU	-	expression tag	UNP O75475
G	437	GLU	-	expression tag	UNP O75475
G	438	VAL	-	expression tag	UNP O75475
G	439	LEU	-	expression tag	UNP O75475
G	440	PHE	-	expression tag	UNP O75475
G	441	GLN	-	expression tag	UNP O75475
H	436	LEU	-	expression tag	UNP O75475
H	437	GLU	-	expression tag	UNP O75475
H	438	VAL	-	expression tag	UNP O75475
H	439	LEU	-	expression tag	UNP O75475
H	440	PHE	-	expression tag	UNP O75475
H	441	GLN	-	expression tag	UNP O75475

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

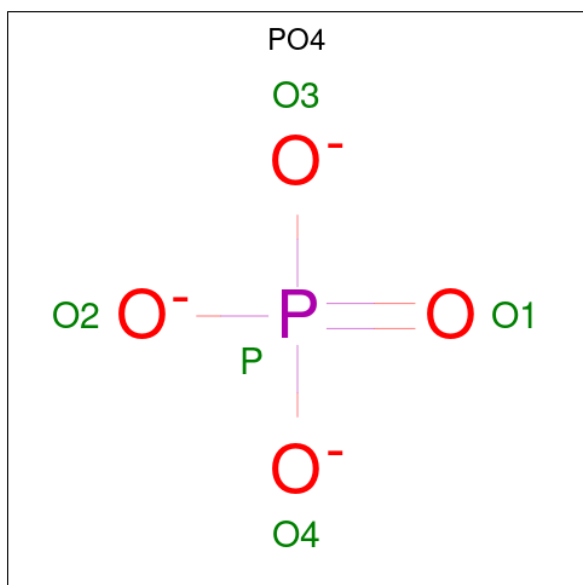
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	1	Total 1	Zn 1	0	0
3	B	1	Total 1	Zn 1	0	0
3	C	1	Total 1	Zn 1	0	0
3	D	1	Total 1	Zn 1	0	0

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C₃H₈O₃).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			6	3	3		
4	A	1	Total	C	O	0	0
			6	3	3		
4	C	1	Total	C	O	0	0
			6	3	3		
4	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 5 is PHOSPHATE ION (three-letter code: PO4) (formula: O₄P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	O	P	0	0
			5	4	1		
5	E	1	Total	O	P	0	0
			5	4	1		
5	H	1	Total	O	P	0	0
			5	4	1		

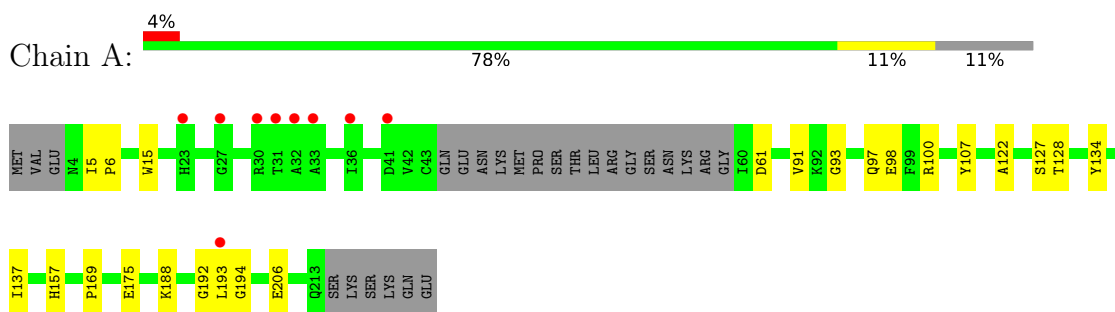
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	18	Total	O	0	0
			18	18		
6	B	28	Total	O	0	0
			28	28		
6	C	34	Total	O	0	0
			34	34		
6	D	14	Total	O	0	0
			14	14		
6	E	6	Total	O	0	0
			6	6		
6	F	1	Total	O	0	0
			1	1		
6	G	2	Total	O	0	0
			2	2		
6	H	7	Total	O	0	0
			7	7		

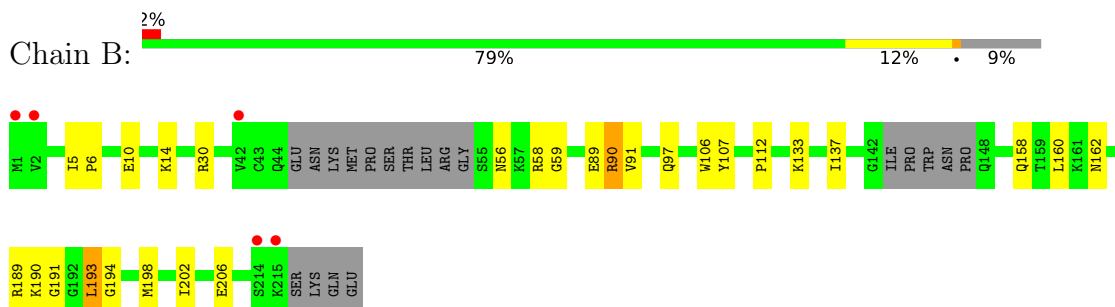
3 Residue-property plots [i](#)

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.

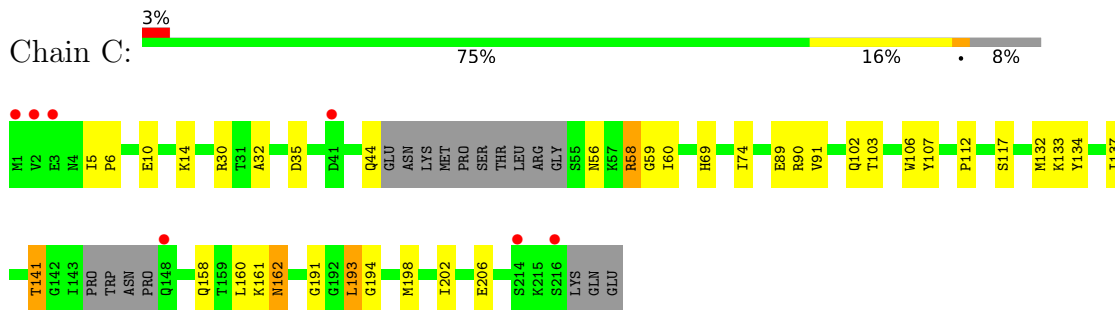
- Molecule 1: Integrase



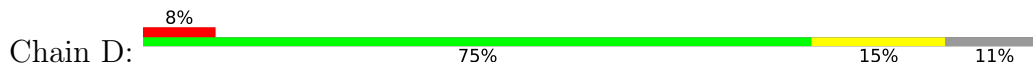
- Molecule 1: Integrase

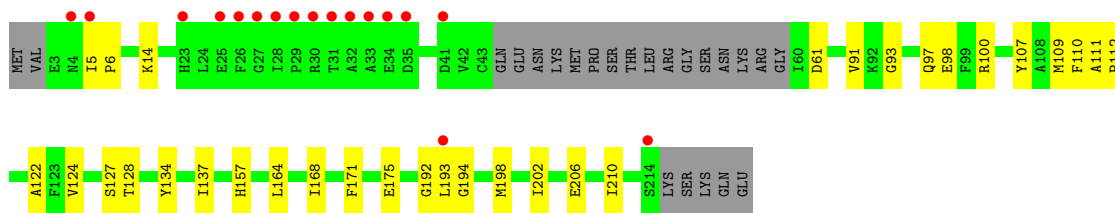


- Molecule 1: Integrase

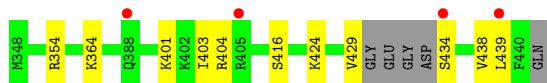
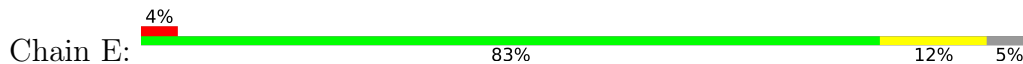


- Molecule 1: Integrase

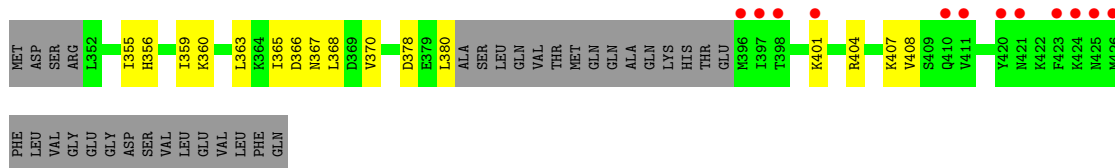




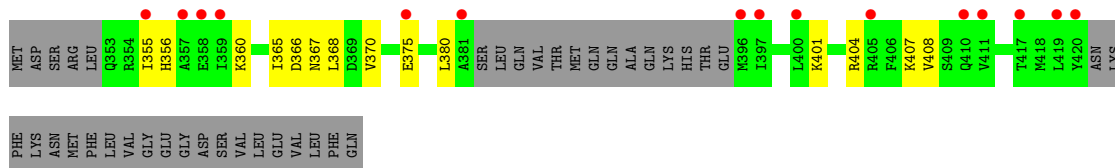
• Molecule 2: PC4 and SFRS1-interacting protein



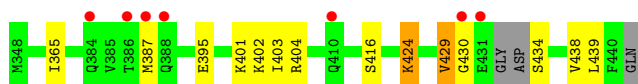
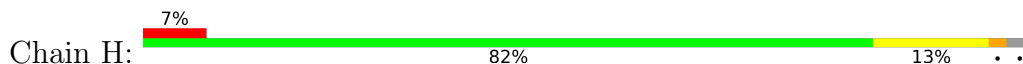
• Molecule 2: PC4 and SFRS1-interacting protein



• Molecule 2: PC4 and SFRS1-interacting protein



• Molecule 2: PC4 and SFRS1-interacting protein



4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α , β , γ	102.87Å 83.15Å 115.30Å 90.00° 101.96° 90.00°	Depositor
Resolution (Å)	39.01 – 2.64 39.01 – 2.64	Depositor EDS
% Data completeness (in resolution range)	99.5 (39.01-2.64) 99.5 (39.01-2.64)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	2.40 (at 2.65Å)	Xtrriage
Refinement program	REFMAC	Depositor
R, R_{free}	0.226 , 0.253 0.237 , 0.263	Depositor DCC
R_{free} test set	2827 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å ²)	44.7	Xtrriage
Anisotropy	0.235	Xtrriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 59.9	EDS
L-test for twinning ²	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
F_o, F_c correlation	0.92	EDS
Total number of atoms	8778	wwPDB-VP
Average B, all atoms (Å ²)	26.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.20% 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: ZN, PO4, 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	0.61	0/1575	0.67	0/2140
1	B	0.74	0/1600	0.79	2/2166 (0.1%)
1	C	0.72	0/1614	0.79	1/2185 (0.0%)
1	D	0.62	0/1590	0.66	0/2160
2	E	0.58	0/735	0.65	0/980
2	F	0.45	0/484	0.54	0/644
2	G	0.45	0/430	0.55	0/574
2	H	0.54	0/744	0.63	0/992
All	All	0.64	0/8772	0.70	3/11841 (0.0%)

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	C	0	1

There are no bond length outliers.

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	193	LEU	CA-CB-CG	5.87	128.79	115.30
1	C	193	LEU	CA-CB-CG	5.31	127.51	115.30
1	B	90	ARG	NE-CZ-NH2	-5.09	117.75	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	141	THR	Peptide

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	1539	0	1476	22	0
1	B	1569	0	1507	19	0
1	C	1583	0	1517	28	0
1	D	1554	0	1489	30	0
2	E	730	0	774	7	0
2	F	482	0	505	11	0
2	G	429	0	452	10	0
2	H	739	0	779	11	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
4	A	12	0	16	0	0
4	C	6	0	8	0	0
4	D	6	0	8	0	0
5	B	5	0	0	0	0
5	E	5	0	0	0	0
5	H	5	0	0	1	0
6	A	18	0	0	1	0
6	B	28	0	0	1	0
6	C	34	0	0	2	0
6	D	14	0	0	1	0
6	E	6	0	0	1	0
6	F	1	0	0	0	0
6	G	2	0	0	0	0
6	H	7	0	0	1	0
All	All	8778	0	8531	119	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 7.

All (119) 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:30:ARG:NH2	1:C:59:GLY:HA2	1.78	0.97
1:D:193:LEU:N	1:D:194:GLY:HA2	1.88	0.89
1:B:30:ARG:NH2	1:B:59:GLY:HA2	1.91	0.85
1:B:90:ARG:HD2	6:B:245:HOH:O	1.76	0.85
1:A:193:LEU:N	1:A:194:GLY:HA2	1.92	0.84
1:C:90:ARG:HD2	6:C:249:HOH:O	1.81	0.81
1:C:30:ARG:HH21	1:C:59:GLY:HA2	1.45	0.80
2:H:387:MET:HE1	2:H:430:GLY:HA2	1.62	0.80
1:B:30:ARG:HH21	1:B:59:GLY:HA2	1.49	0.77
1:A:97:GLN:HE21	2:F:367:ASN:ND2	1.82	0.76
1:D:5:ILE:HG23	1:D:6:PRO:HD3	1.67	0.75
1:A:100:ARG:HH11	1:A:128:THR:HG22	1.53	0.73
1:D:100:ARG:HH11	1:D:128:THR:HG22	1.54	0.73
2:H:387:MET:CE	2:H:430:GLY:HA2	2.20	0.72
1:A:5:ILE:HG23	1:A:6:PRO:HD3	1.73	0.71
1:C:158:GLN:O	1:C:162:ASN:HB2	1.94	0.68
2:H:395:GLU:HB2	6:H:78:HOH:O	1.94	0.68
1:D:97:GLN:HE21	2:G:367:ASN:ND2	1.93	0.67
1:B:97:GLN:HG3	6:E:98:HOH:O	1.95	0.67
1:B:158:GLN:O	1:B:162:ASN:HB2	1.95	0.67
1:A:107:TYR:HD1	1:A:137:ILE:HD11	1.61	0.66
2:E:401:LYS:HA	2:E:404:ARG:HD2	1.78	0.65
1:A:100:ARG:HH11	1:A:128:THR:CG2	2.10	0.64
1:C:191:GLY:O	1:C:194:GLY:HA2	1.99	0.62
1:D:100:ARG:HH11	1:D:128:THR:CG2	2.11	0.62
1:D:107:TYR:CD1	1:D:137:ILE:HD11	2.35	0.62
1:B:193:LEU:N	1:B:194:GLY:HA2	2.15	0.62
2:H:429:VAL:HG12	2:H:429:VAL:O	2.00	0.61
2:H:401:LYS:HA	2:H:404:ARG:HD2	1.82	0.60
1:A:107:TYR:CD1	1:A:137:ILE:HD11	2.36	0.60
1:B:191:GLY:O	1:B:194:GLY:HA2	2.02	0.59
1:D:107:TYR:HD1	1:D:137:ILE:HD11	1.66	0.59
1:C:30:ARG:HH22	1:C:59:GLY:HA2	1.64	0.59
1:A:97:GLN:HE21	2:F:367:ASN:HD21	1.50	0.59
1:D:5:ILE:CG2	1:D:6:PRO:HD3	2.32	0.59
1:C:193:LEU:N	1:C:194:GLY:HA2	2.18	0.57
2:H:429:VAL:O	2:H:429:VAL:CG1	2.53	0.56
1:B:189:ARG:O	1:B:190:LYS:HG2	2.04	0.56
1:A:206:GLU:HG2	1:B:206:GLU:HG2	1.87	0.56
1:D:97:GLN:HE21	2:G:367:ASN:HD21	1.53	0.55
1:A:134:TYR:CZ	1:C:14:LYS:HG2	2.42	0.55
1:C:206:GLU:HG2	1:D:206:GLU:HG2	1.88	0.54

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:5:ILE:CG2	1:A:6:PRO:HD3	2.37	0.54
1:A:98:GLU:HB3	6:A:224:HOH:O	2.07	0.54
1:C:30:ARG:NH2	1:C:59:GLY:CA	2.63	0.53
1:A:91:VAL:HG13	1:A:98:GLU:CD	2.29	0.53
1:D:198:MET:O	1:D:202:ILE:HG12	2.08	0.52
1:B:30:ARG:HH21	1:B:59:GLY:CA	2.21	0.52
1:C:107:TYR:CE1	1:C:137:ILE:HD11	2.44	0.52
2:F:356:HIS:CE1	2:F:360:LYS:HE2	2.45	0.52
1:A:175:GLU:H	1:A:175:GLU:CD	2.13	0.52
1:C:106:TRP:CH2	1:C:112:PRO:HG3	2.45	0.52
1:D:193:LEU:HD12	6:D:232:HOH:O	2.09	0.52
1:B:30:ARG:NH2	1:B:59:GLY:CA	2.68	0.51
2:G:356:HIS:CE1	2:G:360:LYS:HE2	2.46	0.51
2:E:403:ILE:HG13	2:E:416:SER:OG	2.10	0.51
1:B:5:ILE:N	1:B:6:PRO:HD2	2.26	0.50
1:B:106:TRP:CH2	1:B:112:PRO:HG3	2.47	0.50
1:C:5:ILE:N	1:C:6:PRO:HD2	2.26	0.50
1:C:30:ARG:HH21	1:C:60:ILE:H	1.59	0.50
1:C:198:MET:HG2	1:C:202:ILE:HD12	1.95	0.49
1:D:134:TYR:CZ	2:G:365:ILE:HG12	2.47	0.49
2:H:401:LYS:O	2:H:404:ARG:HD3	2.12	0.49
1:B:90:ARG:HG3	1:B:91:VAL:N	2.27	0.48
2:E:429:VAL:HG12	2:E:429:VAL:O	2.11	0.48
2:G:355:ILE:HG21	2:G:380:LEU:HD13	1.95	0.48
1:D:127:SER:OG	2:G:368:LEU:O	2.28	0.48
2:E:401:LYS:O	2:E:404:ARG:HD3	2.13	0.48
1:D:122:ALA:O	1:D:128:THR:HG21	2.13	0.48
1:D:193:LEU:H	1:D:194:GLY:HA2	1.76	0.48
1:C:107:TYR:CD1	1:C:137:ILE:HD11	2.49	0.47
1:C:117:SER:O	1:C:141:THR:HA	2.14	0.47
2:F:355:ILE:HG21	2:F:380:LEU:HD13	1.95	0.47
1:C:30:ARG:HH21	1:C:59:GLY:CA	2.20	0.47
2:E:429:VAL:O	2:E:429:VAL:CG1	2.62	0.47
1:B:14:LYS:HG2	1:D:134:TYR:CZ	2.49	0.46
1:C:90:ARG:HG3	1:C:91:VAL:N	2.30	0.46
1:D:124:VAL:O	1:D:124:VAL:HG12	2.15	0.46
1:D:91:VAL:HG13	1:D:98:GLU:CD	2.36	0.45
1:D:100:ARG:HG3	1:D:128:THR:HG22	1.99	0.45
1:B:107:TYR:CE1	1:B:137:ILE:HD11	2.51	0.45
1:C:102:GLN:NE2	6:C:231:HOH:O	2.50	0.45
1:A:169:PRO:O	2:E:364:LYS:NZ	2.51	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:G:401:LYS:HA	2:G:404:ARG:HD3	2.00	0.44
2:F:359:ILE:O	2:F:363:LEU:HG	2.18	0.44
1:D:175:GLU:H	1:D:175:GLU:CD	2.21	0.44
2:H:403:ILE:HG13	2:H:416:SER:OG	2.18	0.43
2:H:434:SER:O	2:H:438:VAL:HG23	2.18	0.43
1:A:91:VAL:HG12	1:A:93:GLY:O	2.18	0.43
1:D:111:ALA:N	1:D:112:PRO:CD	2.81	0.43
1:A:122:ALA:O	1:A:128:THR:HG21	2.18	0.43
2:E:434:SER:O	2:E:438:VAL:HG23	2.19	0.42
1:D:100:ARG:NE	2:G:366:ASP:O	2.44	0.42
1:B:189:ARG:C	1:B:190:LYS:HG2	2.40	0.42
2:F:366:ASP:OD1	2:F:366:ASP:N	2.50	0.42
1:A:127:SER:OG	2:F:368:LEU:O	2.36	0.42
1:C:10:GLU:OE2	2:F:407:LYS:HE2	2.18	0.42
1:C:69:HIS:O	1:C:161:LYS:HE2	2.19	0.42
1:C:103:THR:HG21	1:C:132:MET:HE2	2.01	0.42
1:D:109:MET:HB3	1:D:110:PHE:CD2	2.55	0.42
1:A:15:TRP:CD1	1:A:188:LYS:HE3	2.55	0.42
1:D:171:PHE:CE1	2:H:365:ILE:HB	2.55	0.42
2:F:401:LYS:HA	2:F:404:ARG:HD3	2.01	0.41
1:A:100:ARG:HG3	1:A:128:THR:HG22	2.02	0.41
1:A:134:TYR:CZ	2:F:365:ILE:HG12	2.55	0.41
1:D:91:VAL:HG12	1:D:93:GLY:O	2.20	0.41
1:C:32:ALA:O	1:C:35:ASP:HB2	2.20	0.41
2:H:424:LYS:NZ	5:H:2:PO4:O3	2.54	0.41
1:C:69:HIS:CD2	1:C:74:ILE:HG12	2.56	0.41
1:C:58:ARG:HH11	1:C:58:ARG:HD2	1.70	0.41
2:G:401:LYS:HA	2:G:404:ARG:CD	2.51	0.40
1:A:134:TYR:CE1	1:C:14:LYS:HG2	2.55	0.40
1:B:10:GLU:OE2	2:G:407:LYS:HE2	2.21	0.40
1:B:198:MET:HG2	1:B:202:ILE:HD12	2.03	0.40
1:D:109:MET:HB3	1:D:110:PHE:CE2	2.57	0.40
1:C:134:TYR:CZ	1:D:14:LYS:HG2	2.57	0.40
1:D:164:LEU:O	1:D:168:ILE:HG13	2.21	0.40
1:D:210:ILE:O	1:D:210:ILE:CG2	2.69	0.40
2:F:378:ASP:C	2:F:380:LEU:H	2.25	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	A	190/219 (87%)	184 (97%)	5 (3%)	1 (0%)	29	43
1	B	194/219 (89%)	188 (97%)	6 (3%)	0	100	100
1	C	196/219 (90%)	190 (97%)	6 (3%)	0	100	100
1	D	192/219 (88%)	186 (97%)	5 (3%)	1 (0%)	29	43
2	E	85/94 (90%)	82 (96%)	3 (4%)	0	100	100
2	F	56/94 (60%)	51 (91%)	5 (9%)	0	100	100
2	G	50/94 (53%)	45 (90%)	4 (8%)	1 (2%)	7	10
2	H	87/94 (93%)	80 (92%)	6 (7%)	1 (1%)	14	20
All	All	1050/1252 (84%)	1006 (96%)	40 (4%)	4 (0%)	34	48

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	192	GLY
1	D	192	GLY
2	G	375	GLU
2	H	429	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	A	163/193 (84%)	161 (99%)	2 (1%)	71	83

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	163/193 (84%)	158 (97%)	5 (3%)	40	58
1	C	164/193 (85%)	157 (96%)	7 (4%)	29	45
1	D	165/193 (86%)	163 (99%)	2 (1%)	71	83
2	E	85/88 (97%)	82 (96%)	3 (4%)	36	53
2	F	53/88 (60%)	51 (96%)	2 (4%)	33	50
2	G	48/88 (54%)	46 (96%)	2 (4%)	30	45
2	H	85/88 (97%)	82 (96%)	3 (4%)	36	53
All	All	926/1124 (82%)	900 (97%)	26 (3%)	43	61

All (26) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	61	ASP
1	A	157	HIS
1	B	56	ASN
1	B	58	ARG
1	B	89	GLU
1	B	133	LYS
1	B	160	LEU
1	C	44	GLN
1	C	56	ASN
1	C	58	ARG
1	C	89	GLU
1	C	133	LYS
1	C	160	LEU
1	C	162	ASN
1	D	61	ASP
1	D	157	HIS
2	E	354	ARG
2	E	424	LYS
2	E	439	LEU
2	F	370	VAL
2	F	408	VAL
2	G	370	VAL
2	G	408	VAL
2	H	402	LYS
2	H	424	LYS
2	H	439	LEU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such

sidechains are listed below:

Mol	Chain	Res	Type
1	A	4	ASN
1	A	157	HIS
1	A	172	ASN
1	C	102	GLN
1	C	204	ASN
2	E	367	ASN
2	F	367	ASN
2	G	367	ASN
2	H	367	ASN

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 11 ligands modelled in this entry, 4 are monoatomic - leaving 7 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	$\# Z > 2$
5	PO4	E	1	-	4,4,4	0.82	0	6,6,6	0.63	0
5	PO4	B	221	-	4,4,4	0.87	0	6,6,6	0.41	0
4	GOL	C	221	-	5,5,5	0.41	0	5,5,5	0.51	0
5	PO4	H	2	-	4,4,4	0.73	0	6,6,6	0.66	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z > 2	Counts	RMSZ	# Z > 2
4	GOL	A	222	-	5,5,5	0.36	0	5,5,5	0.34	0
4	GOL	A	221	-	5,5,5	0.45	0	5,5,5	0.49	0
4	GOL	D	221	-	5,5,5	0.35	0	5,5,5	0.55	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
4	GOL	D	221	-	-	4/4/4/4	-
4	GOL	A	222	-	-	4/4/4/4	-
4	GOL	A	221	-	-	0/4/4/4	-
4	GOL	C	221	-	-	2/4/4/4	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (10) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	222	GOL	O1-C1-C2-C3
4	A	222	GOL	C1-C2-C3-O3
4	C	221	GOL	C1-C2-C3-O3
4	A	222	GOL	O1-C1-C2-O2
4	A	222	GOL	O2-C2-C3-O3
4	D	221	GOL	O1-C1-C2-O2
4	D	221	GOL	O1-C1-C2-C3
4	D	221	GOL	C1-C2-C3-O3
4	C	221	GOL	O2-C2-C3-O3
4	D	221	GOL	O2-C2-C3-O3

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	H	2	PO4	1	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	194/219 (88%)	0.26	9 (4%) 32 28	6, 21, 45, 53	0
1	B	200/219 (91%)	0.13	5 (2%) 57 53	3, 13, 47, 55	0
1	C	202/219 (92%)	0.11	7 (3%) 44 40	3, 13, 47, 56	0
1	D	196/219 (89%)	0.38	17 (8%) 10 8	6, 21, 47, 53	0
2	E	89/94 (94%)	0.24	4 (4%) 33 30	12, 38, 89, 98	0
2	F	60/94 (63%)	1.05	12 (20%) 1 1	17, 44, 57, 62	0
2	G	54/94 (57%)	1.44	15 (27%) 0 0	16, 45, 57, 62	0
2	H	91/94 (96%)	0.50	7 (7%) 13 10	12, 38, 89, 98	0
All	All	1086/1252 (86%)	0.35	76 (6%) 16 13	3, 24, 54, 98	0

All (76) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	27	GLY	6.2
2	G	397	ILE	5.5
2	F	426	MET	5.4
1	D	31	THR	4.9
1	A	31	THR	4.9
2	F	421	ASN	4.4
2	G	396	MET	4.4
1	A	30	ARG	4.4
2	G	420	TYR	4.2
1	B	214	SER	4.1
2	G	411	VAL	4.1
1	A	32	ALA	4.0
1	C	214	SER	4.0
2	H	384	GLN	3.9
1	A	33	ALA	3.7
2	F	420	TYR	3.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
2	F	423	PHE	3.6
1	C	216	SER	3.6
2	G	419	LEU	3.6
1	D	33	ALA	3.6
2	E	434	SER	3.5
2	G	359	ILE	3.5
2	H	387	MET	3.4
2	G	355	ILE	3.4
1	D	23	HIS	3.4
2	F	410	GLN	3.3
1	D	214	SER	3.3
1	D	28	ILE	3.3
2	F	411	VAL	3.2
1	C	148	GLN	3.2
1	A	193	LEU	3.1
1	C	1	MET	3.1
1	A	27	GLY	3.0
1	D	193	LEU	3.0
1	D	29	PRO	3.0
1	D	30	ARG	3.0
1	D	26	PHE	3.0
2	G	417	THR	3.0
2	G	410	GLN	3.0
2	F	401	LYS	3.0
2	G	405	ARG	2.9
2	H	430	GLY	2.9
2	G	357	ALA	2.9
2	H	386	THR	2.9
1	A	41	ASP	2.8
1	D	4	ASN	2.8
2	G	381	ALA	2.8
2	F	424	LYS	2.8
1	D	32	ALA	2.7
1	C	41	ASP	2.7
2	F	397	ILE	2.7
2	H	388	GLN	2.7
2	F	425	ASN	2.6
2	H	410	GLN	2.5
1	B	1	MET	2.5
2	G	400	LEU	2.5
1	A	23	HIS	2.4
1	B	42	VAL	2.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	3	GLU	2.4
1	D	35	ASP	2.4
1	D	25	GLU	2.4
1	D	41	ASP	2.4
2	F	396	MET	2.4
2	G	358	GLU	2.4
1	D	5	ILE	2.4
1	C	2	VAL	2.4
1	B	2	VAL	2.3
2	G	375	GLU	2.3
2	F	398	THR	2.3
2	E	439	LEU	2.2
1	B	215	LYS	2.2
1	D	34	GLU	2.2
2	H	431	GLU	2.1
2	E	388	GLN	2.1
1	A	36	ILE	2.1
2	E	405	ARG	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
4	GOL	D	221	6/6	0.51	0.29	87,89,89,89	0
5	PO4	H	2	5/5	0.81	0.18	111,111,111,112	0
4	GOL	A	222	6/6	0.83	0.18	87,88,88,88	0
5	PO4	B	221	5/5	0.88	0.37	102,103,103,104	0
5	PO4	E	1	5/5	0.90	0.18	105,105,105,106	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(\AA^2)	Q<0.9
4	GOL	A	221	6/6	0.93	0.40	81,83,84,84	0
4	GOL	C	221	6/6	0.94	0.45	60,61,62,62	0
3	ZN	D	220	1/1	0.98	0.06	34,34,34,34	0
3	ZN	C	220	1/1	0.99	0.08	20,20,20,20	0
3	ZN	A	220	1/1	0.99	0.06	30,30,30,30	0
3	ZN	B	220	1/1	0.99	0.10	18,18,18,18	0

6.5 Other polymers [i](#)

There are no such residues in this entry.