Full wwPDB X-ray Structure Validation Report

Feb 19, 2016 – 07:48 PM GMT

PDB ID	:	4U62
Title	:	Trichodysplasia spinulosa-associated polyomavirus (TSPyV) VP1 in complex with 3'-sialyllactose
Authors	:	Stroh, L.J.; Stehle, T.
Deposited on	:	2014-07-26
Resolution	:	1.55 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

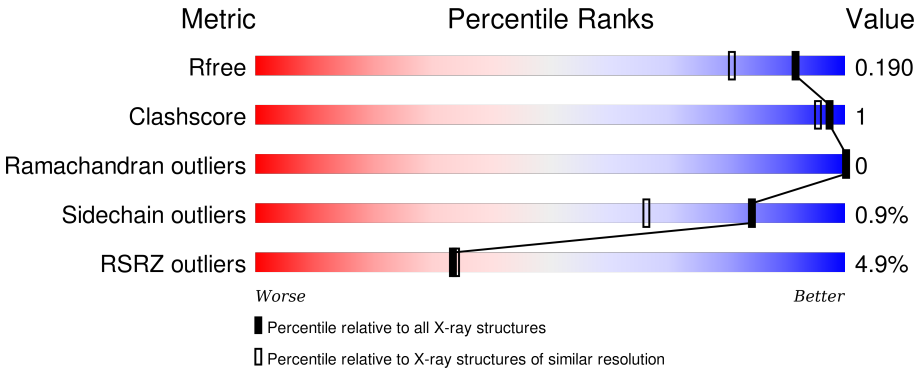
MolProbity	:	4.02b-467
Mogul	:	1.7.1 (RC1), CSD as537be (2016)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026982
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	rb-20026982

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 1.55 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}	91344	1665 (1.58-1.54)
Clashscore	102246	1014 (1.56-1.56)
Ramachandran outliers	100387	1704 (1.58-1.54)
Sidechain outliers	100360	1702 (1.58-1.54)
RSRZ outliers	91569	1668 (1.58-1.54)

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. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=5% The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density. The numeric value is given above the bar.

Mol	Chain	Length	Quality of chain
1	A	280	93%
1	B	280	3% 93%
1	C	280	5% 90% 6%
1	D	280	8% 94%
1	E	280	2% 94%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	280	13% 93% 5%
1	G	280	5% 92% 5%
1	H	280	% 94%
1	I	280	4% 91% 5%
1	J	280	5% 93%

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	GOL	A	401	-	-	-	X
2	GOL	C	401	-	-	-	X
2	GOL	E	401	-	-	-	X
2	GOL	E	402	-	-	-	X
2	GOL	H	401	-	-	-	X
2	GOL	I	401	-	-	-	X
2	GOL	J	401	-	-	-	X
3	SIA	B	401	-	-	-	X
3	SIA	C	402	-	-	-	X
3	SIA	J	402	-	-	-	X
4	EDO	G	401	-	-	-	X

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 23356 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 Structural protein VP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	272	Total	C	N	O	S	0	2	0
			2115	1331	353	417	14			
1	B	272	Total	C	N	O	S	0	3	0
			2123	1336	355	417	15			
1	C	268	Total	C	N	O	S	0	1	0
			2065	1303	342	406	14			
1	D	270	Total	C	N	O	S	0	1	0
			2092	1315	354	409	14			
1	E	271	Total	C	N	O	S	0	0	0
			2106	1323	352	417	14			
1	F	266	Total	C	N	O	S	0	1	0
			2053	1295	343	402	13			
1	G	267	Total	C	N	O	S	0	0	0
			2062	1297	348	404	13			
1	H	272	Total	C	N	O	S	0	2	0
			2119	1333	355	417	14			
1	I	266	Total	C	N	O	S	0	2	0
			2056	1300	343	399	14			
1	J	271	Total	C	N	O	S	0	1	0
			2101	1322	349	416	14			

There are 60 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	24	GLY	-	expression tag	UNP E2ESL7
A	25	SER	-	expression tag	UNP E2ESL7
A	26	HIS	-	expression tag	UNP E2ESL7
A	27	MET	-	expression tag	UNP E2ESL7
A	28	ALA	-	expression tag	UNP E2ESL7
A	29	SER	-	expression tag	UNP E2ESL7
B	24	GLY	-	expression tag	UNP E2ESL7
B	25	SER	-	expression tag	UNP E2ESL7
B	26	HIS	-	expression tag	UNP E2ESL7

Continued on next page...

Continued from previous page...

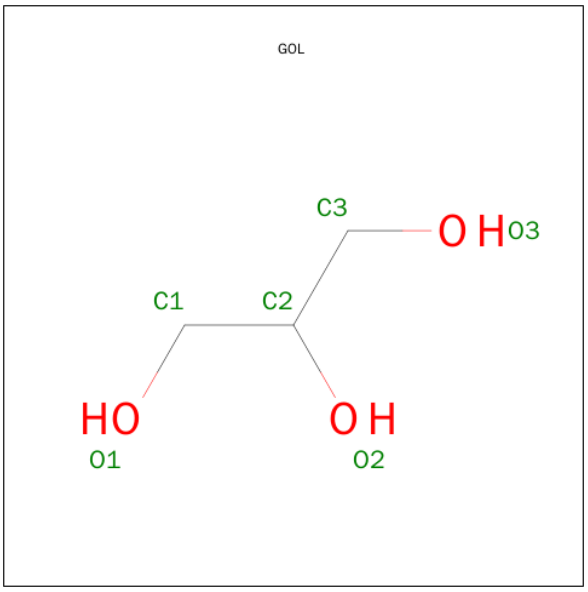
Chain	Residue	Modelled	Actual	Comment	Reference
B	27	MET	-	expression tag	UNP E2ESL7
B	28	ALA	-	expression tag	UNP E2ESL7
B	29	SER	-	expression tag	UNP E2ESL7
C	24	GLY	-	expression tag	UNP E2ESL7
C	25	SER	-	expression tag	UNP E2ESL7
C	26	HIS	-	expression tag	UNP E2ESL7
C	27	MET	-	expression tag	UNP E2ESL7
C	28	ALA	-	expression tag	UNP E2ESL7
C	29	SER	-	expression tag	UNP E2ESL7
D	24	GLY	-	expression tag	UNP E2ESL7
D	25	SER	-	expression tag	UNP E2ESL7
D	26	HIS	-	expression tag	UNP E2ESL7
D	27	MET	-	expression tag	UNP E2ESL7
D	28	ALA	-	expression tag	UNP E2ESL7
D	29	SER	-	expression tag	UNP E2ESL7
E	24	GLY	-	expression tag	UNP E2ESL7
E	25	SER	-	expression tag	UNP E2ESL7
E	26	HIS	-	expression tag	UNP E2ESL7
E	27	MET	-	expression tag	UNP E2ESL7
E	28	ALA	-	expression tag	UNP E2ESL7
E	29	SER	-	expression tag	UNP E2ESL7
F	24	GLY	-	expression tag	UNP E2ESL7
F	25	SER	-	expression tag	UNP E2ESL7
F	26	HIS	-	expression tag	UNP E2ESL7
F	27	MET	-	expression tag	UNP E2ESL7
F	28	ALA	-	expression tag	UNP E2ESL7
F	29	SER	-	expression tag	UNP E2ESL7
G	24	GLY	-	expression tag	UNP E2ESL7
G	25	SER	-	expression tag	UNP E2ESL7
G	26	HIS	-	expression tag	UNP E2ESL7
G	27	MET	-	expression tag	UNP E2ESL7
G	28	ALA	-	expression tag	UNP E2ESL7
G	29	SER	-	expression tag	UNP E2ESL7
H	24	GLY	-	expression tag	UNP E2ESL7
H	25	SER	-	expression tag	UNP E2ESL7
H	26	HIS	-	expression tag	UNP E2ESL7
H	27	MET	-	expression tag	UNP E2ESL7
H	28	ALA	-	expression tag	UNP E2ESL7
H	29	SER	-	expression tag	UNP E2ESL7
I	24	GLY	-	expression tag	UNP E2ESL7
I	25	SER	-	expression tag	UNP E2ESL7
I	26	HIS	-	expression tag	UNP E2ESL7

Continued on next page...

Continued from previous page...

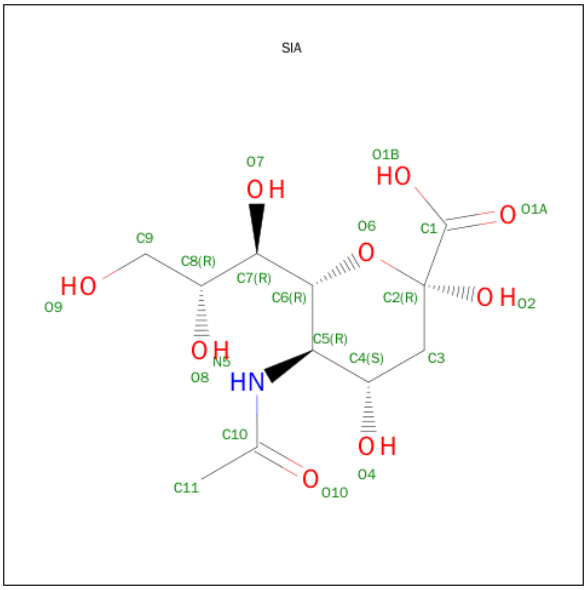
Chain	Residue	Modelled	Actual	Comment	Reference
I	27	MET	-	expression tag	UNP E2ESL7
I	28	ALA	-	expression tag	UNP E2ESL7
I	29	SER	-	expression tag	UNP E2ESL7
J	24	GLY	-	expression tag	UNP E2ESL7
J	25	SER	-	expression tag	UNP E2ESL7
J	26	HIS	-	expression tag	UNP E2ESL7
J	27	MET	-	expression tag	UNP E2ESL7
J	28	ALA	-	expression tag	UNP E2ESL7
J	29	SER	-	expression tag	UNP E2ESL7

- Molecule 2 is GLYCEROL (three-letter code: GOL) (formula: $C_3H_8O_3$).



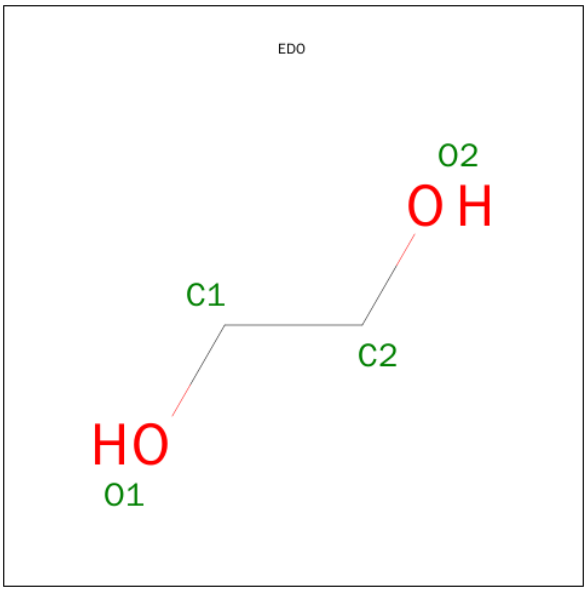
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 6 3 3	0	0
2	C	1	Total C O 6 3 3	0	0
2	E	1	Total C O 6 3 3	0	0
2	E	1	Total C O 6 3 3	0	0
2	H	1	Total C O 6 3 3	0	0
2	I	1	Total C O 6 3 3	0	0
2	J	1	Total C O 6 3 3	0	0

- Molecule 3 is O-SIALIC ACID (three-letter code: SIA) (formula: $C_{11}H_{19}NO_9$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	1	Total C N O 21 11 1 9	0	0
3	C	1	Total C N O 21 11 1 9	0	0
3	J	1	Total C N O 21 11 1 9	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



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

- Molecule 5 is water.

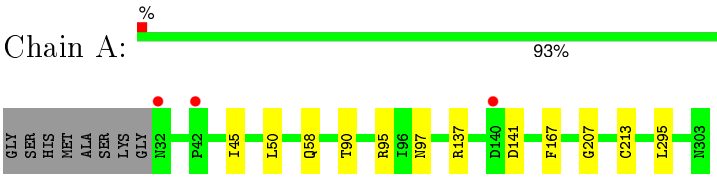
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	276	Total O 276 276	0	0
5	B	264	Total O 264 264	0	0
5	C	225	Total O 225 225	0	0
5	D	207	Total O 207 207	0	0
5	E	292	Total O 292 292	0	0
5	F	179	Total O 179 179	0	0
5	G	222	Total O 222 222	0	0
5	H	253	Total O 253 253	0	0
5	I	224	Total O 224 224	0	0
5	J	213	Total O 213 213	0	0

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 errors 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.

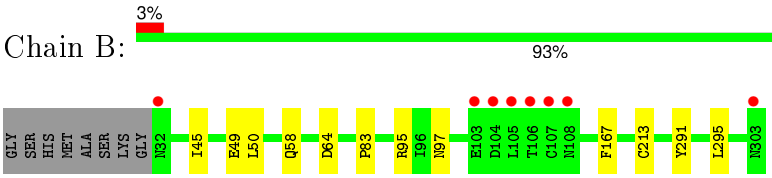
- Molecule 1: Structural protein VP1

Chain A:



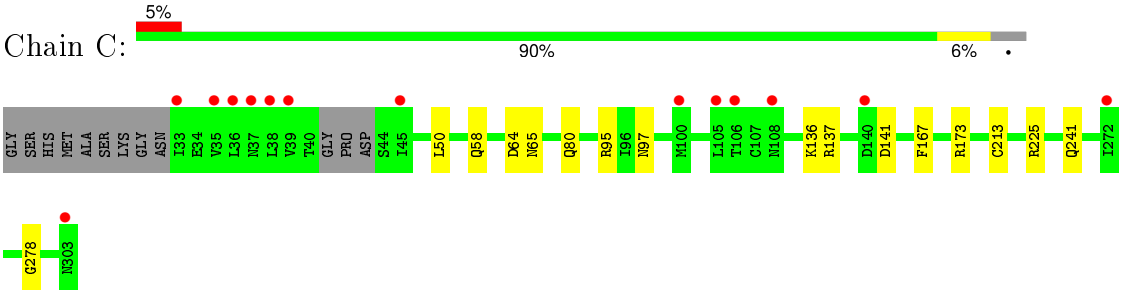
- Molecule 1: Structural protein VP1

Chain B:



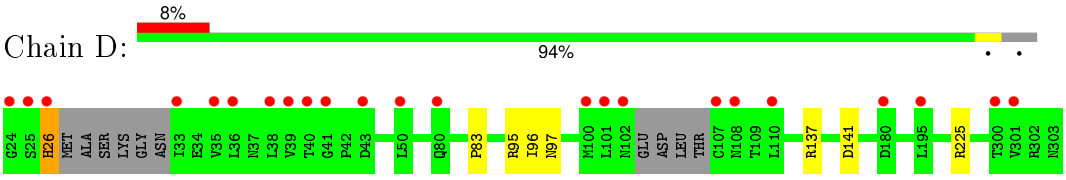
- Molecule 1: Structural protein VP1

Chain C:



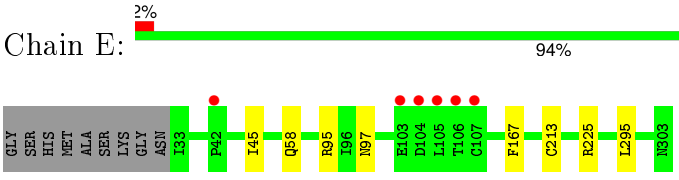
- Molecule 1: Structural protein VP1

Chain D:



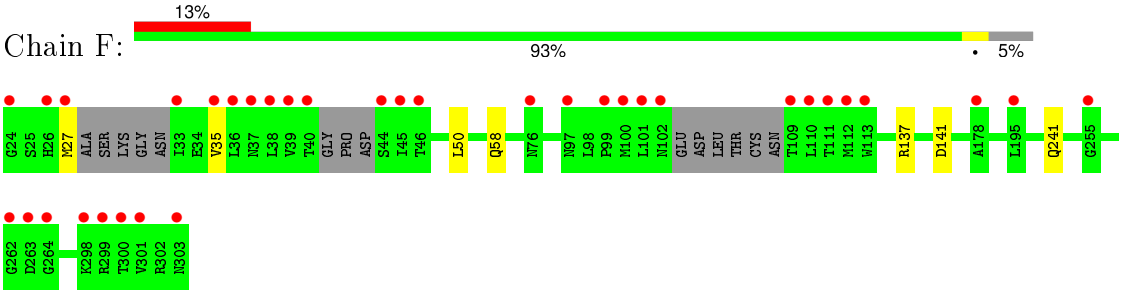
- Molecule 1: Structural protein VP1

Chain E:



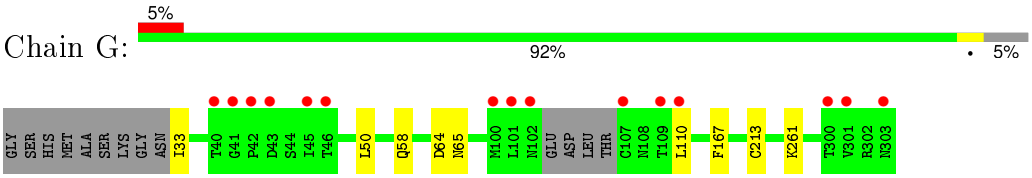
• Molecule 1: Structural protein VP1

Chain F:



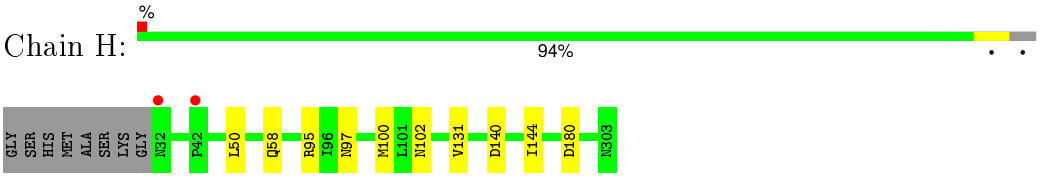
• Molecule 1: Structural protein VP1

Chain G:



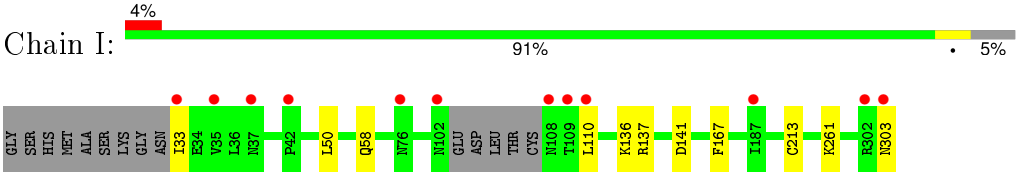
• Molecule 1: Structural protein VP1

Chain H:



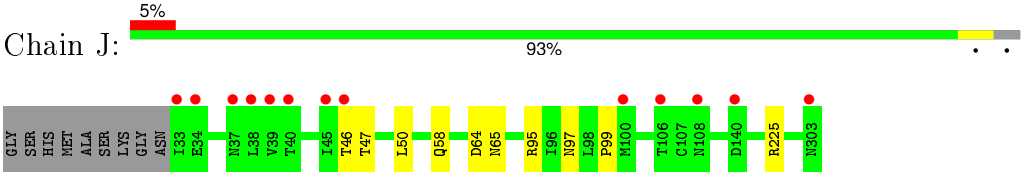
• Molecule 1: Structural protein VP1

Chain I:



• Molecule 1: Structural protein VP1

Chain J:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	66.24Å 152.84Å 147.24Å 90.00° 92.34° 90.00°	Depositor
Resolution (Å)	40.00 – 1.55 39.18 – 1.55	Depositor EDS
% Data completeness (in resolution range)	97.8 (40.00-1.55) 97.8 (39.18-1.55)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.69 (at 1.55Å)	Xtriage
Refinement program	REFMAC 5.8.0049	Depositor
R, R_{free}	0.155 , 0.180 0.167 , 0.190	Depositor DCC
R_{free} test set	20677 reflections (5.27%)	DCC
Wilson B-factor (Å^2)	20.9	Xtriage
Anisotropy	0.380	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 39.8	EDS
Estimated twinning fraction	0.012 for -h,-l,-k 0.006 for -h,l,k 0.039 for h,-k,-l	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 412698 reflections	Xtriage
F_o,F_c correlation	0.97	EDS
Total number of atoms	23356	wwPDB-VP
Average B, all atoms (Å^2)	28.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.19% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: GOL, SIA, EDO

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	0.50	0/2166	0.70	0/2946
1	B	0.47	0/2177	0.68	0/2958
1	C	0.45	0/2111	0.69	1/2871 (0.0%)
1	D	0.72	1/2139 (0.0%)	0.68	2/2904 (0.1%)
1	E	0.50	0/2151	0.71	2/2925 (0.1%)
1	F	0.43	0/2098	0.65	0/2848
1	G	0.43	0/2106	0.65	0/2864
1	H	0.45	0/2170	0.68	0/2949
1	I	0.45	0/2106	0.68	0/2861
1	J	0.46	0/2149	0.70	1/2923 (0.0%)
All	All	0.49	1/21373 (0.0%)	0.68	6/29049 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	26	HIS	C-O	25.64	1.72	1.23

All (6) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	26	HIS	CA-C-O	-6.67	106.10	120.10
1	E	225	ARG	NE-CZ-NH1	6.10	123.35	120.30
1	C	225	ARG	NE-CZ-NH2	-5.32	117.64	120.30
1	E	225	ARG	NE-CZ-NH2	-5.31	117.65	120.30
1	D	225	ARG	NE-CZ-NH1	5.19	122.89	120.30
1	J	225	ARG	NE-CZ-NH1	5.04	122.82	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	2115	0	2061	5	0
1	B	2123	0	2079	5	0
1	C	2065	0	1993	8	0
1	D	2092	0	2038	5	0
1	E	2106	0	2047	3	0
1	F	2053	0	1990	8	0
1	G	2062	0	2001	3	0
1	H	2119	0	2072	6	0
1	I	2056	0	2012	5	0
1	J	2101	0	2041	8	0
2	A	6	0	8	0	0
2	C	6	0	8	0	0
2	E	12	0	16	0	0
2	H	6	0	8	0	0
2	I	6	0	8	0	0
2	J	6	0	8	0	0
3	B	21	0	18	0	0
3	C	21	0	18	0	0
3	J	21	0	18	0	0
4	G	4	0	6	1	0
5	A	276	0	0	0	0
5	B	264	0	0	0	0
5	C	225	0	0	2	0
5	D	207	0	0	0	0
5	E	292	0	0	0	0
5	F	179	0	0	1	0
5	G	222	0	0	1	0
5	H	253	0	0	2	0
5	I	224	0	0	1	0
5	J	213	0	0	0	0
All	All	23356	0	20450	50	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.

All (50) 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:D:26:HIS:C	1:D:26:HIS:O	1.72	1.28
1:E:95:ARG:NH1	1:E:97:ASN:OD1	2.14	0.80
1:C:95:ARG:NH1	1:C:97:ASN:OD1	2.28	0.67
1:J:95:ARG:NH1	1:J:97:ASN:OD1	2.29	0.66
1:B:95:ARG:NH1	1:B:97:ASN:OD1	2.32	0.62
1:F:241:GLN:CG	5:F:540:HOH:O	2.48	0.62
1:F:27:MET:HE3	1:J:99:PRO:HG2	1.83	0.61
1:E:45:ILE:CG2	1:E:295:LEU:HD21	2.39	0.53
1:F:27:MET:HG3	1:J:47:THR:O	2.09	0.52
1:B:64:ASP:OD2	1:C:173:ARG:NH2	2.41	0.52
1:G:33:ILE:HD11	1:G:261:LYS:HE2	1.90	0.52
1:H:95:ARG:NH1	1:H:97:ASN:OD1	2.44	0.50
1:C:64:ASP:O	1:C:65:ASN:HB2	2.11	0.50
1:H:180:ASP:HB3	5:H:529:HOH:O	2.13	0.49
1:F:137:ARG:HG3	1:F:141:ASP:HA	1.93	0.49
1:H:100:MET:CE	1:H:102:ASN:OD1	2.61	0.48
1:F:27:MET:HE3	1:J:99:PRO:CG	2.43	0.47
1:B:45:ILE:CG2	1:B:295:LEU:HD11	2.44	0.47
1:I:137:ARG:HG3	1:I:141:ASP:HA	1.96	0.47
1:E:167:PHE:CE1	1:E:213:CYS:HB2	2.50	0.46
1:I:136:LYS:HE2	5:I:716:HOH:O	2.13	0.46
1:F:27:MET:HE1	1:J:99:PRO:HG3	1.98	0.46
1:C:167:PHE:CE1	1:C:213:CYS:HB2	2.51	0.46
1:A:45:ILE:CG2	1:A:295:LEU:HD11	2.45	0.46
1:C:137:ARG:HG3	1:C:141:ASP:HA	1.98	0.46
1:F:27:MET:CE	1:J:99:PRO:CG	2.94	0.46
1:G:64:ASP:O	1:G:65:ASN:HB2	2.17	0.45
1:J:64:ASP:O	1:J:65:ASN:HB2	2.17	0.45
1:D:26:HIS:CA	1:D:26:HIS:O	2.60	0.45
1:H:100:MET:HE2	1:H:102:ASN:OD1	2.17	0.44
1:G:167:PHE:CE1	1:G:213:CYS:HB2	2.52	0.44
1:C:80:GLN:HG2	5:C:525:HOH:O	2.17	0.44
1:F:27:MET:SD	1:J:46:THR:OG1	2.64	0.43
1:A:137:ARG:HG3	1:A:141:ASP:HA	2.00	0.43
1:D:96:ILE:HD12	1:D:96:ILE:N	2.34	0.42
1:C:241:GLN:CG	5:C:718:HOH:O	2.67	0.42
1:I:110:LEU:HD23	1:I:303:ASN:CB	2.50	0.42
1:A:95:ARG:NH1	1:A:97:ASN:OD1	2.52	0.42
4:G:401:EDO:H22	5:G:588:HOH:O	2.20	0.42
1:I:167:PHE:CE1	1:I:213:CYS:HB2	2.55	0.42
1:D:137:ARG:HG3	1:D:141:ASP:HA	2.00	0.42
1:H:131:VAL:HB	1:H:144:ILE:HB	2.02	0.41

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:I:33:ILE:HD11	1:I:261:LYS:HE2	2.01	0.41
1:C:136:LYS:HG3	1:C:278:GLY:O	2.20	0.41
1:H:140:ASP:HB3	5:H:752:HOH:O	2.21	0.41
1:D:95:ARG:NH1	1:D:97:ASN:OD1	2.52	0.41
1:B:49:GLU:OE2	1:B:291:TYR:OH	2.25	0.41
1:B:167:PHE:CE1	1:B:213:CYS:HB2	2.55	0.41
1:A:167:PHE:CE1	1:A:213:CYS:HB2	2.56	0.40
1:A:90:THR:OG1	1:A:207:GLY:HA2	2.22	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	272/280 (97%)	262 (96%)	10 (4%)	0	100	100
1	B	273/280 (98%)	265 (97%)	8 (3%)	0	100	100
1	C	265/280 (95%)	258 (97%)	7 (3%)	0	100	100
1	D	265/280 (95%)	257 (97%)	8 (3%)	0	100	100
1	E	269/280 (96%)	262 (97%)	7 (3%)	0	100	100
1	F	259/280 (92%)	252 (97%)	7 (3%)	0	100	100
1	G	263/280 (94%)	256 (97%)	7 (3%)	0	100	100
1	H	272/280 (97%)	264 (97%)	8 (3%)	0	100	100
1	I	264/280 (94%)	257 (97%)	7 (3%)	0	100	100
1	J	270/280 (96%)	262 (97%)	8 (3%)	0	100	100
All	All	2672/2800 (95%)	2595 (97%)	77 (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	239/246 (97%)	237 (99%)	2 (1%)	86	71
1	B	241/246 (98%)	238 (99%)	3 (1%)	78	54
1	C	229/246 (93%)	227 (99%)	2 (1%)	84	66
1	D	236/246 (96%)	235 (100%)	1 (0%)	93	84
1	E	238/246 (97%)	237 (100%)	1 (0%)	93	84
1	F	229/246 (93%)	226 (99%)	3 (1%)	76	51
1	G	231/246 (94%)	228 (99%)	3 (1%)	76	51
1	H	240/246 (98%)	238 (99%)	2 (1%)	86	71
1	I	230/246 (94%)	228 (99%)	2 (1%)	84	66
1	J	237/246 (96%)	235 (99%)	2 (1%)	86	71
All	All	2350/2460 (96%)	2329 (99%)	21 (1%)	84	66

All (21) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	50	LEU
1	A	58	GLN
1	B	50	LEU
1	B	58	GLN
1	B	83	PRO
1	C	50	LEU
1	C	58	GLN
1	D	83	PRO
1	E	58	GLN
1	F	35	VAL
1	F	50	LEU
1	F	58	GLN
1	G	50	LEU
1	G	58	GLN
1	G	110	LEU
1	H	50	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	H	58	GLN
1	I	50	LEU
1	I	58	GLN
1	J	50	LEU
1	J	58	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

11 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	$\#\|Z\| > 2$
2	GOL	A	401	-	5,5,5	0.44	0	5,5,5	0.66	0
3	SIA	B	401	-	18,21,21	1.11	1 (5%)	18,31,31	1.06	1 (5%)
2	GOL	C	401	-	5,5,5	0.37	0	5,5,5	0.57	0
3	SIA	C	402	-	18,21,21	1.12	1 (5%)	18,31,31	0.64	0
2	GOL	E	401	-	5,5,5	0.28	0	5,5,5	0.55	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	GOL	E	402	-	5,5,5	0.32	0	5,5,5	0.54	0
4	EDO	G	401	-	3,3,3	0.30	0	2,2,2	0.65	0
2	GOL	H	401	-	5,5,5	0.33	0	5,5,5	0.47	0
2	GOL	I	401	-	5,5,5	0.49	0	5,5,5	0.76	0
2	GOL	J	401	-	5,5,5	0.26	0	5,5,5	0.45	0
3	SIA	J	402	-	18,21,21	1.20	1 (5%)	18,31,31	0.97	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
2	GOL	A	401	-	-	0/4/4/4	0/0/0/0
3	SIA	B	401	-	-	0/14/38/38	0/1/1/1
2	GOL	C	401	-	-	0/4/4/4	0/0/0/0
3	SIA	C	402	-	-	0/14/38/38	0/1/1/1
2	GOL	E	401	-	-	0/4/4/4	0/0/0/0
2	GOL	E	402	-	-	0/4/4/4	0/0/0/0
4	EDO	G	401	-	-	0/1/1/1	0/0/0/0
2	GOL	H	401	-	-	0/4/4/4	0/0/0/0
2	GOL	I	401	-	-	0/4/4/4	0/0/0/0
2	GOL	J	401	-	-	0/4/4/4	0/0/0/0
3	SIA	J	402	-	-	0/14/38/38	0/1/1/1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	402	SIA	O2-C2	3.97	1.44	1.39
3	B	401	SIA	O2-C2	4.03	1.44	1.39
3	J	402	SIA	O2-C2	4.05	1.44	1.39

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	401	SIA	O2-C2-C3	-2.85	105.89	109.46

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	G	401	EDO	1	0

5.7 Other polymers

There are no such residues in this entry.

5.8 Polymer linkage issues

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95^{th} 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	272/280 (97%)	-0.30	3 (1%) 82 84	15, 21, 35, 55	0
1	B	272/280 (97%)	-0.17	8 (2%) 55 59	16, 23, 44, 81	0
1	C	268/280 (95%)	0.01	14 (5%) 31 31	17, 25, 47, 66	0
1	D	270/280 (96%)	0.10	23 (8%) 13 13	17, 26, 62, 83	0
1	E	271/280 (96%)	-0.34	6 (2%) 65 68	15, 22, 39, 56	0
1	F	266/280 (95%)	0.43	35 (13%) 4 4	19, 31, 60, 81	0
1	G	267/280 (95%)	-0.05	15 (5%) 28 28	17, 28, 57, 79	0
1	H	272/280 (97%)	-0.27	2 (0%) 89 90	17, 23, 38, 51	0
1	I	266/280 (95%)	-0.13	12 (4%) 37 38	15, 24, 47, 73	0
1	J	271/280 (96%)	-0.03	13 (4%) 34 35	16, 24, 59, 74	0
All	All	2695/2800 (96%)	-0.08	131 (4%) 33 34	15, 24, 50, 83	0

All (131) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	36	LEU	9.5
1	B	105	LEU	8.4
1	B	107	CYS	8.0
1	F	39	VAL	7.7
1	F	110	LEU	7.5
1	B	106	THR	7.2
1	F	33	ILE	6.5
1	F	40	THR	6.2
1	F	38	LEU	6.2
1	G	110	LEU	6.1
1	F	301	VAL	6.1
1	I	303	ASN	6.0
1	F	101	LEU	6.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	F	102	ASN	5.8
1	D	33	ILE	5.3
1	C	106	THR	5.2
1	F	35	VAL	5.1
1	B	303	ASN	5.0
1	F	109	THR	4.7
1	F	113	TRP	4.7
1	C	105	LEU	4.6
1	D	41	GLY	4.6
1	F	111	THR	4.5
1	F	37	ASN	4.4
1	I	109	THR	4.3
1	G	101	LEU	4.2
1	I	33	ILE	4.2
1	G	107	CYS	4.1
1	G	45	ILE	4.0
1	C	303	ASN	4.0
1	D	26	HIS	4.0
1	F	45	ILE	4.0
1	D	40	THR	4.0
1	C	36	LEU	3.9
1	G	42	PRO	3.9
1	J	37	ASN	3.9
1	J	303	ASN	3.8
1	C	38	LEU	3.8
1	C	35	VAL	3.8
1	J	106	THR	3.6
1	G	43	ASP	3.5
1	I	108	ASN	3.5
1	A	42	PRO	3.5
1	C	39	VAL	3.5
1	D	107	CYS	3.5
1	D	38	LEU	3.5
1	D	110	LEU	3.5
1	F	178	ALA	3.5
1	J	40	THR	3.5
1	C	45	ILE	3.4
1	E	106	THR	3.4
1	F	300	THR	3.4
1	I	110	LEU	3.4
1	D	100	MET	3.4
1	E	103	GLU	3.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	24	GLY	3.4
1	C	33	ILE	3.3
1	D	39	VAL	3.3
1	F	303	ASN	3.3
1	D	50	LEU	3.3
1	I	35	VAL	3.2
1	E	107	CYS	3.2
1	F	262	GLY	3.1
1	J	33	ILE	3.1
1	H	42	PRO	3.0
1	F	27	MET	3.0
1	G	109	THR	3.0
1	E	42	PRO	3.0
1	B	32	ASN	3.0
1	G	300	THR	2.9
1	D	101	LEU	2.9
1	E	105	LEU	2.9
1	B	104	ASP	2.8
1	J	108	ASN	2.8
1	J	45	ILE	2.7
1	G	40	THR	2.7
1	D	25	SER	2.7
1	D	36	LEU	2.7
1	J	39	VAL	2.7
1	I	42	PRO	2.7
1	J	34	GLU	2.7
1	G	100	MET	2.6
1	E	104	ASP	2.6
1	A	140	ASP	2.6
1	J	140	ASP	2.6
1	F	99	PRO	2.6
1	G	102	ASN	2.6
1	F	195	LEU	2.6
1	F	44	SER	2.5
1	H	32	ASN	2.5
1	B	103	GLU	2.5
1	A	32	ASN	2.4
1	I	302	ARG	2.4
1	F	264	GLY	2.4
1	D	43	ASP	2.4
1	D	102	ASN	2.4
1	D	35	VAL	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	G	301	VAL	2.3
1	I	76	ASN	2.3
1	J	100	MET	2.3
1	F	24	GLY	2.3
1	F	255	GLY	2.3
1	C	100	MET	2.3
1	D	195	LEU	2.3
1	F	299	ARG	2.2
1	D	300	THR	2.2
1	D	301	VAL	2.2
1	F	298	LYS	2.2
1	D	80	GLN	2.2
1	C	272	ILE	2.1
1	I	187	ILE	2.1
1	F	100	MET	2.1
1	J	46	THR	2.1
1	G	41	GLY	2.1
1	C	37	ASN	2.1
1	I	102	ASN	2.1
1	D	180	ASP	2.1
1	B	108	ASN	2.1
1	F	97	ASN	2.1
1	F	112	MET	2.1
1	C	108	ASN	2.1
1	C	140	ASP	2.1
1	G	46	THR	2.1
1	J	38	LEU	2.1
1	D	108	ASN	2.0
1	F	26	HIS	2.0
1	G	303	ASN	2.0
1	I	37	ASN	2.0
1	F	263	ASP	2.0
1	F	46	THR	2.0
1	F	76	ASN	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 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. LLDF column lists the quality of electron density of the group with respect to its neighbouring residues in protein, DNA or RNA chains. The B-factors column lists the minimum, median, 95^{th} 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	LLDF	B-factors(Å^2)	Q<0.9
2	GOL	E	402	6/6	0.79	0.18	9.29	34,36,43,52	0
2	GOL	J	401	6/6	0.86	0.18	6.33	34,42,46,49	0
3	SIA	J	402	21/21	0.78	0.27	5.67	27,35,44,53	0
2	GOL	H	401	6/6	0.86	0.12	5.50	27,32,36,38	0
4	EDO	G	401	4/4	0.89	0.12	5.40	29,31,33,48	0
2	GOL	I	401	6/6	0.85	0.12	4.73	29,32,39,47	0
3	SIA	C	402	21/21	0.79	0.26	4.68	33,41,52,53	0
2	GOL	E	401	6/6	0.88	0.11	4.67	26,31,37,38	0
2	GOL	C	401	6/6	0.91	0.12	3.12	33,36,41,59	0
2	GOL	A	401	6/6	0.83	0.11	2.74	28,33,35,37	0
3	SIA	B	401	21/21	0.79	0.21	2.42	27,34,38,44	21

6.5 Other polymers (i)

There are no such residues in this entry.