Full wwPDB X-ray Structure Validation Report (i)

Jan 31, 2016 – 09:55 PM GMT

PDB ID	:	1R8H
Title	:	Comparison of the structure and DNA binding properties of the E2 proteins from an oncogenic and a non-oncogenic human papillomavirus
Authors	:	Dell, G.; Wilkinson, K.W.; Tranter, R.; Parish, J.; Brady, R.L.; Gaston, K.
Deposited on	:	2003-10-24
Resolution	:	1.90 Å(reported)

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

|		
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
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)	:	trunk26865

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 1.90 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.
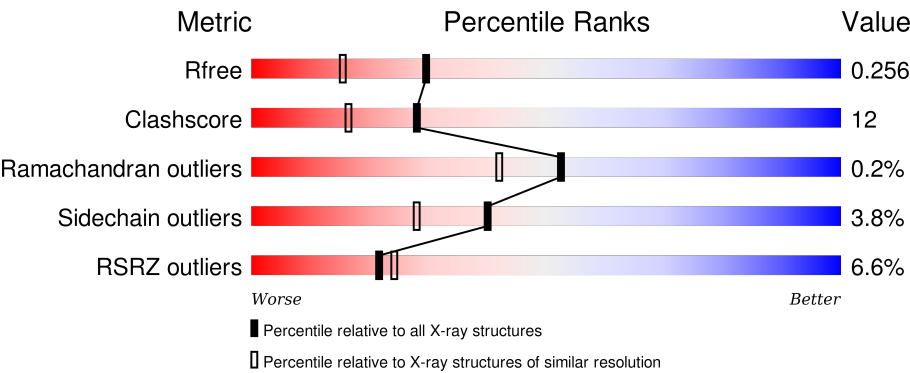
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	4755 (1.90-1.90)
Clashscore	102246	5398 (1.90-1.90)
Ramachandran outliers	100387	5338 (1.90-1.90)
Sidechain outliers	100360	5339 (1.90-1.90)
RSRZ outliers	91569	4766 (1.90-1.90)

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	87	6% / 79% / 16% / 5%
1	B	87	% / 86% / 11% / •
1	C	87	15% / 61% / 32% / • / 5%
1	D	87	% / 85% / 14% / •
1	E	87	2% / 78% / 18% / •

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	87	

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	PO4	E	505	-	-	X	-

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 5125 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 Regulatory protein E2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	87	Total	C	N	O	S	30	4	0
			748	480	139	125	4			
1	B	87	Total	C	N	O	S	20	4	0
			748	480	139	125	4			
1	C	83	Total	C	N	O	S	14	21	0
			866	551	164	145	6			
1	D	87	Total	C	N	O	S	26	0	0
			722	463	134	121	4			
1	E	87	Total	C	N	O	S	23	0	0
			722	463	134	121	4			
1	F	83	Total	C	N	O	S	14	21	0
			866	551	164	145	6			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	363	MET	LEU	SEE REMARK 999	UNP Q84294
B	363	MET	LEU	SEE REMARK 999	UNP Q84294
C	363	MET	LEU	SEE REMARK 999	UNP Q84294
D	363	MET	LEU	SEE REMARK 999	UNP Q84294
E	363	MET	LEU	SEE REMARK 999	UNP Q84294
F	363	MET	LEU	SEE REMARK 999	UNP Q84294

- Molecule 2 is PHOSPHATE ION (three-letter code: PO4) (formula: O_4P).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	D	1	Total 5	O 4	P 1	0	0
2	E	1	Total 5	O 4	P 1	0	0
2	D	1	Total 5	O 4	P 1	0	0
2	D	1	Total 5	O 4	P 1	0	0
2	E	1	Total 5	O 4	P 1	0	0
2	E	1	Total 5	O 4	P 1	0	0
2	F	1	Total 5	O 4	P 1	0	0
2	A	1	Total 5	O 4	P 1	0	0
2	B	1	Total 5	O 4	P 1	0	0
2	A	1	Total 5	O 4	P 1	0	0
2	B	1	Total 5	O 4	P 1	0	0
2	C	1	Total 5	O 4	P 1	0	0
2	F	1	Total 5	O 4	P 1	0	0
2	C	1	Total 5	O 4	P 1	0	0

- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	74	Total 74	O 74	0	0
3	B	69	Total 69	O 69	0	0
3	C	48	Total 48	O 48	0	0
3	D	68	Total 68	O 68	0	0
3	E	75	Total 75	O 75	0	0
3	F	49	Total 49	O 49	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.
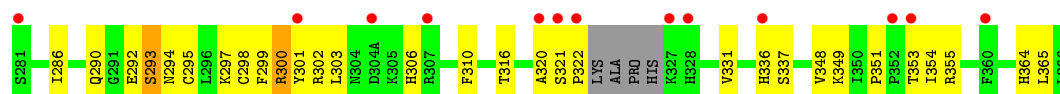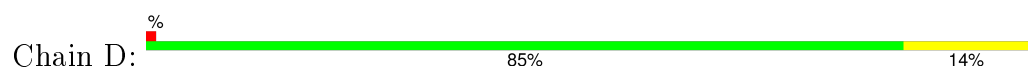
- Molecule 1: Regulatory protein E2
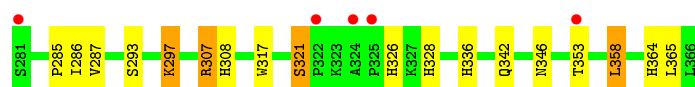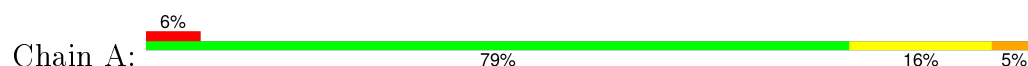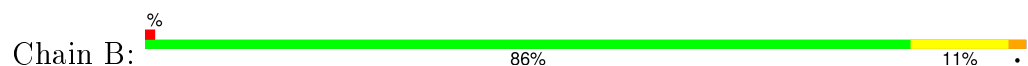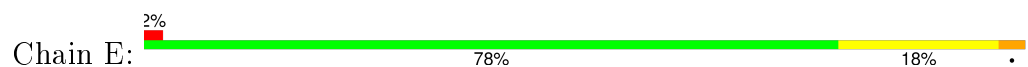

- Molecule 1: Regulatory protein E2

- Molecule 1: Regulatory protein E2

- Molecule 1: Regulatory protein E2

- Molecule 1: Regulatory protein E2

- Molecule 1: Regulatory protein E2

Chain F:
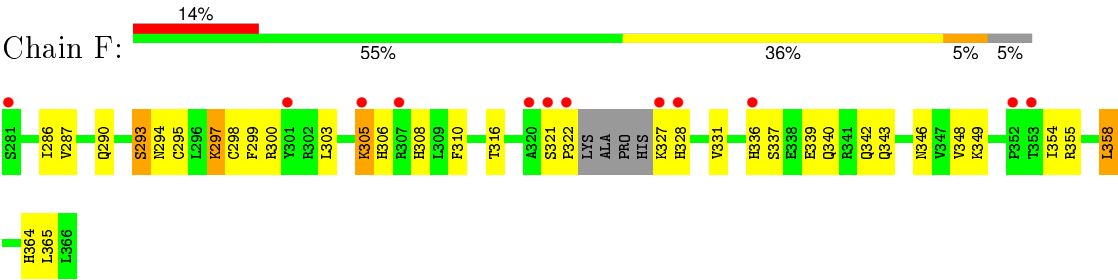
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, α, β, γ	71.66Å　　71.66Å　　195.00Å 90.00°　　90.00°　　120.00°	Depositor
Resolution (Å)	20.00　−　1.90 19.71　−　1.90	Depositor EDS
% Data completeness (in resolution range)	93.4 (20.00-1.90) 93.4 (19.71-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	2.08 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.191　,　0.258 0.201　,　0.256	Depositor DCC
R_{free} test set	2143 reflections (5.44%)	DCC
Wilson B-factor (Å^2)	25.1	Xtriage
Anisotropy	0.627	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 45.8	EDS
Estimated twinning fraction	0.488 for h,-h-k,-l	Xtriage
L-test for twinning[2]	$<|L|> = 0.53$, $<L^2> = 0.37$	Xtriage
Outliers	5 of 42648 reflections (0.012%)	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	5125	wwPDB-VP
Average B, all atoms (Å^2)	37.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 12.85% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: PO4

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	0.93	1/779 (0.1%)	1.45	4/1055 (0.4%)
1	B	1.03	2/779 (0.3%)	0.91	2/1055 (0.2%)
1	C	0.93	2/901 (0.2%)	0.86	1/1211 (0.1%)
1	D	1.11	2/745 (0.3%)	1.08	6/1007 (0.6%)
1	E	1.03	2/745 (0.3%)	0.95	1/1007 (0.1%)
1	F	0.95	3/901 (0.3%)	0.89	2/1211 (0.2%)
All	All	1.00	12/4850 (0.2%)	1.04	16/6546 (0.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	1
1	B	0	1
1	D	0	1
1	E	0	1
All	All	0	4

All (12) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	300	ARG	CD-NE	-19.52	1.13	1.46
1	B	302	ARG	CD-NE	-15.84	1.19	1.46
1	E	307	ARG	CD-NE	-14.71	1.21	1.46
1	A	307	ARG	NE-CZ	-12.12	1.17	1.33
1	F	349	LYS	CG-CD	-11.24	1.14	1.52
1	C	349	LYS	CG-CD	-8.87	1.22	1.52
1	F	327	LYS	CB-CG	-6.71	1.34	1.52
1	B	321	SER	CA-CB	-6.65	1.43	1.52

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	E	355	ARG	CG-CD	-6.25	1.36	1.51
1	C	355	ARG	CD-NE	5.79	1.56	1.46
1	D	288	GLN	CB-CG	-5.56	1.37	1.52
1	F	355	ARG	CD-NE	5.36	1.55	1.46

All (16) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	307	ARG	NE-CZ-NH1	-23.87	108.36	120.30
1	A	307	ARG	NE-CZ-NH2	22.21	131.41	120.30
1	A	307	ARG	CD-NE-CZ	19.25	150.55	123.60
1	D	300	ARG	CD-NE-CZ	-13.39	104.85	123.60
1	E	307	ARG	CG-CD-NE	13.09	139.29	111.80
1	D	307	ARG	CG-CD-NE	-11.27	88.14	111.80
1	D	307	ARG	CB-CG-CD	9.40	136.03	111.60
1	B	321	SER	N-CA-CB	8.62	123.43	110.50
1	B	302	ARG	CG-CD-NE	8.06	128.73	111.80
1	F	355	ARG	CD-NE-CZ	-7.49	113.12	123.60
1	F	355	ARG	CG-CD-NE	-6.94	97.22	111.80
1	D	300	ARG	CG-CD-NE	-6.20	98.77	111.80
1	A	297	LYS	CB-CG-CD	5.98	127.15	111.60
1	D	311	ASP	CB-CG-OD2	5.59	123.33	118.30
1	C	355	ARG	CG-CD-NE	-5.18	100.91	111.80
1	D	304(A)	ASP	CB-CG-OD2	5.05	122.84	118.30

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	307	ARG	Sidechain
1	B	302	ARG	Sidechain
1	D	300	ARG	Sidechain
1	E	307	ARG	Sidechain

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	748	0	738	13	0
1	B	748	0	738	7	0
1	C	866	0	844	44	0
1	D	722	0	713	8	0
1	E	722	0	715	13	0
1	F	866	0	845	42	0
2	A	10	0	0	0	0
2	B	10	0	0	0	0
2	C	10	0	0	1	0
2	D	15	0	0	1	0
2	E	15	0	0	2	0
2	F	10	0	0	0	0
3	A	74	0	0	3	0
3	B	69	0	0	1	0
3	C	48	0	0	6	0
3	D	68	0	0	1	0
3	E	75	0	0	3	0
3	F	49	0	0	4	0
All	All	5125	0	4593	105	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 12.

All (105) 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:301[B]:TYR:HE1	1:F:294[B]:ASN:ND2	1.07	1.42
1:C:301[B]:TYR:CE1	1:F:294[B]:ASN:ND2	1.89	1.37
1:C:321:SER:OG	1:C:322:PRO:HD2	1.15	1.33
1:F:321:SER:OG	1:F:322:PRO:HD2	1.16	1.31
1:C:321:SER:OG	1:C:322:PRO:CD	1.94	1.15
1:F:321:SER:OG	1:F:322:PRO:CD	1.97	1.13
1:C:321:SER:HG	1:C:322:PRO:HD2	1.21	0.94
1:F:321:SER:HG	1:F:322:PRO:HD2	1.14	0.90
1:C:294[A]:ASN:OD1	3:F:534:HOH:O	1.89	0.90
1:F:294[A]:ASN:OD1	3:F:534:HOH:O	1.88	0.89
1:C:298[B]:CYS:HB2	1:F:298[B]:CYS:HB2	1.57	0.85
1:C:320:ALA:HB2	3:C:556:HOH:O	1.76	0.84
1:C:298[B]:CYS:SG	1:F:295[B]:CYS:HA	2.21	0.81
1:A:364:HIS:HE1	3:B:530:HOH:O	1.70	0.75
1:C:290[A]:GLN:NE2	3:C:519:HOH:O	2.21	0.73
1:C:295[B]:CYS:HA	1:F:298[B]:CYS:SG	2.28	0.73

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:F:293[B]:SER:OG	3:F:561:HOH:O	2.06	0.72
1:C:302[B]:ARG:CZ	3:C:544:HOH:O	2.41	0.69
1:D:326:HIS:HD2	1:D:328:HIS:H	1.41	0.69
1:C:299[B]:PHE:HB2	1:C:354:ILE:CD1	2.27	0.64
1:C:302[B]:ARG:NE	3:C:544:HOH:O	2.30	0.63
1:A:308:HIS:HE1	3:A:521:HOH:O	1.81	0.63
1:D:308:HIS:CD2	1:D:308:HIS:H	2.16	0.63
1:E:360:PHE:HB3	3:E:541:HOH:O	1.99	0.61
1:C:294[A]:ASN:O	1:F:294[A]:ASN:HB3	2.01	0.60
1:A:287:VAL:HG22	1:A:358:LEU:HD22	1.84	0.60
1:F:346:ASN:ND2	3:F:552:HOH:O	2.34	0.60
1:B:287:VAL:HG22	1:B:358:LEU:HD22	1.83	0.60
1:A:293:SER:O	1:A:297:LYS:HG3	2.03	0.58
1:F:299[B]:PHE:HB2	1:F:354:ILE:CD1	2.33	0.58
1:D:287:VAL:HG22	1:D:358:LEU:HD22	1.84	0.58
1:E:326:HIS:HE1	3:E:512:HOH:O	1.86	0.58
1:C:299[B]:PHE:CE1	1:C:351:PRO:HD2	2.39	0.57
1:E:326:HIS:HD2	1:E:328:HIS:H	1.51	0.57
1:E:287:VAL:HG22	1:E:358:LEU:HD22	1.86	0.56
1:C:297[B]:LYS:HG2	1:C:316:THR:HG21	1.88	0.56
1:A:286:ILE:HD11	1:B:365:LEU:CD1	2.35	0.56
1:C:294[A]:ASN:HB3	1:F:294[A]:ASN:O	2.05	0.55
1:C:320:ALA:CB	3:C:556:HOH:O	2.44	0.55
1:A:336[A]:HIS:NE2	3:A:538:HOH:O	2.27	0.54
1:F:287:VAL:HG22	1:F:358:LEU:HD22	1.89	0.53
1:E:357:LYS:HG2	1:F:364:HIS:CE1	2.44	0.53
1:B:326:HIS:HE1	1:B:330:ILE:HD11	1.75	0.52
1:C:298[B]:CYS:CB	1:F:298[B]:CYS:HB2	2.36	0.51
1:D:326:HIS:HE1	3:D:507:HOH:O	1.93	0.50
1:C:299[A]:PHE:CE2	1:C:303[A]:LEU:HD11	2.46	0.50
1:A:326:HIS:HD2	1:A:328:HIS:HB2	1.77	0.50
1:C:321:SER:CB	1:C:322:PRO:CD	2.85	0.49
1:F:290[B]:GLN:HE21	1:F:328:HIS:HD2	1.59	0.49
1:C:306[B]:HIS:NE2	2:C:514:PO4:O2	2.44	0.49
1:C:301[B]:TYR:CE1	1:F:294[B]:ASN:CG	2.78	0.48
1:C:299[B]:PHE:HB2	1:C:354:ILE:HD11	1.94	0.48
1:F:305[A]:LYS:HB3	1:F:306[A]:HIS:CD2	2.50	0.47
1:F:300[B]:ARG:HD2	1:F:331:VAL:HG21	1.97	0.47
1:F:342:GLN:HE21	1:F:346:ASN:ND2	2.12	0.47
1:C:306[B]:HIS:ND1	1:C:348:VAL:HG22	2.30	0.47
1:B:342:GLN:HE21	1:B:346:ASN:ND2	2.13	0.47

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:298[B]:CYS:HB2	1:F:298[B]:CYS:CB	2.37	0.47
1:F:336[B]:HIS:CD2	1:F:337:SER:HB3	2.49	0.47
1:C:364:HIS:CE1	1:D:357:LYS:HE2	2.50	0.47
1:A:342:GLN:HE21	1:A:346:ASN:ND2	2.13	0.47
1:C:365:LEU:CD1	1:D:286:ILE:HD11	2.45	0.46
1:E:349:LYS:HB2	2:E:505:PO4:O4	2.15	0.46
1:F:303[A]:LEU:HB3	1:F:310:PHE:CE2	2.50	0.46
1:E:365:LEU:HD13	1:F:286:ILE:HD11	1.98	0.46
1:E:326:HIS:HD2	1:E:328:HIS:N	2.12	0.46
1:A:336[B]:HIS:HD2	3:A:528:HOH:O	1.97	0.46
1:E:350:ILE:HD13	1:E:356:HIS:CD2	2.51	0.45
1:C:300[B]:ARG:HD2	1:C:331:VAL:HG21	1.99	0.45
1:F:336[A]:HIS:CD2	1:F:340:GLN:NE2	2.85	0.44
1:C:299[B]:PHE:CB	1:C:354:ILE:CD1	2.95	0.44
1:A:286:ILE:HD11	1:B:365:LEU:HD11	1.99	0.44
1:C:299[A]:PHE:O	1:C:302[A]:ARG:HB3	2.18	0.44
1:E:286:ILE:HD11	1:F:365:LEU:CD1	2.48	0.44
1:D:326:HIS:HD2	1:D:328:HIS:N	2.10	0.44
1:E:361:MET:C	3:E:541:HOH:O	2.56	0.43
1:F:297[B]:LYS:HG2	1:F:316:THR:HG21	2.01	0.43
1:A:365:LEU:CD1	1:B:286:ILE:HD11	2.48	0.43
1:C:301[B]:TYR:CZ	1:F:294[B]:ASN:ND2	2.73	0.43
1:C:301[B]:TYR:CD1	1:F:294[B]:ASN:HB2	2.54	0.42
1:C:303[A]:LEU:HB3	1:C:310:PHE:CE2	2.54	0.42
1:F:336[A]:HIS:HD2	1:F:340:GLN:NE2	2.16	0.42
1:C:286:ILE:HG21	1:C:286:ILE:HD13	1.89	0.42
1:C:301[B]:TYR:HE1	1:F:294[B]:ASN:HD22	0.45	0.41
1:F:321:SER:HG	1:F:322:PRO:CD	2.03	0.41
1:F:336[A]:HIS:CD2	1:F:340:GLN:HE22	2.38	0.41
1:C:299[B]:PHE:CD1	1:C:351:PRO:HD2	2.54	0.41
1:C:320:ALA:N	3:C:556:HOH:O	2.52	0.41
1:E:306:HIS:HE1	2:E:505:PO4:O2	2.03	0.41
1:F:339:GLU:O	1:F:343:GLN:HG3	2.20	0.41
1:C:336[B]:HIS:CD2	1:C:337:SER:HB3	2.55	0.41
1:F:299[B]:PHE:HB2	1:F:354:ILE:HD11	2.01	0.41
1:F:306[B]:HIS:ND1	1:F:348:VAL:HG22	2.35	0.41
1:A:342:GLN:HE21	1:A:346:ASN:HD21	1.69	0.41
1:E:286:ILE:HG22	1:E:334:THR:HG22	2.03	0.40
1:C:292[B]:GLU:O	1:C:293[B]:SER:C	2.59	0.40
1:A:317:TRP:HB3	1:B:317:TRP:HB3	2.03	0.40
1:D:316:THR:HG23	2:D:504:PO4:O3	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	89/87 (102%)	87 (98%)	1 (1%)	1 (1%)	17	6
1	B	89/87 (102%)	87 (98%)	2 (2%)	0	100	100
1	C	100/87 (115%)	94 (94%)	6 (6%)	0	100	100
1	D	85/87 (98%)	83 (98%)	2 (2%)	0	100	100
1	E	85/87 (98%)	83 (98%)	2 (2%)	0	100	100
1	F	100/87 (115%)	92 (92%)	8 (8%)	0	100	100
All	All	548/522 (105%)	526 (96%)	21 (4%)	1 (0%)	52	42

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	321	SER

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	85/81 (105%)	80 (94%)	5 (6%)	24	12
1	B	85/81 (105%)	82 (96%)	3 (4%)	43	31
1	C	98/81 (121%)	93 (95%)	5 (5%)	29	17

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	81/81 (100%)	81 (100%)	0	100	100
1	E	81/81 (100%)	76 (94%)	5 (6%)	23	11
1	F	98/81 (121%)	91 (93%)	7 (7%)	18	8
All	All	528/486 (109%)	503 (95%)	25 (5%)	40	20

All (25) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	285	PRO
1	A	321	SER
1	A	353[A]	THR
1	A	353[B]	THR
1	A	358	LEU
1	B	353[A]	THR
1	B	353[B]	THR
1	B	358	LEU
1	C	293[A]	SER
1	C	293[B]	SER
1	C	300[A]	ARG
1	C	300[B]	ARG
1	C	353	THR
1	E	281	SER
1	E	293	SER
1	E	301	TYR
1	E	349	LYS
1	E	356	HIS
1	F	293[A]	SER
1	F	293[B]	SER
1	F	297[A]	LYS
1	F	297[B]	LYS
1	F	305[A]	LYS
1	F	305[B]	LYS
1	F	358	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (21) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	343	GLN
1	A	346	ASN
1	A	364	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	326	HIS
1	B	346	ASN
1	C	288	GLN
1	C	328	HIS
1	C	346	ASN
1	D	288	GLN
1	D	290	GLN
1	D	308	HIS
1	D	326	HIS
1	D	336	HIS
1	D	346	ASN
1	E	288	GLN
1	E	290	GLN
1	E	306	HIS
1	E	326	HIS
1	E	346	ASN
1	F	346	ASN
1	F	364	HIS

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 carbohydrates in this entry.

5.6   Ligand geometry (i)

14 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	PO4	A	508	-	4,4,4	0.57	0	6,6,6	0.27	0
2	PO4	A	510	-	4,4,4	0.43	0	6,6,6	0.30	0
2	PO4	B	509	-	4,4,4	0.43	0	6,6,6	0.26	0
2	PO4	B	511	-	4,4,4	0.43	0	6,6,6	0.29	0
2	PO4	C	512	-	4,4,4	0.48	0	6,6,6	0.31	0
2	PO4	C	514	-	4,4,4	0.32	0	6,6,6	0.28	0
2	PO4	D	501	-	4,4,4	0.34	0	6,6,6	0.29	0
2	PO4	D	503	-	4,4,4	0.31	0	6,6,6	0.28	0
2	PO4	D	504	-	4,4,4	0.55	0	6,6,6	0.26	0
2	PO4	E	502	-	4,4,4	0.44	0	6,6,6	0.33	0
2	PO4	E	505	-	4,4,4	0.34	0	6,6,6	0.30	0
2	PO4	E	506	-	4,4,4	0.40	0	6,6,6	0.26	0
2	PO4	F	507	-	4,4,4	0.44	0	6,6,6	0.25	0
2	PO4	F	513	-	4,4,4	0.36	0	6,6,6	0.28	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	PO4	A	508	-	-	0/0/0/0	0/0/0/0
2	PO4	A	510	-	-	0/0/0/0	0/0/0/0
2	PO4	B	509	-	-	0/0/0/0	0/0/0/0
2	PO4	B	511	-	-	0/0/0/0	0/0/0/0
2	PO4	C	512	-	-	0/0/0/0	0/0/0/0
2	PO4	C	514	-	-	0/0/0/0	0/0/0/0
2	PO4	D	501	-	-	0/0/0/0	0/0/0/0
2	PO4	D	503	-	-	0/0/0/0	0/0/0/0
2	PO4	D	504	-	-	0/0/0/0	0/0/0/0
2	PO4	E	502	-	-	0/0/0/0	0/0/0/0
2	PO4	E	505	-	-	0/0/0/0	0/0/0/0
2	PO4	E	506	-	-	0/0/0/0	0/0/0/0
2	PO4	F	507	-	-	0/0/0/0	0/0/0/0
2	PO4	F	513	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	C	514	PO4	1	0
2	D	504	PO4	1	0
2	E	505	PO4	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(Å²)	Q<0.9
1	A	87/87 (100%)	-0.10	5 (5%)	27 30	23, 33, 52, 55	14 (16%)
1	B	87/87 (100%)	-0.14	1 (1%)	82 84	23, 33, 52, 56	14 (16%)
1	C	83/87 (95%)	0.87	13 (15%)	3 3	23, 35, 52, 57	7 (8%)
1	D	87/87 (100%)	-0.27	1 (1%)	82 84	24, 35, 49, 54	7 (8%)
1	E	87/87 (100%)	-0.25	2 (2%)	64 67	24, 35, 48, 54	6 (6%)
1	F	83/87 (95%)	0.92	12 (14%)	3 4	24, 35, 51, 57	7 (8%)
All	All	514/522 (98%)	0.16	34 (6%)	22 24	23, 34, 52, 57	55 (10%)

All (34) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	F	321	SER	19.7
1	C	281	SER	19.2
1	F	281	SER	19.0
1	F	322	PRO	17.4
1	C	322	PRO	16.7
1	C	321	SER	16.6
1	A	281	SER	8.0
1	F	301[A]	TYR	5.4
1	C	328	HIS	4.5
1	C	353	THR	4.0
1	A	322	PRO	3.7
1	C	301[A]	TYR	3.7
1	B	322	PRO	3.6
1	F	328	HIS	3.4
1	F	327	LYS	2.8
1	F	320	ALA	2.8
1	C	327	LYS	2.8
1	F	352	PRO	2.8
1	F	353	THR	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	E	352	PRO	2.6
1	C	304(A)[A]	ASP	2.5
1	E	301	TYR	2.5
1	F	305[A]	LYS	2.5
1	A	353[A]	THR	2.4
1	F	336[A]	HIS	2.3
1	C	352	PRO	2.3
1	D	352	PRO	2.3
1	C	307[A]	ARG	2.3
1	C	320	ALA	2.3
1	A	324	ALA	2.2
1	F	307[A]	ARG	2.1
1	C	360[A]	PHE	2.1
1	A	325	PRO	2.1
1	C	336[A]	HIS	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, 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	LLDF	B-factors(Å^2)	Q<0.9
2	PO4	B	511	5/5	0.93	0.14	2.00	58,61,61,63	0
2	PO4	F	513	5/5	0.89	0.25	0.83	70,72,73,74	0
2	PO4	C	514	5/5	0.92	0.20	0.37	68,68,69,70	0
2	PO4	D	501	5/5	0.96	0.11	0.35	45,46,49,51	0
2	PO4	E	502	5/5	0.98	0.07	-0.44	43,43,45,45	0
2	PO4	F	507	5/5	0.90	0.12	-0.57	47,57,58,59	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors(Å²)	Q<0.9
2	PO4	C	512	5/5	0.89	0.12	-0.71	44,56,58,59	0
2	PO4	A	510	5/5	0.90	0.14	-	48,58,59,60	0
2	PO4	E	505	5/5	0.94	0.20	-	65,66,68,68	0
2	PO4	B	509	5/5	0.93	0.14	-	51,60,60,61	0
2	PO4	D	504	5/5	0.92	0.17	-	51,58,60,60	0
2	PO4	E	506	5/5	0.95	0.11	-	48,55,55,55	0
2	PO4	D	503	5/5	0.94	0.19	-	64,64,67,68	0
2	PO4	A	508	5/5	0.95	0.14	-	66,66,67,68	0

6.5 Other polymers (i)

There are no such residues in this entry.