Full wwPDB X-ray Structure Validation Report (i)

Feb 19, 2016 – 10:39 PM GMT

PDB ID	:	5DS5
Title	:	Crystal structure the Escherichia coli Cas1-Cas2 complex bound to protospacer DNA and Mg
Authors	:	Nunez, J.K.; Harrington, L.B.; Kranzusch, P.J.; Engelman, A.N.; Doudna, J.A.
Deposited on	:	2015-09-16
Resolution	:	2.95 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:

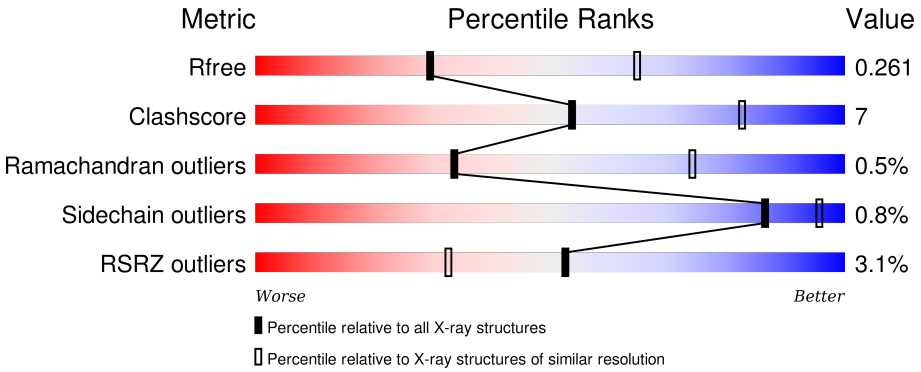
MolProbity	:	4.02b-467
Mogul	:	unknown
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026982
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
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 2.95 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	2184 (3.00-2.92)
Clashscore	102246	2552 (3.00-2.92)
Ramachandran outliers	100387	2468 (3.00-2.92)
Sidechain outliers	100360	2471 (3.00-2.92)
RSRZ outliers	91569	2201 (3.00-2.92)

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	306	4% 68% 13% 18%
1	B	306	% 74% 16% 9%
1	C	306	5% 73% 9% 18%
1	D	306	3% 74% 15% • 10%
2	E	104	% 71% 19% 10%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	104	
3	G	28	
4	H	28	

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
5	MG	G	101	-	-	-	X
5	MG	H	101	-	-	-	X

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 10719 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 CRISPR-associated endonuclease Cas1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	250	Total	C	N	O	S	0	0	0
			1907	1218	337	345	7			
1	B	277	Total	C	N	O	S	0	0	0
			2136	1367	379	383	7			
1	C	252	Total	C	N	O	S	0	0	0
			1926	1230	342	347	7			
1	D	275	Total	C	N	O	S	0	0	0
			2125	1362	377	379	7			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	0	SER	-	expression tag	UNP Q46896
B	0	SER	-	expression tag	UNP Q46896
C	0	SER	-	expression tag	UNP Q46896
D	0	SER	-	expression tag	UNP Q46896

- Molecule 2 is a protein called CRISPR-associated endoribonuclease Cas2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	E	94	Total	C	N	O	S	0	0	0
			739	475	128	132	4			
2	F	94	Total	C	N	O	S	0	0	0
			739	475	128	132	4			

There are 20 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	0	MET	-	initiating methionine	UNP P45956
E	95	GLY	-	expression tag	UNP P45956
E	96	SER	-	expression tag	UNP P45956
E	97	SER	-	expression tag	UNP P45956

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	98	GLU	-	expression tag	UNP P45956
E	99	ASN	-	expression tag	UNP P45956
E	100	LEU	-	expression tag	UNP P45956
E	101	TYR	-	expression tag	UNP P45956
E	102	PHE	-	expression tag	UNP P45956
E	103	GLN	-	expression tag	UNP P45956
F	0	MET	-	initiating methionine	UNP P45956
F	95	GLY	-	expression tag	UNP P45956
F	96	SER	-	expression tag	UNP P45956
F	97	SER	-	expression tag	UNP P45956
F	98	GLU	-	expression tag	UNP P45956
F	99	ASN	-	expression tag	UNP P45956
F	100	LEU	-	expression tag	UNP P45956
F	101	TYR	-	expression tag	UNP P45956
F	102	PHE	-	expression tag	UNP P45956
F	103	GLN	-	expression tag	UNP P45956

- Molecule 3 is a DNA chain called DNA (28-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	G	28	Total	C	N	O	P	0	0	0
			578	275	118	158	27			

- Molecule 4 is a DNA chain called DNA (28-MER).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	H	28	Total	C	N	O	P	0	0	0
			564	274	86	177	27			

- Molecule 5 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

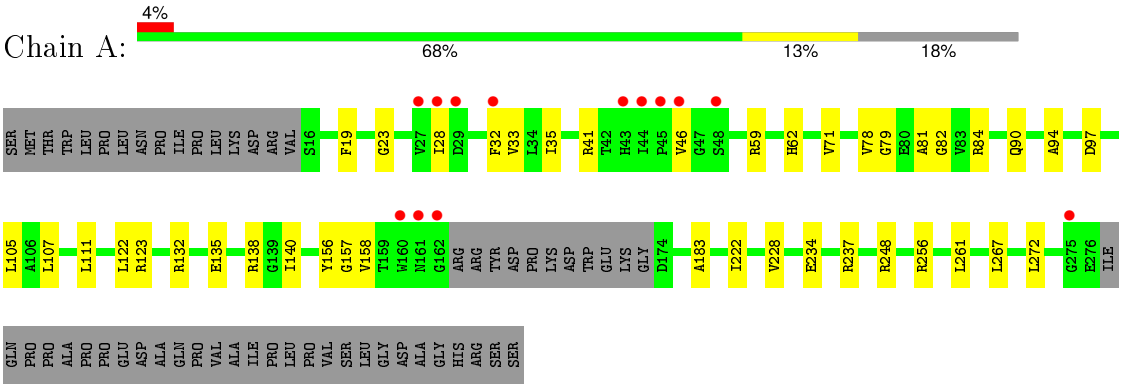
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	H	1	Total	Mg	0	0
			1	1		
5	G	1	Total	Mg	0	0
			1	1		
5	B	2	Total	Mg	0	0
			2	2		
5	D	1	Total	Mg	0	0
			1	1		

3 Residue-property plots (i)

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.

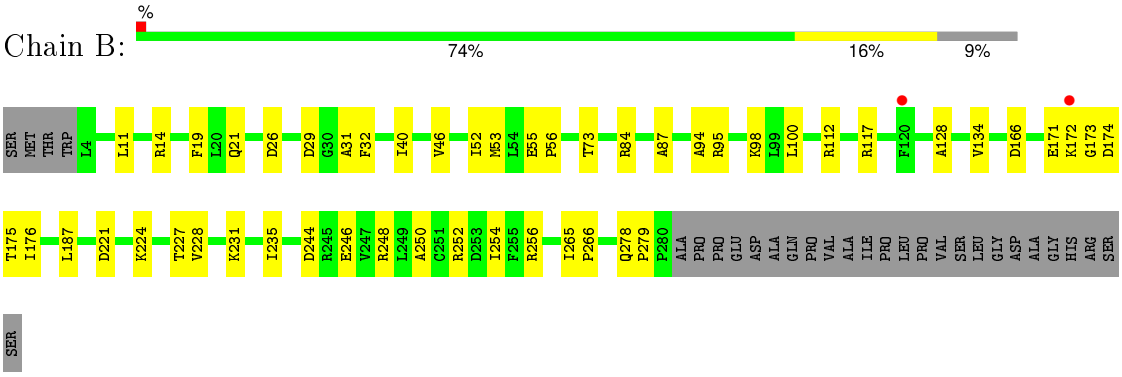
- Molecule 1: CRISPR-associated endonuclease Cas1

Chain A:



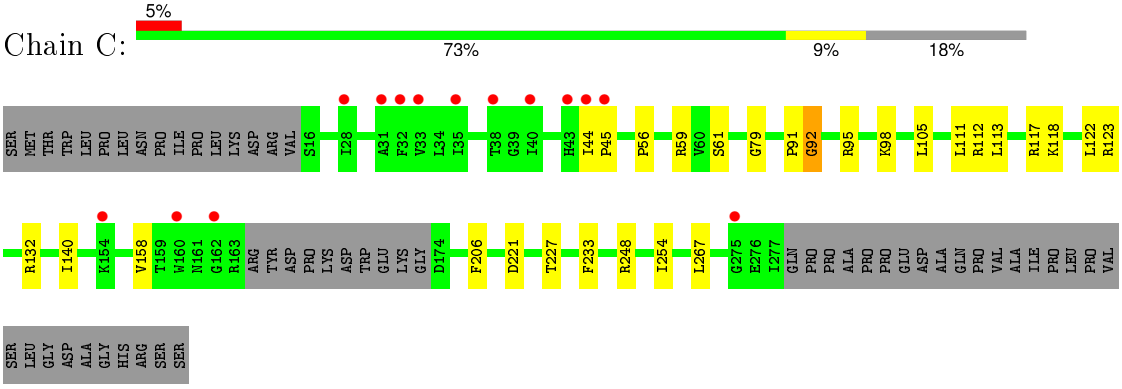
- Molecule 1: CRISPR-associated endonuclease Cas1

Chain B:



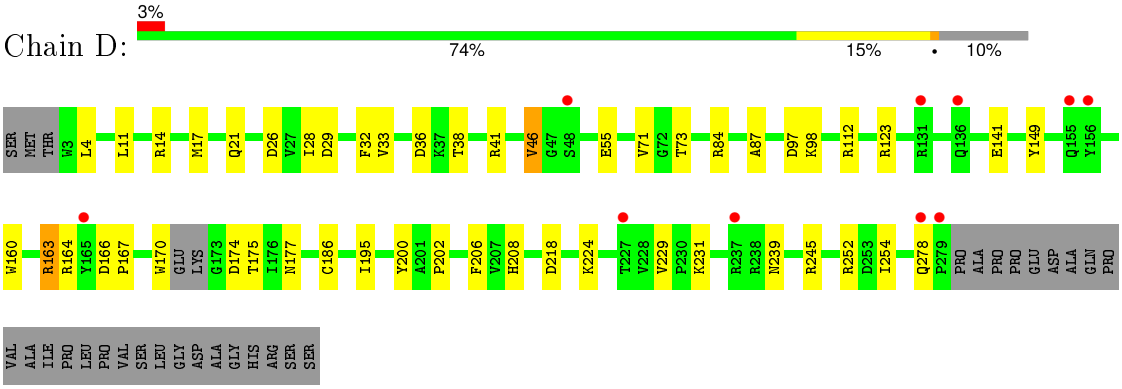
- Molecule 1: CRISPR-associated endonuclease Cas1

Chain C:



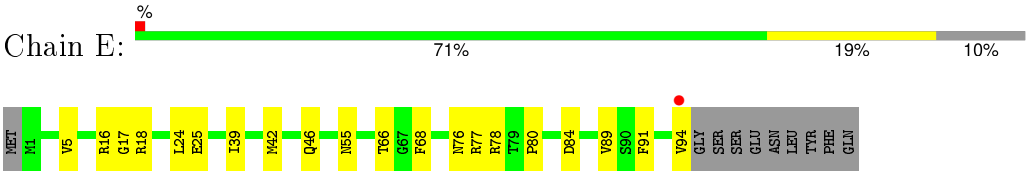
• Molecule 1: CRISPR-associated endonuclease Cas1

Chain D:



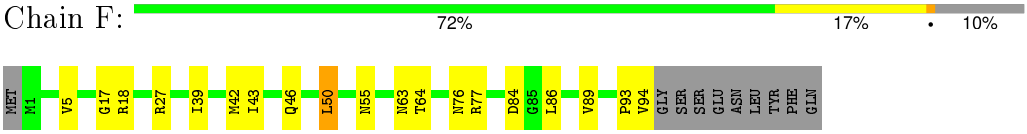
• Molecule 2: CRISPR-associated endoribonuclease Cas2

Chain E:



• Molecule 2: CRISPR-associated endoribonuclease Cas2

Chain F:

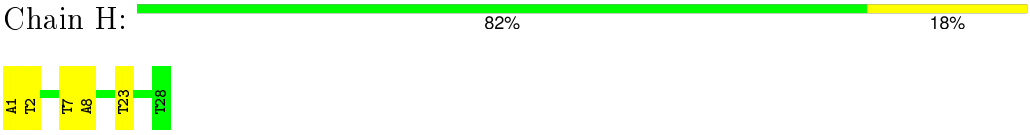


• Molecule 3: DNA (28-MER)

Chain G:

• Molecule 4: DNA (28-MER)

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	75.66Å 165.93Å 167.26Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.48 – 2.95 46.48 – 2.95	Depositor EDS
% Data completeness (in resolution range)	99.7 (46.48-2.95) 99.7 (46.48-2.95)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.71 (at 2.96Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.9_1692)	Depositor
R, R_{free}	0.234 , 0.256 0.238 , 0.261	Depositor DCC
R_{free} test set	2641 reflections (6.24%)	DCC
Wilson B-factor (Å^2)	65.7	Xtriage
Anisotropy	0.129	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 33.8	EDS
Estimated twinning fraction	0.109 for -h,l,k	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.25$	Xtriage
Outliers	0 of 44960 reflections	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	10719	wwPDB-VP
Average B, all atoms (Å^2)	66.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.37% 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: MG

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	0.22	0/1940	0.43	0/2628
1	B	0.22	0/2179	0.43	0/2957
1	C	0.22	0/1959	0.43	0/2653
1	D	0.21	0/2168	0.44	0/2942
2	E	0.21	0/753	0.42	0/1024
2	F	0.20	0/753	0.39	0/1024
3	G	0.46	0/652	0.79	0/1005
4	H	0.52	0/627	1.04	0/966
All	All	0.26	0/11031	0.52	0/15199

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

5.2 Too-close contacts

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	1907	0	1963	27	0
1	B	2136	0	2202	35	0
1	C	1926	0	1987	17	0
1	D	2125	0	2185	34	0
2	E	739	0	756	17	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	739	0	756	21	0
3	G	578	0	314	10	0
4	H	564	0	324	4	0
5	B	2	0	0	0	0
5	D	1	0	0	0	0
5	G	1	0	0	0	0
5	H	1	0	0	0	0
All	All	10719	0	10487	136	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 (136) 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:E:5:VAL:HG21	2:F:5:VAL:HG21	1.54	0.86
1:A:59:ARG:NH2	1:B:26:ASP:OD1	2.13	0.81
1:C:59:ARG:NH2	1:D:26:ASP:OD1	2.14	0.81
1:A:28:ILE:HD13	1:A:41:ARG:HE	1.51	0.75
2:E:55:ASN:HD22	2:F:27:ARG:HH21	1.33	0.73
1:D:123:ARG:NH1	1:D:141:GLU:OE1	2.22	0.73
1:A:234:GLU:HG3	1:A:237:ARG:HH21	1.56	0.69
1:A:59:ARG:NH1	4:H:2:DT:OP1	2.26	0.68
1:A:78:VAL:HB	1:A:82:GLY:HA2	1.77	0.67
1:B:252:ARG:NH2	2:F:84:ASP:O	2.29	0.66
1:A:84:ARG:NH1	3:G:25:DG:OP1	2.30	0.64
1:A:28:ILE:HD11	1:A:35:ILE:HD11	1.79	0.64
1:A:28:ILE:HG13	1:A:33:VAL:HG21	1.80	0.64
1:B:244:ASP:O	1:B:248:ARG:HG2	1.97	0.63
2:E:66:THR:OG1	2:E:68:PHE:O	2.17	0.63
1:B:256:ARG:HD2	2:F:64:THR:HA	1.81	0.62
1:B:174:ASP:O	1:B:176:ILE:N	2.31	0.62
1:C:158:VAL:HG21	1:C:233:PHE:HD2	1.65	0.61
2:E:76:ASN:HD22	2:E:78:ARG:H	1.47	0.61
1:C:112:ARG:HG3	1:C:206:PHE:HD2	1.68	0.59
1:B:171:GLU:O	1:B:173:GLY:N	2.34	0.59
1:D:17:MET:O	1:D:252:ARG:NH1	2.34	0.59
1:A:105:LEU:HD23	1:A:111:LEU:HD23	1.84	0.58
1:A:123:ARG:HD3	1:A:222:ILE:HA	1.85	0.58
2:E:18:ARG:NH1	2:E:46:GLN:OE1	2.37	0.58
1:A:156:TYR:O	1:A:158:VAL:HG23	2.04	0.57

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:252:ARG:NH1	2:F:86:LEU:HB2	2.19	0.57
1:C:248:ARG:NH1	4:H:23:DT:OP1	2.36	0.57
1:C:105:LEU:HD23	1:C:111:LEU:HD23	1.88	0.56
2:E:16:ARG:NH2	2:E:25:GLU:OE2	2.36	0.55
1:D:11:LEU:HD12	1:D:14:ARG:HD2	1.89	0.55
1:A:90:GLN:HB3	1:A:94:ALA:HB2	1.88	0.55
4:H:7:DT:H2"	4:H:8:DA:C8	2.42	0.54
1:D:4:LEU:HG	2:F:50:LEU:HD11	1.89	0.54
1:B:252:ARG:HH12	2:F:86:LEU:HB2	1.73	0.54
1:B:227:THR:HB	1:B:254:ILE:HG21	1.90	0.54
2:E:77:ARG:HH12	2:E:94:VAL:HG13	1.72	0.54
2:F:18:ARG:NH1	2:F:46:GLN:OE1	2.42	0.53
2:E:76:ASN:ND2	2:E:78:ARG:H	2.06	0.52
1:D:41:ARG:HB2	2:E:91:PHE:HB3	1.92	0.52
1:B:231:LYS:HG3	1:B:250:ALA:HB1	1.91	0.52
3:G:20:DA:H2'	3:G:21:DA:C8	2.43	0.52
1:B:46:VAL:HG12	1:B:52:ILE:HD11	1.92	0.52
1:A:97:ASP:N	1:A:97:ASP:OD1	2.42	0.52
1:A:135:GLU:OE1	1:A:138:ARG:NH1	2.43	0.51
1:D:29:ASP:O	2:F:17:GLY:HA3	2.10	0.51
1:B:166:ASP:OD1	1:B:166:ASP:N	2.41	0.51
2:F:39:ILE:O	2:F:43:ILE:HG12	2.11	0.51
1:B:235:ILE:HD12	1:B:246:GLU:HB3	1.93	0.51
1:D:36:ASP:OD2	1:D:38:THR:OG1	2.27	0.51
2:F:77:ARG:HH22	2:F:94:VAL:HG13	1.75	0.51
1:C:123:ARG:NH2	1:C:221:ASP:HB3	2.26	0.51
1:C:158:VAL:HG21	1:C:233:PHE:CD2	2.46	0.50
1:D:195:ILE:HG23	1:D:200:TYR:HB2	1.94	0.50
3:G:7:DC:H2"	3:G:8:DA:C8	2.46	0.50
1:B:95:ARG:HD2	1:B:98:LYS:HD2	1.93	0.50
2:E:39:ILE:O	2:E:42:MET:HG3	2.10	0.50
1:B:84:ARG:HH21	1:B:87:ALA:HB3	1.76	0.50
1:C:59:ARG:NH1	3:G:2:DA:OP1	2.45	0.49
1:B:256:ARG:NH1	2:F:63:ASN:O	2.45	0.49
1:B:19:PHE:CE1	1:B:248:ARG:HD2	2.48	0.49
1:D:167:PRO:HA	1:D:170:TRP:CE2	2.48	0.49
1:C:132:ARG:NH1	1:C:140:ILE:HD11	2.28	0.49
1:D:4:LEU:HB3	2:F:18:ARG:HH22	1.78	0.48
1:A:79:GLY:N	1:A:84:ARG:O	2.45	0.48
1:A:248:ARG:NH1	3:G:23:DT:OP1	2.40	0.48
1:A:19:PHE:HZ	1:A:261:LEU:HD23	1.78	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:112:ARG:NH2	1:D:206:PHE:O	2.46	0.48
1:D:149:TYR:CE2	1:D:163:ARG:HG3	2.48	0.48
2:F:55:ASN:OD1	2:F:76:ASN:ND2	2.42	0.47
1:D:4:LEU:HB3	2:F:18:ARG:NH2	2.30	0.47
3:G:25:DG:H4'	3:G:26:DG:H5"	1.97	0.46
1:B:221:ASP:HA	1:B:224:LYS:HG2	1.97	0.46
1:D:46:VAL:O	1:D:73:THR:OG1	2.29	0.46
1:B:29:ASP:O	2:E:17:GLY:HA3	2.15	0.46
1:C:95:ARG:HD2	1:C:98:LYS:HD2	1.97	0.46
1:B:19:PHE:CZ	1:B:53:MET:HG3	2.50	0.46
1:C:118:LYS:HE3	1:C:122:LEU:HD11	1.98	0.46
2:E:55:ASN:HD22	2:F:27:ARG:NH2	2.07	0.46
1:D:160:TRP:HH2	1:D:163:ARG:HG2	1.79	0.46
1:C:227:THR:HG22	1:C:254:ILE:HD13	1.97	0.46
1:A:19:PHE:CZ	1:A:261:LEU:HD23	2.51	0.46
1:C:56:PRO:HD3	1:C:79:GLY:HA2	1.98	0.45
1:A:81:ALA:HB2	1:A:256:ARG:HE	1.82	0.45
1:D:164:ARG:HH22	1:D:174:ASP:HB3	1.81	0.45
1:D:224:LYS:HG2	1:D:229:VAL:HG23	1.98	0.45
1:D:21:GLN:HB3	1:D:55:GLU:HB2	1.99	0.45
1:B:21:GLN:HB3	1:B:55:GLU:HB2	1.99	0.45
1:D:112:ARG:HG3	1:D:206:PHE:HD2	1.81	0.45
1:D:164:ARG:HH21	1:D:177:ASN:HD22	1.64	0.45
1:B:11:LEU:HD12	1:B:14:ARG:HD2	1.98	0.45
2:F:42:MET:O	2:F:46:GLN:HG2	2.17	0.44
1:A:62:HIS:CG	1:B:56:PRO:HA	2.52	0.44
1:D:164:ARG:HE	1:D:177:ASN:ND2	2.15	0.44
3:G:9:DG:H2"	3:G:10:DA:C8	2.52	0.44
1:B:46:VAL:O	1:B:73:THR:OG1	2.34	0.44
1:B:40:ILE:HG22	2:F:93:PRO:HD3	2.00	0.44
1:A:132:ARG:NH2	1:A:140:ILE:HD11	2.33	0.44
1:A:23:GLY:HA2	4:H:1:DA:H4'	2.00	0.44
1:D:245:ARG:HG3	2:E:84:ASP:HA	1.99	0.44
2:E:55:ASN:ND2	2:E:76:ASN:OD1	2.50	0.44
1:C:92:GLY:HA3	1:D:202:PRO:HD2	2.00	0.44
2:F:46:GLN:O	2:F:50:LEU:HB2	2.19	0.43
1:B:117:ARG:NH2	1:B:128:ALA:O	2.51	0.43
1:A:84:ARG:HA	1:A:84:ARG:HD3	1.80	0.43
1:B:19:PHE:CZ	1:B:248:ARG:HD2	2.53	0.43
1:D:28:ILE:HG13	1:D:33:VAL:HG21	1.99	0.43
1:D:208:HIS:ND1	1:D:218:ASP:OD1	2.38	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:278:GLN:CD	1:B:278:GLN:H	2.22	0.43
2:E:80:PRO:HA	2:E:89:VAL:HA	2.00	0.43
1:D:231:LYS:HB2	1:D:254:ILE:HD11	2.01	0.43
1:C:122:LEU:HD13	1:C:267:LEU:HD11	2.00	0.42
1:D:17:MET:HE1	1:D:186:CYS:HB3	2.00	0.42
1:B:227:THR:HB	1:B:254:ILE:CG2	2.49	0.42
1:A:107:LEU:HD21	1:B:100:LEU:HB3	2.01	0.42
1:D:97:ASP:OD1	1:D:98:LYS:N	2.53	0.42
1:B:31:ALA:HB2	2:E:17:GLY:O	2.20	0.41
3:G:26:DG:H8	3:G:27:DG:C8	2.38	0.41
3:G:12:DC:H2"	3:G:13:DG:C8	2.54	0.41
1:B:187:LEU:HD22	1:B:228:VAL:HG21	2.03	0.41
1:C:44:ILE:HA	1:C:45:PRO:HD3	1.89	0.41
1:A:46:VAL:HB	1:A:71:VAL:HG21	2.02	0.41
2:E:24:LEU:HD11	2:F:89:VAL:HG12	2.03	0.41
1:A:183:ALA:HB1	1:A:228:VAL:HG12	2.03	0.41
2:F:77:ARG:HH12	2:F:94:VAL:HG13	1.84	0.41
1:D:84:ARG:HH21	1:D:87:ALA:HB3	1.86	0.41
1:C:113:LEU:O	1:C:117:ARG:HG3	2.20	0.41
1:A:122:LEU:HD13	1:A:267:LEU:HD11	2.03	0.41
1:D:166:ASP:HA	1:D:167:PRO:HD3	1.90	0.41
1:D:175:THR:HG21	1:D:239:ASN:HA	2.02	0.41
1:B:278:GLN:HA	1:B:279:PRO:HD3	1.92	0.41
1:D:278:GLN:CD	1:D:278:GLN:H	2.24	0.41
1:B:112:ARG:NH1	1:B:134:VAL:HG21	2.36	0.40
1:D:46:VAL:HB	1:D:71:VAL:HG21	2.03	0.40
3:G:13:DG:H2"	3:G:14:DA:C8	2.56	0.40
1:B:265:ILE:HB	1:B:266:PRO:HD3	2.03	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	246/306 (80%)	236 (96%)	9 (4%)	1 (0%)	39 78
1	B	275/306 (90%)	261 (95%)	11 (4%)	3 (1%)	17 56
1	C	248/306 (81%)	235 (95%)	11 (4%)	2 (1%)	24 64
1	D	271/306 (89%)	262 (97%)	9 (3%)	0	100 100
2	E	92/104 (88%)	87 (95%)	5 (5%)	0	100 100
2	F	92/104 (88%)	87 (95%)	5 (5%)	0	100 100
All	All	1224/1432 (86%)	1168 (95%)	50 (4%)	6 (0%)	34 74

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	175	THR
1	C	91	PRO
1	A	157	GLY
1	B	94	ALA
1	C	92	GLY
1	B	172	LYS

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	197/246 (80%)	195 (99%)	2 (1%)	82 94
1	B	223/246 (91%)	222 (100%)	1 (0%)	93 98
1	C	199/246 (81%)	198 (100%)	1 (0%)	92 97
1	D	221/246 (90%)	218 (99%)	3 (1%)	74 92
2	E	79/88 (90%)	79 (100%)	0	100 100
2	F	79/88 (90%)	78 (99%)	1 (1%)	76 92
All	All	998/1160 (86%)	990 (99%)	8 (1%)	86 95

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	32	PHE
1	A	272	LEU
1	B	32	PHE
1	C	61	SER
1	D	32	PHE
1	D	46	VAL
1	D	163	ARG
2	F	50	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	E	55	ASN
2	E	76	ASN

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

Of 5 ligands modelled in this entry, 5 are monoatomic - leaving 0 for Mogul analysis.

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.

No monomer is involved in short contacts.

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

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	250/306 (81%)	0.27	13 (5%) 31 17	42, 70, 97, 113	0
1	B	277/306 (90%)	0.14	2 (0%) 89 76	38, 60, 91, 108	0
1	C	252/306 (82%)	0.39	14 (5%) 28 15	47, 70, 103, 115	0
1	D	275/306 (89%)	0.29	10 (3%) 46 28	43, 63, 88, 121	0
2	E	94/104 (90%)	0.06	1 (1%) 82 65	41, 54, 74, 93	0
2	F	94/104 (90%)	0.06	0 100 100	39, 53, 76, 85	0
3	G	28/28 (100%)	-0.41	0 100 100	49, 61, 116, 126	0
4	H	28/28 (100%)	-0.44	0 100 100	46, 61, 101, 111	0
All	All	1298/1488 (87%)	0.21	40 (3%) 52 33	38, 63, 96, 126	0

All (40) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	48	SER	6.5
1	C	28	ILE	4.3
1	C	162	GLY	4.1
1	C	43	HIS	3.9
1	D	279	PRO	3.8
1	C	160	TRP	3.8
2	E	94	VAL	3.7
1	A	46	VAL	3.6
1	A	28	ILE	3.5
1	C	40	ILE	3.3
1	A	161	ASN	3.0
1	D	278	GLN	2.9
1	D	136	GLN	2.8
1	D	237	ARG	2.7
1	D	131	ARG	2.7
1	A	32	PHE	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	33	VAL	2.6
1	B	172	LYS	2.5
1	D	155	GLN	2.5
1	D	48	SER	2.4
1	A	43	HIS	2.4
1	A	160	TRP	2.4
1	D	156	TYR	2.4
1	A	29	ASP	2.4
1	D	227	THR	2.4
1	C	38	THR	2.3
1	A	44	ILE	2.3
1	D	165	TYR	2.2
1	C	45	PRO	2.2
1	A	275	GLY	2.1
1	A	45	PRO	2.1
1	C	32	PHE	2.1
1	C	35	ILE	2.1
1	C	154	LYS	2.1
1	C	275	GLY	2.1
1	A	27	VAL	2.1
1	B	120	PHE	2.1
1	C	44	ILE	2.0
1	A	162	GLY	2.0
1	C	31	ALA	2.0

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

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

6.3 Carbohydrates

There are no carbohydrates in this entry.

6.4 Ligands

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
5	MG	G	101	1/1	0.92	0.60	27.01	44,44,44,44	0
5	MG	H	101	1/1	0.94	0.60	22.79	63,63,63,63	0
5	MG	B	402	1/1	0.81	0.22	1.59	30,30,30,30	0
5	MG	D	401	1/1	0.90	0.19	-	53,53,53,53	0
5	MG	B	401	1/1	0.94	0.17	-	44,44,44,44	0

6.5 Other polymers (i)

There are no such residues in this entry.