Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 12:43 PM GMT

PDB ID	:	3ROE
Title	:	Crystal Structure of Mouse Apolipoprotein A-I Binding Protein in Complex with Thymidine
Authors	:	Shumilin, I.A.; Jha, K.N.; Cymborowski, M.; Herr, J.C.; Minor, W.
Deposited on	:	2011-04-25
Resolution	:	2.11 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 ⓘ symbol.

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

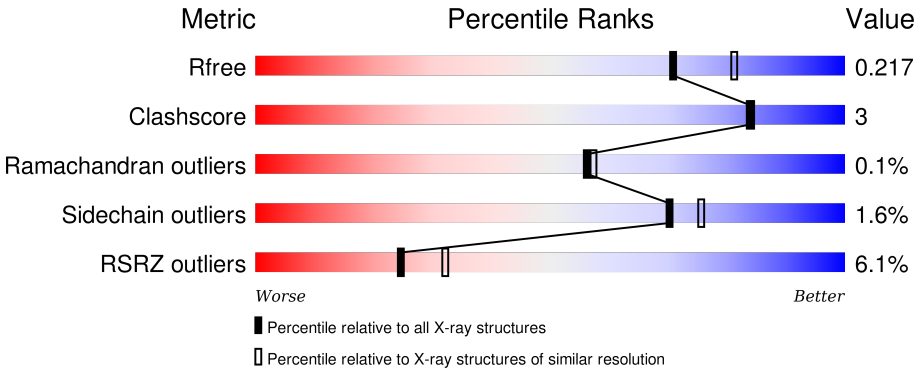
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 2.11 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.



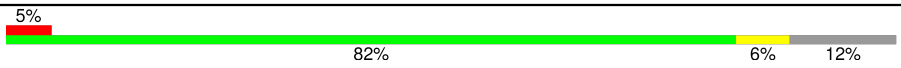
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	4587 (2.14-2.10)
Clashscore	102246	5132 (2.14-2.10)
Ramachandran outliers	100387	5080 (2.14-2.10)
Sidechain outliers	100360	5081 (2.14-2.10)
RSRZ outliers	91569	4597 (2.14-2.10)

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	265	8% 77% 11% 12%
1	B	265	4% 83% 5% 12%
1	C	265	7% 82% 6% 12%
1	D	265	4% 80% 8% 12%
1	E	265	4% 78% 9% 12%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	F	265	

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 11437 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 Apolipoprotein A-I-binding protein.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	233	Total	C	N	O	S	Se	0	0	0
			1808	1170	290	337	5	6			
1	B	233	Total	C	N	O	S	Se	0	0	0
			1808	1170	290	337	5	6			
1	C	233	Total	C	N	O	S	Se	0	0	0
			1808	1170	290	337	5	6			
1	D	233	Total	C	N	O	S	Se	0	1	0
			1813	1175	290	337	5	6			
1	E	233	Total	C	N	O	S	Se	0	1	0
			1813	1173	290	339	5	6			
1	F	233	Total	C	N	O	S	Se	0	0	0
			1808	1170	290	337	5	6			

There are 42 discrepancies between the modelled and reference sequences:

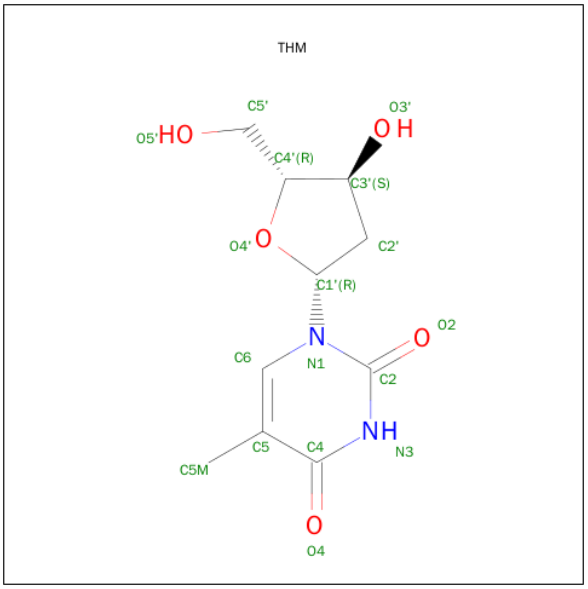
Chain	Residue	Modelled	Actual	Comment	Reference
A	0	MSE	-	EXPRESSION TAG	UNP Q8K4Z3
A	259	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
A	260	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
A	261	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
A	262	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
A	263	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
A	264	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
B	0	MSE	-	EXPRESSION TAG	UNP Q8K4Z3
B	259	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
B	260	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
B	261	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
B	262	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
B	263	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
B	264	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
C	0	MSE	-	EXPRESSION TAG	UNP Q8K4Z3
C	259	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
C	260	HIS	-	EXPRESSION TAG	UNP Q8K4Z3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	261	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
C	262	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
C	263	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
C	264	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
D	0	MSE	-	EXPRESSION TAG	UNP Q8K4Z3
D	259	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
D	260	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
D	261	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
D	262	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
D	263	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
D	264	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
E	0	MSE	-	EXPRESSION TAG	UNP Q8K4Z3
E	259	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
E	260	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
E	261	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
E	262	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
E	263	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
E	264	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
F	0	MSE	-	EXPRESSION TAG	UNP Q8K4Z3
F	259	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
F	260	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
F	261	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
F	262	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
F	263	HIS	-	EXPRESSION TAG	UNP Q8K4Z3
F	264	HIS	-	EXPRESSION TAG	UNP Q8K4Z3

- Molecule 2 is THYMIDINE (three-letter code: THM) (formula: $C_{10}H_{14}N_2O_5$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			17	10	2	5		
2	B	1	Total	C	N	O	0	0
			17	10	2	5		
2	C	1	Total	C	N	O	0	0
			17	10	2	5		
2	D	1	Total	C	N	O	0	0
			17	10	2	5		
2	E	1	Total	C	N	O	0	0
			17	10	2	5		
2	F	1	Total	C	N	O	0	0
			17	10	2	5		

- Molecule 3 is water.

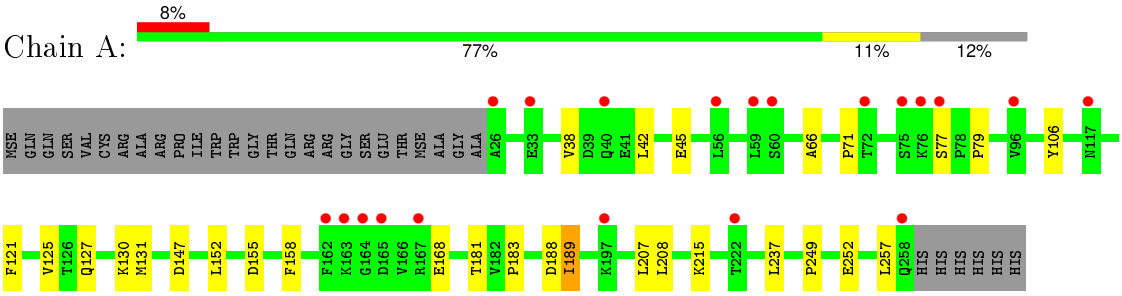
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	55	Total	O	0	0
			55	55		
3	B	109	Total	O	0	0
			109	109		
3	C	46	Total	O	0	0
			46	46		
3	D	91	Total	O	0	0
			91	91		
3	E	88	Total	O	0	0
			88	88		
3	F	88	Total	O	0	0
			88	88		

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.

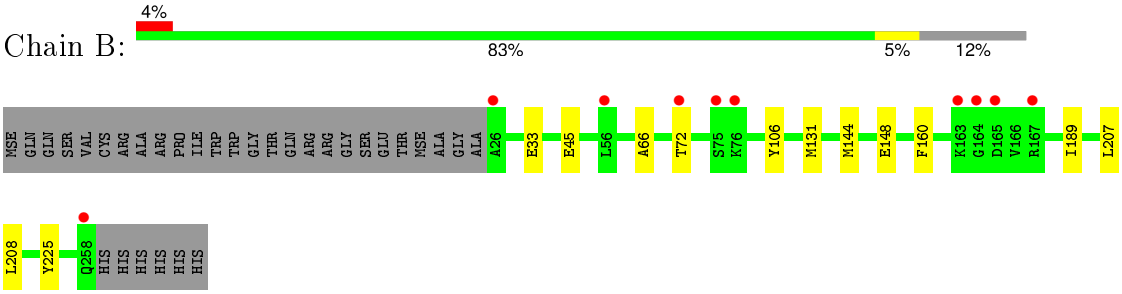
- Molecule 1: Apolipoprotein A-I-binding protein

Chain A:



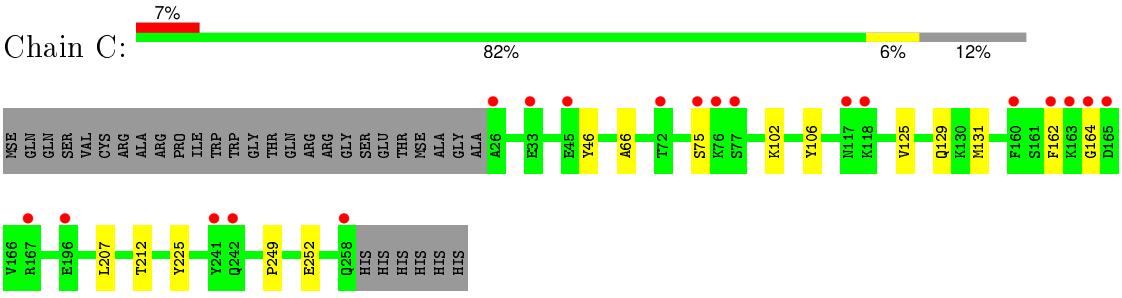
- Molecule 1: Apolipoprotein A-I-binding protein

Chain B:



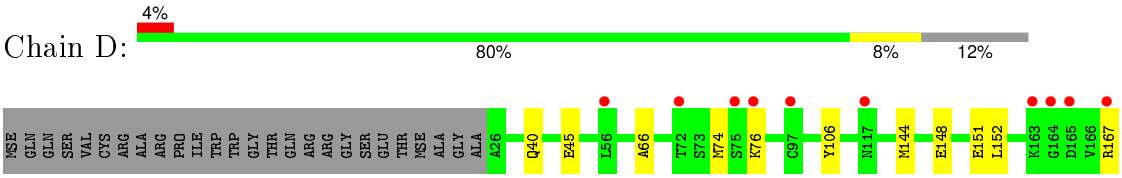
- Molecule 1: Apolipoprotein A-I-binding protein

Chain C:

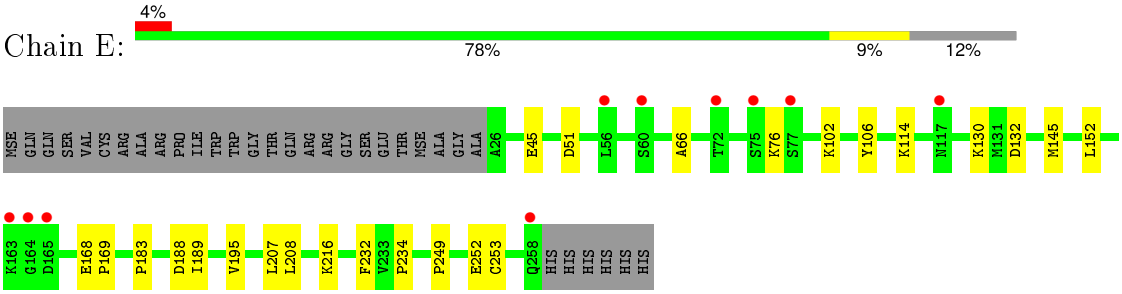


- Molecule 1: Apolipoprotein A-I-binding protein

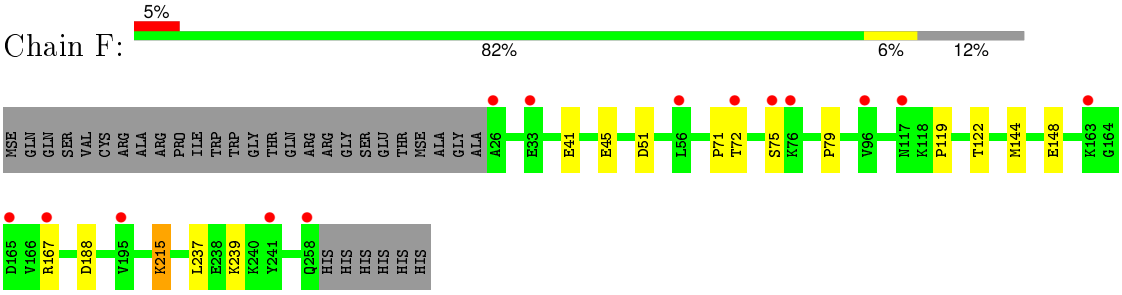
Chain D:



• Molecule 1: Apolipoprotein A-I-binding protein



• Molecule 1: Apolipoprotein A-I-binding protein



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	106.21Å 124.45Å 164.25Å 90.00° 102.71° 90.00°	Depositor
Resolution (Å)	50.00 – 2.11 35.72 – 2.11	Depositor EDS
% Data completeness (in resolution range)	97.8 (50.00-2.11) 97.4 (35.72-2.11)	Depositor EDS
R_{merge}	0.09	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	1.60 (at 2.10Å)	Xtriage
Refinement program	REFMAC 5.5.0109	Depositor
R, R_{free}	0.177 , 0.211 0.184 , 0.217	Depositor DCC
R_{free} test set	5864 reflections (5.28%)	DCC
Wilson B-factor (Å^2)	32.6	Xtriage
Anisotropy	0.320	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.35 , 41.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 116947 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	11437	wwPDB-VP
Average B, all atoms (Å^2)	42.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 26.74 % of the origin peak, indicating pseudo translational symmetry.The chance of finding a peak of this or larger height randomly in a structure without pseudo translational symmetry is equal to 2.4927e-03.The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: THM

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	0.76	0/1853	0.75	0/2514
1	B	1.04	1/1853 (0.1%)	0.80	0/2514
1	C	0.80	0/1853	0.74	0/2514
1	D	1.02	1/1861 (0.1%)	0.86	1/2525 (0.0%)
1	E	0.91	2/1861 (0.1%)	0.83	3/2525 (0.1%)
1	F	0.93	0/1853	0.83	0/2514
All	All	0.92	4/11134 (0.0%)	0.80	4/15106 (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	A	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	45	GLU	CB-CG	-5.86	1.41	1.52
1	B	33	GLU	CG-CD	5.59	1.60	1.51
1	E	253	CYS	CB-SG	-5.30	1.73	1.81
1	E	195	VAL	CB-CG2	5.17	1.63	1.52

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	E	102	LYS	CD-CE-NZ	-5.57	98.88	111.70
1	E	51	ASP	CB-CG-OD1	-5.40	113.44	118.30
1	E	145	MSE	CG-SE-CE	-5.22	87.41	98.90

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	D	256	ARG	NE-CZ-NH1	5.17	122.88	120.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	257	LEU	Peptide

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	1808	0	1798	14	0
1	B	1808	0	1798	12	0
1	C	1808	0	1798	8	0
1	D	1813	0	1809	14	0
1	E	1813	0	1802	15	0
1	F	1808	0	1798	6	0
2	A	17	0	14	1	0
2	B	17	0	14	0	0
2	C	17	0	14	0	0
2	D	17	0	14	2	0
2	E	17	0	14	0	0
2	F	17	0	14	0	0
3	A	55	0	0	0	0
3	B	109	0	0	0	0
3	C	46	0	0	0	0
3	D	91	0	0	2	0
3	E	88	0	0	1	0
3	F	88	0	0	0	0
All	All	11437	0	10887	65	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 3.

All (65) 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:151:GLU:O	1:D:152[B]:LEU:HD23	1.90	0.72
1:B:45:GLU:CD	1:E:45:GLU:HG3	2.09	0.71
1:E:130:LYS:NZ	1:F:45:GLU:OE2	2.23	0.68
1:B:45:GLU:OE2	1:E:45:GLU:HG3	1.97	0.65
2:D:265:THM:H5'2	3:D:371:HOH:O	1.97	0.64
1:A:127:GLN:O	1:A:131:MSE:HG2	1.98	0.64
1:D:207:LEU:HD23	1:D:208:LEU:N	2.12	0.63
1:A:147:ASP:OD2	1:A:181:THR:HB	2.01	0.60
1:D:74:MSE:HE1	1:D:152[B]:LEU:HD21	1.86	0.56
1:D:207:LEU:HD23	1:D:207:LEU:C	2.26	0.56
1:B:45:GLU:HG3	1:E:45:GLU:CD	2.28	0.54
1:A:66:ALA:HA	1:A:106:TYR:OH	2.08	0.54
1:A:121:PHE:O	1:A:125:VAL:HG23	2.08	0.53
1:A:38:VAL:HG13	1:A:237:LEU:HD21	1.90	0.53
1:E:152:LEU:HD12	1:E:183:PRO:O	2.09	0.53
1:C:125:VAL:O	1:C:129:GLN:HG3	2.09	0.53
1:B:207:LEU:HD12	1:B:208:LEU:N	2.24	0.52
1:E:249:PRO:HG2	1:E:252:GLU:HG3	1.91	0.52
1:E:66:ALA:HA	1:E:106:TYR:OH	2.10	0.52
1:F:41:GLU:HG3	1:F:45:GLU:HG2	1.93	0.51
1:E:188:ASP:O	1:E:189:ILE:HD13	2.11	0.51
1:E:207:LEU:HD12	1:E:208:LEU:N	2.26	0.51
1:A:158:PHE:O	1:A:189:ILE:HD11	2.10	0.50
1:C:249:PRO:HG2	1:C:252:GLU:HG3	1.93	0.50
1:D:144:MSE:O	1:D:148:GLU:HG2	2.11	0.49
1:B:72:THR:HG22	1:B:72:THR:O	2.12	0.49
1:B:160:PHE:HA	1:B:189:ILE:HD13	1.93	0.48
1:C:129:GLN:NE2	1:E:132[A]:ASP:OD2	2.47	0.48
1:F:144:MSE:O	1:F:148:GLU:HG3	2.14	0.47
1:E:45:GLU:OE1	3:E:458:HOH:O	2.20	0.47
1:A:168:GLU:HA	1:A:168:GLU:OE1	2.15	0.47
1:E:168:GLU:HA	1:E:168:GLU:OE1	2.15	0.47
2:D:265:THM:C5'	3:D:371:HOH:O	2.58	0.46
1:D:249:PRO:HG2	1:D:252:GLU:HG3	1.98	0.46
1:B:131:MSE:HG2	1:D:234:PRO:HB3	1.97	0.45
1:D:238:GLU:HG3	1:D:243:LEU:HB2	1.98	0.45
1:D:74:MSE:HE1	1:D:152[B]:LEU:CD2	2.46	0.45
1:B:207:LEU:HD13	1:B:225:TYR:HB3	1.98	0.45
1:A:188:ASP:OD1	1:A:215:LYS:NZ	2.42	0.45
1:F:119:PRO:HA	1:F:122:THR:HB	1.99	0.45
1:D:207:LEU:HD21	1:D:209:ILE:HG13	2.00	0.44
1:A:207:LEU:HD12	1:A:208:LEU:N	2.32	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:207:LEU:HD13	1:B:225:TYR:CB	2.48	0.44
1:C:102:LYS:HG2	1:C:131:MSE:HE3	2.00	0.43
1:A:42:LEU:HD21	1:A:237:LEU:CD2	2.49	0.43
1:B:144:MSE:O	1:B:148:GLU:HG2	2.18	0.43
1:A:152:LEU:HD12	1:A:183:PRO:O	2.18	0.43
1:C:66:ALA:HA	1:C:106:TYR:OH	2.19	0.43
1:D:168:GLU:HA	1:D:168:GLU:OE1	2.17	0.43
1:B:207:LEU:HD12	1:B:208:LEU:H	1.83	0.43
1:E:114:LYS:HE3	1:E:169:PRO:HD3	2.01	0.43
1:E:232:PHE:O	1:E:234:PRO:HD3	2.18	0.43
2:A:265:THM:O2	2:A:265:THM:C5'	2.67	0.42
1:A:130:LYS:HB3	1:C:46:TYR:CE2	2.54	0.42
1:D:66:ALA:HA	1:D:106:TYR:OH	2.19	0.42
1:C:207:LEU:HD13	1:C:225:TYR:CB	2.49	0.41
1:F:188:ASP:OD1	1:F:215:LYS:NZ	2.45	0.41
1:A:71:PRO:HA	1:A:79:PRO:HG3	2.02	0.41
1:D:76:LYS:HB2	1:D:151:GLU:OE1	2.20	0.41
1:F:71:PRO:HA	1:F:79:PRO:HG3	2.03	0.41
1:A:249:PRO:HG2	1:A:252:GLU:HG3	2.04	0.40
1:D:144:MSE:O	1:D:148:GLU:CG	2.70	0.40
1:C:207:LEU:HD13	1:C:225:TYR:HB2	2.03	0.40
1:B:66:ALA:HA	1:B:106:TYR:OH	2.20	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	231/265 (87%)	222 (96%)	9 (4%)	0	100	100
1	B	231/265 (87%)	223 (96%)	8 (4%)	0	100	100
1	C	231/265 (87%)	220 (95%)	10 (4%)	1 (0%)	39	36

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	D	232/265 (88%)	225 (97%)	7 (3%)	0	100	100
1	E	232/265 (88%)	225 (97%)	7 (3%)	0	100	100
1	F	231/265 (87%)	223 (96%)	8 (4%)	0	100	100
All	All	1388/1590 (87%)	1338 (96%)	49 (4%)	1 (0%)	56	57

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	164	GLY

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	201/219 (92%)	197 (98%)	4 (2%)	63	67
1	B	201/219 (92%)	201 (100%)	0	100	100
1	C	201/219 (92%)	198 (98%)	3 (2%)	72	77
1	D	202/219 (92%)	199 (98%)	3 (2%)	72	77
1	E	202/219 (92%)	200 (99%)	2 (1%)	82	86
1	F	201/219 (92%)	194 (96%)	7 (4%)	43	44
All	All	1208/1314 (92%)	1189 (98%)	19 (2%)	70	75

All (19) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	45	GLU
1	A	77	SER
1	A	155	ASP
1	A	189	ILE
1	C	75	SER
1	C	162	PHE
1	C	212	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	40	GLN
1	D	167	ARG
1	D	177	LEU
1	E	76	LYS
1	E	216	LYS
1	F	51	ASP
1	F	72	THR
1	F	75	SER
1	F	167	ARG
1	F	215	LYS
1	F	237	LEU
1	F	239	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	129	GLN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no carbohydrates in this entry.

5.6 Ligand geometry (i)

6 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	THM	A	265	-	13,18,18	0.56	0	16,26,26	2.49	3 (18%)
2	THM	B	265	-	13,18,18	0.76	0	16,26,26	2.92	5 (31%)
2	THM	C	265	-	13,18,18	0.53	0	16,26,26	2.68	3 (18%)
2	THM	D	265	-	13,18,18	0.90	0	16,26,26	3.69	7 (43%)
2	THM	E	265	-	13,18,18	0.79	0	16,26,26	2.83	2 (12%)
2	THM	F	265	-	13,18,18	0.56	0	16,26,26	2.80	4 (25%)

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	THM	A	265	-	-	0/2/18/18	0/2/2/2
2	THM	B	265	-	-	0/2/18/18	0/2/2/2
2	THM	C	265	-	-	0/2/18/18	0/2/2/2
2	THM	D	265	-	-	0/2/18/18	0/2/2/2
2	THM	E	265	-	-	0/2/18/18	0/2/2/2
2	THM	F	265	-	-	0/2/18/18	0/2/2/2

There are no bond length outliers.

All (24) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	265	THM	C5-C4-N3	-6.32	118.10	125.14
2	C	265	THM	C5-C4-N3	-5.91	118.56	125.14
2	D	265	THM	C3'-C2'-C1'	-5.85	88.32	102.40
2	A	265	THM	C5-C4-N3	-5.74	118.75	125.14
2	F	265	THM	C5-C4-N3	-5.32	119.21	125.14
2	B	265	THM	C5-C4-N3	-4.81	119.78	125.14
2	D	265	THM	C5-C4-N3	-3.82	120.88	125.14
2	B	265	THM	C5'-C4'-C3'	-3.43	105.67	114.80
2	D	265	THM	C4'-O4'-C1'	-2.59	102.93	109.47
2	B	265	THM	O3'-C3'-C4'	-2.27	100.89	110.05
2	C	265	THM	O4'-C4'-C5'	2.04	113.59	109.17
2	F	265	THM	O4'-C1'-N1	2.13	111.40	107.72
2	A	265	THM	O4'-C4'-C5'	2.28	114.10	109.17
2	D	265	THM	O4'-C4'-C5'	2.58	114.77	109.17

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	D	265	THM	C5M-C5-C4	4.03	125.25	120.05
2	F	265	THM	O4'-C4'-C5'	4.10	118.05	109.17
2	B	265	THM	O4'-C4'-C5'	5.42	120.91	109.17
2	A	265	THM	C4-N3-C2	6.94	121.25	115.25
2	B	265	THM	C4-N3-C2	7.34	121.59	115.25
2	F	265	THM	C4-N3-C2	7.63	121.84	115.25
2	D	265	THM	O4'-C1'-N1	7.80	121.23	107.72
2	C	265	THM	C4-N3-C2	7.81	122.00	115.25
2	D	265	THM	C4-N3-C2	8.43	122.53	115.25
2	E	265	THM	C4-N3-C2	8.46	122.56	115.25

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	265	THM	1	0
2	D	265	THM	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

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	227/265 (85%)	0.27	20 (8%)	12	17	30, 45, 68, 74	0
1	B	227/265 (85%)	-0.03	10 (4%)	38	47	23, 33, 53, 67	0
1	C	227/265 (85%)	0.30	19 (8%)	14	18	29, 47, 72, 86	0
1	D	227/265 (85%)	-0.03	10 (4%)	38	47	23, 34, 55, 67	0
1	E	227/265 (85%)	0.01	10 (4%)	38	47	25, 37, 57, 69	0
1	F	227/265 (85%)	0.05	14 (6%)	24	31	24, 37, 58, 69	0
All	All	1362/1590 (85%)	0.10	83 (6%)	25	32	23, 39, 64, 86	0

All (83) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	72	THR	5.4
1	F	75	SER	5.3
1	C	164	GLY	5.2
1	B	72	THR	4.4
1	D	165	ASP	4.4
1	C	165	ASP	4.3
1	F	72	THR	4.1
1	A	72	THR	3.9
1	E	163	LYS	3.8
1	C	117	ASN	3.7
1	B	165	ASP	3.7
1	D	163	LYS	3.7
1	F	258	GLN	3.6
1	B	164	GLY	3.6
1	C	258	GLN	3.6
1	D	164	GLY	3.6
1	B	167	ARG	3.5
1	E	165	ASP	3.5
1	A	75	SER	3.5

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	163	LYS	3.3
1	B	258	GLN	3.3
1	D	72	THR	3.2
1	A	258	GLN	3.2
1	C	167	ARG	3.1
1	C	26	ALA	3.1
1	C	75	SER	3.1
1	F	26	ALA	3.1
1	A	162	PHE	3.1
1	B	26	ALA	3.0
1	D	76	LYS	3.0
1	D	117	ASN	3.0
1	A	77	SER	3.0
1	E	72	THR	3.0
1	A	164	GLY	3.0
1	E	258	GLN	2.9
1	B	75	SER	2.9
1	F	33	GLU	2.9
1	A	40	GLN	2.9
1	F	167	ARG	2.8
1	E	164	GLY	2.8
1	A	167	ARG	2.8
1	A	56	LEU	2.8
1	C	77	SER	2.8
1	C	118	LYS	2.7
1	A	165	ASP	2.7
1	C	196	GLU	2.7
1	F	165	ASP	2.7
1	A	33	GLU	2.7
1	F	163	LYS	2.6
1	E	77	SER	2.6
1	A	96	VAL	2.5
1	A	163	LYS	2.5
1	C	160	PHE	2.5
1	C	45	GLU	2.5
1	B	163	LYS	2.4
1	A	59	LEU	2.4
1	E	117	ASN	2.4
1	C	242	GLN	2.3
1	E	56	LEU	2.3
1	C	162	PHE	2.3
1	B	56	LEU	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	76	LYS	2.3
1	C	241	TYR	2.3
1	D	167	ARG	2.3
1	A	26	ALA	2.3
1	D	56	LEU	2.2
1	E	75	SER	2.2
1	F	195	VAL	2.2
1	A	197	LYS	2.2
1	C	76	LYS	2.2
1	A	117	ASN	2.2
1	F	76	LYS	2.2
1	F	117	ASN	2.2
1	F	56	LEU	2.1
1	B	76	LYS	2.1
1	A	60	SER	2.1
1	D	97	CYS	2.1
1	E	60	SER	2.1
1	A	222	THR	2.0
1	F	96	VAL	2.0
1	F	241	TYR	2.0
1	C	33	GLU	2.0
1	D	75	SER	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	THM	B	265	17/17	0.90	0.15	1.03	45,53,66,66	0
2	THM	C	265	17/17	0.90	0.17	1.01	64,68,75,77	0
2	THM	D	265	17/17	0.88	0.13	0.57	44,50,59,62	0
2	THM	E	265	17/17	0.90	0.14	0.53	42,50,63,64	0
2	THM	F	265	17/17	0.88	0.12	0.13	48,56,62,62	0
2	THM	A	265	17/17	0.92	0.13	-0.03	62,67,69,70	0

6.5 Other polymers (i)

There are no such residues in this entry.