Full wwPDB X-ray Structure Validation Report (i)

Mar 20, 2021 – 08:01 am GMT

PDB ID	:	6XXS
Title	:	Crystal structure of an NCoR1BBD2-BCL6BTB chimera in complex with the NcoR1 BBD1 corepressor peptide.
Authors	:	Zacharchenko, T.; Wright, S.C.
Deposited on	:	2020-01-28
Resolution	:	3.25 Å(reported)

This is a Full wwPDB X-ray Structure Validation Report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
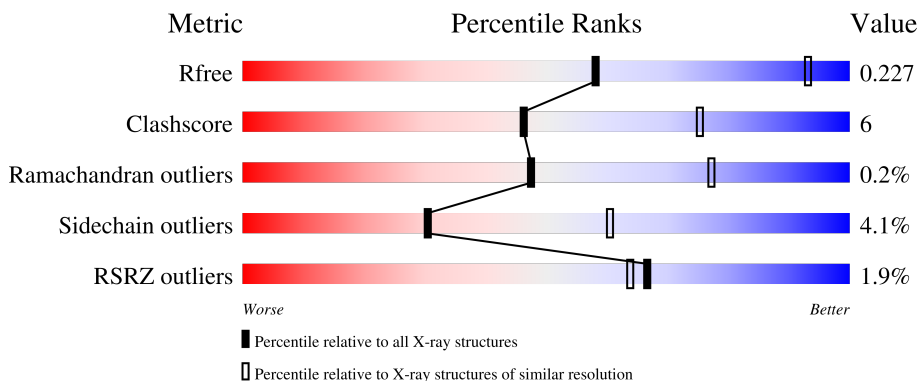with specific help available everywhere you see the (i) symbol.

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 3.25 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.
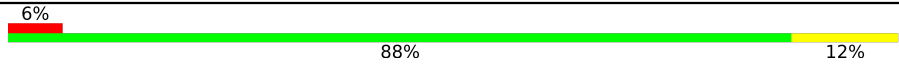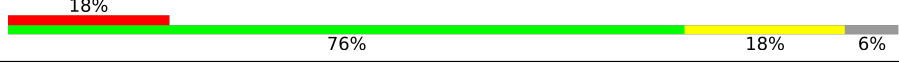
Metric	Percentile Ranks	Value
Rfree		0.227
Clashscore		6
Ramachandran outliers		0.2%
Sidechain outliers		4.1%
RSRZ outliers		1.9%

Worse　　　　　　　　　　　　　　　*Better*

■ Percentile relative to all X-ray structures

▯ Percentile relative to X-ray structures of similar resolution

Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	1191 (3.30-3.22)
Clashscore	141614	1251 (3.30-3.22)
Ramachandran outliers	138981	1229 (3.30-3.22)
Sidechain outliers	138945	1228 (3.30-3.22)
RSRZ outliers	127900	1154 (3.30-3.22)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=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	135	76% · · 19% · ·
1	B	135	% 79% 19% ·
1	F	135	4% 79% 16% · ·
2	C	17	82% 18%
2	D	17	82% 12% 6%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	17	6% · · · 88% · · · 12%
2	H	17	18% · · · 76% · · · 18% · · · 6%
3	E	126	% · · · 75% · · · 21% · ·

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 4693 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 B-cell lymphoma 6 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	130	Total	C	N	O	S	0	0	0
			1051	666	186	191	8			
1	B	131	Total	C	N	O	S	0	2	0
			1070	678	190	194	8			
1	F	129	Total	C	N	O	S	0	1	0
			1049	665	185	191	8			

There are 42 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-5	GLY	-	expression tag	UNP P41182
A	-4	PRO	-	expression tag	UNP P41182
A	-3	SER	-	expression tag	UNP P41182
A	-2	SER	-	expression tag	UNP P41182
A	-1	ASP	-	expression tag	UNP P41182
A	0	LEU	-	expression tag	UNP P41182
A	1	TYR	-	expression tag	UNP P41182
A	2	LEU	-	expression tag	UNP P41182
A	3	ARG	-	expression tag	UNP P41182
A	4	PRO	-	expression tag	UNP P41182
A	5	GLY	-	expression tag	UNP P41182
A	8	GLN	CYS	engineered mutation	UNP P41182
A	67	ARG	CYS	engineered mutation	UNP P41182
A	84	ASN	CYS	engineered mutation	UNP P41182
B	-5	GLY	-	expression tag	UNP P41182
B	-4	PRO	-	expression tag	UNP P41182
B	-3	SER	-	expression tag	UNP P41182
B	-2	SER	-	expression tag	UNP P41182
B	-1	ASP	-	expression tag	UNP P41182
B	0	LEU	-	expression tag	UNP P41182
B	1	TYR	-	expression tag	UNP P41182
B	2	LEU	-	expression tag	UNP P41182
B	3	ARG	-	expression tag	UNP P41182

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
B	4	PRO	-	expression tag	UNP P41182
B	5	GLY	-	expression tag	UNP P41182
B	8	GLN	CYS	engineered mutation	UNP P41182
B	67	ARG	CYS	engineered mutation	UNP P41182
B	84	ASN	CYS	engineered mutation	UNP P41182
F	-5	GLY	-	expression tag	UNP P41182
F	-4	PRO	-	expression tag	UNP P41182
F	-3	SER	-	expression tag	UNP P41182
F	-2	SER	-	expression tag	UNP P41182
F	-1	ASP	-	expression tag	UNP P41182
F	0	LEU	-	expression tag	UNP P41182
F	1	TYR	-	expression tag	UNP P41182
F	2	LEU	-	expression tag	UNP P41182
F	3	ARG	-	expression tag	UNP P41182
F	4	PRO	-	expression tag	UNP P41182
F	5	GLY	-	expression tag	UNP P41182
F	8	GLN	CYS	engineered mutation	UNP P41182
F	67	ARG	CYS	engineered mutation	UNP P41182
F	84	ASN	CYS	engineered mutation	UNP P41182

- Molecule 2 is a protein called Nuclear receptor corepressor 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	17	Total 135	C 83	N 26	O 25	S 1	0	0	0
2	D	16	Total 131	C 81	N 25	O 24	S 1	0	0	0
2	G	17	Total 135	C 83	N 26	O 25	S 1	0	0	0
2	H	16	Total 131	C 81	N 25	O 24	S 1	0	0	0

- Molecule 3 is a protein called B-cell lymphoma 6 protein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	E	122	Total 991	C 629	N 175	O 179	S 8	0	1	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
E	4	GLY	-	expression tag	UNP P41182

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
E	5	PRO	-	expression tag	UNP P41182
E	8	GLN	CYS	engineered mutation	UNP P41182
E	67	ARG	CYS	engineered mutation	UNP P41182
E	84	ASN	CYS	engineered mutation	UNP P41182

3 Residue-property plots ⓘ

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.
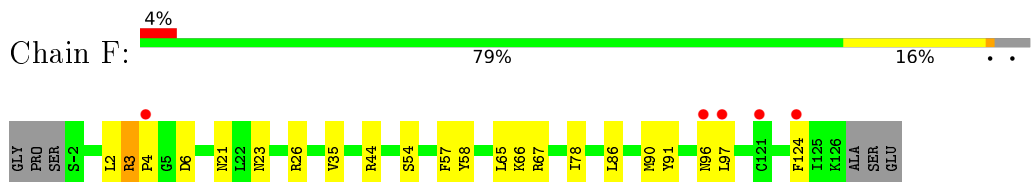
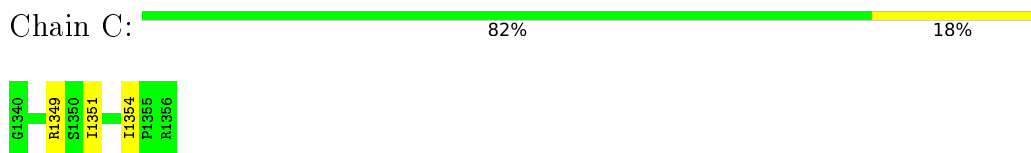- Molecule 1: B-cell lymphoma 6 protein
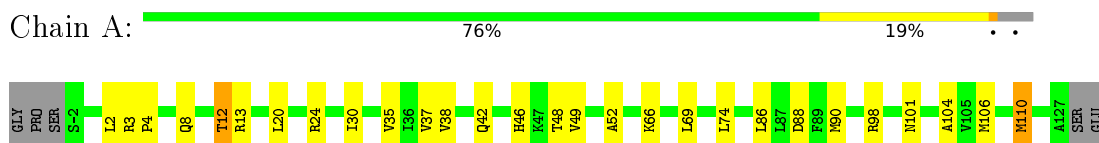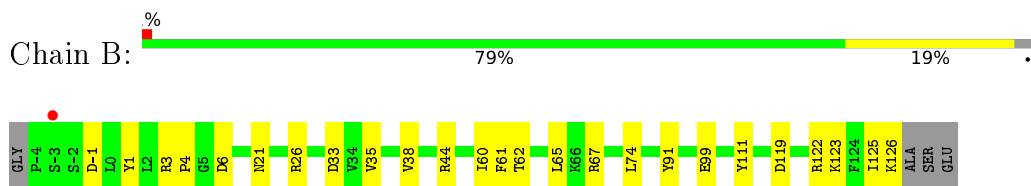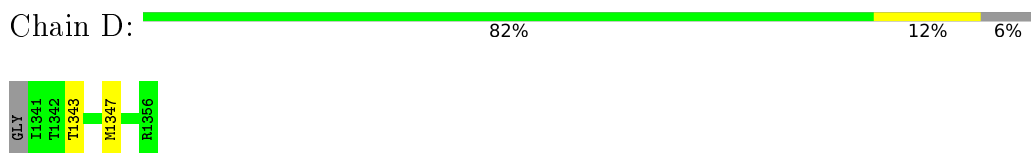
Chain A:

- Molecule 1: B-cell lymphoma 6 protein

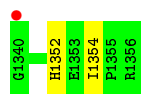Chain B:

- Molecule 1: B-cell lymphoma 6 protein
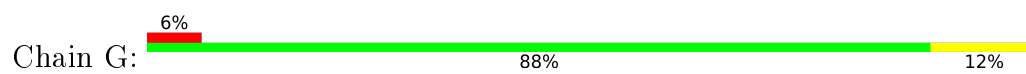
Chain F:

- Molecule 2: Nuclear receptor corepressor 1

Chain C:

- Molecule 2: Nuclear receptor corepressor 1

Chain D:

- Molecule 2: Nuclear receptor corepressor 1

Chain G:

6%   88%   12%

G1340 H1352 E1353 I1354 P1355 R1356

• Molecule 2: Nuclear receptor corepressor 1

Chain H:

18%   76%   18%   6%

GLY I1341 T1342 T1343 I1344 K1345 I1354 P1355 R1356
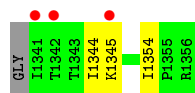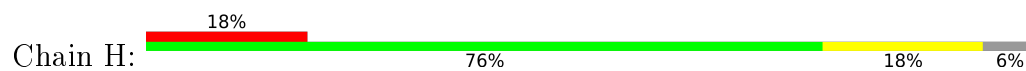
• Molecule 3: B-cell lymphoma 6 protein

Chain E:

%   75%   21%

GLY PRO D6 S7 Q8 L19 L20 N21 L22 R26 S27 R28 D29 I30 V35 S39 R44 K47 F57 I60 F61 L65 K66 R67 N68 L69 Y91 R94 R98 I102 M106 M110 D119 F124 I125 K126 A127 SER GLU
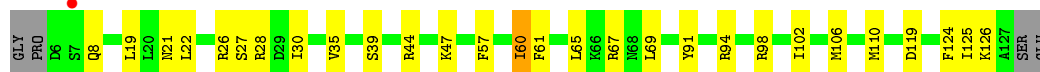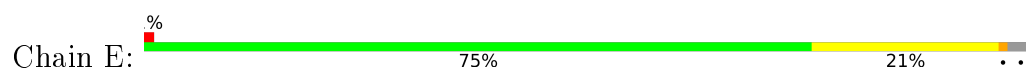
4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 65 2 2	Depositor
Cell constants a, b, c, α, β, γ	165.38Å　　165.38Å　　244.80Å 90.00°　　　90.00°　　　120.00°	Depositor
Resolution (Å)	82.69　–　3.25 82.69　–　3.25	Depositor EDS
% Data completeness (in resolution range)	99.9 (82.69-3.25) 100.0 (82.69-3.25)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$<I/\sigma(I)>$ [1]	1.33 (at 3.26Å)	Xtriage
Refinement program	PHENIX 1.17.1_3660, PHENIX 1.17.1_3660	Depositor
R, R_{free}	0.197　,　0.225 0.198　,　0.227	Depositor DCC
R_{free} test set	1610 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	70.0	Xtriage
Anisotropy	0.049	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 56.7	EDS
L-test for twinning [2]	$<|L|> = 0.46, <L^2> = 0.29$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.93	EDS
Total number of atoms	4693	wwPDB-VP
Average B, all atoms (Å^2)	78.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.00% 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, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

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	0.27	0/1067	0.44	0/1440
1	B	0.27	0/1093	0.45	0/1474
1	F	0.26	0/1068	0.43	0/1441
2	C	0.27	0/136	0.46	0/179
2	D	0.26	0/132	0.47	0/174
2	G	0.25	0/136	0.47	0/179
2	H	0.24	0/132	0.50	0/174
3	E	0.27	0/1008	0.44	0/1360
All	All	0.27	0/4772	0.44	0/6421

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 (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	1051	0	1073	16	0
1	B	1070	0	1098	15	0
1	F	1049	0	1073	19	0
2	C	135	0	142	3	0
2	D	131	0	139	1	0
2	G	135	0	142	1	0
2	H	131	0	139	3	0
3	E	991	0	1016	21	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	4693	0	4822	60	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 6.

All (60) 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:B:35:VAL:HG12	1:B:44:ARG:HG2	1.74	0.69
1:A:38:VAL:HG22	1:A:74:LEU:HD12	1.78	0.65
3:E:35[B]:VAL:HG12	3:E:44:ARG:HG2	1.77	0.65
3:E:102:ILE:HD11	3:E:125:ILE:HG22	1.79	0.65
1:F:124:PHE:HE2	2:H:1344:ILE:HD12	1.67	0.59
3:E:47:LYS:NZ	3:E:67:ARG:O	2.36	0.58
1:A:2:LEU:HD21	1:B:125:ILE:HG23	1.85	0.57
1:A:101:ASN:HB2	1:A:104:ALA:HB3	1.85	0.57
3:E:35[A]:VAL:HG22	3:E:44:ARG:HG2	1.86	0.56
3:E:30:ILE:HG23	1:F:67:ARG:HH12	1.71	0.55
1:B:33:ASP:OD2	1:B:33:ASP:N	2.40	0.55
1:A:46:HIS:HB2	1:A:49:VAL:HG23	1.89	0.54
1:B:3:ARG:HB3	1:B:4:PRO:HD3	1.89	0.53
1:B:119:ASP:O	1:B:123:LYS:HD3	2.09	0.52
1:F:3:ARG:CZ	1:F:4:PRO:HD2	2.42	0.50
1:A:37:VAL:HG22	1:A:42:GLN:HG2	1.93	0.50
1:B:26:ARG:HD3	1:B:91:TYR:CZ	2.48	0.49
1:A:106:MET:O	1:A:110:MET:HG2	2.13	0.49
1:B:21:ASN:ND2	2:C:1354:ILE:HG13	2.28	0.49
1:F:3:ARG:HB3	1:F:3:ARG:NH2	2.28	0.48
3:E:125:ILE:O	3:E:126:LYS:NZ	2.33	0.48
1:B:-1:ASP:HB3	1:B:1:TYR:CE2	2.50	0.47
1:A:52:ALA:O	2:C:1351:ILE:HD12	2.14	0.47
3:E:98:ARG:HH12	1:F:6:ASP:HB3	1.79	0.47
1:F:6:ASP:N	1:F:6:ASP:OD1	2.48	0.46
3:E:30:ILE:HG23	1:F:67:ARG:NH1	2.31	0.46
1:A:30:ILE:HA	1:B:67:ARG:HH12	1.80	0.46
1:B:122:ARG:O	1:B:126:LYS:HG3	2.16	0.45
1:B:65:LEU:HD12	1:B:65:LEU:H	1.81	0.44
2:G:1354:ILE:HG13	1:F:21:ASN:ND2	2.33	0.44
3:E:102:ILE:HD12	3:E:124:PHE:HE2	1.83	0.44
1:B:21:ASN:HB3	2:C:1354:ILE:HD12	2.00	0.43
1:F:124:PHE:CE2	2:H:1344:ILE:HD12	2.50	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:E:22:LEU:HD23	3:E:22:LEU:HA	1.91	0.43
1:F:54:SER:HB3	1:F:57:PHE:HB2	2.01	0.43
1:A:98:ARG:HH22	1:F:44:ARG:HH22	1.67	0.42
3:E:19:LEU:HD12	3:E:19:LEU:HA	1.89	0.42
3:E:28:ARG:NH1	1:F:58:TYR:OH	2.46	0.42
3:E:98:ARG:HH11	3:E:98:ARG:HA	1.84	0.42
1:B:38:VAL:HG22	1:B:74:LEU:HD12	2.01	0.42
1:A:8:GLN:HE21	2:D:1343:THR:HG22	1.84	0.42
1:A:86:LEU:O	1:A:90:MET:HG3	2.20	0.42
1:A:20:LEU:O	1:A:24:ARG:HG2	2.19	0.42
1:A:66:LYS:HA	1:A:69:LEU:HD13	2.01	0.42
1:A:88:ASP:OD1	1:F:66:LYS:NZ	2.39	0.41
3:E:125:ILE:HB	1:F:2:LEU:HD21	2.01	0.41
3:E:26:ARG:HD3	3:E:91:TYR:CZ	2.55	0.41
1:F:78:ILE:HD13	1:F:78:ILE:HA	1.81	0.41
1:B:74:LEU:HD22	1:B:111:TYR:HE2	1.85	0.41
3:E:57:PHE:HD1	3:E:60:ILE:HD11	1.86	0.41
1:F:26:ARG:HD3	1:F:91:TYR:CZ	2.55	0.41
1:A:12:THR:O	1:A:13:ARG:HG2	2.20	0.41
3:E:102:ILE:HD13	1:F:2:LEU:HD13	2.03	0.41
1:F:86:LEU:O	1:F:90:MET:HG3	2.21	0.41
1:B:60:ILE:HG22	1:B:61:PHE:CD1	2.56	0.40
3:E:47:LYS:HE2	3:E:61:PHE:CE1	2.56	0.40
3:E:8:GLN:HG3	1:F:96:ASN:OD1	2.21	0.40
3:E:106:MET:O	3:E:110:MET:HG3	2.21	0.40
3:E:21:ASN:ND2	2:H:1354:ILE:HG13	2.36	0.40
1:A:3:ARG:NH1	1:A:4:PRO:HD2	2.36	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	128/135 (95%)	119 (93%)	9 (7%)	0	100	100
1	B	131/135 (97%)	125 (95%)	6 (5%)	0	100	100
1	F	128/135 (95%)	120 (94%)	7 (6%)	1 (1%)	19	52
2	C	15/17 (88%)	15 (100%)	0	0	100	100
2	D	14/17 (82%)	14 (100%)	0	0	100	100
2	G	15/17 (88%)	14 (93%)	1 (7%)	0	100	100
2	H	14/17 (82%)	14 (100%)	0	0	100	100
3	E	121/126 (96%)	116 (96%)	5 (4%)	0	100	100
All	All	566/599 (94%)	537 (95%)	28 (5%)	1 (0%)	47	77

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	F	3	ARG

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	119/123 (97%)	115 (97%)	4 (3%)	37	64
1	B	123/123 (100%)	120 (98%)	3 (2%)	49	72
1	F	120/123 (98%)	116 (97%)	4 (3%)	38	65
2	C	15/15 (100%)	14 (93%)	1 (7%)	16	45
2	D	15/15 (100%)	14 (93%)	1 (7%)	16	45
2	G	15/15 (100%)	14 (93%)	1 (7%)	16	45
2	H	15/15 (100%)	14 (93%)	1 (7%)	16	45
3	E	113/115 (98%)	106 (94%)	7 (6%)	18	48
All	All	535/544 (98%)	513 (96%)	22 (4%)	30	60

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	12	THR
1	A	35	VAL
1	A	48	THR
1	A	110	MET
1	B	6	ASP
1	B	62	THR
1	B	99	GLU
2	C	1349	ARG
2	D	1347	MET
3	E	27	SER
3	E	39	SER
3	E	60	ILE
3	E	65	LEU
3	E	69	LEU
3	E	94	ARG
3	E	119	ASP
2	G	1352	HIS
1	F	23	ASN
1	F	35	VAL
1	F	65	LEU
1	F	97	LEU
2	H	1345	LYS

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

Mol	Chain	Res	Type
1	A	8	GLN
1	B	96	ASN

5.3.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

There are no ligands in this entry.

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data (i)

6.1 Protein, DNA and RNA chains (i)

In the following table, the column labelled '#RSRZ> 2' contains the number (and percentage) of RSRZ outliers, followed by percent RSRZ outliers for the chain as percentile scores relative to all X-ray entries and entries of similar resolution. The OWAB column contains the minimum, median, 95th percentile and maximum values of the occupancy-weighted average B-factor per residue. The column labelled 'Q< 0.9' lists the number of (and percentage) of residues with an average occupancy less than 0.9.

Mol	Chain	Analysed	<RSRZ>	#RSRZ>2			OWAB(Å^2)	Q<0.9
1	A	130/135 (96%)	-0.00	0	100	100	38, 62, 119, 143	1 (0%)
1	B	131/135 (97%)	0.18	1 (0%)	86	86	47, 63, 122, 158	0
1	F	129/135 (95%)	0.33	5 (3%)	39	36	55, 82, 132, 154	0
2	C	17/17 (100%)	0.22	0	100	100	45, 68, 89, 96	0
2	D	16/17 (94%)	-0.13	0	100	100	47, 68, 96, 103	0
2	G	17/17 (100%)	0.28	1 (5%)	22	21	67, 87, 148, 154	0
2	H	16/17 (94%)	1.05	3 (18%)	1	1	68, 106, 142, 142	0
3	E	122/126 (96%)	0.08	1 (0%)	86	86	46, 71, 134, 169	0
All	All	578/599 (96%)	0.17	11 (1%)	66	64	38, 71, 129, 169	1 (0%)

All (11) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	-3	SER	4.1
3	E	7	SER	4.0
2	H	1341	ILE	3.6
2	H	1342	THR	3.1
1	F	97	LEU	2.7
1	F	4	PRO	2.5
2	G	1340	GLY	2.5
2	H	1345	LYS	2.2
1	F	121	CYS	2.1
1	F	124	PHE	2.1
1	F	96	ASN	2.1

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

There are no ligands in this entry.

6.5 Other polymers (i)

There are no such residues in this entry.