wwPDB X-ray Structure Validation Summary Report (i)

Sep 6, 2023 – 07:31 AM EDT

PDB ID	:	4E41
Title	:	Structural basis for the recognition of mutant self by a tumor-specific, MHC class II-restricted T cell receptor G4
Authors	:	Deng, L.; Langley, R.J.; Wang, Q.; Topalian, S.L.; Mariuzza, R.A.
Deposited on	:	2012-03-11
Resolution	:	2.60 Å(reported)

This is a wwPDB X-ray Structure Validation Summary 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.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

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

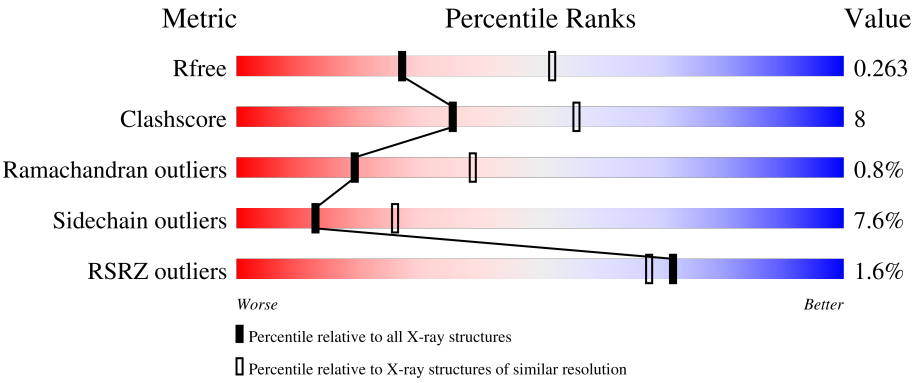
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

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 2.60 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.



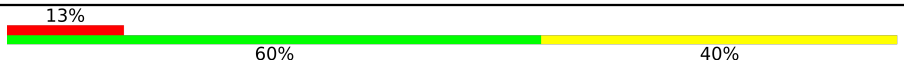
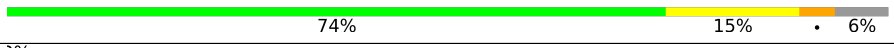
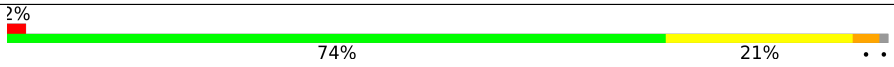
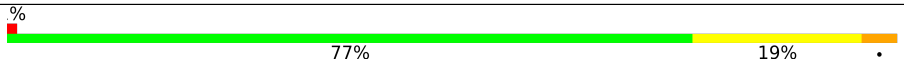
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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	182	2% 82% 15%
1	F	182	2% 75% 20%
2	B	190	3% 73% 19% 6%
2	G	190	3% 68% 23% 7%
3	C	15	7% 67% 27% 7%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	H	15	
4	D	203	
4	I	203	
5	E	239	
5	J	239	

2 Entry composition (i)

There are 7 unique types of molecules in this entry. The entry contains 12983 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 HLA class II histocompatibility antigen, DR alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	179	Total	C	N	O	S	0	0	0
			1473	954	239	275	5			
1	F	179	Total	C	N	O	S	0	0	0
			1473	954	239	275	5			

- Molecule 2 is a protein called HLA class II histocompatibility antigen, DRB1-1 beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	179	Total	C	N	O	S	0	0	0
			1468	926	261	275	6			
2	G	177	Total	C	N	O	S	0	0	0
			1450	915	256	273	6			

- Molecule 3 is a protein called Triosephosphate isomerase.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	C	15	Total	C	N	O	0	0	0
			104	66	17	21			
3	H	15	Total	C	N	O	0	0	0
			104	66	17	21			

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	28	ILE	THR	conflict	UNP P60174
H	28	ILE	THR	conflict	UNP P60174

- Molecule 4 is a protein called T cell receptor G4 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	D	181	Total	C	N	O	S	0	1	0
			1432	903	240	282	7			

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	I	190	Total	C	N	O	S	0	0	0
			1497	940	252	298	7			

- Molecule 5 is a protein called T cell receptor G4 beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	E	236	Total	C	N	O	S	0	0	0
			1898	1200	334	359	5			
5	J	238	Total	C	N	O	S	0	0	0
			1910	1208	336	361	5			

- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	A	1	Total	Na	0	0
			1	1		
6	B	1	Total	Na	0	0
			1	1		

- Molecule 7 is water.

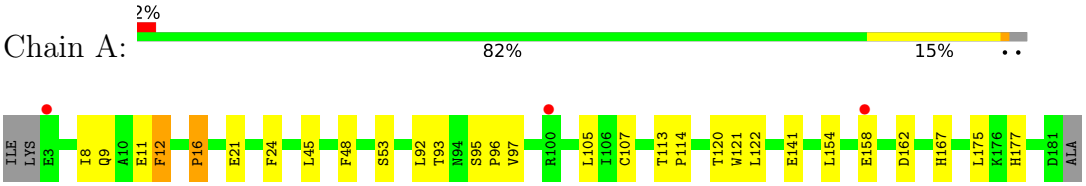
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	26	Total	O	0	0
			26	26		
7	B	20	Total	O	0	0
			20	20		
7	C	1	Total	O	0	0
			1	1		
7	D	22	Total	O	0	0
			22	22		
7	E	13	Total	O	0	0
			13	13		
7	F	21	Total	O	0	0
			21	21		
7	G	21	Total	O	0	0
			21	21		
7	H	4	Total	O	0	0
			4	4		
7	I	26	Total	O	0	0
			26	26		
7	J	18	Total	O	0	0
			18	18		

3 Residue-property plots (i)

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.

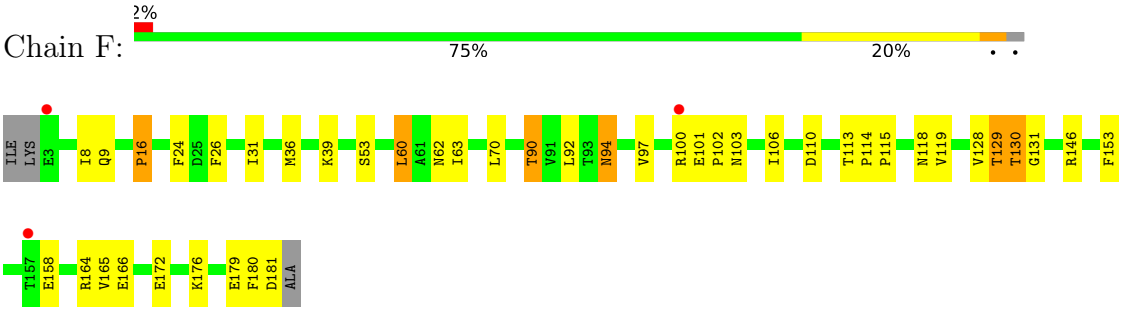
• Molecule 1: HLA class II histocompatibility antigen, DR alpha chain

Chain A:



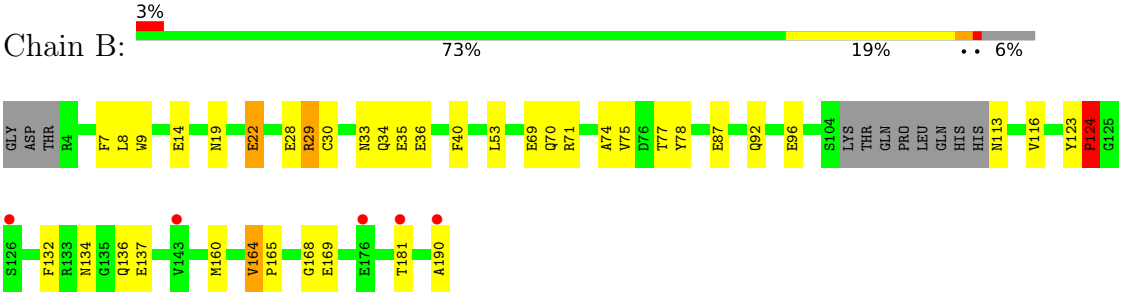
• Molecule 1: HLA class II histocompatibility antigen, DR alpha chain

Chain F:



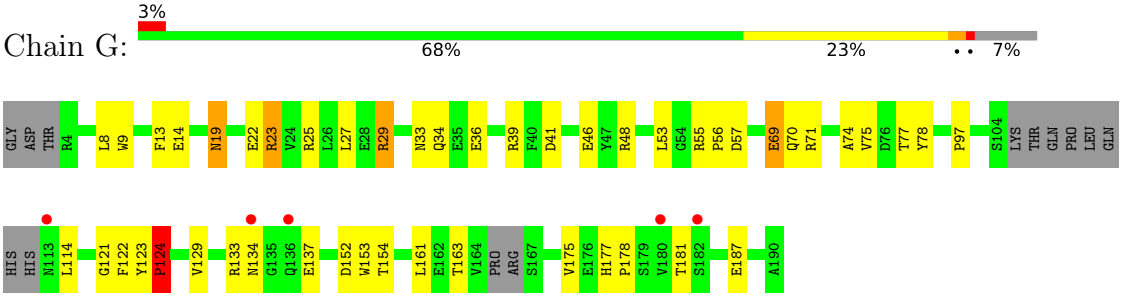
• Molecule 2: HLA class II histocompatibility antigen, DRB1-1 beta chain

Chain B:



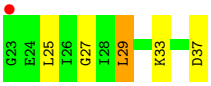
• Molecule 2: HLA class II histocompatibility antigen, DRB1-1 beta chain

Chain G:



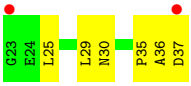
• Molecule 3: Triosephosphate isomerase

Chain C:



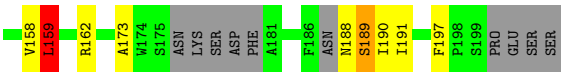
• Molecule 3: Triosephosphate isomerase

Chain H:



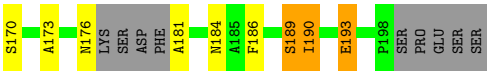
• Molecule 4: T cell receptor G4 alpha chain

Chain D:



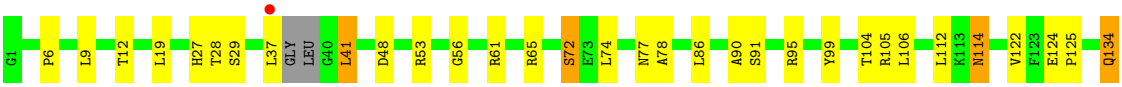
• Molecule 4: T cell receptor G4 alpha chain

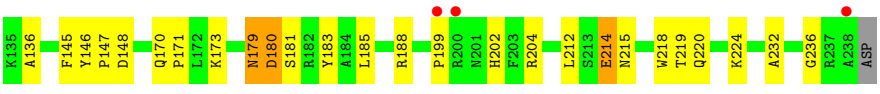
Chain I:



• Molecule 5: T cell receptor G4 beta chain

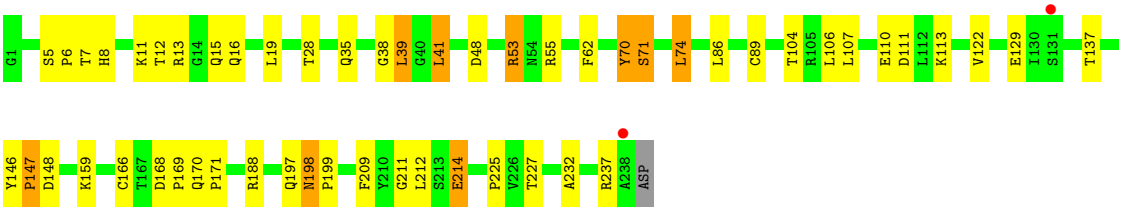
Chain E:





• Molecule 5: T cell receptor G4 beta chain

Chain J:


G1
S5
P6
T7
H8
K11
T12
R13
G14
Q15
Q16
L19
T28
Q35
G38
L39
G40
L41
D48
R53
N54
R55
F62
Y70
S71
L74
L86
C89
T104
R105
L106
L107
E110
D111
L112
K113
V122
E129
I130
S131
T137

Y146
P147
D148
K159
C166
T167
D168
P169
Q170
P171
R188
Q197
N198
P199
F209
Y210
G211
L212
S213
E214
P225
V226
T227
A232
R237
A238
ASP

4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, α, β, γ	126.13Å 175.61Å 88.65Å 90.00° 110.75° 90.00°	Depositor
Resolution (Å)	30.00 – 2.60 27.35 – 2.41	Depositor EDS
% Data completeness (in resolution range)	97.2 (30.00-2.60) 88.8 (27.35-2.41)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.72 (at 2.42Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.197 , 0.264 0.198 , 0.263	Depositor DCC
R_{free} test set	3102 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å^2)	44.2	Xtriage
Anisotropy	0.266	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 40.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	12983	wwPDB-VP
Average B, all atoms (Å^2)	48.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.48% 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)

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

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	0.60	1/1518 (0.1%)	0.84	2/2070 (0.1%)
1	F	0.68	1/1518 (0.1%)	0.72	0/2070
2	B	0.62	1/1504 (0.1%)	0.69	0/2041
2	G	0.61	0/1484	0.66	1/2012 (0.0%)
3	C	0.59	0/104	0.74	0/139
3	H	0.64	0/104	0.65	0/139
4	D	0.64	1/1468 (0.1%)	0.75	1/1995 (0.1%)
4	I	0.61	0/1530	0.72	0/2081
5	E	0.64	2/1951 (0.1%)	0.68	1/2658 (0.0%)
5	J	0.67	3/1964 (0.2%)	0.73	3/2677 (0.1%)
All	All	0.63	9/13145 (0.1%)	0.72	8/17882 (0.0%)

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	16	PRO	N-CD	5.66	1.55	1.47
4	D	7	PRO	N-CD	5.50	1.55	1.47
1	F	16	PRO	N-CD	5.42	1.55	1.47
5	E	147	PRO	N-CD	5.25	1.55	1.47
5	E	6	PRO	N-CD	5.24	1.55	1.47

The worst 5 of 8 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	113	THR	C-N-CD	-18.40	80.12	120.60
1	A	113	THR	C-N-CA	11.75	171.35	122.00
4	D	159	LEU	CA-CB-CG	5.46	127.87	115.30
2	G	124	PRO	CA-N-CD	-5.25	104.15	111.50
5	J	5	SER	C-N-CD	5.25	139.42	128.40

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	1473	0	1407	13	0
1	F	1473	0	1407	26	0
2	B	1468	0	1402	21	0
2	G	1450	0	1381	39	0
3	C	104	0	109	5	0
3	H	104	0	109	6	0
4	D	1432	0	1357	25	0
4	I	1497	0	1420	32	0
5	E	1898	0	1805	31	0
5	J	1910	0	1820	30	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
7	A	26	0	0	1	0
7	B	20	0	0	0	0
7	C	1	0	0	0	0
7	D	22	0	0	1	0
7	E	13	0	0	2	0
7	F	21	0	0	0	0
7	G	21	0	0	0	0
7	H	4	0	0	0	0
7	I	26	0	0	4	0
7	J	18	0	0	1	0
All	All	12983	0	12217	203	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 8.

The worst 5 of 203 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)
4:I:36:GLN:HE22	5:J:35:GLN:HE22	1.15	0.89
2:G:8:LEU:H	2:G:33:ASN:ND2	1.74	0.85

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:J:237:ARG:HB2	5:J:237:ARG:HH11	1.42	0.84
4:I:109:PRO:HD3	4:I:158:VAL:HG11	1.61	0.83
4:I:158:VAL:HG13	4:I:169:ASN:ND2	1.99	0.78

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	177/182 (97%)	169 (96%)	8 (4%)	0	100	100
1	F	177/182 (97%)	169 (96%)	8 (4%)	0	100	100
2	B	175/190 (92%)	166 (95%)	8 (5%)	1 (1%)	25	47
2	G	171/190 (90%)	158 (92%)	11 (6%)	2 (1%)	13	27
3	C	13/15 (87%)	13 (100%)	0	0	100	100
3	H	13/15 (87%)	12 (92%)	1 (8%)	0	100	100
4	D	172/203 (85%)	159 (92%)	11 (6%)	2 (1%)	13	27
4	I	184/203 (91%)	179 (97%)	4 (2%)	1 (0%)	29	52
5	E	232/239 (97%)	215 (93%)	14 (6%)	3 (1%)	12	24
5	J	236/239 (99%)	221 (94%)	12 (5%)	3 (1%)	12	24
All	All	1550/1658 (94%)	1461 (94%)	77 (5%)	12 (1%)	19	39

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	D	190	ILE
5	E	180	ASP
5	J	39	LEU
5	J	71	SER

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	E	214	GLU

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	164/166 (99%)	155 (94%)	9 (6%)	21 43
1	F	164/166 (99%)	153 (93%)	11 (7%)	16 33
2	B	161/171 (94%)	148 (92%)	13 (8%)	11 23
2	G	159/171 (93%)	151 (95%)	8 (5%)	24 47
3	C	10/10 (100%)	7 (70%)	3 (30%)	0 0
3	H	10/10 (100%)	10 (100%)	0	100 100
4	D	163/184 (89%)	149 (91%)	14 (9%)	10 20
4	I	171/184 (93%)	152 (89%)	19 (11%)	6 11
5	E	206/208 (99%)	189 (92%)	17 (8%)	11 22
5	J	207/208 (100%)	194 (94%)	13 (6%)	18 36
All	All	1415/1478 (96%)	1308 (92%)	107 (8%)	13 26

5 of 107 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	F	60	LEU
2	G	69	GLU
5	J	106	LEU
1	F	90	THR
1	F	172	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 52 such sidechains are listed below:

Mol	Chain	Res	Type
1	F	62	ASN
2	G	150	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
5	J	114	ASN
1	F	94	ASN
2	G	70	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 2 ligands modelled in this entry, 2 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	179/182 (98%)	-0.32	3 (1%) 70 66	30, 43, 63, 72	0
1	F	179/182 (98%)	-0.31	3 (1%) 70 66	33, 44, 63, 77	0
2	B	179/190 (94%)	-0.31	5 (2%) 53 46	27, 45, 72, 75	0
2	G	177/190 (93%)	-0.17	5 (2%) 53 46	31, 48, 80, 91	0
3	C	15/15 (100%)	0.43	1 (6%) 17 13	32, 38, 56, 58	0
3	H	15/15 (100%)	0.07	2 (13%) 3 2	33, 38, 55, 61	0
4	D	181/203 (89%)	-0.30	0 100 100	26, 42, 78, 81	0
4	I	190/203 (93%)	-0.34	0 100 100	27, 40, 74, 77	0
5	E	236/239 (98%)	-0.08	4 (1%) 70 66	31, 55, 72, 85	0
5	J	238/239 (99%)	-0.23	2 (0%) 86 84	28, 49, 66, 76	0
All	All	1589/1658 (95%)	-0.24	25 (1%) 72 68	26, 47, 73, 91	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	190	ALA	4.9
3	C	23	GLY	4.3
5	E	238	ALA	3.4
5	E	37	LEU	3.1
1	F	3	GLU	2.9

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 (i)

There are no monosaccharides 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. 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	B-factors(Å^2)	Q<0.9
6	NA	A	201	1/1	0.75	0.19	72,72,72,72	0
6	NA	B	201	1/1	0.85	0.34	53,53,53,53	0

6.5 Other polymers (i)

There are no such residues in this entry.