



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 30, 2023 – 07:19 pm BST

PDB ID : 7Q9B  
Title : MHC Class I A02 Allele presenting EAAGIGILTV, in complex with Mel8 TCR  
Authors : Rizkallah, P.J.; Sewell, A.K.; Wall, A.; Fuller, A.  
Deposited on : 2021-11-12  
Resolution : 3.24 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](mailto: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 ⓘ 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 ⓘ](#)) were used in the production of this report:

MolProbity : 4.02b-467  
Mogul : 1.8.4, CSD as541be (2020)  
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

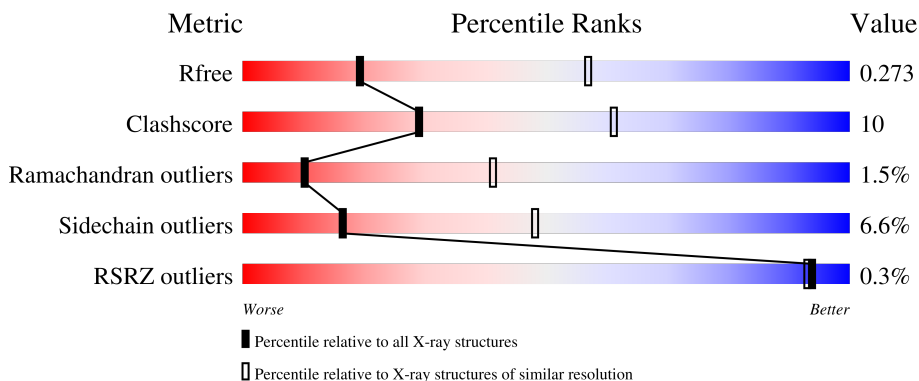
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.24 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.




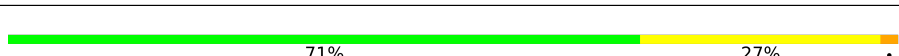
Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1619 (3.28-3.20)
Clashscore	141614	1755 (3.28-3.20)
Ramachandran outliers	138981	1728 (3.28-3.20)
Sidechain outliers	138945	1727 (3.28-3.20)
RSRZ outliers	127900	1567 (3.28-3.20)

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  $\geq 3$ , 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  $\leq 5\%$ . 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	AAA	275	77% (green), 20% (yellow), 3% (orange), 0% (red)
1	FFF	275	76% (green), 23% (yellow), 1% (orange), 0% (red)
2	BBB	100	81% (green), 16% (yellow), 3% (orange), 0% (red)
2	GGG	100	80% (green), 16% (yellow), 4% (orange), 0% (red)
3	CCC	10	30% (green), 60% (yellow), 10% (orange), 0% (red)

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	HHH	10	 70% 10% 10% 10%
4	DDD	193	 % 75% 22% .
4	III	193	 2% 70% 26% . .
5	EEE	245	 67% 31% .
5	JJJ	245	 71% 27% .

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 13221 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 MHC class I antigen.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	AAA	275	2247	1403	409	426	9	0	0	0
1	FFF	275	2247	1403	409	426	9	0	0	0

- Molecule 2 is a protein called Beta-2-microglobulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	BBB	100	837	533	141	159	4	0	0	0
2	GGG	100	837	533	141	159	4	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
BBB	0	MET	-	initiating methionine	UNP P61769
GGG	0	MET	-	initiating methionine	UNP P61769

- Molecule 3 is a protein called GLU-ALA-ALA-GLY-ILE-GLY-ILE-LEU-THR-VAL.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	CCC	10	66	42	10	14	0	0	0
3	HHH	10	66	42	10	14	0	0	0

- Molecule 4 is a protein called Human T Cell Receptor Mel8, Alpha Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	DDD	193	1501	932	251	310	8	0	0	0

*Continued on next page...*

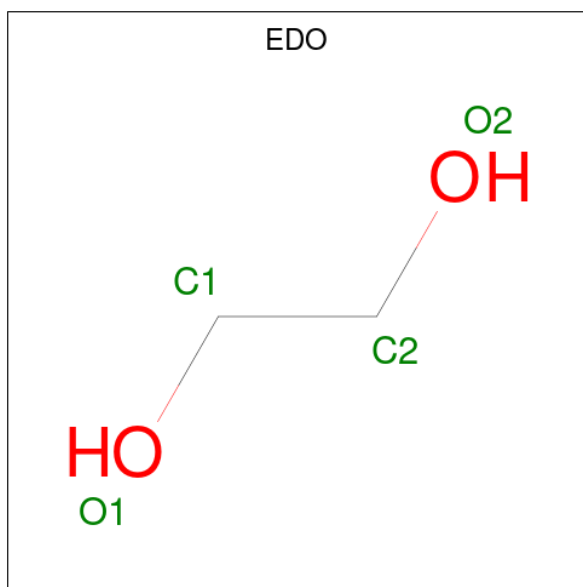
Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	III	190	Total	C	N	O	S	0	0	0
			1475	915	248	304	8			

- Molecule 5 is a protein called Human T Cell Receptor Mel8, Beta Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
5	EEE	245	Total	C	N	O	S	0	0	0
			1942	1221	332	380	9			
5	JJJ	245	Total	C	N	O	S	0	0	0
			1942	1221	332	380	9			

- Molecule 6 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	AAA	1	Total	C	O	0	0
			4	2	2		
6	GGG	1	Total	C	O	0	0
			4	2	2		

- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	AAA	11	Total	O	0	0
			11	11		
7	BBB	7	Total	O	0	0
			7	7		

Continued on next page...


*Continued from previous page...*

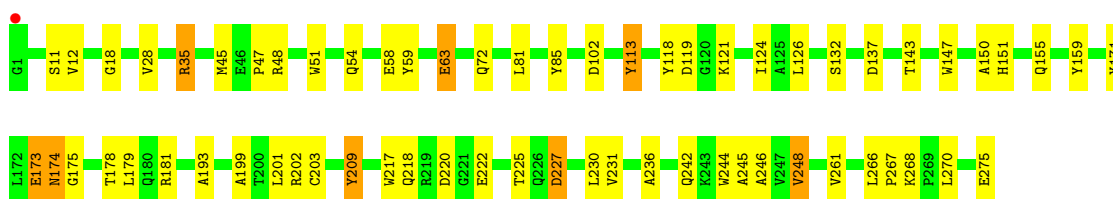
<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>	<b>ZeroOcc</b>	<b>AltConf</b>
7	CCC	1	Total O 1 1	0	0
7	DDD	11	Total O 11 11	0	0
7	EEE	5	Total O 5 5	0	0
7	FFF	7	Total O 7 7	0	0
7	GGG	6	Total O 6 6	0	0
7	III	2	Total O 2 2	0	0
7	JJJ	3	Total O 3 3	0	0

### 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.

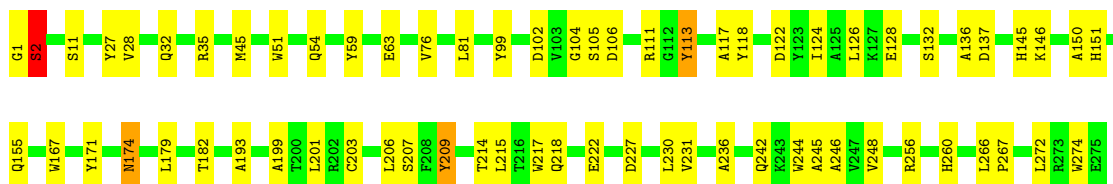
- Molecule 1: MHC class I antigen

Chain AAA:  77% 20%




- Molecule 1: MHC class I antigen

Chain FFF:  76% 23%




- Molecule 2: Beta-2-microglobulin

Chain BBB:  81% 16%



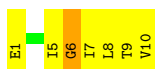
- Molecule 2: Beta-2-microglobulin

Chain GGG:  80% 16%



- Molecule 3: GLU-ALA-ALA-GLY-ILE-GLY-ILE-LEU-THR-VAL

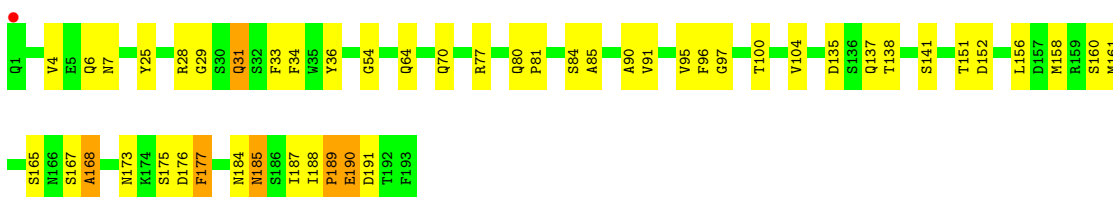
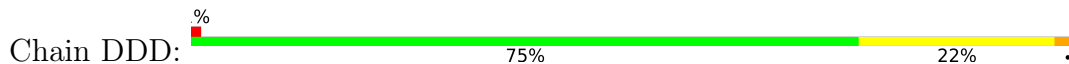
Chain CCC:  30% 60% 10%



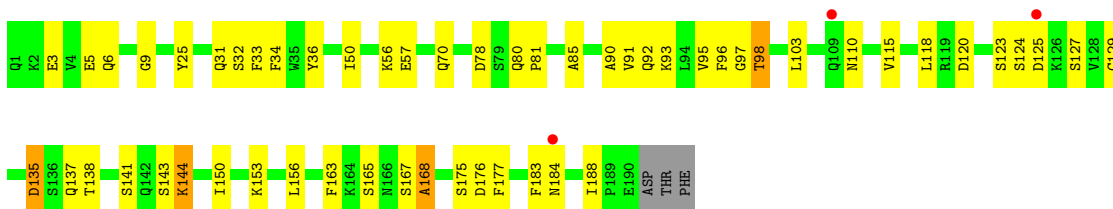
- Molecule 3: GLU-ALA-ALA-GLY-ILE-GLY-ILE-LEU-THR-VAL



- Molecule 4: Human T Cell Receptor Mel8, Alpha Chain



- Molecule 4: Human T Cell Receptor Mel8, Alpha Chain



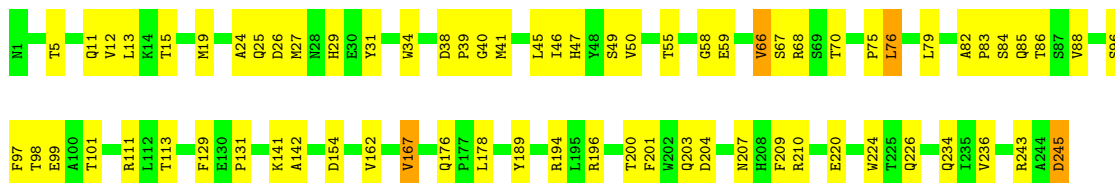
- Molecule 5: Human T Cell Receptor Mel8, Beta Chain



- Molecule 5: Human T Cell Receptor Mel8, Beta Chain







## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	99.97Å 53.65Å 203.51Å 90.00° 94.33° 90.00°	Depositor
Resolution (Å)	99.69 – 3.24 99.69 – 3.24	Depositor EDS
% Data completeness (in resolution range)	99.9 (99.69-3.24) 99.9 (99.69-3.24)	Depositor EDS
$R_{merge}$	0.28	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.49 (at 3.26Å)	Xtrriage
Refinement program	REFMAC 5.8.0267	Depositor
R, $R_{free}$	0.226 , 0.277 0.227 , 0.273	Depositor DCC
$R_{free}$ test set	1743 reflections (4.98%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	46.0	Xtrriage
Anisotropy	1.109	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 40.7	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.90	EDS
Total number of atoms	13221	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	60.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.10% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical values of  $\langle |L| \rangle$ ,  $\langle L^2 \rangle$  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: EDO

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	AAA	0.86	3/2312 (0.1%)	0.83	0/3137
1	FFF	0.81	1/2312 (0.0%)	0.84	0/3137
2	BBB	0.83	0/860	0.85	0/1162
2	GGG	0.85	0/860	0.85	0/1162
3	CCC	1.18	1/65 (1.5%)	0.95	0/86
3	HHH	1.36	1/65 (1.5%)	0.98	0/86
4	DDD	0.83	0/1530	0.86	0/2072
4	III	0.80	0/1503	0.86	0/2035
5	EEE	0.79	1/1996 (0.1%)	0.81	0/2721
5	JJJ	0.78	0/1996	0.79	0/2721
All	All	0.82	7/13499 (0.1%)	0.83	0/18319

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	FFF	0	1
4	DDD	0	1
5	EEE	0	1
5	JJJ	0	1
All	All	0	4

The worst 5 of 7 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	HHH	1	GLU	CD-OE1	7.18	1.33	1.25
1	FFF	63	GLU	CD-OE2	6.55	1.32	1.25
1	AAA	58	GLU	CD-OE2	6.36	1.32	1.25
3	CCC	6	GLY	C-O	5.28	1.32	1.23

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	AAA	173	GLU	CD-OE2	5.27	1.31	1.25

There are no bond angle outliers.

There are no chirality outliers.

All (4) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	DDD	29	GLY	Peptide
5	EEE	96	SER	Peptide
1	FFF	274	TRP	Peptide
5	JJJ	96	SER	Peptide

## 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	AAA	2247	0	2096	40	0
1	FFF	2247	0	2096	37	0
2	BBB	837	0	803	15	0
2	GGG	837	0	803	12	0
3	CCC	66	0	73	7	0
3	HHH	66	0	73	4	0
4	DDD	1501	0	1429	30	0
4	III	1475	0	1409	34	0
5	EEE	1942	0	1826	54	0
5	JJJ	1942	0	1826	49	0
6	AAA	4	0	6	0	0
6	GGG	4	0	6	0	0
7	AAA	11	0	0	0	0
7	BBB	7	0	0	1	0
7	CCC	1	0	0	0	0
7	DDD	11	0	0	1	0
7	EEE	5	0	0	0	0
7	FFF	7	0	0	1	0
7	GGG	6	0	0	0	0
7	III	2	0	0	0	0
7	JJJ	3	0	0	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
All	All	13221	0	12446	247	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 10.

The worst 5 of 247 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:DDD:187:ILE:O	4:DDD:188:ILE:HG23	1.61	1.01
4:DDD:151:THR:HG22	4:DDD:152:ASP:O	1.85	0.76
5:JJJ:66:VAL:HG11	5:JJJ:76:LEU:HD12	1.68	0.75
5:EEE:13:LEU:HD12	5:EEE:112:LEU:HD11	1.69	0.75
1:AAA:35:ARG:HD3	1:AAA:48:ARG:NH1	2.01	0.75

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	AAA	273/275 (99%)	247 (90%)	23 (8%)	3 (1%)	14	48
1	FFF	273/275 (99%)	246 (90%)	25 (9%)	2 (1%)	22	58
2	BBB	98/100 (98%)	89 (91%)	7 (7%)	2 (2%)	7	34
2	GGG	98/100 (98%)	87 (89%)	11 (11%)	0	100	100
3	CCC	8/10 (80%)	7 (88%)	0	1 (12%)	0	1
3	HHH	8/10 (80%)	7 (88%)	1 (12%)	0	100	100
4	DDD	191/193 (99%)	160 (84%)	24 (13%)	7 (4%)	3	20
4	III	188/193 (97%)	156 (83%)	27 (14%)	5 (3%)	5	27
5	EEE	243/245 (99%)	219 (90%)	21 (9%)	3 (1%)	13	46

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
5	JJJ	243/245 (99%)	222 (91%)	20 (8%)	1 (0%)	34	68
All	All	1623/1646 (99%)	1440 (89%)	159 (10%)	24 (2%)	10	41

5 of 24 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
3	CCC	9	THR
4	DDD	176	ASP
4	DDD	185	ASN
4	DDD	190	GLU
1	FFF	2	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	AAA	231/231 (100%)	220 (95%)	11 (5%)	25	59
1	FFF	231/231 (100%)	213 (92%)	18 (8%)	12	41
2	BBB	95/95 (100%)	90 (95%)	5 (5%)	22	56
2	GGG	95/95 (100%)	87 (92%)	8 (8%)	11	37
3	CCC	6/6 (100%)	6 (100%)	0	100	100
3	HHH	6/6 (100%)	4 (67%)	2 (33%)	0	0
4	DDD	173/173 (100%)	164 (95%)	9 (5%)	23	56
4	III	170/173 (98%)	157 (92%)	13 (8%)	13	42
5	EEE	212/212 (100%)	200 (94%)	12 (6%)	20	54
5	JJJ	212/212 (100%)	195 (92%)	17 (8%)	12	40
All	All	1431/1434 (100%)	1336 (93%)	95 (7%)	16	48

5 of 95 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	GGG	75	LYS

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
4	III	110	ASN
2	GGG	86	THR
4	III	3	GLU
4	III	135	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 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](#)

2 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	$\# Z  > 2$
6	EDO	AAA	301	-	3,3,3	0.23	0	2,2,2	0.27	0
6	EDO	GGG	101	-	3,3,3	0.10	0	2,2,2	0.17	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
6	EDO	AAA	301	-	-	0/1/1/1	-
6	EDO	GGG	101	-	-	0/1/1/1	-

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 [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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	AAA	275/275 (100%)	-0.10	1 (0%) 92   90	34, 51, 89, 126	0
1	FFF	275/275 (100%)	0.02	0 100   100	35, 55, 122, 145	0
2	BBB	100/100 (100%)	-0.00	0 100   100	30, 44, 59, 69	0
2	GGG	100/100 (100%)	-0.02	0 100   100	37, 46, 55, 62	0
3	CCC	10/10 (100%)	-0.19	0 100   100	34, 37, 44, 46	0
3	HHH	10/10 (100%)	-0.10	0 100   100	37, 39, 46, 48	0
4	DDD	193/193 (100%)	-0.12	1 (0%) 91   87	38, 51, 87, 99	0
4	III	190/193 (98%)	0.10	3 (1%) 72   63	47, 73, 115, 132	0
5	EEE	245/245 (100%)	-0.06	0 100   100	35, 58, 83, 102	0
5	JJJ	245/245 (100%)	0.01	0 100   100	45, 71, 96, 110	0
All	All	1643/1646 (99%)	-0.03	5 (0%) 94   93	30, 56, 102, 145	0

All (5) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
4	III	184	ASN	3.8
1	AAA	1	GLY	2.3
4	III	109	GLN	2.2
4	III	125	ASP	2.2
4	DDD	1	GLN	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](#)

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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
6	EDO	AAA	301	4/4	0.95	0.19	39,39,40,40	0
6	EDO	GGG	101	4/4	0.95	0.31	37,37,38,38	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.