



# wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 03:33 PM GMT

PDB ID : 3B9K  
Title : Crystal structure of CD8alpha-beta in complex with YTS 156.7 FAB  
Authors : Shore, D.; Wilson, I.A.  
Deposited on : 2007-11-05  
Resolution : 2.70 Å(reported)

This is a wwPDB 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 <http://wwpdb.org/ValidationPDFNotes.html>

---

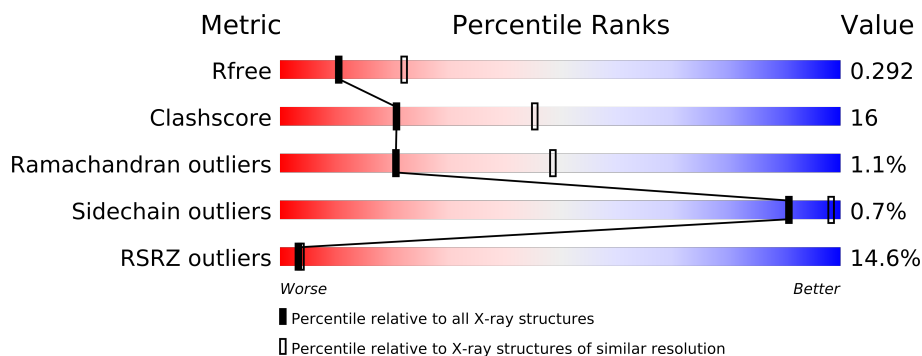
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.70 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))
$R_{free}$	66092	1557 (2.70-2.70)
Clashscore	79885	1939 (2.70-2.70)
Ramachandran outliers	78287	1905 (2.70-2.70)
Sidechain outliers	78261	1905 (2.70-2.70)
RSRZ outliers	66119	1559 (2.70-2.70)

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  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	C	213	
1	L	213	
2	D	214	
2	H	214	
3	A	131	
3	E	131	
4	B	125	
4	F	125	

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 10298 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Fab Heavy Chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	L	213	Total	C	N	O	S	0	0	0
			1645	1027	273	337	8			
1	C	213	Total	C	N	O	S	0	0	0
			1645	1027	273	337	8			

- Molecule 2 is a protein called Fab Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	214	Total	C	N	O	S	0	0	0
			1605	1006	264	327	8			
2	D	214	Total	C	N	O	S	0	0	0
			1605	1006	264	327	8			

- Molecule 3 is a protein called T-cell surface glycoprotein CD8 alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	A	118	Total	C	N	O	S	0	0	0
			938	602	152	177	7			
3	E	118	Total	C	N	O	S	0	0	0
			938	602	152	177	7			

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	125	SER	-	EXPRESSION TAG	UNP P01731
A	126	ALA	-	EXPRESSION TAG	UNP P01731
A	127	ASP	-	EXPRESSION TAG	UNP P01731
A	128	LEU	-	EXPRESSION TAG	UNP P01731
A	129	VAL	-	EXPRESSION TAG	UNP P01731
A	130	PRO	-	EXPRESSION TAG	UNP P01731
A	131	ARG	-	EXPRESSION TAG	UNP P01731
E	125	SER	-	EXPRESSION TAG	UNP P01731

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
E	126	ALA	-	EXPRESSION TAG	UNP P01731
E	127	ASP	-	EXPRESSION TAG	UNP P01731
E	128	LEU	-	EXPRESSION TAG	UNP P01731
E	129	VAL	-	EXPRESSION TAG	UNP P01731
E	130	PRO	-	EXPRESSION TAG	UNP P01731
E	131	ARG	-	EXPRESSION TAG	UNP P01731

- Molecule 4 is a protein called T-cell surface glycoprotein CD8 beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
4	B	117	Total	C	N	O	S	0	0	0
			929	593	153	178	5			
4	F	117	Total	C	N	O	S	0	0	0
			929	593	153	178	5			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	118	SER	-	EXPRESSION TAG	UNP P10300
B	119	SER	-	EXPRESSION TAG	UNP P10300
B	120	ALA	-	EXPRESSION TAG	UNP P10300
B	121	ASP	-	EXPRESSION TAG	UNP P10300
B	122	LEU	-	EXPRESSION TAG	UNP P10300
B	123	VAL	-	EXPRESSION TAG	UNP P10300
B	124	PRO	-	EXPRESSION TAG	UNP P10300
B	125	ARG	-	EXPRESSION TAG	UNP P10300
F	118	SER	-	EXPRESSION TAG	UNP P10300
F	119	SER	-	EXPRESSION TAG	UNP P10300
F	120	ALA	-	EXPRESSION TAG	UNP P10300
F	121	ASP	-	EXPRESSION TAG	UNP P10300
F	122	LEU	-	EXPRESSION TAG	UNP P10300
F	123	VAL	-	EXPRESSION TAG	UNP P10300
F	124	PRO	-	EXPRESSION TAG	UNP P10300
F	125	ARG	-	EXPRESSION TAG	UNP P10300

- Molecule 5 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	E	1	Total	C	N	O	0	0
			14	8	1	5		

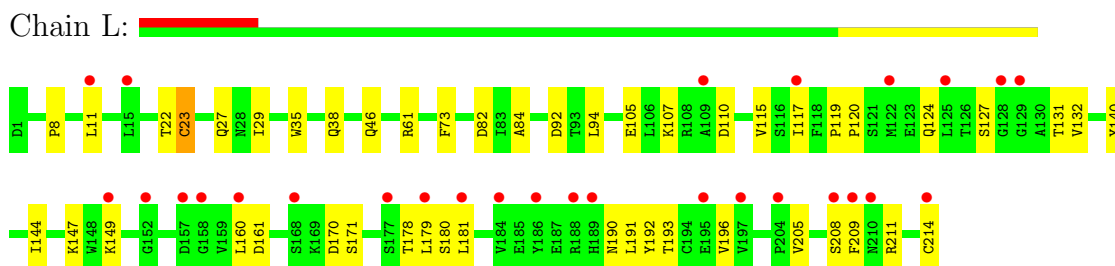
- Molecule 6 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	L	7	Total	O	0	0
			7	7		
6	H	6	Total	O	0	0
			6	6		
6	A	1	Total	O	0	0
			1	1		
6	B	1	Total	O	0	0
			1	1		
6	C	11	Total	O	0	0
			11	11		
6	D	5	Total	O	0	0
			5	5		
6	E	1	Total	O	0	0
			1	1		
6	F	4	Total	O	0	0
			4	4		

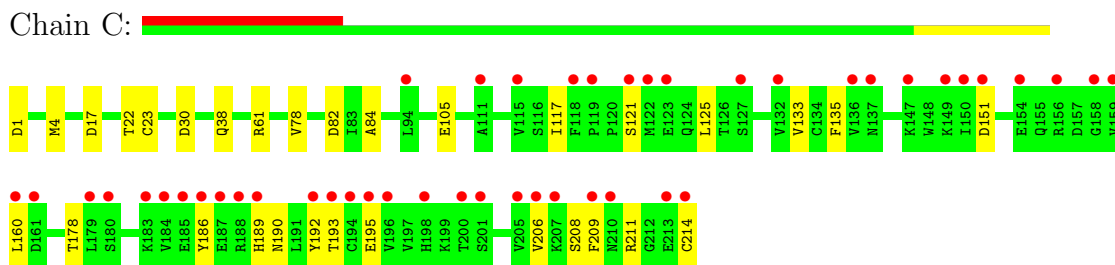
### 3 Residue-property plots

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.

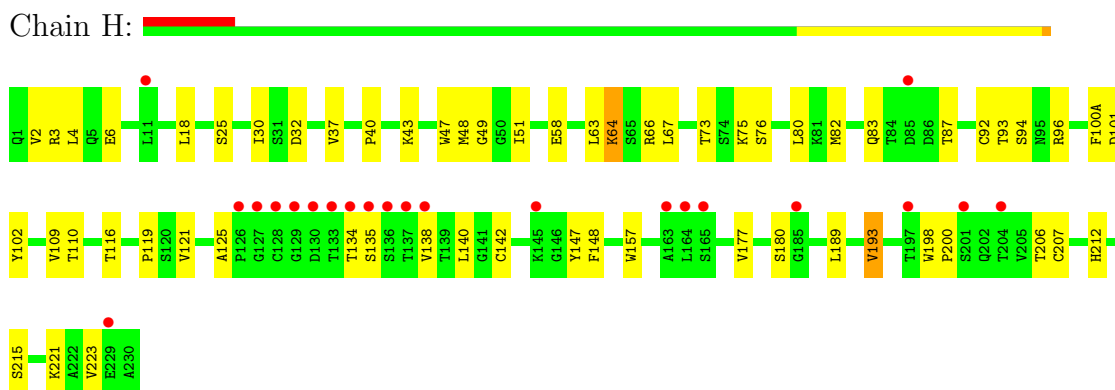
- Molecule 1: Fab Heavy Chain



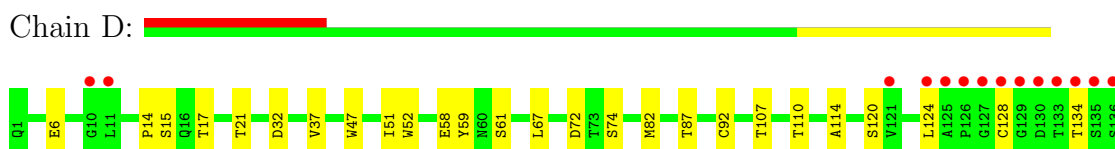
- Molecule 1: Fab Heavy Chain

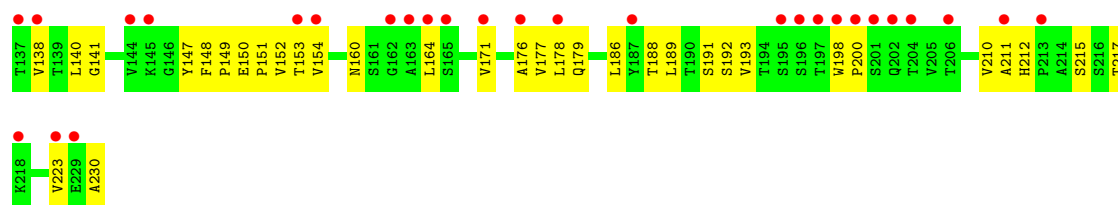


- Molecule 2: Fab Light chain



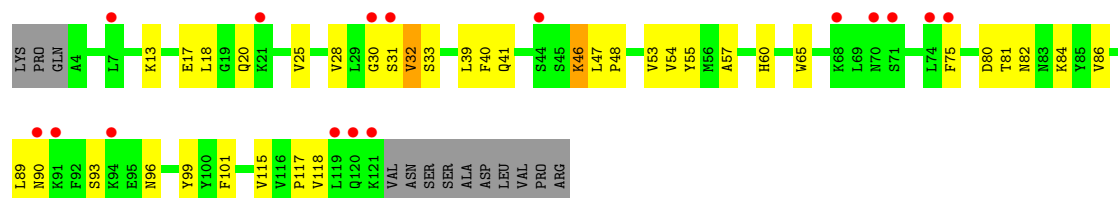
- Molecule 2: Fab Light chain





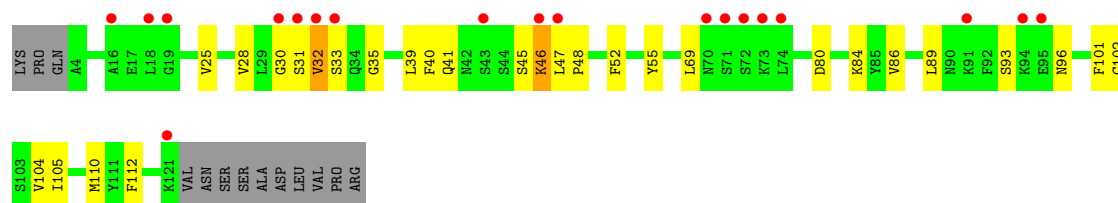
• Molecule 3: T-cell surface glycoprotein CD8 alpha chain

Chain A:



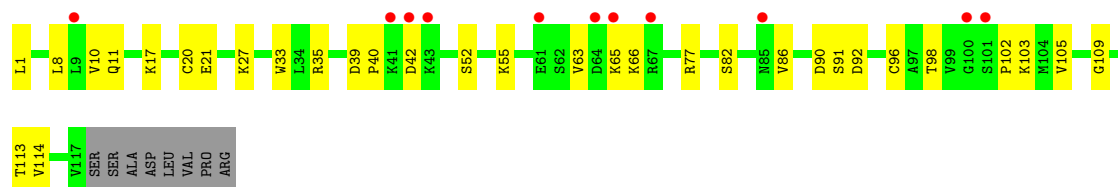
• Molecule 3: T-cell surface glycoprotein CD8 alpha chain

Chain E:



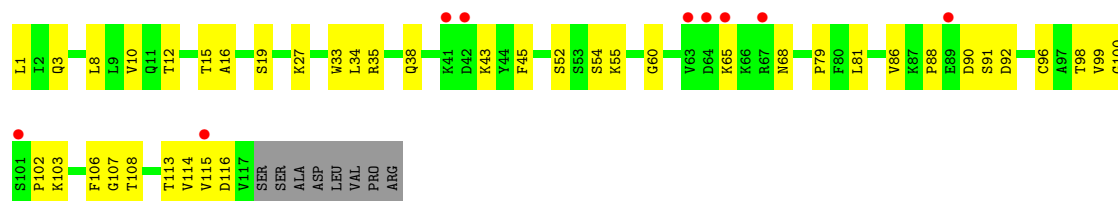
• Molecule 4: T-cell surface glycoprotein CD8 beta chain

Chain B:



• Molecule 4: T-cell surface glycoprotein CD8 beta chain

Chain F:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	91.78Å 92.75Å 190.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	42.33 – 2.70 42.33 – 2.70	Depositor EDS
% Data completeness (in resolution range)	97.6 (42.33-2.70) 97.6 (42.33-2.70)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.20 (at 2.69Å)	Xtriage
Refinement program	REFMAC 5.2.0019	Depositor
R, $R_{free}$	0.235 , 0.284 0.251 , 0.292	Depositor DCC
$R_{free}$ test set	2213 reflections (5.00%)	DCC
Wilson B-factor (Å <sup>2</sup> )	52.5	Xtriage
Anisotropy	0.206	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 32.2	EDS
Estimated twinning fraction	0.000 for k,h,-l	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 44293 reflections	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	10298	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	73.0	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 32.26 % 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 9.6787e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

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



## 5 Model quality ⓘ

### 5.1 Standard geometry ⓘ

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

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	C	0.73	0/1679	0.76	0/2277
1	L	0.80	1/1679 (0.1%)	0.79	0/2277
2	D	0.73	1/1645 (0.1%)	0.74	0/2251
2	H	0.79	1/1645 (0.1%)	0.79	1/2251 (0.0%)
3	A	0.66	0/959	0.73	0/1295
3	E	0.65	0/959	0.73	1/1295 (0.1%)
4	B	0.78	1/948 (0.1%)	0.84	0/1279
4	F	0.74	0/948	0.80	0/1279
All	All	0.74	4/10462 (0.0%)	0.77	2/14204 (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
4	B	0	1

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	H	58	GLU	CD-OE2	6.94	1.33	1.25
4	B	20	CYS	CB-SG	5.37	1.91	1.82
1	L	23	CYS	CB-SG	-5.36	1.73	1.81
2	D	120	SER	CB-OG	5.22	1.49	1.42

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	66	ARG	NE-CZ-NH2	-5.51	117.54	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	102	CYS	CA-CB-SG	-5.22	104.61	114.00

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
4	B	109	GLY	Peptide

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	C	1645	0	1595	33	0
1	L	1645	0	1595	50	0
2	D	1605	0	1545	58	0
2	H	1605	0	1547	48	0
3	A	938	0	935	35	0
3	E	938	0	932	26	0
4	B	929	0	940	35	0
4	F	929	0	940	54	0
5	A	14	0	13	0	0
5	E	14	0	13	3	0
6	A	1	0	0	0	0
6	B	1	0	0	0	0
6	C	11	0	0	1	0
6	D	5	0	0	1	0
6	E	1	0	0	3	0
6	F	4	0	0	2	0
6	H	6	0	0	1	0
6	L	7	0	0	0	0
All	All	10298	0	10055	315	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 16.

The worst 5 of 315 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
3:A:81:THR:HG22	3:A:82:ASN:H	1.06	1.13
2:D:140:LEU:HD11	2:D:198:TRP:CD1	1.84	1.11
2:H:93:THR:HG21	2:H:100(A):PHE:HB3	1.37	1.03
4:F:8:LEU:CD2	4:F:10:VAL:HG23	1.88	1.03
1:L:117:ILE:HG12	1:L:209:PHE:CE1	1.94	1.01

There are no symmetry-related clashes.

## 5.3 Torsion angles

### 5.3.1 Protein backbone ⓘ

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	C	211/213 (99%)	197 (93%)	14 (7%)	0	100	100
1	L	211/213 (99%)	198 (94%)	11 (5%)	2 (1%)	25	55
2	D	212/214 (99%)	198 (93%)	13 (6%)	1 (0%)	38	70
2	H	212/214 (99%)	195 (92%)	15 (7%)	2 (1%)	25	55
3	A	116/131 (88%)	102 (88%)	10 (9%)	4 (3%)	6	12
3	E	116/131 (88%)	97 (84%)	15 (13%)	4 (3%)	6	12
4	B	115/125 (92%)	107 (93%)	7 (6%)	1 (1%)	25	55
4	F	115/125 (92%)	107 (93%)	7 (6%)	1 (1%)	25	55
All	All	1308/1366 (96%)	1201 (92%)	92 (7%)	15 (1%)	21	49

5 of 15 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
4	B	103	LYS
4	F	103	LYS
3	A	46	LYS
3	E	46	LYS
2	H	64	LYS

### 5.3.2 Protein sidechains ⓘ

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	C	189/189 (100%)	188 (100%)	1 (0%)	94	99
1	L	189/189 (100%)	188 (100%)	1 (0%)	94	99
2	D	185/185 (100%)	185 (100%)	0	100	100
2	H	185/185 (100%)	183 (99%)	2 (1%)	84	96
3	A	110/122 (90%)	109 (99%)	1 (1%)	87	97
3	E	110/122 (90%)	109 (99%)	1 (1%)	87	97
4	B	109/116 (94%)	108 (99%)	1 (1%)	87	97
4	F	109/116 (94%)	108 (99%)	1 (1%)	87	97
All	All	1186/1224 (97%)	1178 (99%)	8 (1%)	91	98

5 of 8 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	A	40	PHE
4	F	92	ASP
1	C	105	GLU
2	H	83	GLN
4	B	92	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 9 such sidechains are listed below:

Mol	Chain	Res	Type
2	D	179	GLN
4	F	68	ASN
3	E	96	ASN
3	A	96	ASN
3	E	90	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

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	$\# Z  > 2$
5	NAG	A	201	3	12,14,15	0.62	0	15,19,21	1.06	1 (6%)
5	NAG	E	201	3	12,14,15	0.79	1 (8%)	15,19,21	1.67	3 (20%)

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
5	NAG	A	201	3	-	0/6/23/26	0/1/1/1
5	NAG	E	201	3	-	0/6/23/26	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	E	201	NAG	O5-C5	-2.20	1.41	1.45

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	E	201	NAG	O5-C5-C4	-3.97	105.62	110.65

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	E	201	NAG	C3-C2-N2	-3.01	107.17	111.76
5	E	201	NAG	O5-C5-C6	2.83	109.95	106.98
5	A	201	NAG	O5-C5-C6	2.38	109.48	106.98

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.7 Other polymers ⓘ

There are no such residues in this entry.

## 5.8 Polymer linkage issues

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<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	C	213/213 (100%)	1.01	46 (21%) 1 2	30, 56, 163, 201	0
1	L	213/213 (100%)	0.69	28 (13%) 4 4	31, 57, 137, 162	0
2	D	214/214 (100%)	1.32	42 (19%) 2 2	34, 64, 185, 240	0
2	H	214/214 (100%)	0.61	22 (10%) 7 7	35, 56, 134, 174	0
3	A	118/131 (90%)	0.86	16 (13%) 4 4	51, 76, 117, 126	0
3	E	118/131 (90%)	0.89	19 (16%) 2 3	47, 75, 121, 133	0
4	B	117/125 (93%)	0.65	11 (9%) 9 8	38, 56, 113, 127	0
4	F	117/125 (93%)	0.69	9 (7%) 13 14	39, 55, 109, 124	0
All	All	1324/1366 (96%)	0.86	193 (14%) 3 4	30, 62, 145, 240	0

The worst 5 of 193 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	129	GLY	13.3
2	D	130	ASP	11.0
2	D	133	THR	10.7
2	D	136	SER	10.0
2	D	126	PRO	8.8

### 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<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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
5	NAG	E	201	14/15	0.23	-	124,131,138,142	0
5	NAG	A	201	14/15	0.19	-	90,98,103,103	0

## 6.5 Other polymers

There are no such residues in this entry.