



# wwPDB X-ray Structure Validation Summary Report ⓘ

Sep 19, 2022 – 02:19 PM EDT

PDB ID : 7TUF  
Title : Crystal structure of Tapasin in complex with PaSta1-Fab  
Authors : Jiang, J.; Natarajan, K.; Taylor, D.K.; Boyd, L.F.; Margulies, D.H.  
Deposited on : 2022-02-02  
Resolution : 2.80 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:

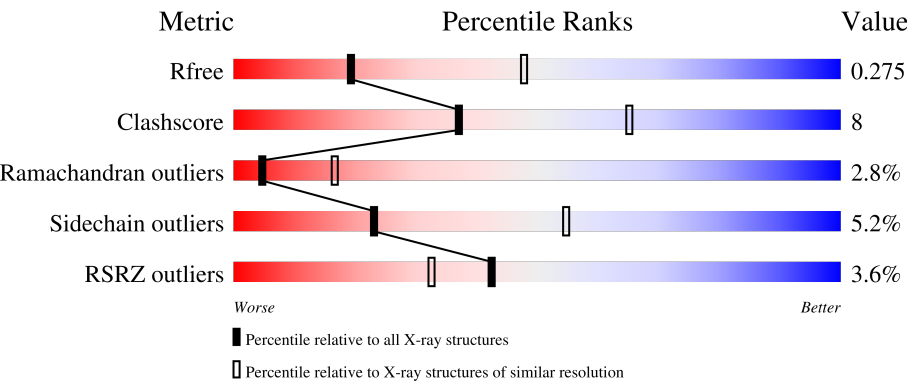
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.29
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.29

# 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 2.80 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 <sub>free</sub>	130704	3140 (2.80-2.80)
Clashscore	141614	3569 (2.80-2.80)
Ramachandran outliers	138981	3498 (2.80-2.80)
Sidechain outliers	138945	3500 (2.80-2.80)
RSRZ outliers	127900	3078 (2.80-2.80)

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	C	416	<div><div>4%</div><div>66%</div><div>16%</div><div>•</div><div>17%</div></div>
1	D	416	<div><div>4%</div><div>66%</div><div>16%</div><div>•</div><div>18%</div></div>
2	A	233	<div><div>2%</div><div>78%</div><div>15%</div><div>•</div><div>5%</div></div>
2	H	233	<div><div>67%</div><div>23%</div><div>5%</div><div>•</div><div>5%</div></div>
3	B	213	<div><div>5%</div><div>78%</div><div>17%</div><div>••</div></div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
3	L	213	<div><div></div><div>4%</div><div>78%</div><div>20%</div><div></div></div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 11341 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 Tapasin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	D	343	Total	C	N	O	S	0	0	0
			2480	1596	428	448	8			
1	C	344	Total	C	N	O	S	2	0	0
			2429	1566	420	434	9			

There are 70 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
D	382	GLY	-	expression tag	UNP O15533
D	383	GLY	-	expression tag	UNP O15533
D	384	LEU	-	expression tag	UNP O15533
D	385	GLU	-	expression tag	UNP O15533
D	386	VAL	-	expression tag	UNP O15533
D	387	LEU	-	expression tag	UNP O15533
D	388	PHE	-	expression tag	UNP O15533
D	389	GLN	-	expression tag	UNP O15533
D	390	GLY	-	expression tag	UNP O15533
D	391	PRO	-	expression tag	UNP O15533
D	392	GLY	-	expression tag	UNP O15533
D	393	GLY	-	expression tag	UNP O15533
D	394	GLY	-	expression tag	UNP O15533
D	395	LEU	-	expression tag	UNP O15533
D	396	ASN	-	expression tag	UNP O15533
D	397	ASP	-	expression tag	UNP O15533
D	398	ILE	-	expression tag	UNP O15533
D	399	PHE	-	expression tag	UNP O15533
D	400	GLU	-	expression tag	UNP O15533
D	401	ALA	-	expression tag	UNP O15533
D	402	GLN	-	expression tag	UNP O15533
D	403	LYS	-	expression tag	UNP O15533
D	404	ILE	-	expression tag	UNP O15533
D	405	GLU	-	expression tag	UNP O15533
D	406	TRP	-	expression tag	UNP O15533

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	407	HIS	-	expression tag	UNP O15533
D	408	GLU	-	expression tag	UNP O15533
D	409	GLY	-	expression tag	UNP O15533
D	410	GLY	-	expression tag	UNP O15533
D	411	HIS	-	expression tag	UNP O15533
D	412	HIS	-	expression tag	UNP O15533
D	413	HIS	-	expression tag	UNP O15533
D	414	HIS	-	expression tag	UNP O15533
D	415	HIS	-	expression tag	UNP O15533
D	416	HIS	-	expression tag	UNP O15533
C	382	GLY	-	expression tag	UNP O15533
C	383	GLY	-	expression tag	UNP O15533
C	384	LEU	-	expression tag	UNP O15533
C	385	GLU	-	expression tag	UNP O15533
C	386	VAL	-	expression tag	UNP O15533
C	387	LEU	-	expression tag	UNP O15533
C	388	PHE	-	expression tag	UNP O15533
C	389	GLN	-	expression tag	UNP O15533
C	390	GLY	-	expression tag	UNP O15533
C	391	PRO	-	expression tag	UNP O15533
C	392	GLY	-	expression tag	UNP O15533
C	393	GLY	-	expression tag	UNP O15533
C	394	GLY	-	expression tag	UNP O15533
C	395	LEU	-	expression tag	UNP O15533
C	396	ASN	-	expression tag	UNP O15533
C	397	ASP	-	expression tag	UNP O15533
C	398	ILE	-	expression tag	UNP O15533
C	399	PHE	-	expression tag	UNP O15533
C	400	GLU	-	expression tag	UNP O15533
C	401	ALA	-	expression tag	UNP O15533
C	402	GLN	-	expression tag	UNP O15533
C	403	LYS	-	expression tag	UNP O15533
C	404	ILE	-	expression tag	UNP O15533
C	405	GLU	-	expression tag	UNP O15533
C	406	TRP	-	expression tag	UNP O15533
C	407	HIS	-	expression tag	UNP O15533
C	408	GLU	-	expression tag	UNP O15533
C	409	GLY	-	expression tag	UNP O15533
C	410	GLY	-	expression tag	UNP O15533
C	411	HIS	-	expression tag	UNP O15533
C	412	HIS	-	expression tag	UNP O15533
C	413	HIS	-	expression tag	UNP O15533

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	414	HIS	-	expression tag	UNP O15533
C	415	HIS	-	expression tag	UNP O15533
C	416	HIS	-	expression tag	UNP O15533

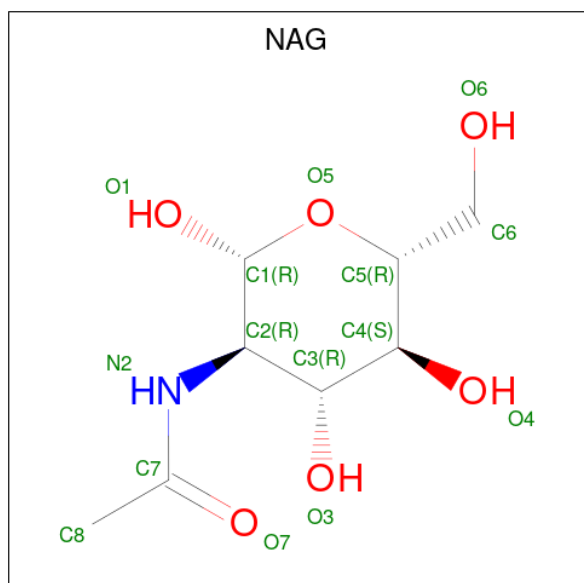
- Molecule 2 is a protein called PaSta1 Fab heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	221	Total	C	N	O	S	0	0	0
			1658	1053	270	330	5			
2	A	221	Total	C	N	O	S	0	0	0
			1645	1045	271	324	5			

- Molecule 3 is a protein called PaSta1 Fab kappa light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	210	Total	C	N	O	S	7	0	0
			1515	951	254	304	6			
3	B	210	Total	C	N	O	S	0	0	0
			1533	961	256	310	6			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	D	1	Total	C	N	O		0	0
			14	8	1	5			

*Continued on next page...*

*Continued from previous page...*

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

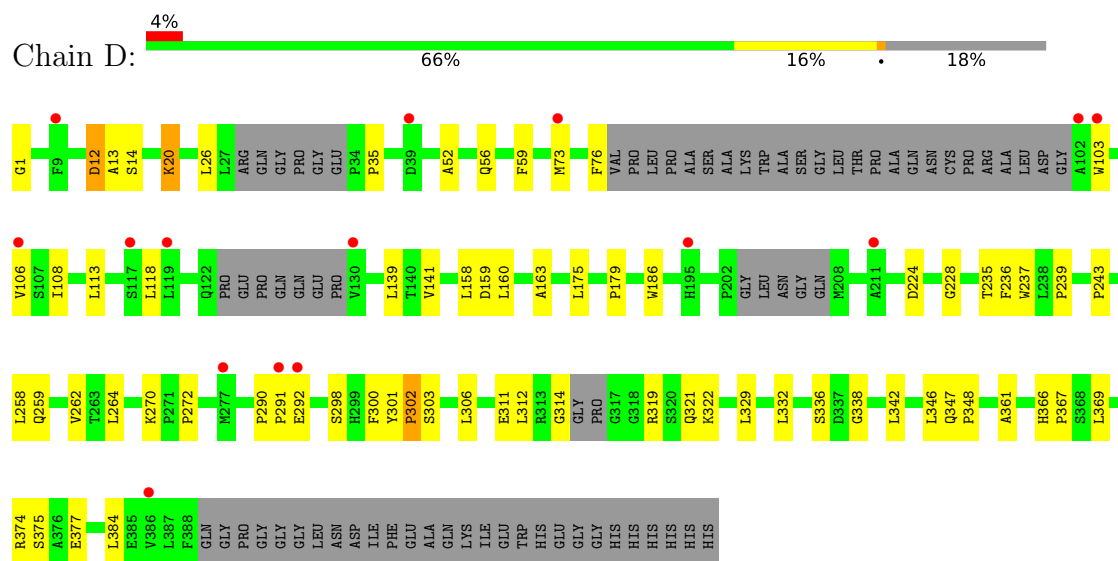
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	D	6	Total	O	0	0
			6	6		
5	H	9	Total	O	0	0
			9	9		
5	L	9	Total	O	0	0
			9	9		
5	C	6	Total	O	0	0
			6	6		
5	A	11	Total	O	0	0
			11	11		
5	B	12	Total	O	0	0
			12	12		

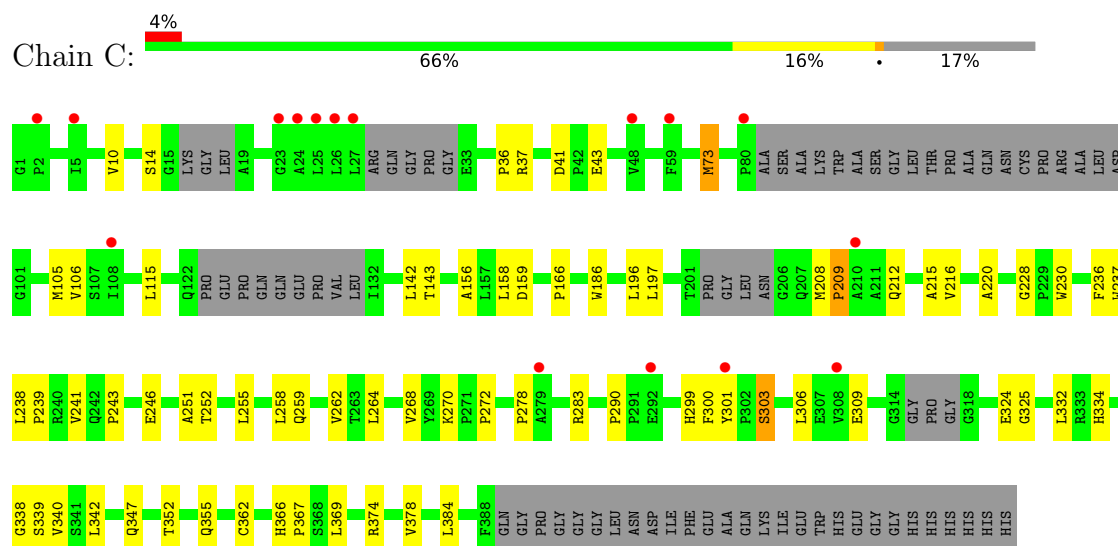
### 3 Residue-property plots [i](#)

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: Tapasin



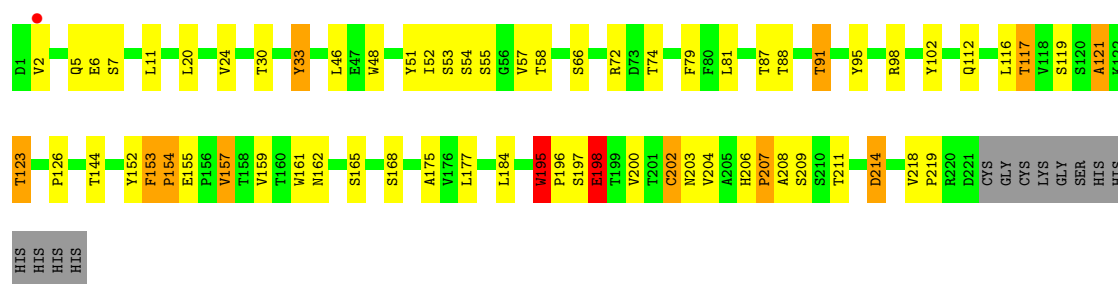
#### • Molecule 1: Tapasin



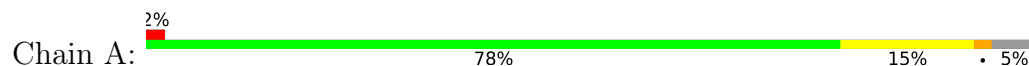
#### • Molecule 2: PaSta1 Fab heavy chain



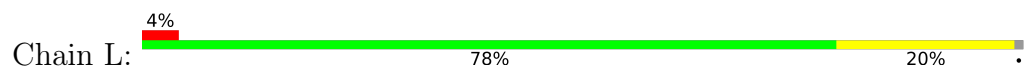




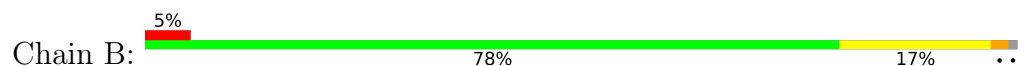
• Molecule 2: PaSta1 Fab heavy chain



• Molecule 3: PaSta1 Fab kappa light chain



• Molecule 3: PaSta1 Fab kappa light chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	117.42Å 68.99Å 142.98Å 90.00° 108.42° 90.00°	Depositor
Resolution (Å)	47.45 – 2.80 75.23 – 2.80	Depositor EDS
% Data completeness (in resolution range)	97.7 (47.45-2.80) 97.7 (75.23-2.80)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.36 (at 2.82Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, $R_{free}$	0.231 , 0.274 0.231 , 0.275	Depositor DCC
$R_{free}$ test set	2000 reflections (3.79%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	51.2	Xtriage
Anisotropy	0.613	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.29 , 54.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	11341	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	69.0	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 22.97 % 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 5.1173e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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

### 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	0.26	0/2508	0.50	0/3457
1	D	0.25	0/2562	0.47	0/3522
2	A	0.26	0/1691	0.53	0/2329
2	H	0.25	0/1704	0.53	0/2347
3	B	0.25	0/1574	0.48	0/2159
3	L	0.26	0/1555	0.50	0/2133
All	All	0.26	0/11594	0.50	0/15947

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
2	A	0	1
2	H	0	1
All	All	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	A	153	PHE	Peptide
2	H	195	TRP	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	C	2429	0	2223	36	0
1	D	2480	0	2330	37	0
2	A	1645	0	1545	23	0
2	H	1658	0	1571	34	0
3	B	1533	0	1392	22	0
3	L	1515	0	1374	21	0
4	C	14	0	13	0	0
4	D	14	0	13	0	0
5	A	11	0	0	0	0
5	B	12	0	0	1	0
5	C	6	0	0	0	0
5	D	6	0	0	0	0
5	H	9	0	0	0	0
5	L	9	0	0	0	0
All	All	11341	0	10461	169	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 169 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:C:228:GLY:HA3	2:A:103:ARG:HG2	1.60	0.81
1:C:243:PRO:HG2	1:C:338:GLY:HA3	1.61	0.79
3:L:47:ILE:HG22	3:L:53:LEU:HA	1.69	0.74
3:B:21:MET:HG3	3:B:72:LEU:HB3	1.72	0.71
1:D:314:GLY:H	1:D:319:ARG:HA	1.59	0.68

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	330/416 (79%)	299 (91%)	27 (8%)	4 (1%)	13	39
1	D	331/416 (80%)	307 (93%)	22 (7%)	2 (1%)	25	56
2	A	219/233 (94%)	197 (90%)	15 (7%)	7 (3%)	4	13
2	H	219/233 (94%)	187 (85%)	16 (7%)	16 (7%)	1	2
3	B	208/213 (98%)	186 (89%)	15 (7%)	7 (3%)	3	13
3	L	208/213 (98%)	187 (90%)	14 (7%)	7 (3%)	3	13
All	All	1515/1724 (88%)	1363 (90%)	109 (7%)	43 (3%)	5	17

5 of 43 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	153	PHE
2	H	154	PRO
2	H	195	TRP
2	H	196	PRO
2	H	219	PRO

### 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	223/328 (68%)	212 (95%)	11 (5%)	25	57
1	D	242/328 (74%)	228 (94%)	14 (6%)	20	50
2	A	179/207 (86%)	173 (97%)	6 (3%)	37	71

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	H	185/207 (89%)	169 (91%)	16 (9%)	10	30
3	B	162/188 (86%)	155 (96%)	7 (4%)	29	62
3	L	157/188 (84%)	151 (96%)	6 (4%)	33	67
All	All	1148/1446 (79%)	1088 (95%)	60 (5%)	23	55

5 of 60 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	202	CYS
3	B	150	ASP
3	L	179	THR
3	B	92	SER
3	B	208	PHE

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 6 such sidechains are listed below:

Mol	Chain	Res	Type
3	L	33	HIS
1	C	195	HIS
1	C	212	GLN
1	D	254	HIS
1	D	195	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules 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 monosaccharides 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	$\# Z  > 2$
4	NAG	D	501	1	14,14,15	0.24	0	17,19,21	0.45	0
4	NAG	C	501	1	14,14,15	0.29	0	17,19,21	0.36	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
4	NAG	D	501	1	-	2/6/23/26	0/1/1/1
4	NAG	C	501	1	-	2/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	D	501	NAG	O5-C5-C6-O6
4	C	501	NAG	O5-C5-C6-O6
4	C	501	NAG	C4-C5-C6-O6
4	D	501	NAG	C4-C5-C6-O6

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	C	344/416 (82%)	0.34	16 (4%) 31 22	36, 76, 142, 167	2 (0%)
1	D	343/416 (82%)	0.28	15 (4%) 34 24	35, 70, 117, 138	3 (0%)
2	A	221/233 (94%)	0.24	4 (1%) 68 61	23, 57, 119, 151	0
2	H	221/233 (94%)	0.17	1 (0%) 91 88	23, 55, 95, 137	0
3	B	210/213 (98%)	0.41	10 (4%) 30 21	30, 68, 113, 148	1 (0%)
3	L	210/213 (98%)	0.33	9 (4%) 35 25	22, 63, 116, 147	4 (1%)
All	All	1549/1724 (89%)	0.30	55 (3%) 42 32	22, 67, 125, 167	10 (0%)

The worst 5 of 55 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	279	ALA	4.2
1	C	24	ALA	4.1
1	D	130	VAL	4.0
2	A	132	ALA	3.8
1	C	80	PRO	3.7

### 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( $\text{\AA}^2$ )	Q<0.9
4	NAG	C	501	14/15	0.88	0.20	43,65,88,95	0
4	NAG	D	501	14/15	0.91	0.19	44,64,80,93	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.