



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jun 17, 2024 – 02:32 PM EDT

PDB ID : 3OHX  
Title : Molecular Basis for Complement Recognition and Inhibition Determined by Crystallographic Studies of the Staphylococcal Complement Inhibitor (SCIN) Bound to C3c and C3b  
Authors : Geisbrecht, B.V.; Garcia, B.L.  
Deposited on : 2010-08-18  
Resolution : 3.50 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 : 2022.3.0, CSD as543be (2022)  
Xtriage (Phenix) : 1.20.1  
EDS : 2.37.1  
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) : 2.37.1

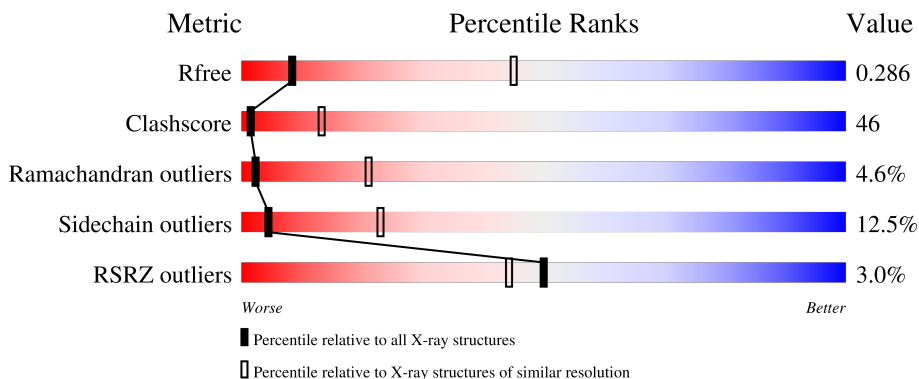
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.50 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.



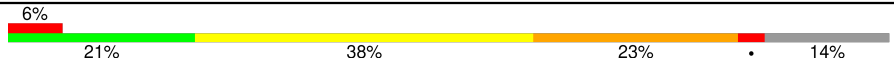


Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
$R_{free}$	130704	1659 (3.60-3.40)
Clashscore	141614	1036 (3.58-3.42)
Ramachandran outliers	138981	1005 (3.58-3.42)
Sidechain outliers	138945	1006 (3.58-3.42)
RSRZ outliers	127900	1559 (3.60-3.40)

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	A	645	
1	D	645	
2	B	206	
2	E	206	
3	C	343	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
3	F	343	
4	M	88	
4	P	88	

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 19119 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 Complement C3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	641	Total	C	N	O	S	0	0	0
			4989	3175	846	953	15			
1	D	641	Total	C	N	O	S	0	0	0
			4989	3175	846	953	15			

- Molecule 2 is a protein called Complement C3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	184	Total	C	N	O	S	0	0	0
			1488	956	250	277	5			
2	E	184	Total	C	N	O	S	0	0	0
			1488	956	250	277	5			

- Molecule 3 is a protein called Complement C3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	293	Total	C	N	O	S	0	0	0
			2377	1495	390	472	20			
3	F	295	Total	C	N	O	S	0	0	0
			2396	1510	392	474	20			

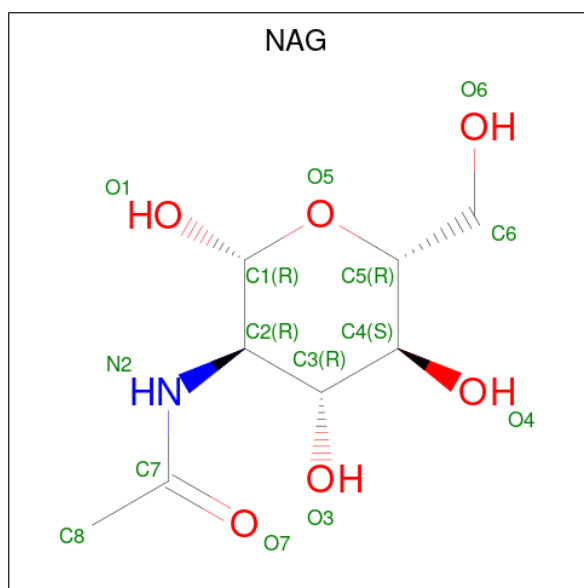
- Molecule 4 is a protein called Staphylococcal complement inhibitor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	M	84	Total	C	N	O	S	0	0	0
			682	432	111	137	2			
4	P	84	Total	C	N	O	S	0	0	0
			682	432	111	137	2			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
M	-2	GLY	-	expression tag	UNP Q931M7
M	-1	THR	-	expression tag	UNP Q931M7
M	0	SER	-	expression tag	UNP Q931M7
P	-2	GLY	-	expression tag	UNP Q931M7
P	-1	THR	-	expression tag	UNP Q931M7
P	0	SER	-	expression tag	UNP Q931M7

- Molecule 5 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula:  $C_8H_{15}NO_6$ ).

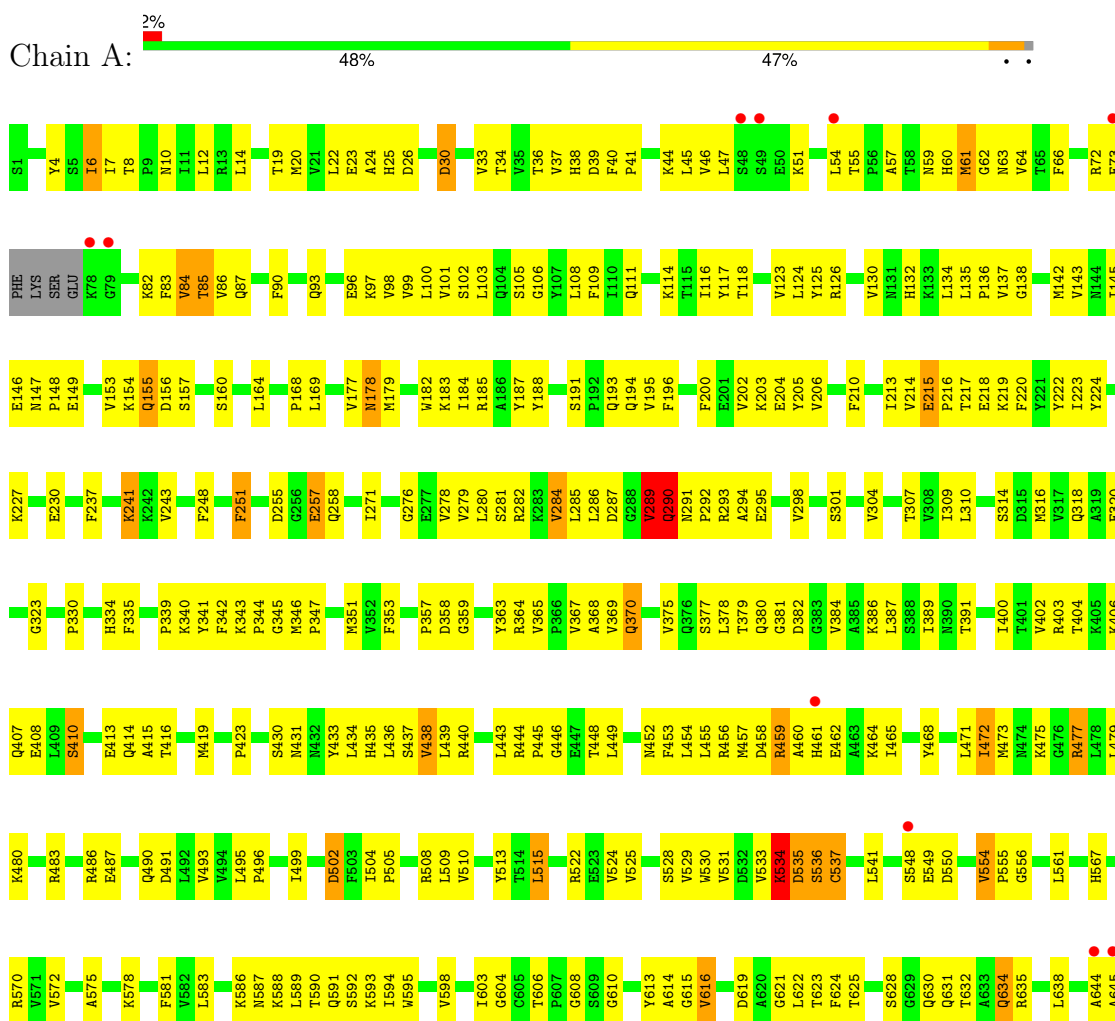


Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
5	A	1	Total	C	N	O	0	0
			14	8	1	5		
5	D	1	Total	C	N	O	0	0
			14	8	1	5		

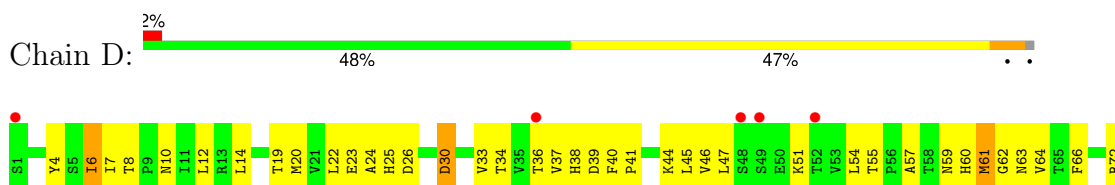
### 3 Residue-property plots

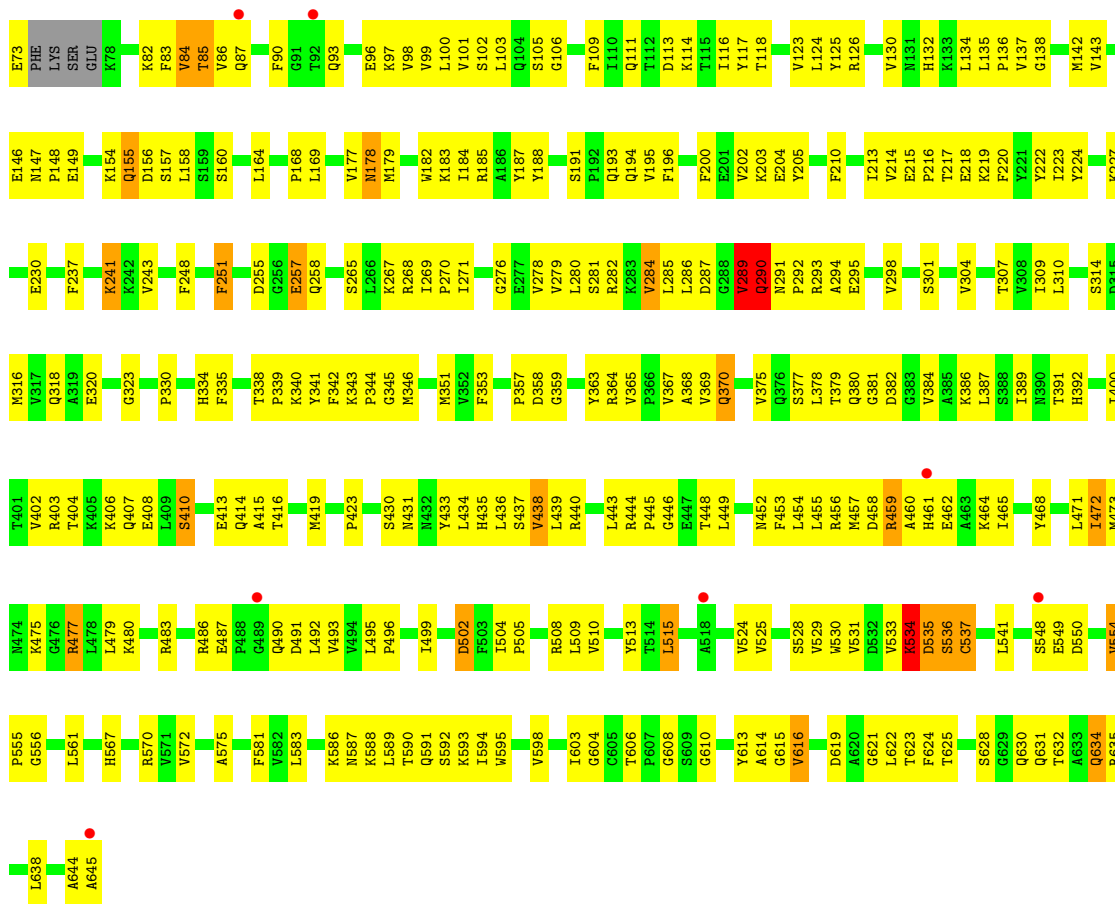
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Complement C3

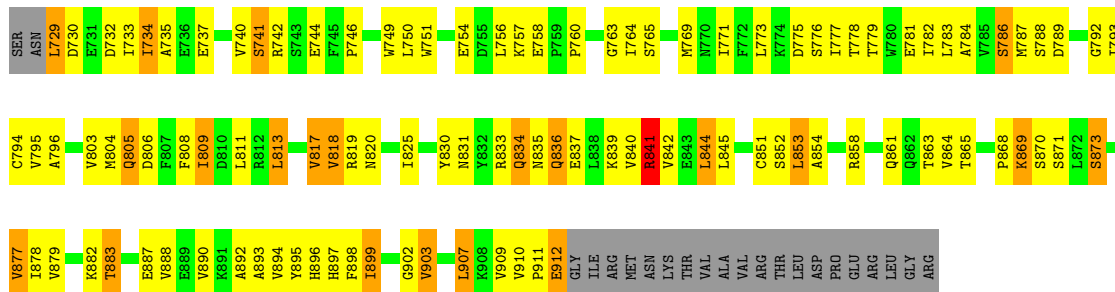


#### • Molecule 1: Complement C3

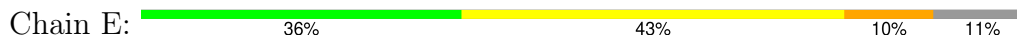


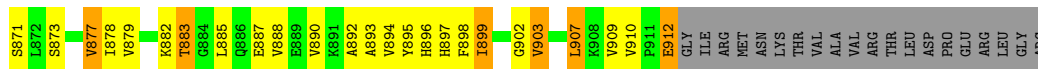


• Molecule 2: Complement C3

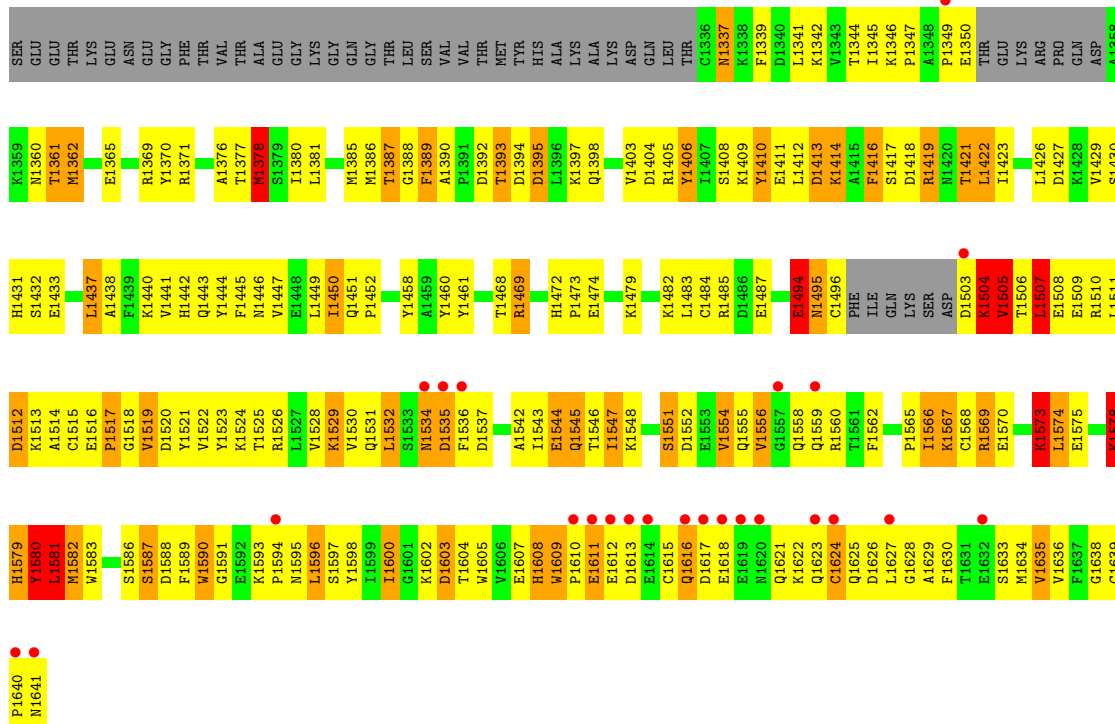


• Molecule 2: Complement C3

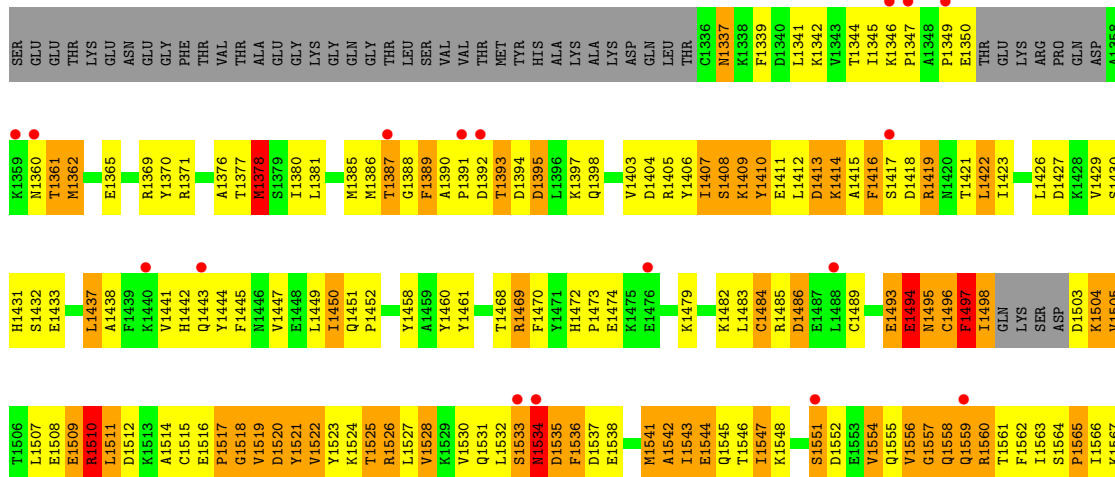
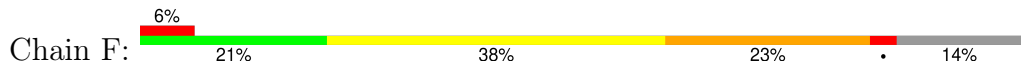




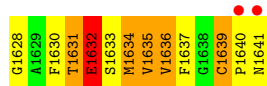
• Molecule 3: Complement C3



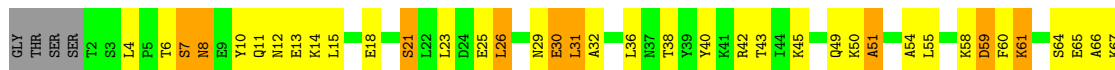
• Molecule 3: Complement C3



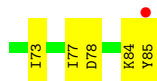




● Molecule 4: Staphylococcal complement inhibitor



● Molecule 4: Staphylococcal complement inhibitor



## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 21 21 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	231.16Å 231.50Å 68.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	46.27 – 3.50 46.96 – 3.50	Depositor EDS
% Data completeness (in resolution range)	93.1 (46.27-3.50) 93.2 (46.96-3.50)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.26 (at 3.48Å)	Xtrriage
Refinement program	PHENIX 1.5_2	Depositor
R, $R_{free}$	0.274 , 0.294 0.268 , 0.286	Depositor DCC
$R_{free}$ test set	2017 reflections (4.55%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	68.7	Xtrriage
Anisotropy	0.494	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.28 , 39.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.47$ , $\langle L^2 \rangle = 0.30$	Xtrriage
Estimated twinning fraction	0.000 for k,h,-l	Xtrriage
$F_o, F_c$ correlation	0.87	EDS
Total number of atoms	19119	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	69.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 26.98 % 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 2.3878e-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 [i](#)

### 5.1 Standard geometry [i](#)

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	A	0.45	0/5089	0.63	0/6916
1	D	0.45	0/5089	0.63	0/6916
2	B	0.51	0/1520	0.70	1/2066 (0.0%)
2	E	0.51	0/1520	0.70	1/2066 (0.0%)
3	C	0.49	0/2422	0.73	2/3264 (0.1%)
3	F	0.56	0/2442	0.74	1/3291 (0.0%)
4	M	0.39	0/690	0.59	0/923
4	P	0.39	0/690	0.59	0/923
All	All	0.48	0/19462	0.67	5/26365 (0.0%)

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	C	1609	TRP	N-CA-C	-5.93	94.98	111.00
3	C	1581	LEU	CA-CB-CG	-5.75	102.08	115.30
2	E	841	ARG	NE-CZ-NH1	5.09	122.84	120.30
2	B	841	ARG	NE-CZ-NH1	5.07	122.84	120.30
3	F	1536	PHE	N-CA-C	-5.05	97.36	111.00

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	4989	0	5051	367	0
1	D	4989	0	5051	372	0
2	B	1488	0	1512	125	0
2	E	1488	0	1512	119	0
3	C	2377	0	2279	319	0
3	F	2396	0	2299	474	0
4	M	682	0	697	56	0
4	P	682	0	697	50	0
5	A	14	0	13	4	0
5	D	14	0	13	3	0
All	All	19119	0	19124	1763	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 46.

The worst 5 of 1763 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)
3:F:1497:PHE:CZ	3:F:1498:ILE:HG13	1.39	1.57
3:C:1639:CYS:SG	3:C:1640:PRO:HD2	1.45	1.56
3:F:1497:PHE:CE2	3:F:1498:ILE:HG13	1.52	1.44
3:F:1571:ALA:C	3:F:1572:LEU:HD22	1.41	1.39
3:F:1569:ARG:CG	3:F:1570:GLU:OE1	1.73	1.33

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	637/645 (99%)	557 (87%)	69 (11%)	11 (2%)	9	42
1	D	637/645 (99%)	557 (87%)	69 (11%)	11 (2%)	9	42

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	B	182/206 (88%)	162 (89%)	18 (10%)	2 (1%)	14	52
2	E	182/206 (88%)	162 (89%)	18 (10%)	2 (1%)	14	52
3	C	287/343 (84%)	216 (75%)	44 (15%)	27 (9%)	0	8
3	F	289/343 (84%)	196 (68%)	42 (14%)	51 (18%)	0	2
4	M	82/88 (93%)	60 (73%)	19 (23%)	3 (4%)	3	26
4	P	82/88 (93%)	60 (73%)	19 (23%)	3 (4%)	3	26
All	All	2378/2564 (93%)	1970 (83%)	298 (12%)	110 (5%)	2	21

5 of 110 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	537	CYS
3	C	1505	VAL
3	C	1545	GLN
3	C	1578	LYS
3	C	1579	HIS

### 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	563/567 (99%)	529 (94%)	34 (6%)	19	52
1	D	563/567 (99%)	529 (94%)	34 (6%)	19	52
2	B	172/191 (90%)	142 (83%)	30 (17%)	2	11
2	E	172/191 (90%)	142 (83%)	30 (17%)	2	11
3	C	266/309 (86%)	210 (79%)	56 (21%)	1	5
3	F	268/309 (87%)	202 (75%)	66 (25%)	0	4
4	M	76/79 (96%)	67 (88%)	9 (12%)	5	25
4	P	76/79 (96%)	66 (87%)	10 (13%)	4	21
All	All	2156/2292 (94%)	1887 (88%)	269 (12%)	4	23

5 of 269 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	F	1627	LEU
1	D	84	VAL
1	D	472	ILE
3	C	1507	LEU
3	C	1503	ASP

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
3	F	1616	GLN
1	D	291	ASN
1	D	634	GLN
1	D	392	HIS
3	C	1608	HIS

### 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
5	NAG	A	646	-	14,14,15	0.44	0	17,19,21	1.46	1 (5%)
5	NAG	D	646	-	14,14,15	0.43	0	17,19,21	1.45	1 (5%)

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	646	-	-	4/6/23/26	0/1/1/1
5	NAG	D	646	-	-	4/6/23/26	0/1/1/1

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	646	NAG	C1-O5-C5	4.95	118.82	112.19
5	D	646	NAG	C1-O5-C5	4.92	118.78	112.19

There are no chirality outliers.

5 of 8 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
5	A	646	NAG	O5-C5-C6-O6
5	D	646	NAG	O5-C5-C6-O6
5	A	646	NAG	C8-C7-N2-C2
5	A	646	NAG	O7-C7-N2-C2
5	D	646	NAG	C8-C7-N2-C2

There are no ring outliers.

2 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	646	NAG	4	0
5	D	646	NAG	3	0

## 5.7 Other polymers [i](#)

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 [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	A	641/645 (99%)	-0.08	10 (1%) 72 66	18, 53, 106, 172	0
1	D	641/645 (99%)	-0.03	12 (1%) 66 61	18, 53, 106, 172	0
2	B	184/206 (89%)	-0.36	0 100 100	19, 50, 80, 114	0
2	E	184/206 (89%)	-0.31	0 100 100	25, 57, 87, 120	0
3	C	293/343 (85%)	0.28	24 (8%) 11 12	35, 87, 198, 230	0
3	F	295/343 (86%)	0.41	22 (7%) 14 14	31, 93, 143, 194	0
4	M	84/88 (95%)	-0.23	1 (1%) 79 73	48, 68, 105, 175	0
4	P	84/88 (95%)	-0.20	2 (2%) 59 53	43, 65, 101, 184	0
All	All	2406/2564 (93%)	-0.01	71 (2%) 50 44	18, 63, 125, 230	0

The worst 5 of 71 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	1613	ASP	8.7
3	C	1624	CYS	7.3
4	P	85	TYR	6.1
3	F	1640	PRO	5.9
3	C	1610	PRO	5.3

### 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
5	NAG	D	646	14/15	0.70	0.32	62,87,108,108	0
5	NAG	A	646	14/15	0.74	0.34	62,87,108,108	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.