



# wwPDB X-ray Structure Validation Summary Report

Mar 1, 2014 – 03:48 AM GMT

PDB ID : 3DMK  
Title : Crystal structure of Down Syndrome Cell Adhesion Molecule (DSCAM) isoform 1.30.30, N-terminal eight Ig domains  
Authors : Sawaya, M.R.; Wojtowicz, W.M.; Eisenberg, D.; Zipursky, S.L.  
Deposited on : 2008-07-01  
Resolution : 4.19 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>

---

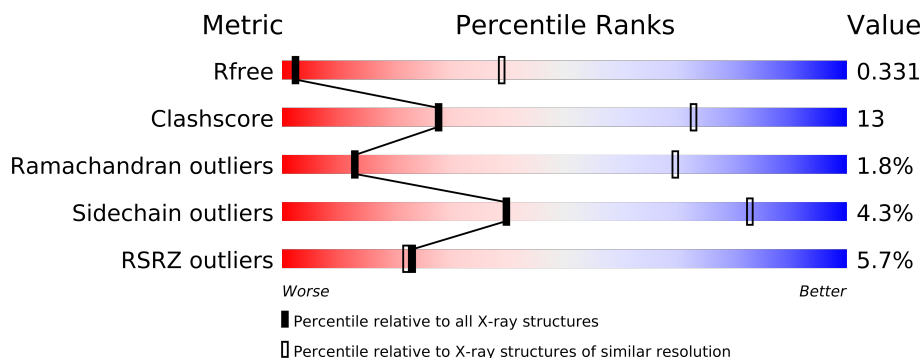
The following versions of software and data (see [references](#)) were used in the production of this report:

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 4.19 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	1001 (4.84-3.50)
Clashscore	79885	1259 (4.84-3.50)
Ramachandran outliers	78287	1192 (4.84-3.50)
Sidechain outliers	78261	1175 (4.84-3.50)
RSRZ outliers	66119	1001 (4.84-3.50)

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	A	816	
1	B	816	
1	C	816	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
6	GOL	A	786	-	X
6	GOL	B	794	-	X

## 2 Entry composition

There are 6 unique types of molecules in this entry. The entry contains 16740 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 Down Syndrome Cell Adhesion Molecule (DSCAM) isoform 1.30.30, N-terminal eight Ig domains.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	775	Total	C	N	O	S	0	0	0
			5985	3778	1035	1147	25			
1	B	676	Total	C	N	O	S	0	0	0
			5223	3295	908	998	22			
1	C	676	Total	C	N	O	S	0	0	0
			5223	3295	908	998	22			

There are 18 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	776	HIS	-	EXPRESSION TAG	PDB 3DMK
A	777	HIS	-	EXPRESSION TAG	PDB 3DMK
A	778	HIS	-	EXPRESSION TAG	PDB 3DMK
A	779	HIS	-	EXPRESSION TAG	PDB 3DMK
A	780	HIS	-	EXPRESSION TAG	PDB 3DMK
A	781	HIS	-	EXPRESSION TAG	PDB 3DMK
B	776	HIS	-	EXPRESSION TAG	PDB 3DMK
B	777	HIS	-	EXPRESSION TAG	PDB 3DMK
B	778	HIS	-	EXPRESSION TAG	PDB 3DMK
B	779	HIS	-	EXPRESSION TAG	PDB 3DMK
B	780	HIS	-	EXPRESSION TAG	PDB 3DMK
B	781	HIS	-	EXPRESSION TAG	PDB 3DMK
C	776	HIS	-	EXPRESSION TAG	PDB 3DMK
C	777	HIS	-	EXPRESSION TAG	PDB 3DMK
C	778	HIS	-	EXPRESSION TAG	PDB 3DMK
C	779	HIS	-	EXPRESSION TAG	PDB 3DMK
C	780	HIS	-	EXPRESSION TAG	PDB 3DMK
C	781	HIS	-	EXPRESSION TAG	PDB 3DMK

- Molecule 2 is a polymer of unknown type called SUGAR (2-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	2	Total	C	N	O	0	0
			28	16	2	10		
2	A	2	Total	C	N	O	0	0
			28	16	2	10		
2	B	2	Total	C	N	O	0	0
			28	16	2	10		
2	B	2	Total	C	N	O	0	0
			28	16	2	10		

There are 24 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	776	HIS	-	EXPRESSION TAG	PDB 3DMK
A	777	HIS	-	EXPRESSION TAG	PDB 3DMK
A	778	HIS	-	EXPRESSION TAG	PDB 3DMK
A	779	HIS	-	EXPRESSION TAG	PDB 3DMK
A	780	HIS	-	EXPRESSION TAG	PDB 3DMK
A	781	HIS	-	EXPRESSION TAG	PDB 3DMK
A	776	HIS	-	EXPRESSION TAG	PDB 3DMK
A	777	HIS	-	EXPRESSION TAG	PDB 3DMK
A	778	HIS	-	EXPRESSION TAG	PDB 3DMK
A	779	HIS	-	EXPRESSION TAG	PDB 3DMK
A	780	HIS	-	EXPRESSION TAG	PDB 3DMK
A	781	HIS	-	EXPRESSION TAG	PDB 3DMK
B	776	HIS	-	EXPRESSION TAG	PDB 3DMK
B	777	HIS	-	EXPRESSION TAG	PDB 3DMK
B	778	HIS	-	EXPRESSION TAG	PDB 3DMK
B	779	HIS	-	EXPRESSION TAG	PDB 3DMK
B	780	HIS	-	EXPRESSION TAG	PDB 3DMK
B	781	HIS	-	EXPRESSION TAG	PDB 3DMK
B	776	HIS	-	EXPRESSION TAG	PDB 3DMK
B	777	HIS	-	EXPRESSION TAG	PDB 3DMK
B	778	HIS	-	EXPRESSION TAG	PDB 3DMK
B	779	HIS	-	EXPRESSION TAG	PDB 3DMK
B	780	HIS	-	EXPRESSION TAG	PDB 3DMK
B	781	HIS	-	EXPRESSION TAG	PDB 3DMK

- Molecule 3 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	C	3	Total	C	N	O	0	0
			42	24	3	15		

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
3	B	3	Total	C	N	O	0	0
			42	24	3	15		

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	776	HIS	-	EXPRESSION TAG	PDB 3DMK
C	777	HIS	-	EXPRESSION TAG	PDB 3DMK
C	778	HIS	-	EXPRESSION TAG	PDB 3DMK
C	779	HIS	-	EXPRESSION TAG	PDB 3DMK
C	780	HIS	-	EXPRESSION TAG	PDB 3DMK
C	781	HIS	-	EXPRESSION TAG	PDB 3DMK
B	776	HIS	-	EXPRESSION TAG	PDB 3DMK
B	777	HIS	-	EXPRESSION TAG	PDB 3DMK
B	778	HIS	-	EXPRESSION TAG	PDB 3DMK
B	779	HIS	-	EXPRESSION TAG	PDB 3DMK
B	780	HIS	-	EXPRESSION TAG	PDB 3DMK
B	781	HIS	-	EXPRESSION TAG	PDB 3DMK

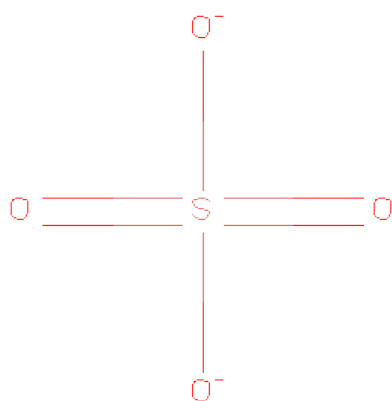
- Molecule 4 is a polymer of unknown type called SUGAR (3-MER).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	B	3	Total	C	N	O	0	0
			42	24	3	15		
4	C	3	Total	C	N	O	0	0
			42	24	3	15		

There are 12 discrepancies between the modelled and reference sequences:

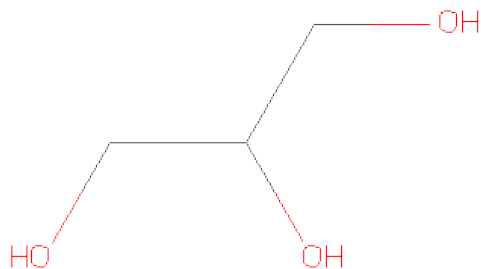
Chain	Residue	Modelled	Actual	Comment	Reference
B	776	HIS	-	EXPRESSION TAG	PDB 3DMK
B	777	HIS	-	EXPRESSION TAG	PDB 3DMK
B	778	HIS	-	EXPRESSION TAG	PDB 3DMK
B	779	HIS	-	EXPRESSION TAG	PDB 3DMK
B	780	HIS	-	EXPRESSION TAG	PDB 3DMK
B	781	HIS	-	EXPRESSION TAG	PDB 3DMK
C	776	HIS	-	EXPRESSION TAG	PDB 3DMK
C	777	HIS	-	EXPRESSION TAG	PDB 3DMK
C	778	HIS	-	EXPRESSION TAG	PDB 3DMK
C	779	HIS	-	EXPRESSION TAG	PDB 3DMK
C	780	HIS	-	EXPRESSION TAG	PDB 3DMK
C	781	HIS	-	EXPRESSION TAG	PDB 3DMK

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 6 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	A	1	Total	C	O	0	0
			6	3	3		
6	A	1	Total	C	O	0	0
			6	3	3		

*Continued on next page...*

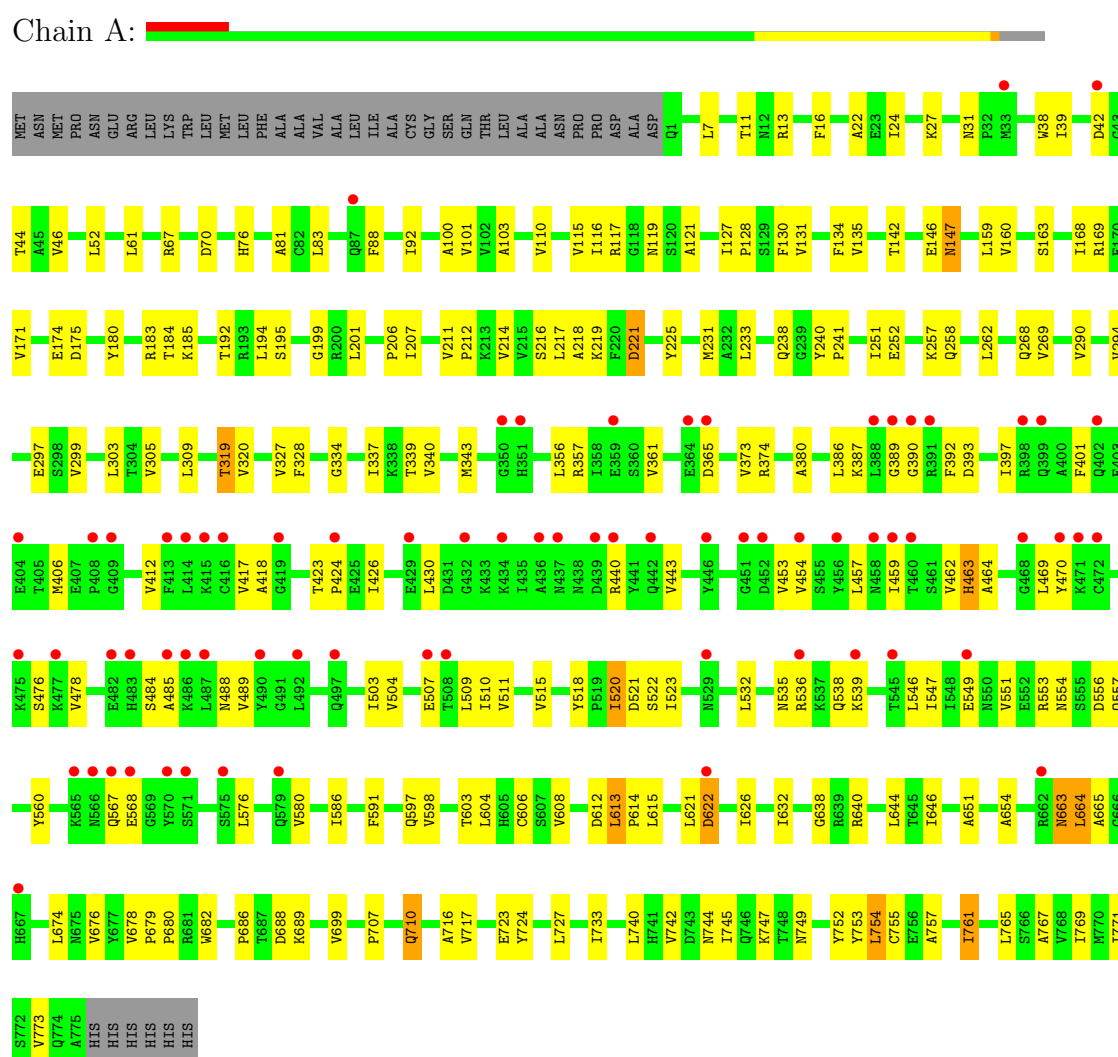
*Continued from previous page...*

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total	C	O	0	0
			6	3	3		
6	B	1	Total	C	O	0	0
			6	3	3		

### 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.

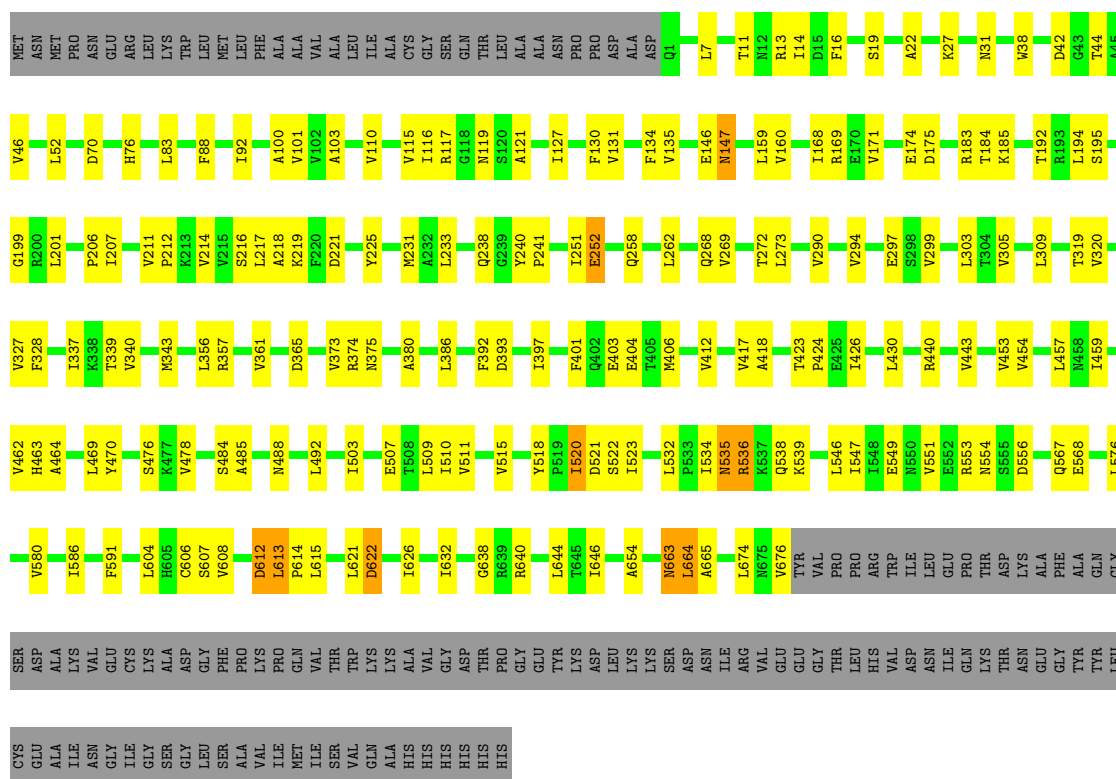
- Molecule 1: Down Syndrome Cell Adhesion Molecule (DSCAM) isoform 1.30.30, N-terminal eight Ig domains



- Molecule 1: Down Syndrome Cell Adhesion Molecule (DSCAM) isoform 1.30.30, N-terminal eight Ig domains

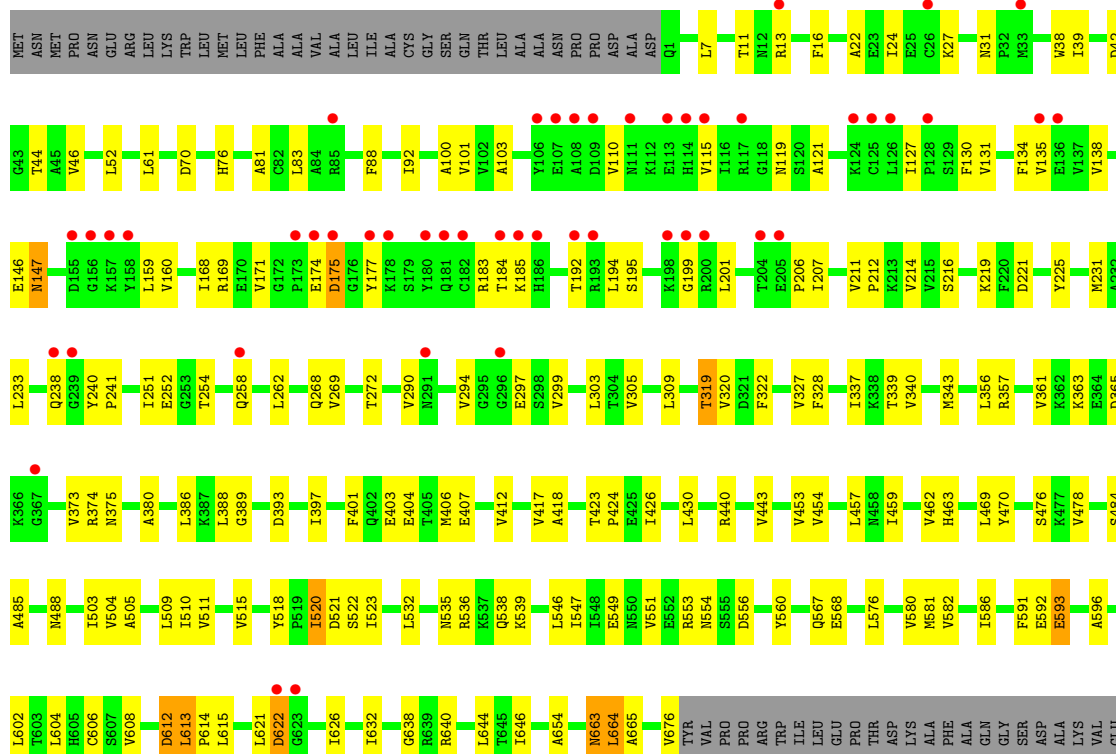
Chain B: 





- Molecule 1: Down Syndrome Cell Adhesion Molecule (DSCAM) isoform 1.30.30, N-terminal eight Ig domains

Chain C:



CYS	LYS	ALA	ASP	PHE	PRO	LYS	PRO	GLN	VAL	THR	TRP	LYS	LYS	ALA	VAL	GLY	ASP	THR	PRO	GLY	GLU	TYR	LYS	ASP	ASN	ILE	ARG	VAL	GLU	GLU	GLY	THR	THR	HIS	LEU	LEU	VAL	VAL	ASP	ASN	ILE	GLN	LYS	LYS	THR	THR	TYR	TYR	GLY	GLU	CYS	GLU	ALA	ALA	ASN	GLY
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

ILE	GLY	SER	GLY	LEU	SER	ALA	VAL	ILE	MET	ILE	SER	VAL	GLN	ALA	HIS	HIS	HIS	HIS	HIS	HIS
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	I 2 2 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	118.46Å 177.61Å 434.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	12.00 – 4.19 11.99 – 4.19	Depositor EDS
% Data completeness (in resolution range)	99.4 (12.00-4.19) 99.4 (11.99-4.19)	Depositor EDS
$R_{merge}$	0.07	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.72 (at 4.19Å)	Xtriage
Refinement program	REFMAC 5.4.0061	Depositor
R, $R_{free}$	0.280 , 0.327 0.284 , 0.331	Depositor DCC
$R_{free}$ test set	1618 reflections (5.24%)	DCC
Wilson B-factor (Å <sup>2</sup> )	99.4	Xtriage
Anisotropy	0.178	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.19 , 89.7	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 32525 reflections	Xtriage
$F_o, F_c$ correlation	0.82	EDS
Total number of atoms	16740	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.38% of the height of the origin peak. No significant pseudotranslation is detected.*

<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: GOL, NAG, NDG, SO4

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	0.33	0/6113	0.52	0/8303
1	B	0.38	0/5334	0.54	0/7246
1	C	0.33	0/5334	0.52	0/7246
All	All	0.35	0/16781	0.53	0/22795

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
3	B	1	0

There are no bond length outliers.

There are no bond angle outliers.

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	B	784	NAG	C1

There are no planarity outliers.

### 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	A	5985	0	5928	174	0
1	B	5223	0	5172	128	1
1	C	5223	0	5174	131	1
2	A	56	0	50	3	0
2	B	56	0	50	1	0
3	B	42	0	37	0	0
3	C	42	0	37	1	0
4	B	42	0	37	0	0
4	C	42	0	37	0	0
5	B	5	0	0	0	0
6	A	12	0	16	0	0
6	B	12	0	16	0	0
All	All	16740	0	16554	425	1

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 13.

The worst 5 of 425 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:654:ALA:HB2	1:A:676:VAL:HG12	1.38	1.00
1:A:733:ILE:HG23	1:A:742:VAL:HG22	1.46	0.97
1:B:412:VAL:HB	1:B:459:ILE:HG22	1.53	0.90
1:C:412:VAL:HB	1:C:459:ILE:HG22	1.56	0.87
1:C:504:VAL:HG12	1:C:581:MET:HB2	1.56	0.87

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:492:LEU:N	1:C:407:GLU:OE1[8_455]	2.04	0.16

## 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	A	773/816 (95%)	696 (90%)	64 (8%)	13 (2%)	14	71
1	B	674/816 (83%)	605 (90%)	56 (8%)	13 (2%)	12	68
1	C	674/816 (83%)	600 (89%)	62 (9%)	12 (2%)	13	69
All	All	2121/2448 (87%)	1901 (90%)	182 (9%)	38 (2%)	13	69

5 of 38 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	613	LEU
1	A	626	ILE
1	B	613	LEU
1	B	626	ILE
1	C	613	LEU

### 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	A	654/687 (95%)	625 (96%)	29 (4%)	39	82
1	B	572/687 (83%)	548 (96%)	24 (4%)	40	83
1	C	572/687 (83%)	548 (96%)	24 (4%)	40	83
All	All	1798/2061 (87%)	1721 (96%)	77 (4%)	40	83

5 of 77 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	175	ASP
1	B	509	LEU
1	C	535	ASN
1	B	195	SER
1	B	374	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 18 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	402	GLN
1	B	535	ASN
1	C	167	HIS
1	B	31	ASN
1	B	147	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 ⓘ

20 carbohydrates 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	$\# Z  > 2$
2	NAG	A	782	1,2	12,14,15	0.56	0	15,19,21	0.98	1 (6%)
2	NAG	A	783	2	12,14,15	0.56	0	15,19,21	1.03	1 (6%)
2	NAG	A	784	1,2	12,14,15	0.54	0	15,19,21	1.06	1 (6%)
2	NAG	A	785	2	12,14,15	0.59	0	15,19,21	1.12	1 (6%)
2	NAG	B	782	1,2	12,14,15	0.81	1 (8%)	15,19,21	0.83	0
2	NAG	B	783	2	12,14,15	0.80	1 (8%)	15,19,21	1.22	1 (6%)
3	NAG	B	784	1,3	12,14,15	0.66	0	15,19,21	0.94	0
3	NAG	B	785	3	12,14,15	0.62	0	15,19,21	1.01	2 (13%)
3	NAG	B	786	3	12,14,15	0.62	0	15,19,21	0.89	0
4	NDG	B	787	1,4	12,14,15	0.87	0	15,19,21	1.90	4 (26%)
4	NAG	B	788	4	12,14,15	0.72	0	15,19,21	1.35	3 (20%)
4	NAG	B	789	4	12,14,15	0.61	0	15,19,21	0.90	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	NAG	B	790	1,2	12,14,15	0.75	0	15,19,21	1.35	3 (20%)
2	NAG	B	791	2	12,14,15	0.67	1 (8%)	15,19,21	1.32	3 (20%)
3	NAG	C	782	1,3	12,14,15	0.53	0	15,19,21	1.09	2 (13%)
3	NAG	C	783	3	12,14,15	0.71	1 (8%)	15,19,21	1.28	2 (13%)
3	NAG	C	784	3	12,14,15	0.62	0	15,19,21	0.79	0
4	NDG	C	785	1,4	12,14,15	0.55	0	15,19,21	0.76	0
4	NAG	C	786	4	12,14,15	0.72	1 (8%)	15,19,21	1.15	1 (6%)
4	NAG	C	787	4	12,14,15	0.69	0	15,19,21	1.04	1 (6%)

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
2	NAG	A	782	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	783	2	-	0/6/23/26	0/1/1/1
2	NAG	A	784	1,2	-	0/6/23/26	0/1/1/1
2	NAG	A	785	2	-	2/6/23/26	0/1/1/1
2	NAG	B	782	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	783	2	-	0/6/23/26	0/1/1/1
3	NAG	B	784	1,3	1/1/5/7	0/6/23/26	0/1/1/1
3	NAG	B	785	3	-	0/6/23/26	0/1/1/1
3	NAG	B	786	3	-	0/6/23/26	0/1/1/1
4	NDG	B	787	1,4	-	0/6/23/26	0/1/1/1
4	NAG	B	788	4	-	0/6/23/26	0/1/1/1
4	NAG	B	789	4	-	0/6/23/26	0/1/1/1
2	NAG	B	790	1,2	-	0/6/23/26	0/1/1/1
2	NAG	B	791	2	-	0/6/23/26	0/1/1/1
3	NAG	C	782	1,3	-	0/6/23/26	0/1/1/1
3	NAG	C	783	3	-	0/6/23/26	0/1/1/1
3	NAG	C	784	3	-	0/6/23/26	0/1/1/1
4	NDG	C	785	1,4	-	0/6/23/26	0/1/1/1
4	NAG	C	786	4	-	0/6/23/26	0/1/1/1
4	NAG	C	787	4	-	0/6/23/26	0/1/1/1

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	782	NAG	O5-C5	-2.58	1.40	1.45
2	B	783	NAG	O5-C5	-2.27	1.41	1.45

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	C	786	NAG	O5-C5	-2.22	1.41	1.45
3	C	783	NAG	O5-C5	-2.12	1.41	1.45
2	B	791	NAG	O5-C5	-2.04	1.41	1.45

The worst 5 of 26 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	B	787	NDG	O-C5-C6	3.66	110.83	106.98
4	B	787	NDG	C3-C4-C5	3.43	116.33	110.20
3	C	783	NAG	O5-C5-C4	-3.35	106.40	110.65
2	B	791	NAG	O5-C5-C4	-3.30	106.47	110.65
4	B	788	NAG	O5-C5-C6	3.28	110.42	106.98

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	B	784	NAG	C1

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	785	NAG	O7-C7-N2-C2
2	A	785	NAG	C8-C7-N2-C2

There are no ring outliers.

## 5.6 Ligand geometry

5 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	GOL	A	786	-	5,5,5	0.33	0	5,5,5	0.36	0
6	GOL	A	787	-	5,5,5	0.35	0	5,5,5	0.31	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	SO4	B	792	-	4,4,4	0.26	0	6,6,6	0.12	0
6	GOL	B	793	-	5,5,5	0.36	0	5,5,5	0.21	0
6	GOL	B	794	-	5,5,5	0.38	0	5,5,5	0.33	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	GOL	A	786	-	-	0/4/4/4	0/0/0/0
6	GOL	A	787	-	-	0/4/4/4	0/0/0/0
5	SO4	B	792	-	-	0/0/0/0	0/0/0/0
6	GOL	B	793	-	-	0/4/4/4	0/0/0/0
6	GOL	B	794	-	-	0/4/4/4	0/0/0/0

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.

## 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	A	775/816 (94%)	0.55	72 (9%) 9 12	27, 55, 89, 101	0
1	B	676/816 (82%)	-0.13	0 100 100	16, 42, 82, 93	0
1	C	676/816 (82%)	0.55	49 (7%) 15 17	25, 53, 101, 108	0
All	All	2127/2448 (86%)	0.33	121 (5%) 23 22	16, 50, 93, 108	0

The worst 5 of 121 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	567	GLN	6.6
1	A	439	ASP	5.8
1	A	477	LYS	5.7
1	A	492	LEU	5.0
1	A	570	TYR	5.0

### 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 ⓘ

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
2	NAG	B	791	14/15	0.42	5.99	71,72,73,73	0

*Continued on next page...*

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
4	NAG	B	788	14/15	0.36	5.15	77,80,81,81	0
2	NAG	B	790	14/15	0.45	3.12	64,66,68,70	0
3	NAG	B	785	14/15	0.28	3.00	44,50,51,51	0
2	NAG	A	783	14/15	0.38	2.95	52,53,53,54	0
2	NAG	A	784	14/15	0.37	2.68	62,65,66,66	0
4	NAG	B	789	14/15	0.47	2.51	83,84,85,86	0
3	NAG	B	784	14/15	0.27	1.99	38,44,48,48	0
2	NAG	A	785	14/15	0.34	1.95	66,66,67,67	0
4	NDG	B	787	14/15	0.27	1.05	60,64,68,72	0
4	NDG	C	785	14/15	0.25	0.78	63,64,67,67	0
2	NAG	B	782	14/15	0.21	0.33	46,47,49,51	0
3	NAG	C	784	14/15	0.36	0.15	53,53,55,55	0
3	NAG	C	782	14/15	0.24	0.04	42,45,46,46	0
2	NAG	A	782	14/15	0.20	-0.16	50,52,52,53	0
3	NAG	C	783	14/15	0.21	-1.17	46,49,50,51	0
4	NAG	C	787	14/15	0.36	-	66,67,67,67	0
2	NAG	B	783	14/15	0.41	-	54,56,59,59	0
4	NAG	C	786	14/15	0.31	-	66,66,67,67	0
3	NAG	B	786	14/15	0.58	-	52,53,54,54	0

## 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( $\text{\AA}^2$ )	Q<0.9
6	GOL	A	786	6/6	0.27	3.80	25,27,27,27	0
6	GOL	B	794	6/6	0.27	2.83	41,41,42,44	0
6	GOL	B	793	6/6	0.19	1.86	23,23,24,24	0
6	GOL	A	787	6/6	0.24	1.20	22,22,22,23	0
5	SO4	B	792	5/5	0.14	-1.07	20,21,21,24	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.