



# Full wwPDB X-ray Structure Validation Report ⓘ

Nov 14, 2023 – 01:24 AM JST

PDB ID : 5YQ5  
Title : Crystal structure of human osteomodulin  
Authors : Caaveiro, J.M.M.; Tashima, T.; Tsumoto, K.  
Deposited on : 2017-11-05  
Resolution : 2.17 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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)  
Xtrriage (Phenix) : 1.13  
EDS : 2.36  
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) : 2.36

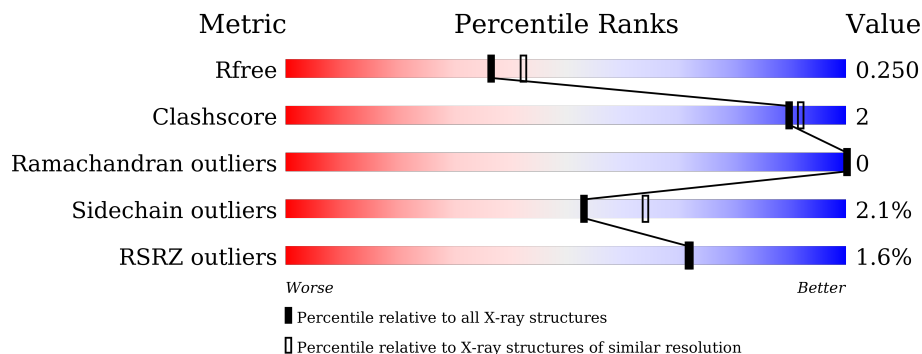
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.17 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	6864 (2.20-2.16)
Clashscore	141614	7689 (2.20-2.16)
Ramachandran outliers	138981	7564 (2.20-2.16)
Sidechain outliers	138945	7564 (2.20-2.16)
RSRZ outliers	127900	6738 (2.20-2.16)

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	430	 66% 30%
1	B	430	 65% 6% 29%
1	C	430	 65% 6% 29%
1	D	430	 66% 29%

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 10521 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 Osteomodulin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	302	2483	1603	413	449	18	0	0	0
1	B	306	2512	1622	417	454	19	0	1	0
1	C	304	2507	1618	416	454	19	0	2	0
1	D	305	2512	1621	416	456	19	0	2	0

There are 116 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	3	MET	-	expression tag	UNP Q99983
A	4	ARG	-	expression tag	UNP Q99983
A	5	VAL	-	expression tag	UNP Q99983
A	6	LEU	-	expression tag	UNP Q99983
A	7	VAL	-	expression tag	UNP Q99983
A	8	LEU	-	expression tag	UNP Q99983
A	9	LEU	-	expression tag	UNP Q99983
A	10	ALA	-	expression tag	UNP Q99983
A	11	CYS	-	expression tag	UNP Q99983
A	12	LEU	-	expression tag	UNP Q99983
A	13	ALA	-	expression tag	UNP Q99983
A	14	ALA	-	expression tag	UNP Q99983
A	15	ALA	-	expression tag	UNP Q99983
A	16	SER	-	expression tag	UNP Q99983
A	17	ASN	-	expression tag	UNP Q99983
A	18	ALA	-	expression tag	UNP Q99983
A	19	GLY	-	expression tag	UNP Q99983
A	20	SER	-	expression tag	UNP Q99983
A	422	ALA	-	expression tag	UNP Q99983
A	423	ALA	-	expression tag	UNP Q99983
A	424	ALA	-	expression tag	UNP Q99983

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
A	425	ASP	-	expression tag	UNP Q99983
A	426	TYR	-	expression tag	UNP Q99983
A	427	LYS	-	expression tag	UNP Q99983
A	428	ASP	-	expression tag	UNP Q99983
A	429	ASP	-	expression tag	UNP Q99983
A	430	ASP	-	expression tag	UNP Q99983
A	431	ASP	-	expression tag	UNP Q99983
A	432	LYS	-	expression tag	UNP Q99983
B	3	MET	-	expression tag	UNP Q99983
B	4	ARG	-	expression tag	UNP Q99983
B	5	VAL	-	expression tag	UNP Q99983
B	6	LEU	-	expression tag	UNP Q99983
B	7	VAL	-	expression tag	UNP Q99983
B	8	LEU	-	expression tag	UNP Q99983
B	9	LEU	-	expression tag	UNP Q99983
B	10	ALA	-	expression tag	UNP Q99983
B	11	CYS	-	expression tag	UNP Q99983
B	12	LEU	-	expression tag	UNP Q99983
B	13	ALA	-	expression tag	UNP Q99983
B	14	ALA	-	expression tag	UNP Q99983
B	15	ALA	-	expression tag	UNP Q99983
B	16	SER	-	expression tag	UNP Q99983
B	17	ASN	-	expression tag	UNP Q99983
B	18	ALA	-	expression tag	UNP Q99983
B	19	GLY	-	expression tag	UNP Q99983
B	20	SER	-	expression tag	UNP Q99983
B	422	ALA	-	expression tag	UNP Q99983
B	423	ALA	-	expression tag	UNP Q99983
B	424	ALA	-	expression tag	UNP Q99983
B	425	ASP	-	expression tag	UNP Q99983
B	426	TYR	-	expression tag	UNP Q99983
B	427	LYS	-	expression tag	UNP Q99983
B	428	ASP	-	expression tag	UNP Q99983
B	429	ASP	-	expression tag	UNP Q99983
B	430	ASP	-	expression tag	UNP Q99983
B	431	ASP	-	expression tag	UNP Q99983
B	432	LYS	-	expression tag	UNP Q99983
C	3	MET	-	expression tag	UNP Q99983
C	4	ARG	-	expression tag	UNP Q99983
C	5	VAL	-	expression tag	UNP Q99983
C	6	LEU	-	expression tag	UNP Q99983
C	7	VAL	-	expression tag	UNP Q99983

*Continued on next page...*

*Continued from previous page...*

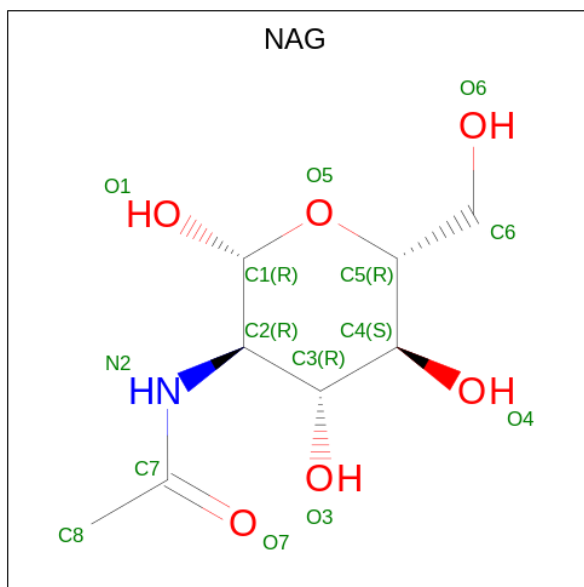
Chain	Residue	Modelled	Actual	Comment	Reference
C	8	LEU	-	expression tag	UNP Q99983
C	9	LEU	-	expression tag	UNP Q99983
C	10	ALA	-	expression tag	UNP Q99983
C	11	CYS	-	expression tag	UNP Q99983
C	12	LEU	-	expression tag	UNP Q99983
C	13	ALA	-	expression tag	UNP Q99983
C	14	ALA	-	expression tag	UNP Q99983
C	15	ALA	-	expression tag	UNP Q99983
C	16	SER	-	expression tag	UNP Q99983
C	17	ASN	-	expression tag	UNP Q99983
C	18	ALA	-	expression tag	UNP Q99983
C	19	GLY	-	expression tag	UNP Q99983
C	20	SER	-	expression tag	UNP Q99983
C	422	ALA	-	expression tag	UNP Q99983
C	423	ALA	-	expression tag	UNP Q99983
C	424	ALA	-	expression tag	UNP Q99983
C	425	ASP	-	expression tag	UNP Q99983
C	426	TYR	-	expression tag	UNP Q99983
C	427	LYS	-	expression tag	UNP Q99983
C	428	ASP	-	expression tag	UNP Q99983
C	429	ASP	-	expression tag	UNP Q99983
C	430	ASP	-	expression tag	UNP Q99983
C	431	ASP	-	expression tag	UNP Q99983
C	432	LYS	-	expression tag	UNP Q99983
D	3	MET	-	expression tag	UNP Q99983
D	4	ARG	-	expression tag	UNP Q99983
D	5	VAL	-	expression tag	UNP Q99983
D	6	LEU	-	expression tag	UNP Q99983
D	7	VAL	-	expression tag	UNP Q99983
D	8	LEU	-	expression tag	UNP Q99983
D	9	LEU	-	expression tag	UNP Q99983
D	10	ALA	-	expression tag	UNP Q99983
D	11	CYS	-	expression tag	UNP Q99983
D	12	LEU	-	expression tag	UNP Q99983
D	13	ALA	-	expression tag	UNP Q99983
D	14	ALA	-	expression tag	UNP Q99983
D	15	ALA	-	expression tag	UNP Q99983
D	16	SER	-	expression tag	UNP Q99983
D	17	ASN	-	expression tag	UNP Q99983
D	18	ALA	-	expression tag	UNP Q99983
D	19	GLY	-	expression tag	UNP Q99983
D	20	SER	-	expression tag	UNP Q99983

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	422	ALA	-	expression tag	UNP Q99983
D	423	ALA	-	expression tag	UNP Q99983
D	424	ALA	-	expression tag	UNP Q99983
D	425	ASP	-	expression tag	UNP Q99983
D	426	TYR	-	expression tag	UNP Q99983
D	427	LYS	-	expression tag	UNP Q99983
D	428	ASP	-	expression tag	UNP Q99983
D	429	ASP	-	expression tag	UNP Q99983
D	430	ASP	-	expression tag	UNP Q99983
D	431	ASP	-	expression tag	UNP Q99983
D	432	LYS	-	expression tag	UNP Q99983

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



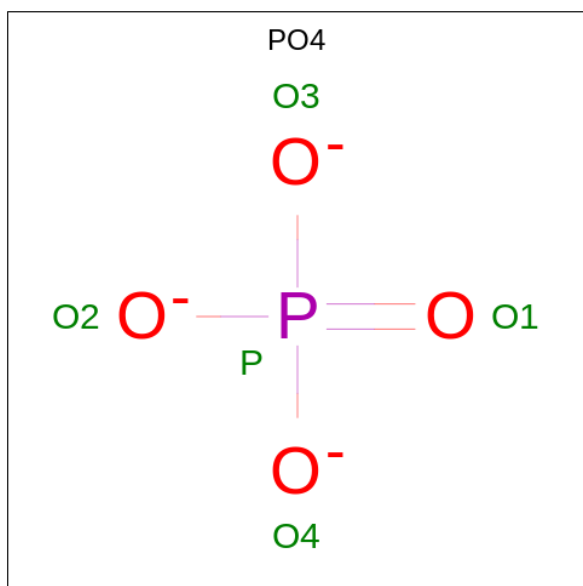
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

Continued from previous page...

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

- Molecule 3 is PHOSPHATE ION (three-letter code: PO4) (formula: O<sub>4</sub>P).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	O	P		
3	B	1	Total 5	O 4	P 1	0	0
3	D	1	Total 5	O 4	P 1	0	0

- Molecule 4 is water.

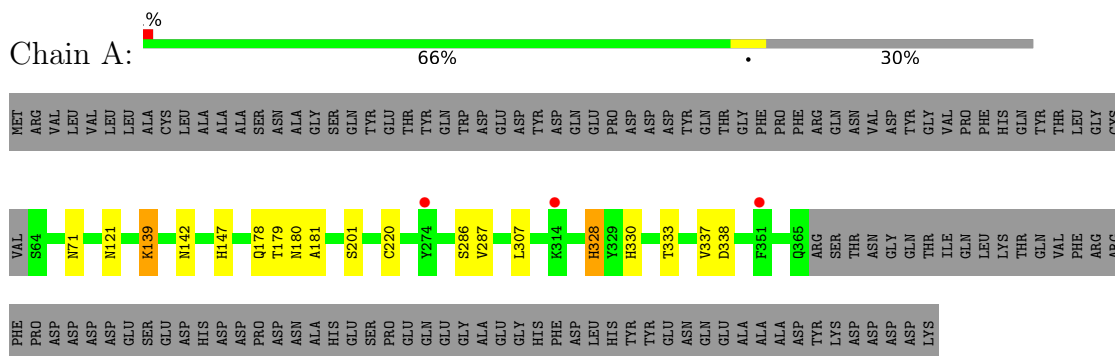
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	54	Total O 54 54	0	0
4	B	54	Total O 54 54	0	0
4	C	97	Total O 97 97	0	0
4	D	109	Total O 110 110	0	2



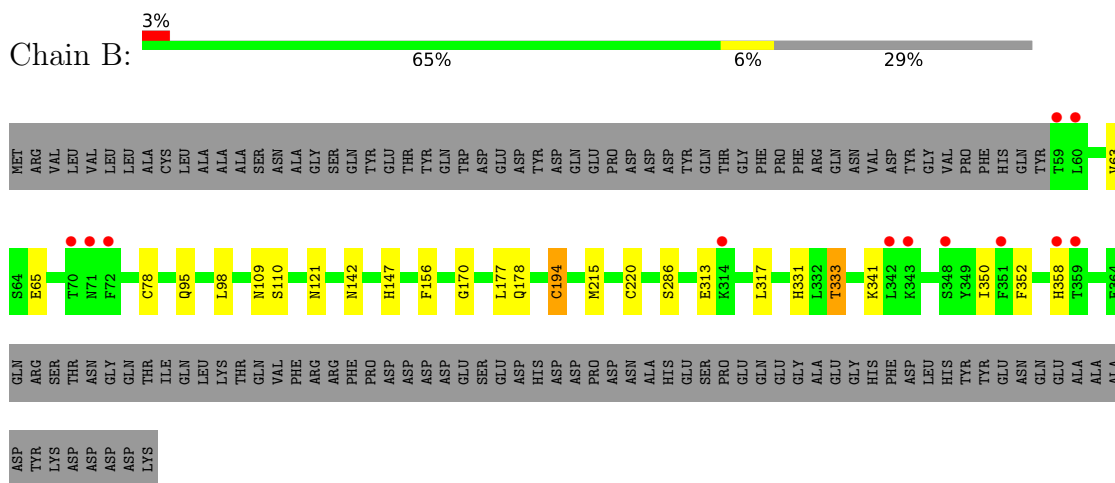
### 3 Residue-property plots i

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.

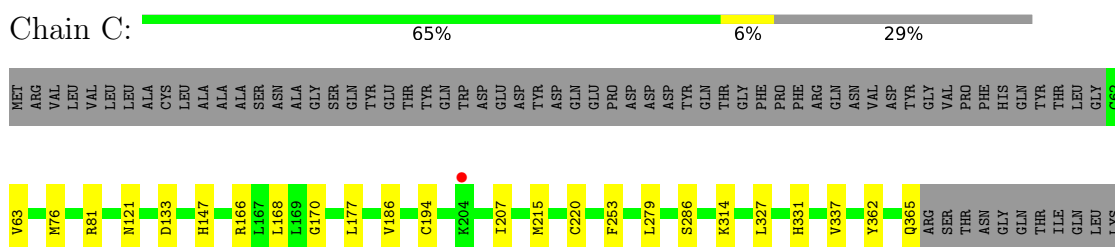
- Molecule 1: Osteomodulin



- Molecule 1: Osteomodulin



- Molecule 1: Osteomodulin



THR  
GLN  
VAL  
PHE  
ARG  
ARG  
PHE  
PRO  
ASP  
ASP  
ASP  
GLU  
SER  
GLU  
HIS  
ASP  
ASP  
PRO  
ASP  
ASN  
ALA  
HIS  
GLU  
SER  
PRO  
ASP  
GLU  
GLY  
ALA  
GLU  
GLY  
PHE  
HIS  
ASP  
LEU  
HIS  
TYR  
TYR  
GLU  
ASN  
GLN  
GLU  
ALA  
ALA  
ASP  
ASP  
TYR  
TYR  
THR  
LYS  
ASP  
ASP  
ASP  
LYS

● Molecule 1: Osteomodulin



MET  
ARG  
VAL  
LEU  
VAL  
LEU  
ALA  
CYS  
LEU  
ALA  
ALA  
SER  
ASN  
ALA  
GLY  
SER  
GLN  
TYR  
GLU  
THR  
TYR  
GLN  
TRP  
ASP  
GLU  
ASP  
TYR  
ASP  
GLN  
GLU  
PRO  
ASP  
ASP  
TYR  
GLN  
THR  
GLY  
PHE  
PRO  
PHE  
ARG  
GLN  
GLN  
ASN  
VAL  
ASP  
TYR  
GLY  
VAL  
PRO  
PHE  
HIS  
HIS  
TYR  
THR  
LEU  
C61  
C62



PHE  
PRO  
ASP  
ASP  
ASP  
GLU  
SER  
GLU  
ASP  
HIS  
ASP  
PRO  
ASP  
ASN  
ALA  
HIS  
SER  
PRO  
GLU  
GLN  
GLY  
ALA  
GLY  
PHE  
ASP  
LEU  
HIS  
TYR  
GLU  
ASN  
GLN  
GLU  
ALA  
ALA  
ASP  
LYS  
ASP  
ASP  
ASP  
LYS

## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	75.64Å 110.70Å 122.15Å 90.00° 106.97° 90.00°	Depositor
Resolution (Å)	35.19 – 2.17 35.19 – 2.17	Depositor EDS
% Data completeness (in resolution range)	99.0 (35.19-2.17) 99.1 (35.19-2.17)	Depositor EDS
$R_{merge}$	0.08	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.93 (at 2.16Å)	Xtrriage
Refinement program	REFMAC 5.8.0158	Depositor
R, $R_{free}$	0.217 , 0.245 0.221 , 0.250	Depositor DCC
$R_{free}$ test set	5191 reflections (5.12%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	37.7	Xtrriage
Anisotropy	0.384	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 23.3	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.46$ , $\langle L^2 \rangle = 0.29$	Xtrriage
Estimated twinning fraction	0.073 for h,-k,-h-l	Xtrriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	10521	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	45.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.60% 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: NAG, PO4

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	0.74	0/2547	0.83	2/3451 (0.1%)
1	B	0.68	0/2579	0.80	2/3495 (0.1%)
1	C	0.73	2/2577 (0.1%)	0.81	4/3492 (0.1%)
1	D	0.77	0/2582	0.83	3/3498 (0.1%)
All	All	0.73	2/10285 (0.0%)	0.82	11/13936 (0.1%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	133[A]	ASP	CB-CG	6.24	1.64	1.51
1	C	133[B]	ASP	CB-CG	6.24	1.64	1.51

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	194	CYS	CA-CB-SG	-10.05	95.91	114.00
1	D	81	ARG	NE-CZ-NH2	-9.66	115.47	120.30
1	C	166	ARG	NE-CZ-NH1	6.82	123.71	120.30
1	C	81	ARG	NE-CZ-NH1	6.80	123.70	120.30
1	A	328	HIS	N-CA-CB	6.14	121.66	110.60
1	C	215	MET	CG-SD-CE	5.98	109.76	100.20
1	B	215	MET	CG-SD-CE	5.58	109.13	100.20
1	A	338	ASP	CB-CG-OD2	-5.40	113.44	118.30
1	D	166	ARG	NE-CZ-NH1	5.37	122.98	120.30
1	C	166	ARG	NE-CZ-NH2	-5.26	117.67	120.30
1	D	251	LYS	CA-CB-CG	5.02	124.45	113.40

There are no chirality outliers.

There are no planarity outliers.

## 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	A	2483	0	2481	6	0
1	B	2512	0	2515	13	0
1	C	2507	0	2506	11	0
1	D	2512	0	2509	7	0
2	A	56	0	52	0	0
2	B	42	0	39	0	0
2	C	42	0	39	0	0
2	D	42	0	39	0	0
3	B	5	0	0	1	0
3	D	5	0	0	0	0
4	A	54	0	0	1	0
4	B	54	0	0	1	0
4	C	97	0	0	1	0
4	D	110	0	0	0	0
All	All	10521	0	10180	37	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 2.

All (37) 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:B:156:PHE:HB2	1:B:177:LEU:HD12	1.69	0.75
1:C:253:PHE:CB	1:C:279:LEU:HD11	2.28	0.64
1:C:253:PHE:HB3	1:C:279:LEU:HD11	1.81	0.61
1:C:177:LEU:HD22	1:C:207:ILE:HD13	1.83	0.59
1:C:63:VAL:HG22	1:C:76:MET:CE	2.35	0.56
1:B:331:HIS:NE2	3:B:504:PO4:O1	2.36	0.55
1:A:142:ASN:HB3	4:A:641:HOH:O	2.07	0.54
1:B:194:CYS:SG	1:B:220:CYS:N	2.81	0.53
1:B:317:LEU:HD21	1:B:350:ILE:N	2.26	0.51
1:D:333:THR:HG23	1:D:358:HIS:CD2	2.47	0.50
1:D:63:VAL:HG12	1:D:65:GLU:HG2	1.96	0.47
1:B:333:THR:HG23	1:B:358:HIS:CD2	2.50	0.46
1:C:147:HIS:CE1	1:C:168:LEU:HD12	2.50	0.46

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:147:HIS:ND1	1:C:168:LEU:HD12	2.31	0.46
1:B:313:GLU:HG2	1:B:341:LYS:HG3	1.98	0.46
1:C:362:TYR:O	1:C:365:GLN:NE2	2.48	0.45
1:D:62:CYS:SG	1:D:76:MET:HE2	2.56	0.45
1:B:109:ASN:OD1	1:B:110:SER:N	2.50	0.45
1:B:63:VAL:HG12	1:B:65:GLU:HG2	1.99	0.44
1:B:121:ASN:HA	1:B:147:HIS:HB2	2.00	0.44
1:D:121:ASN:HA	1:D:147:HIS:HB2	2.01	0.43
1:D:170:GLY:HA2	1:D:194:CYS:O	2.19	0.42
1:A:178:GLN:HG2	1:A:181:ALA:HB2	2.02	0.42
1:C:121:ASN:HA	1:C:147:HIS:HB2	2.02	0.42
1:A:179:THR:O	1:A:180:ASN:HB2	2.20	0.42
1:A:121:ASN:HA	1:A:147:HIS:HB2	2.01	0.41
1:C:186:VAL:HG12	1:C:186:VAL:O	2.19	0.41
1:B:78:CYS:HB2	1:B:98:LEU:HD23	2.02	0.41
1:D:194:CYS:HA	1:D:220:CYS:O	2.19	0.41
1:D:362:TYR:O	1:D:365:GLN:NE2	2.51	0.41
1:B:170:GLY:HA2	1:B:194:CYS:O	2.20	0.41
1:B:317:LEU:HD21	1:B:350:ILE:CA	2.51	0.41
1:B:142:ASN:HB3	4:B:639:HOH:O	2.21	0.41
1:A:287:VAL:HG23	1:A:307:LEU:HD23	2.03	0.41
1:C:170:GLY:HA2	1:C:194:CYS:O	2.21	0.41
1:C:194:CYS:HA	1:C:220:CYS:O	2.21	0.40
1:A:139:LYS:NZ	4:C:609:HOH:O	2.54	0.40

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	300/430 (70%)	277 (92%)	23 (8%)	0	<a href="#">100</a> <a href="#">100</a>

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	B	305/430 (71%)	283 (93%)	22 (7%)	0	100	100
1	C	304/430 (71%)	282 (93%)	22 (7%)	0	100	100
1	D	305/430 (71%)	282 (92%)	23 (8%)	0	100	100
All	All	1214/1720 (71%)	1124 (93%)	90 (7%)	0	100	100

There are no Ramachandran outliers to report.

### 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	289/400 (72%)	280 (97%)	9 (3%)	40	48
1	B	293/400 (73%)	288 (98%)	5 (2%)	60	72
1	C	293/400 (73%)	288 (98%)	5 (2%)	60	72
1	D	293/400 (73%)	288 (98%)	5 (2%)	60	72
All	All	1168/1600 (73%)	1144 (98%)	24 (2%)	53	64

All (24) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	71	ASN
1	A	139	LYS
1	A	201	SER
1	A	220	CYS
1	A	286	SER
1	A	328	HIS
1	A	330	HIS
1	A	333	THR
1	A	337	VAL
1	B	95	GLN
1	B	178	GLN
1	B	286	SER
1	B	333	THR

Continued on next page...

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	352	PHE
1	C	286	SER
1	C	314	LYS
1	C	327	LEU
1	C	331	HIS
1	C	337	VAL
1	D	271	ASP
1	D	286	SER
1	D	317	LEU
1	D	333	THR
1	D	337	VAL

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	95	GLN
1	B	113	ASN
1	B	150	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](#)

15 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
2	NAG	C	501	1	14,14,15	0.75	0	17,19,21	1.66	4 (23%)
2	NAG	D	503	1	14,14,15	0.63	0	17,19,21	1.91	3 (17%)
2	NAG	A	504	1	14,14,15	0.62	0	17,19,21	1.61	4 (23%)
2	NAG	B	502	1	14,14,15	1.05	1 (7%)	17,19,21	2.20	5 (29%)
2	NAG	A	503	1	14,14,15	0.61	0	17,19,21	1.38	2 (11%)
2	NAG	B	501	1	14,14,15	0.71	0	17,19,21	1.66	3 (17%)
2	NAG	D	502	1	14,14,15	0.71	0	17,19,21	1.96	2 (11%)
2	NAG	D	501	1	14,14,15	1.00	1 (7%)	17,19,21	2.37	6 (35%)
2	NAG	C	502	1	14,14,15	0.67	0	17,19,21	1.60	3 (17%)
2	NAG	C	503	1	14,14,15	0.77	0	17,19,21	2.07	5 (29%)
3	PO4	D	504	-	4,4,4	1.05	0	6,6,6	0.85	0
3	PO4	B	504	-	4,4,4	0.81	0	6,6,6	0.75	0
2	NAG	A	501	1	14,14,15	0.57	0	17,19,21	1.55	4 (23%)
2	NAG	A	502	1	14,14,15	0.67	0	17,19,21	1.99	3 (17%)
2	NAG	B	503	1	14,14,15	0.87	0	17,19,21	2.02	4 (23%)

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	C	501	1	-	2/6/23/26	0/1/1/1
2	NAG	D	503	1	-	4/6/23/26	0/1/1/1
2	NAG	A	504	1	-	4/6/23/26	0/1/1/1
2	NAG	B	502	1	-	0/6/23/26	0/1/1/1
2	NAG	A	503	1	-	2/6/23/26	0/1/1/1
2	NAG	B	501	1	-	0/6/23/26	0/1/1/1
2	NAG	D	502	1	-	1/6/23/26	0/1/1/1
2	NAG	D	501	1	-	4/6/23/26	0/1/1/1
2	NAG	C	502	1	-	2/6/23/26	0/1/1/1
2	NAG	C	503	1	-	2/6/23/26	0/1/1/1
2	NAG	A	501	1	-	3/6/23/26	0/1/1/1
2	NAG	A	502	1	-	0/6/23/26	0/1/1/1

Continued on next page...

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	B	503	1	-	3/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	502	NAG	O5-C5	2.06	1.47	1.43
2	D	501	NAG	C3-C2	2.04	1.56	1.52

All (48) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	502	NAG	C1-O5-C5	6.81	121.42	112.19
2	D	502	NAG	C1-O5-C5	6.30	120.73	112.19
2	D	503	NAG	C1-O5-C5	5.40	119.51	112.19
2	D	501	NAG	C4-C3-C2	4.99	118.34	111.02
2	B	501	NAG	C1-O5-C5	4.77	118.66	112.19
2	C	502	NAG	C1-O5-C5	4.69	118.55	112.19
2	B	502	NAG	C1-O5-C5	4.69	118.55	112.19
2	B	502	NAG	C1-C2-N2	-4.29	103.16	110.49
2	D	501	NAG	C1-C2-N2	4.26	117.77	110.49
2	B	503	NAG	C8-C7-N2	4.17	123.16	116.10
2	D	501	NAG	C1-O5-C5	4.11	117.75	112.19
2	A	504	NAG	C1-O5-C5	4.04	117.66	112.19
2	C	503	NAG	C2-N2-C7	3.74	128.22	122.90
2	D	501	NAG	C2-N2-C7	3.72	128.20	122.90
2	B	503	NAG	C1-O5-C5	3.69	117.19	112.19
2	C	503	NAG	C8-C7-N2	3.64	122.27	116.10
2	D	503	NAG	C2-N2-C7	3.61	128.05	122.90
2	C	501	NAG	C1-C2-N2	-3.57	104.39	110.49
2	C	503	NAG	O7-C7-N2	-3.53	115.47	121.95
2	B	503	NAG	C2-N2-C7	3.40	127.75	122.90
2	C	503	NAG	C1-C2-N2	-3.37	104.73	110.49
2	C	501	NAG	O5-C5-C4	-3.29	102.82	110.83
2	A	501	NAG	C1-C2-N2	-3.11	105.17	110.49
2	D	503	NAG	C1-C2-N2	-3.03	105.31	110.49
2	A	501	NAG	C2-N2-C7	3.00	127.17	122.90
2	B	502	NAG	O5-C1-C2	2.90	115.87	111.29
2	B	503	NAG	O5-C5-C6	2.89	111.73	107.20
2	C	501	NAG	C1-O5-C5	2.86	116.07	112.19
2	D	502	NAG	C6-C5-C4	-2.85	106.33	113.00
2	C	501	NAG	C6-C5-C4	2.83	119.64	113.00
2	B	502	NAG	O3-C3-C2	2.77	115.19	109.47

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	501	NAG	C1-C2-N2	-2.75	105.80	110.49
2	B	502	NAG	O4-C4-C5	2.59	115.72	109.30
2	A	503	NAG	O3-C3-C2	-2.58	104.12	109.47
2	D	501	NAG	C3-C4-C5	2.58	114.84	110.24
2	A	504	NAG	C2-N2-C7	2.57	126.57	122.90
2	A	504	NAG	C8-C7-N2	2.48	120.29	116.10
2	C	502	NAG	C1-C2-N2	-2.46	106.28	110.49
2	D	501	NAG	O5-C1-C2	-2.44	107.44	111.29
2	A	501	NAG	C8-C7-N2	2.37	120.11	116.10
2	A	503	NAG	C1-C2-N2	2.35	114.50	110.49
2	A	501	NAG	O3-C3-C2	2.32	114.27	109.47
2	A	502	NAG	C6-C5-C4	-2.26	107.71	113.00
2	C	503	NAG	C3-C4-C5	-2.26	106.22	110.24
2	A	504	NAG	O3-C3-C2	-2.25	104.82	109.47
2	B	501	NAG	O4-C4-C3	2.18	115.38	110.35
2	A	502	NAG	O7-C7-N2	2.05	125.73	121.95
2	C	502	NAG	C3-C4-C5	-2.02	106.64	110.24

There are no chirality outliers.

All (27) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	D	503	NAG	O5-C5-C6-O6
2	D	501	NAG	C4-C5-C6-O6
2	D	503	NAG	C4-C5-C6-O6
2	D	501	NAG	O5-C5-C6-O6
2	A	501	NAG	C8-C7-N2-C2
2	A	501	NAG	O7-C7-N2-C2
2	A	504	NAG	C8-C7-N2-C2
2	A	504	NAG	O7-C7-N2-C2
2	B	503	NAG	C8-C7-N2-C2
2	B	503	NAG	O7-C7-N2-C2
2	C	503	NAG	C8-C7-N2-C2
2	C	503	NAG	O7-C7-N2-C2
2	D	503	NAG	C8-C7-N2-C2
2	D	503	NAG	O7-C7-N2-C2
2	C	502	NAG	O5-C5-C6-O6
2	C	501	NAG	O5-C5-C6-O6
2	A	504	NAG	C4-C5-C6-O6
2	A	503	NAG	O5-C5-C6-O6
2	A	504	NAG	O5-C5-C6-O6
2	C	501	NAG	C4-C5-C6-O6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms
2	C	502	NAG	C4-C5-C6-O6
2	D	502	NAG	C4-C5-C6-O6
2	D	501	NAG	C1-C2-N2-C7
2	A	501	NAG	C4-C5-C6-O6
2	D	501	NAG	C3-C2-N2-C7
2	A	503	NAG	C4-C5-C6-O6
2	B	503	NAG	C4-C5-C6-O6

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	B	504	PO4	1	0

## 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

### 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	302/430 (70%)	-0.08	3 (0%) 82 82	24, 43, 71, 101	0
1	B	306/430 (71%)	0.14	12 (3%) 39 40	27, 50, 87, 112	0
1	C	304/430 (70%)	-0.14	1 (0%) 94 94	23, 39, 63, 80	0
1	D	305/430 (70%)	-0.15	3 (0%) 82 82	22, 37, 64, 94	0
All	All	1217/1720 (70%)	-0.06	19 (1%) 72 72	22, 42, 74, 112	0

All (19) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	343	LYS	7.8
1	B	71	ASN	4.8
1	A	274	TYR	3.7
1	D	343	LYS	3.7
1	B	72	PHE	3.5
1	C	204	LYS	2.7
1	B	70	THR	2.5
1	B	59	THR	2.5
1	B	60	LEU	2.5
1	B	342	LEU	2.5
1	A	314	LYS	2.4
1	A	351	PHE	2.3
1	B	358	HIS	2.3
1	B	348	SER	2.3
1	B	351	PHE	2.3
1	B	314	LYS	2.2
1	B	359	THR	2.1
1	D	351	PHE	2.1
1	D	109	ASN	2.0

## 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
2	NAG	A	501	14/15	0.72	0.19	98,107,118,120	0
2	NAG	B	503	14/15	0.72	0.20	101,108,117,118	0
2	NAG	D	501	14/15	0.77	0.19	72,89,101,107	0
2	NAG	C	501	14/15	0.79	0.27	80,95,103,112	0
2	NAG	A	502	14/15	0.87	0.18	58,74,83,84	0
2	NAG	D	503	14/15	0.87	0.17	58,76,84,90	0
2	NAG	C	503	14/15	0.88	0.10	71,77,82,82	0
3	PO4	B	504	5/5	0.88	0.22	50,53,58,61	5
2	NAG	B	501	14/15	0.89	0.14	58,65,71,72	0
2	NAG	B	502	14/15	0.90	0.17	52,61,85,95	0
2	NAG	C	502	14/15	0.91	0.14	46,59,62,67	0
2	NAG	D	502	14/15	0.93	0.15	41,54,72,77	0
2	NAG	A	504	14/15	0.93	0.11	68,80,85,92	0
2	NAG	A	503	14/15	0.93	0.12	59,63,69,71	0
3	PO4	D	504	5/5	0.98	0.06	65,68,72,77	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.