



# wwPDB X-ray Structure Validation Summary Report ⓘ

Aug 22, 2020 – 02:53 AM BST

PDB ID : 5EIV  
Title : Crystal structure of complex of osteoclast-associated immunoglobulin-like receptor (OSCAR) and a synthetic collagen consensus peptide  
Authors : Zhou, L.; Blaszczyk, M.; Chirgadze, D.; Bihan, D.; Farndale, R.W.  
Deposited on : 2015-10-30  
Resolution : 2.41 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 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 : 2.13.1  
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) : 2.13.1

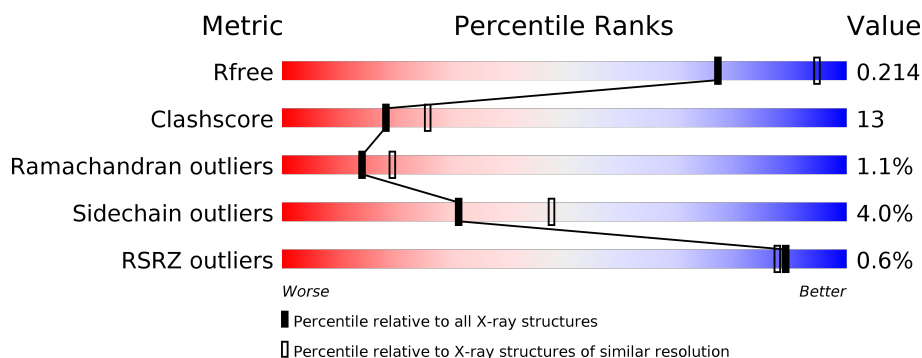
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.41 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	4647 (2.44-2.40)
Clashscore	141614	5161 (2.44-2.40)
Ramachandran outliers	138981	5073 (2.44-2.40)
Sidechain outliers	138945	5074 (2.44-2.40)
RSRZ outliers	127900	4543 (2.44-2.40)

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 on 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	203	<div> <div>5%</div> <div>61% 27% 11%</div> </div>
1	B	203	<div> <div>63% 24% 11%</div> </div>
2	C	21	<div> <div>57% 38% 5%</div> </div>
2	D	21	<div> <div>48% 43% 5% 5%</div> </div>
2	E	21	<div> <div>52% 48%</div> </div>
2	F	21	<div> <div>5% 57% 38% 5%</div> </div>

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	G	21	<div><div></div><div>48%48%5%</div></div>
2	H	21	<div><div></div><div>52%33%14%</div></div>
3	I	7	<div><div></div><div>29%43%14%14%</div></div>
3	J	7	<div><div></div><div>29%14%57%</div></div>
3	K	7	<div><div></div><div>14%57%14%14%</div></div>

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 4106 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 Osteoclast-associated immunoglobulin-like receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	181	Total	C	N	O	S	0	0	0
			1404	907	243	248	6			
1	B	181	Total	C	N	O	S	0	0	0
			1417	916	246	249	6			

There are 28 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-3	ALA	-	expression tag	UNP Q8IYS5
A	-2	PRO	-	expression tag	UNP Q8IYS5
A	-1	LEU	-	expression tag	UNP Q8IYS5
A	0	ALA	-	expression tag	UNP Q8IYS5
A	72	SER	ILE	variant	UNP Q8IYS5
A	191	ALA	-	expression tag	UNP Q8IYS5
A	192	ALA	-	expression tag	UNP Q8IYS5
A	193	ALA	-	expression tag	UNP Q8IYS5
A	194	HIS	-	expression tag	UNP Q8IYS5
A	195	HIS	-	expression tag	UNP Q8IYS5
A	196	HIS	-	expression tag	UNP Q8IYS5
A	197	HIS	-	expression tag	UNP Q8IYS5
A	198	HIS	-	expression tag	UNP Q8IYS5
A	199	HIS	-	expression tag	UNP Q8IYS5
B	-3	ALA	-	expression tag	UNP Q8IYS5
B	-2	PRO	-	expression tag	UNP Q8IYS5
B	-1	LEU	-	expression tag	UNP Q8IYS5
B	0	ALA	-	expression tag	UNP Q8IYS5
B	72	SER	ILE	variant	UNP Q8IYS5
B	191	ALA	-	expression tag	UNP Q8IYS5
B	192	ALA	-	expression tag	UNP Q8IYS5
B	193	ALA	-	expression tag	UNP Q8IYS5
B	194	HIS	-	expression tag	UNP Q8IYS5
B	195	HIS	-	expression tag	UNP Q8IYS5
B	196	HIS	-	expression tag	UNP Q8IYS5

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	197	HIS	-	expression tag	UNP Q8IYS5
B	198	HIS	-	expression tag	UNP Q8IYS5
B	199	HIS	-	expression tag	UNP Q8IYS5

- Molecule 2 is a protein called GLY-PRO-HYP-GLY-PRO-HYP-GLY-PRO-HYP-GLY-PRO-ALA-GLY-PHE-HYP-GLY-PRO-HYP-GLY-PRO-HYP.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	21	Total	C	N	O	0	0	0
			134	86	21	27			
2	D	20	Total	C	N	O	0	0	0
			130	84	20	26			
2	E	21	Total	C	N	O	0	0	0
			134	86	21	27			
2	F	20	Total	C	N	O	0	0	0
			126	81	20	25			
2	G	20	Total	C	N	O	0	0	0
			130	84	20	26			
2	H	21	Total	C	N	O	0	0	0
			134	86	21	27			

- Molecule 3 is a protein called PRO-GLY-PRO-PRO-GLY-PRO-PRO.

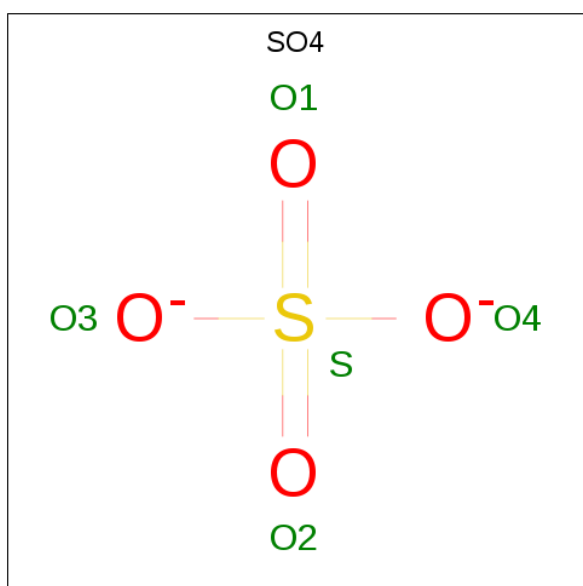
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
3	I	6	Total	C	N	O	0	0	0
			38	24	6	8			
3	J	7	Total	C	N	O	0	0	0
			46	29	7	10			
3	K	6	Total	C	N	O	0	0	0
			38	24	6	8			

- Molecule 4 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: C<sub>8</sub>H<sub>15</sub>NO<sub>6</sub>).



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

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



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

*Continued on next page...*

*Continued from previous page...*

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

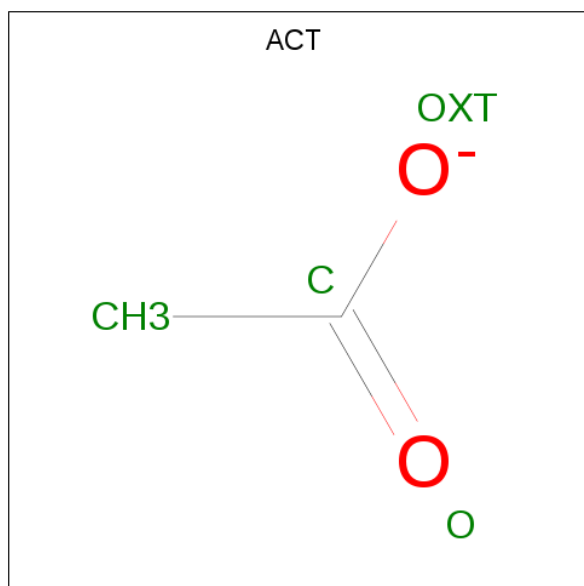
- Molecule 6 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
6	I	1	Total	Cl	0	0
			1	1		
6	A	2	Total	Cl	0	0
			2	2		

- Molecule 7 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	B	2	Total	Na	0	0
			2	2		

- Molecule 8 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	F	1	Total	C	O	0	0
			3	2	1		

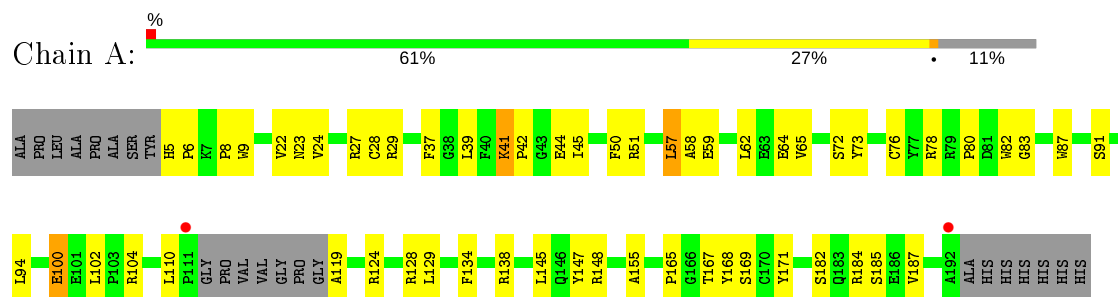
- Molecule 9 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	114	Total 114	O 114	0	0
9	B	106	Total 106	O 106	0	0
9	C	7	Total 7	O 7	0	0
9	D	19	Total 19	O 19	0	0
9	E	16	Total 16	O 16	0	0
9	F	7	Total 7	O 7	0	0
9	G	6	Total 6	O 6	0	0
9	H	21	Total 21	O 21	0	0
9	I	5	Total 5	O 5	0	0
9	J	11	Total 11	O 11	0	0
9	K	7	Total 7	O 7	0	0

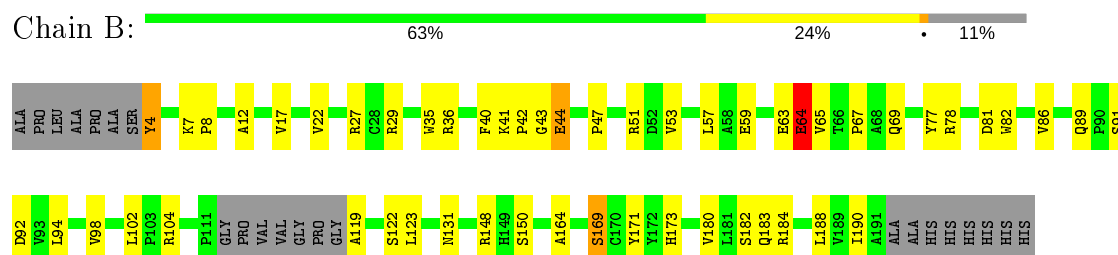
### 3 Residue-property plots

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: Osteoclast-associated immunoglobulin-like receptor



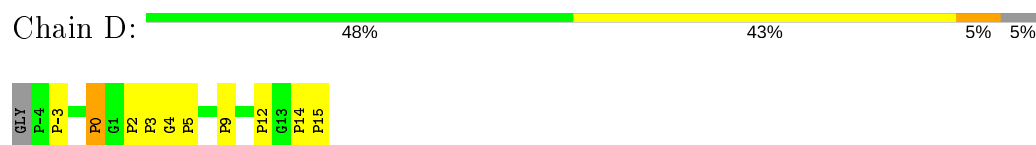
- Molecule 1: Osteoclast-associated immunoglobulin-like receptor



- Molecule 2: GLY-PRO-HYP-GLY-PRO-HYP-GLY-PRO-HYP-GLY-PRO-ALA-GLY-PHE-HYP-GLY-PRO-HYP-GLY-PRO-HYP



- Molecule 2: GLY-PRO-HYP-GLY-PRO-HYP-GLY-PRO-HYP-GLY-PRO-ALA-GLY-PHE-HYP-GLY-PRO-HYP-GLY-PRO-HYP



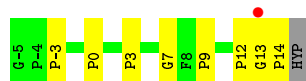
- Molecule 2: GLY-PRO-HYP-GLY-PRO-HYP-GLY-PRO-HYP-GLY-PRO-ALA-GLY-PHE-HYP-GLY-PRO-HYP-GLY-PRO-HYP

Chain E:  52% 48%



- Molecule 2: GLY-PRO-HYP-GLY-PRO-HYP-GLY-PRO-HYP-GLY-PRO-ALA-GLY-PHE-HYP-GLY-PRO-HYP-GLY-PRO-HYP

Chain F:  5% 57% 38% 5%



- Molecule 2: GLY-PRO-HYP-GLY-PRO-HYP-GLY-PRO-HYP-GLY-PRO-ALA-GLY-PHE-HYP-GLY-PRO-HYP-GLY-PRO-HYP

Chain G:  48% 48% 5%



- Molecule 2: GLY-PRO-HYP-GLY-PRO-HYP-GLY-PRO-HYP-GLY-PRO-ALA-GLY-PHE-HYP-GLY-PRO-HYP-GLY-PRO-HYP

Chain H:  52% 33% 14%



- Molecule 3: PRO-GLY-PRO-PRO-GLY-PRO-PRO

Chain I:  29% 43% 14% 14%

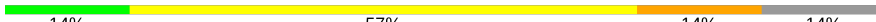


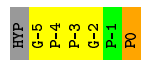
- Molecule 3: PRO-GLY-PRO-PRO-GLY-PRO-PRO

Chain J:  29% 14% 57%



- Molecule 3: PRO-GLY-PRO-PRO-GLY-PRO-PRO

Chain K:  14% 57% 14% 14%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	38.71Å 136.59Å 70.58Å 90.00° 90.05° 90.00°	Depositor
Resolution (Å)	70.58 – 2.41 70.58 – 2.41	Depositor EDS
% Data completeness (in resolution range)	99.9 (70.58-2.41) 99.7 (70.58-2.41)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.13	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.17 (at 2.42Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.174 , 0.220 0.175 , 0.214	Depositor DCC
$R_{free}$ test set	1415 reflections (5.03%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	33.8	Xtriage
Anisotropy	0.534	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 27.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.27$	Xtriage
Estimated twinning fraction	0.339 for h,-k,-l	Xtriage
Reported twinning fraction	0.330 for h,-k,-l	Depositor
Outliers	0 of 28132 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	4106	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	33.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.43% 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

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, CL, NA, HYP, SO4, ACT

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	0.43	0/1450	0.57	0/1989
1	B	0.42	0/1464	0.56	0/2007
2	C	0.32	0/87	0.48	0/112
2	D	0.35	0/83	0.50	0/106
2	E	0.33	0/87	0.50	0/112
2	F	0.34	0/87	0.53	0/112
2	G	0.32	0/83	0.53	0/106
2	H	0.41	0/87	0.58	0/112
3	I	0.33	0/22	0.55	0/28
3	J	0.50	0/22	0.55	0/28
3	K	0.37	0/22	0.52	0/28
All	All	0.41	0/3494	0.56	0/4740

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	H	0	1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	H	14	PRO	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	A	1404	0	1371	40	0
1	B	1417	0	1386	34	0
2	C	134	0	121	5	0
2	D	130	0	118	5	0
2	E	134	0	121	5	0
2	F	126	0	111	2	0
2	G	130	0	118	4	0
2	H	134	0	121	7	0
3	I	38	0	36	4	0
3	J	46	0	42	6	0
3	K	38	0	36	5	0
4	A	14	0	13	0	0
4	B	14	0	13	0	0
5	A	5	0	0	0	0
5	B	10	0	0	0	0
5	F	5	0	0	0	0
6	A	2	0	0	0	0
6	I	1	0	0	0	0
7	B	2	0	0	0	0
8	F	3	0	3	0	0
9	A	114	0	0	10	0
9	B	106	0	0	5	0
9	C	7	0	0	0	0
9	D	19	0	0	1	0
9	E	16	0	0	0	0
9	F	7	0	0	0	0
9	G	6	0	0	0	0
9	H	21	0	0	4	0
9	I	5	0	0	1	0
9	J	11	0	0	1	0
9	K	7	0	0	1	0
All	All	4106	0	3610	98	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 13.

The worst 5 of 98 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:78:ARG:NH1	1:B:86:VAL:O	2.11	0.83
1:B:12:ALA:HB2	1:B:94:LEU:HD21	1.67	0.77
1:A:138:ARG:NH2	1:A:165:PRO:O	2.23	0.71
2:D:0:HYP:O	9:D:101:HOH:O	2.07	0.71
1:A:119:ALA:N	9:A:304:HOH:O	2.24	0.70

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	177/203 (87%)	170 (96%)	7 (4%)	0	100	100
1	B	177/203 (87%)	166 (94%)	8 (4%)	3 (2%)	9	11
2	C	14/21 (67%)	14 (100%)	0	0	100	100
2	D	13/21 (62%)	13 (100%)	0	0	100	100
2	E	14/21 (67%)	14 (100%)	0	0	100	100
2	F	13/21 (62%)	12 (92%)	1 (8%)	0	100	100
2	G	13/21 (62%)	13 (100%)	0	0	100	100
2	H	14/21 (67%)	13 (93%)	0	1 (7%)	1	0
3	I	3/7 (43%)	3 (100%)	0	0	100	100
3	J	4/7 (57%)	3 (75%)	0	1 (25%)	0	0
3	K	3/7 (43%)	3 (100%)	0	0	100	100
All	All	445/553 (80%)	424 (95%)	16 (4%)	5 (1%)	14	19

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	64	GLU
1	B	42	PRO

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	H	14	PRO
3	J	2	PRO
1	B	164	ALA

### 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	148/165 (90%)	141 (95%)	7 (5%)	26	41
1	B	150/165 (91%)	143 (95%)	7 (5%)	26	41
2	C	7/7 (100%)	7 (100%)	0	100	100
2	D	7/7 (100%)	7 (100%)	0	100	100
2	E	7/7 (100%)	7 (100%)	0	100	100
2	F	7/7 (100%)	7 (100%)	0	100	100
2	G	7/7 (100%)	7 (100%)	0	100	100
2	H	7/7 (100%)	7 (100%)	0	100	100
3	I	2/2 (100%)	2 (100%)	0	100	100
3	J	2/2 (100%)	2 (100%)	0	100	100
3	K	2/2 (100%)	2 (100%)	0	100	100
All	All	346/378 (92%)	332 (96%)	14 (4%)	31	48

5 of 14 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	171	TYR
1	B	4	TYR
1	B	122	SER
1	A	100	GLU
1	B	92	ASP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

42 non-standard protein/DNA/RNA residues 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	$\# Z  > 2$
3	HYP	I	-3	3	6,8,9	0.74	0	5,10,12	1.67	1 (20%)
2	HYP	H	3	2	6,8,9	0.81	0	5,10,12	2.12	1 (20%)
2	HYP	D	-3	2	6,8,9	0.69	0	5,10,12	2.09	2 (40%)
2	HYP	F	-3	2	6,8,9	0.73	0	5,10,12	2.21	2 (40%)
2	HYP	D	3	2	6,8,9	0.67	0	5,10,12	1.44	1 (20%)
2	HYP	E	9	2	6,8,9	0.62	0	5,10,12	1.70	1 (20%)
3	HYP	J	0	3	6,8,9	0.78	0	5,10,12	2.15	1 (20%)
2	HYP	C	3	2	6,8,9	0.76	0	5,10,12	1.82	1 (20%)
2	HYP	C	15	2	6,8,9	0.69	0	5,10,12	1.92	2 (40%)
2	HYP	C	-3	2	6,8,9	0.70	0	5,10,12	2.07	2 (40%)
2	HYP	F	0	2	6,8,9	0.77	0	5,10,12	1.78	1 (20%)
2	HYP	G	0	2	6,8,9	0.63	0	5,10,12	1.97	2 (40%)
2	HYP	E	0	2	6,8,9	0.76	0	5,10,12	1.68	1 (20%)
2	HYP	E	15	2	6,8,9	0.68	0	5,10,12	2.06	2 (40%)
2	HYP	F	3	2	6,8,9	0.80	0	5,10,12	2.05	1 (20%)
2	HYP	F	9	2	6,8,9	0.63	0	5,10,12	2.01	2 (40%)
2	HYP	G	12	2	6,8,9	0.56	0	5,10,12	1.98	2 (40%)
2	HYP	D	9	2	6,8,9	0.78	0	5,10,12	1.64	1 (20%)
2	HYP	H	12	2	6,8,9	0.80	0	5,10,12	2.25	1 (20%)
3	HYP	K	0	3	6,8,9	0.67	0	5,10,12	2.37	2 (40%)
2	HYP	G	-3	2	6,8,9	0.60	0	5,10,12	2.27	2 (40%)
2	HYP	E	3	2	6,8,9	1.00	0	5,10,12	1.81	3 (60%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	HYP	G	3	2	6,8,9	0.59	0	5,10,12	2.08	3 (60%)
2	HYP	C	0	2	6,8,9	0.64	0	5,10,12	2.13	1 (20%)
2	HYP	C	12	2	6,8,9	0.75	0	5,10,12	1.83	2 (40%)
2	HYP	D	12	2	6,8,9	0.74	0	5,10,12	2.40	4 (80%)
2	HYP	H	-3	2	6,8,9	0.71	0	5,10,12	1.82	1 (20%)
2	HYP	F	12	2	6,8,9	0.69	0	5,10,12	2.02	2 (40%)
2	HYP	H	15	2	6,8,9	0.77	0	5,10,12	2.30	2 (40%)
2	HYP	G	9	2	6,8,9	0.60	0	5,10,12	1.57	1 (20%)
3	HYP	K	-3	3	6,8,9	0.72	0	5,10,12	1.63	1 (20%)
2	HYP	E	-3	2	6,8,9	0.71	0	5,10,12	2.18	2 (40%)
2	HYP	H	0	2	6,8,9	0.73	0	5,10,12	1.57	1 (20%)
2	HYP	H	9	2	6,8,9	0.66	0	5,10,12	1.95	2 (40%)
2	HYP	G	15	2	6,8,9	0.69	0	5,10,12	2.12	2 (40%)
3	HYP	J	3	3	6,8,9	0.64	0	5,10,12	1.91	2 (40%)
2	HYP	E	12	2	6,8,9	0.67	0	5,10,12	1.91	2 (40%)
2	HYP	D	0	2	6,8,9	0.74	0	5,10,12	1.69	1 (20%)
3	HYP	J	-3	3	6,8,9	0.66	0	5,10,12	2.03	2 (40%)
2	HYP	D	15	2	6,8,9	0.65	0	5,10,12	2.16	2 (40%)
3	HYP	I	0	3	6,8,9	0.57	0	5,10,12	2.23	2 (40%)
2	HYP	C	9	2	6,8,9	0.73	0	5,10,12	1.87	1 (20%)

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	HYP	I	-3	3	-	0/0/11/13	0/1/1/1
2	HYP	H	3	2	-	0/0/11/13	0/1/1/1
2	HYP	D	-3	2	-	0/0/11/13	0/1/1/1
2	HYP	F	-3	2	-	0/0/11/13	0/1/1/1
2	HYP	D	3	2	-	0/0/11/13	0/1/1/1
2	HYP	E	9	2	-	0/0/11/13	0/1/1/1
3	HYP	J	0	3	-	0/0/11/13	0/1/1/1
2	HYP	C	3	2	-	0/0/11/13	0/1/1/1
2	HYP	C	15	2	-	0/0/11/13	0/1/1/1
2	HYP	C	-3	2	-	0/0/11/13	0/1/1/1
2	HYP	F	0	2	-	0/0/11/13	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	HYP	G	0	2	-	0/0/11/13	0/1/1/1
2	HYP	E	0	2	-	0/0/11/13	0/1/1/1
2	HYP	E	15	2	-	0/0/11/13	0/1/1/1
2	HYP	F	3	2	-	0/0/11/13	0/1/1/1
2	HYP	F	9	2	-	0/0/11/13	0/1/1/1
2	HYP	G	12	2	-	0/0/11/13	0/1/1/1
2	HYP	D	9	2	-	0/0/11/13	0/1/1/1
2	HYP	H	12	2	-	0/0/11/13	0/1/1/1
3	HYP	K	0	3	-	0/0/11/13	0/1/1/1
2	HYP	G	-3	2	-	0/0/11/13	0/1/1/1
2	HYP	E	3	2	-	0/0/11/13	0/1/1/1
2	HYP	G	3	2	-	0/0/11/13	0/1/1/1
2	HYP	C	0	2	-	0/0/11/13	0/1/1/1
2	HYP	C	12	2	-	0/0/11/13	0/1/1/1
2	HYP	D	12	2	-	0/0/11/13	0/1/1/1
2	HYP	H	-3	2	-	0/0/11/13	0/1/1/1
2	HYP	F	12	2	-	0/0/11/13	0/1/1/1
2	HYP	H	15	2	-	0/0/11/13	0/1/1/1
2	HYP	G	9	2	-	0/0/11/13	0/1/1/1
3	HYP	K	-3	3	-	0/0/11/13	0/1/1/1
2	HYP	E	-3	2	-	0/0/11/13	0/1/1/1
2	HYP	H	0	2	-	0/0/11/13	0/1/1/1
2	HYP	H	9	2	-	0/0/11/13	0/1/1/1
2	HYP	G	15	2	-	0/0/11/13	0/1/1/1
3	HYP	J	3	3	-	0/0/11/13	0/1/1/1
2	HYP	E	12	2	-	0/0/11/13	0/1/1/1
2	HYP	D	0	2	-	0/0/11/13	0/1/1/1
3	HYP	J	-3	3	-	0/0/11/13	0/1/1/1
2	HYP	D	15	2	-	0/0/11/13	0/1/1/1
3	HYP	I	0	3	-	0/0/11/13	0/1/1/1
2	HYP	C	9	2	-	0/0/11/13	0/1/1/1

There are no bond length outliers.

The worst 5 of 70 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	J	0	HYP	CB-CG-CD	4.05	108.23	103.27
2	E	-3	HYP	CB-CG-CD	3.99	108.16	103.27
2	D	15	HYP	CB-CG-CD	3.79	107.92	103.27
2	C	0	HYP	CB-CG-CD	3.79	107.91	103.27
2	H	3	HYP	CB-CG-CD	3.76	107.87	103.27

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

9 monomers are involved in 12 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	J	0	HYP	1	0
2	C	3	HYP	2	0
3	K	0	HYP	1	0
2	H	15	HYP	2	0
2	H	9	HYP	1	0
3	J	3	HYP	2	0
2	D	0	HYP	1	0
3	J	-3	HYP	1	0
3	I	0	HYP	1	0

## 5.5 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 5.6 Ligand geometry [i](#)

Of 12 ligands modelled in this entry, 5 are monoatomic - leaving 7 for Mogul analysis.

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	SO4	B	202	-	4,4,4	0.17	0	6,6,6	0.16	0
5	SO4	F	101	-	4,4,4	0.17	0	6,6,6	0.25	0
4	NAG	A	201	1	14,14,15	0.29	0	17,19,21	0.44	0
4	NAG	B	201	1	14,14,15	0.28	0	17,19,21	0.50	0
5	SO4	B	204	-	4,4,4	0.23	0	6,6,6	0.37	0
8	ACT	F	102	2	1,2,3	1.12	0	1,1,3	0.48	0
5	SO4	A	202	-	4,4,4	0.22	0	6,6,6	0.12	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	A	201	1	-	0/6/23/26	0/1/1/1
4	NAG	B	201	1	-	0/6/23/26	0/1/1/1

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.

No monomer is involved in short contacts.

## 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 [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	181/203 (89%)	-0.37	2 (1%) 80 78	21, 32, 46, 58	0
1	B	181/203 (89%)	-0.33	0 100 100	23, 31, 46, 66	0
2	C	15/21 (71%)	-0.27	0 100 100	24, 29, 50, 57	0
2	D	14/21 (66%)	-0.38	0 100 100	24, 30, 43, 43	0
2	E	15/21 (71%)	-0.43	0 100 100	24, 29, 50, 52	0
2	F	15/21 (71%)	-0.11	1 (6%) 17 15	24, 31, 57, 60	0
2	G	14/21 (66%)	-0.47	0 100 100	26, 31, 52, 54	0
2	H	15/21 (71%)	-0.41	0 100 100	24, 31, 42, 47	0
3	I	4/7 (57%)	-0.48	0 100 100	35, 36, 36, 39	0
3	J	4/7 (57%)	0.11	0 100 100	35, 38, 40, 43	0
3	K	4/7 (57%)	-0.42	0 100 100	33, 33, 36, 38	0
All	All	462/553 (83%)	-0.35	3 (0%) 89 88	21, 32, 47, 66	0

All (3) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	111	PRO	2.4
2	F	13	GLY	2.2
1	A	192	ALA	2.0

### 6.2 Non-standard residues in protein, DNA, RNA chains [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	HYP	H	15	8/9	0.76	0.19	37,51,55,60	0
2	HYP	C	15	8/9	0.82	0.19	50,57,59,62	0
2	HYP	G	15	8/9	0.85	0.16	50,53,58,74	0
3	HYP	J	3	8/9	0.87	0.16	40,43,46,47	0
2	HYP	H	-3	8/9	0.90	0.13	32,34,40,44	0
3	HYP	I	0	8/9	0.90	0.17	32,33,40,43	0
2	HYP	D	15	8/9	0.91	0.15	46,48,51,59	0
2	HYP	E	12	8/9	0.92	0.14	43,46,48,50	0
2	HYP	E	15	8/9	0.92	0.18	41,43,49,49	0
2	HYP	C	0	8/9	0.92	0.11	31,36,37,43	0
3	HYP	J	-3	8/9	0.93	0.11	29,34,37,38	0
2	HYP	G	0	8/9	0.93	0.11	32,33,36,38	0
2	HYP	F	12	8/9	0.93	0.11	41,46,49,53	0
2	HYP	G	12	8/9	0.94	0.15	40,44,50,51	0
2	HYP	G	-3	8/9	0.94	0.12	34,36,38,39	0
2	HYP	D	3	8/9	0.94	0.12	27,28,30,30	0
2	HYP	H	0	8/9	0.94	0.12	28,34,35,36	0
2	HYP	D	12	8/9	0.94	0.11	33,39,39,45	0
2	HYP	E	-3	8/9	0.95	0.12	32,34,36,38	0
2	HYP	H	12	8/9	0.95	0.12	30,35,38,39	0
2	HYP	H	9	8/9	0.95	0.12	29,30,31,32	0
2	HYP	F	0	8/9	0.95	0.14	26,28,30,31	0
2	HYP	F	-3	8/9	0.95	0.12	31,33,34,37	0
3	HYP	I	-3	8/9	0.95	0.15	34,36,38,40	0
2	HYP	D	0	8/9	0.95	0.12	30,33,33,39	0
2	HYP	F	3	8/9	0.95	0.12	21,24,25,26	0
2	HYP	F	9	8/9	0.95	0.17	30,30,33,37	0
2	HYP	D	-3	8/9	0.95	0.12	34,34,36,38	0
2	HYP	G	9	8/9	0.96	0.12	31,32,33,34	0
3	HYP	K	-3	8/9	0.96	0.09	32,36,38,38	0
3	HYP	J	0	8/9	0.96	0.15	35,37,40,40	0
3	HYP	K	0	8/9	0.96	0.11	36,40,42,42	0
2	HYP	E	9	8/9	0.96	0.12	28,31,33,34	0
2	HYP	C	-3	8/9	0.96	0.08	35,38,40,41	0
2	HYP	C	3	8/9	0.97	0.12	23,25,27,28	0
2	HYP	E	3	8/9	0.97	0.14	23,23,25,25	0
2	HYP	G	3	8/9	0.97	0.13	25,25,29,30	0
2	HYP	E	0	8/9	0.97	0.10	27,29,32,32	0
2	HYP	C	9	8/9	0.97	0.08	23,27,31,31	0
2	HYP	C	12	8/9	0.98	0.10	36,40,42,42	0
2	HYP	H	3	8/9	0.98	0.10	26,28,30,35	0
2	HYP	D	9	8/9	0.98	0.11	22,25,28,33	0

### 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
6	CL	I	101	1/1	0.83	0.15	58,58,58,58	0
6	CL	A	203	1/1	0.89	0.11	57,57,57,57	0
7	NA	B	203	1/1	0.91	0.08	38,38,38,38	0
5	SO4	F	101	5/5	0.91	0.29	40,43,57,64	0
5	SO4	B	202	5/5	0.92	0.21	39,45,58,60	0
6	CL	A	204	1/1	0.92	0.10	47,47,47,47	0
4	NAG	B	201	14/15	0.93	0.16	27,30,32,33	0
5	SO4	B	204	5/5	0.94	0.15	38,38,54,62	0
4	NAG	A	201	14/15	0.94	0.11	24,30,31,33	0
5	SO4	A	202	5/5	0.95	0.14	36,39,46,52	0
7	NA	B	205	1/1	0.96	0.11	32,32,32,32	0
8	ACT	F	102	3/4	0.97	0.13	30,30,31,33	0

### 6.5 Other polymers [i](#)

There are no such residues in this entry.