



# Full wwPDB X-ray Structure Validation Report ⓘ

Jan 31, 2016 – 09:46 PM GMT

PDB ID : 1QH1  
Title : NITROGENASE MOFE PROTEIN FROM KLEBSIELLA PNEUMONIAE,  
PHENOSAFRANIN OXIDIZED STATE  
Authors : Mayer, S.M.; Lawson, D.M.; Gormal, C.A.; Roe, S.M.; Smith, B.E.  
Deposited on : 1999-05-10  
Resolution : 1.60 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  
<http://wwpdb.org/validation/2016/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 : 1.7 (RC4), CSD as536be (2015)  
Xtriage (Phenix) : 1.9-1692  
EDS : rb-20026688  
Percentile statistics : 20151230.v01 (using entries in the PDB archive December 30th 2015)  
Refmac : 5.8.0135  
CCP4 : 6.5.0  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk26865

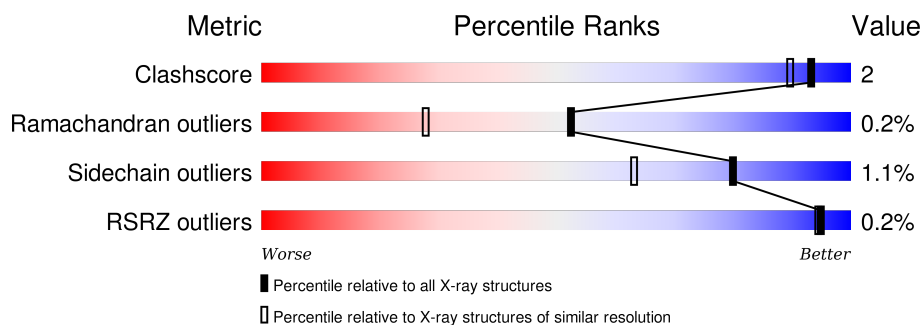
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.60 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))
Clashscore	102246	2732 (1.60-1.60)
Ramachandran outliers	100387	2654 (1.60-1.60)
Sidechain outliers	100360	2653 (1.60-1.60)
RSRZ outliers	91569	2479 (1.60-1.60)

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. 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	478	 88% 12%
1	C	478	 88% 11% .
2	B	519	 89% 10%
2	D	519	 92% 8% .

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
8	EDO	A	2774	-	-	-	X
8	EDO	A	2775	-	-	-	X
8	EDO	A	2777	-	-	-	X
8	EDO	B	2778	-	-	-	X
8	EDO	B	2779	-	-	-	X
8	EDO	B	2781	-	-	-	X
8	EDO	D	2773	-	-	-	X
8	EDO	D	2780	-	-	-	X
8	EDO	D	2784	-	-	-	X

## 2 Entry composition

There are 9 unique types of molecules in this entry. The entry contains 18303 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 PROTEIN (NITROGENASE MOLYBDENUM IRON PROTEIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	478	Total	C	N	O	S	0	3	0
			3698	2355	630	690	23			
1	C	478	Total	C	N	O	S	0	1	0
			3706	2360	635	687	24			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	85	VAL	ALA	CONFLICT	UNP P00466
A	94	GLY	GLU	CONFLICT	UNP P09772
C	85	VAL	ALA	CONFLICT	UNP P00466
C	94	GLY	GLU	CONFLICT	UNP P09772

- Molecule 2 is a protein called PROTEIN (NITROGENASE MOLYBDENUM IRON PROTEIN).

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	519	Total	C	N	O	S	0	10	0
			4067	2601	675	760	31			
2	D	519	Total	C	N	O	S	0	3	0
			4057	2588	678	762	29			

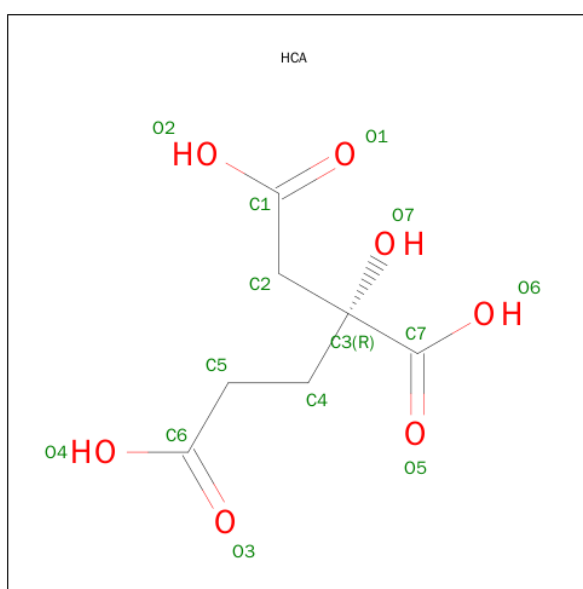
- Molecule 3 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	B	1	Total	Cl	0	0
			1	1		
3	D	1	Total	Cl	0	0
			1	1		

- Molecule 4 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

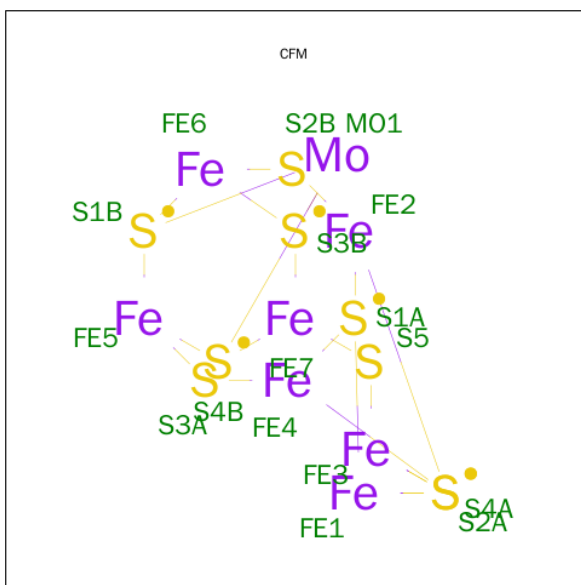
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	3	Total Mg 3 3	0	0
4	A	1	Total Mg 1 1	0	0
4	D	2	Total Mg 2 2	0	0
4	C	1	Total Mg 1 1	0	0

- Molecule 5 is 3-HYDROXY-3-CARBOXY-ADIPIC ACID (three-letter code: HCA) (formula:  $C_7H_{10}O_7$ ).



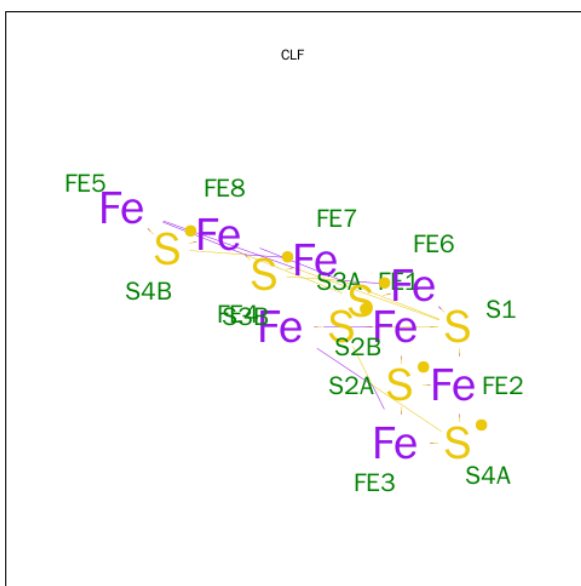
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total C O 14 7 7	0	0
5	C	1	Total C O 14 7 7	0	0

- Molecule 6 is FE-MO-S CLUSTER (three-letter code: CFM) (formula:  $Fe_7MoS_9$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
6	A	1	Total 17	Fe 7	Mo 1	S 9	0	0
6	C	1	Total 17	Fe 7	Mo 1	S 9	0	0

- Molecule 7 is FE(8)-S(7) CLUSTER (three-letter code: CLF) (formula:  $\text{Fe}_8\text{S}_7$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
7	A	1	Total 15	Fe 8	S 7	0	0
7	C	1	Total 15	Fe 8	S 7	0	0

- Molecule 8 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $C_2H_6O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	D	1	Total	C	O	0	0
			4	2	2		
8	A	1	Total	C	O	0	0
			4	2	2		
8	A	1	Total	C	O	0	0
			4	2	2		
8	A	1	Total	C	O	0	0
			4	2	2		
8	B	1	Total	C	O	0	0
			4	2	2		
8	B	1	Total	C	O	0	0
			4	2	2		
8	D	1	Total	C	O	0	0
			4	2	2		
8	B	1	Total	C	O	0	0
			4	2	2		
8	C	1	Total	C	O	0	0
			4	2	2		
8	D	1	Total	C	O	0	0
			4	2	2		
8	D	1	Total	C	O	0	0
			4	2	2		
8	D	1	Total	C	O	0	0
			4	2	2		

- Molecule 9 is water.

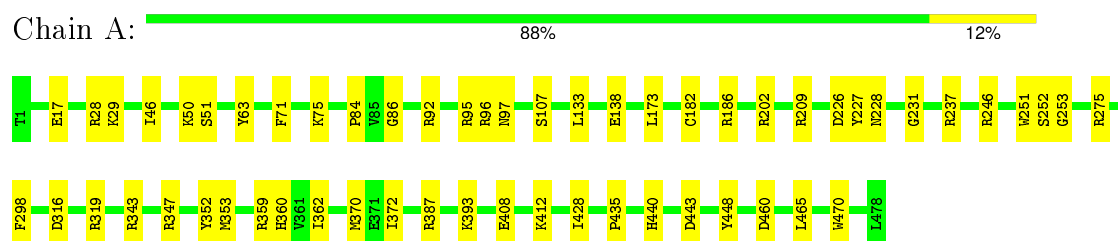
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
9	A	589	Total 589	O 589	0	0
9	B	737	Total 737	O 737	0	0
9	C	538	Total 538	O 538	0	0
9	D	758	Total 758	O 758	0	0



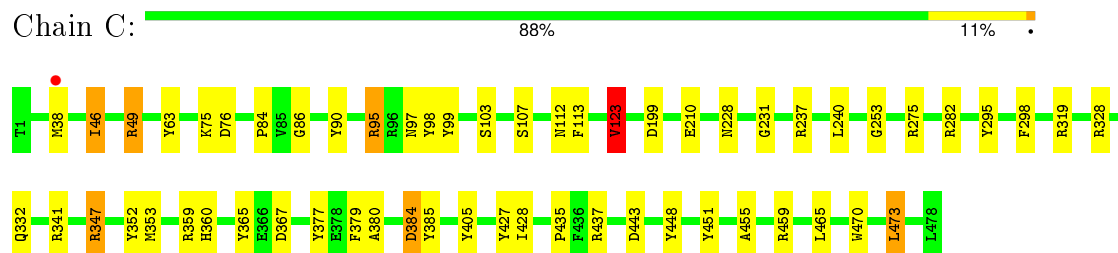
### 3 Residue-property plots

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.

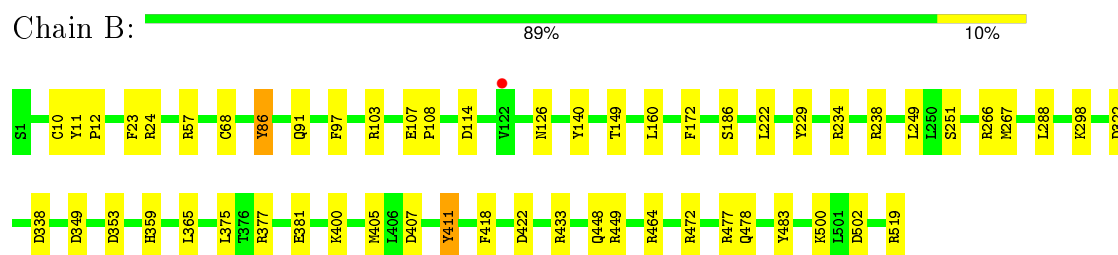
- Molecule 1: PROTEIN (NITROGENASE MOLYBDENUM IRON PROTEIN)



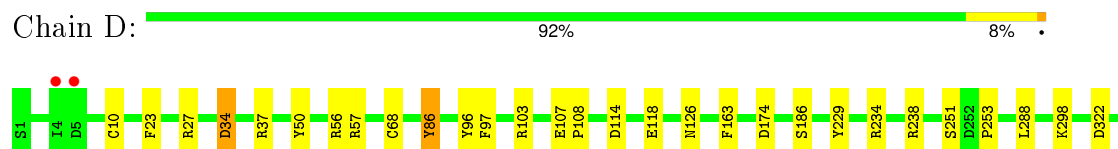
- Molecule 1: PROTEIN (NITROGENASE MOLYBDENUM IRON PROTEIN)



- Molecule 2: PROTEIN (NITROGENASE MOLYBDENUM IRON PROTEIN)



- Molecule 2: PROTEIN (NITROGENASE MOLYBDENUM IRON PROTEIN)



R344	G345	D353	F371	R427	Q434	F446	R449	F458	R464	Q478	Y493	D512	F515	R519

## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	204.18Å 75.22Å 162.84Å 90.00° 122.88° 90.00°	Depositor
Resolution (Å)	50.00 – 1.60 136.75 – 1.60	Depositor EDS
% Data completeness (in resolution range)	98.9 (50.00-1.60) 98.7 (136.75-1.60)	Depositor EDS
$R_{merge}$	0.05	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.91 (at 1.60Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.158 , 0.199 0.154 , (Not available)	Depositor DCC
$R_{free}$ test set	No test flags present.	DCC
Wilson B-factor (Å <sup>2</sup> )	15.0	Xtriage
Anisotropy	0.184	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 54.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 271201 reflections	Xtriage
$F_o, F_c$ correlation	0.97	EDS
Total number of atoms	18303	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	20.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.38% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: MG, CFM, CL, CLF, EDO, HCA

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	0.71	0/3795	1.38	35/5142 (0.7%)
1	C	0.70	0/3795	1.40	46/5138 (0.9%)
2	B	0.73	0/4210	1.39	44/5714 (0.8%)
2	D	0.75	0/4166	1.37	44/5653 (0.8%)
All	All	0.72	0/15966	1.39	169/21647 (0.8%)

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
1	A	0	3
1	C	0	6
All	All	0	9

There are no bond length outliers.

All (169) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	407	ASP	CB-CG-OD2	17.77	134.29	118.30
1	C	49	ARG	NE-CZ-NH2	-14.28	113.16	120.30
2	B	449	ARG	CD-NE-CZ	14.08	143.31	123.60
2	B	103	ARG	NE-CZ-NH2	-13.49	113.55	120.30
2	B	103	ARG	NE-CZ-NH1	12.27	126.44	120.30
2	D	103	ARG	NE-CZ-NH2	-12.14	114.23	120.30
1	C	352	TYR	CB-CG-CD2	11.97	128.19	121.00
2	D	37	ARG	NE-CZ-NH2	-11.09	114.76	120.30
1	A	275	ARG	NE-CZ-NH2	-10.79	114.90	120.30
1	C	359	ARG	NE-CZ-NH2	10.79	125.70	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	103	ARG	NE-CZ-NH1	10.71	125.65	120.30
2	D	27	ARG	CD-NE-CZ	10.27	137.98	123.60
1	A	95	ARG	NE-CZ-NH2	-10.01	115.30	120.30
1	A	202	ARG	NE-CZ-NH2	-9.95	115.32	120.30
2	B	377	ARG	NE-CZ-NH2	-9.48	115.56	120.30
1	A	63	TYR	CB-CG-CD1	-9.47	115.32	121.00
1	A	319	ARG	NE-CZ-NH2	-8.98	115.81	120.30
2	D	519	ARG	NE-CZ-NH1	8.98	124.79	120.30
2	D	234	ARG	NE-CZ-NH2	-8.86	115.87	120.30
2	D	434	GLN	CG-CD-OE1	-8.75	104.10	121.60
1	A	353	MET	CA-CB-CG	8.58	127.88	113.30
1	C	385	TYR	CB-CG-CD1	-8.54	115.87	121.00
2	B	411	TYR	CB-CG-CD1	-8.51	115.89	121.00
2	B	472	ARG	NE-CZ-NH2	-8.46	116.07	120.30
2	D	464	ARG	NE-CZ-NH1	8.43	124.52	120.30
2	D	27	ARG	NE-CZ-NH2	-8.34	116.13	120.30
2	B	349	ASP	CB-CG-OD2	-8.29	110.84	118.30
2	D	449	ARG	NE-CZ-NH1	8.29	124.44	120.30
1	A	237	ARG	NE-CZ-NH2	-8.28	116.16	120.30
2	B	238	ARG	NE-CZ-NH1	8.25	124.42	120.30
1	A	387	ARG	NE-CZ-NH2	-8.24	116.18	120.30
1	A	237	ARG	NE-CZ-NH1	8.12	124.36	120.30
2	D	449	ARG	NE-CZ-NH2	-8.12	116.24	120.30
1	A	352	TYR	CB-CG-CD1	-8.11	116.14	121.00
1	A	352	TYR	CB-CG-CD2	8.04	125.83	121.00
2	D	427	ARG	NE-CZ-NH2	-7.91	116.35	120.30
2	B	57	ARG	NE-CZ-NH1	7.88	124.24	120.30
1	C	63	TYR	CB-CG-CD1	-7.87	116.28	121.00
2	D	353	ASP	CB-CG-OD1	7.77	125.29	118.30
2	D	483	TYR	CB-CG-CD2	-7.63	116.42	121.00
2	D	353	ASP	CB-CG-OD2	-7.39	111.65	118.30
1	A	343	ARG	NE-CZ-NH1	7.36	123.98	120.30
2	D	449	ARG	CD-NE-CZ	7.35	133.89	123.60
1	C	237	ARG	NE-CZ-NH1	7.33	123.97	120.30
2	D	434	GLN	CA-CB-CG	-7.25	97.45	113.40
2	B	234	ARG	NE-CZ-NH1	7.20	123.90	120.30
1	C	49	ARG	NE-CZ-NH1	7.19	123.89	120.30
1	C	63	TYR	CB-CG-CD2	7.18	125.31	121.00
1	C	437	ARG	NE-CZ-NH1	7.16	123.88	120.30
1	C	359	ARG	CD-NE-CZ	7.14	133.59	123.60
1	C	459	ARG	NE-CZ-NH1	7.12	123.86	120.30
2	B	266	ARG	NE-CZ-NH2	7.06	123.83	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	57	ARG	NE-CZ-NH2	-7.01	116.79	120.30
2	D	114	ASP	CB-CG-OD2	6.99	124.59	118.30
1	C	90	TYR	CB-CG-CD2	6.94	125.16	121.00
2	D	238	ARG	NE-CZ-NH2	-6.92	116.84	120.30
1	C	328	ARG	NE-CZ-NH1	6.91	123.76	120.30
2	B	483	TYR	CB-CG-CD2	-6.86	116.89	121.00
2	B	234	ARG	NE-CZ-NH2	-6.85	116.87	120.30
2	D	234	ARG	NE-CZ-NH1	6.78	123.69	120.30
1	C	405	TYR	CB-CG-CD2	-6.77	116.94	121.00
2	B	238	ARG	NE-CZ-NH2	-6.71	116.94	120.30
2	B	349	ASP	CB-CG-OD1	6.70	124.33	118.30
2	B	107	GLU	OE1-CD-OE2	-6.65	115.32	123.30
2	B	97	PHE	CB-CG-CD2	-6.61	116.17	120.80
2	B	24	ARG	NE-CZ-NH2	-6.60	117.00	120.30
2	B	229	TYR	CB-CG-CD2	6.57	124.94	121.00
1	C	384	ASP	CB-CG-OD2	6.57	124.21	118.30
1	A	92	ARG	NE-CZ-NH2	-6.51	117.04	120.30
2	B	23	PHE	CB-CG-CD2	-6.49	116.26	120.80
1	C	367	ASP	CB-CG-OD2	6.47	124.12	118.30
1	A	316	ASP	CB-CG-OD1	6.43	124.09	118.30
2	D	322	ASP	CB-CG-OD1	6.43	124.08	118.30
2	D	464	ARG	NE-CZ-NH2	-6.42	117.09	120.30
1	A	246	ARG	NE-CZ-NH2	-6.42	117.09	120.30
2	B	86	TYR	CB-CG-CD2	-6.41	117.15	121.00
1	A	86	GLY	O-C-N	-6.39	112.48	122.70
2	B	472	ARG	NE-CZ-NH1	6.38	123.49	120.30
1	C	405	TYR	CB-CG-CD1	6.38	124.83	121.00
2	B	353	ASP	CB-CG-OD1	6.30	123.97	118.30
2	B	422	ASP	CB-CG-OD1	6.29	123.96	118.30
2	B	11	TYR	CB-CG-CD1	-6.28	117.23	121.00
2	B	114	ASP	CB-CG-OD1	6.27	123.94	118.30
2	B	433	ARG	NE-CZ-NH2	-6.26	117.17	120.30
1	C	275	ARG	NE-CZ-NH2	-6.25	117.17	120.30
1	A	448	TYR	CB-CG-CD2	-6.17	117.30	121.00
1	A	319	ARG	CD-NE-CZ	6.16	132.22	123.60
2	D	434	GLN	OE1-CD-NE2	6.11	135.96	121.90
1	A	316	ASP	CB-CG-OD2	-6.11	112.81	118.30
2	D	344	ARG	NE-CZ-NH2	-6.09	117.25	120.30
2	B	407	ASP	OD1-CG-OD2	-6.04	111.81	123.30
1	C	451	TYR	CB-CG-CD2	-6.03	117.38	121.00
1	C	448	TYR	CB-CG-CD1	6.02	124.61	121.00
2	D	446	PHE	CB-CG-CD2	-6.01	116.59	120.80

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	328	ARG	NE-CZ-NH2	-5.99	117.31	120.30
1	C	319	ARG	NE-CZ-NH2	5.99	123.29	120.30
2	D	238	ARG	NE-CZ-NH1	5.99	123.29	120.30
1	C	90	TYR	CB-CG-CD1	-5.98	117.41	121.00
2	D	96	TYR	CB-CG-CD2	-5.98	117.41	121.00
2	B	519	ARG	NE-CZ-NH2	-5.97	117.31	120.30
1	A	71	PHE	CB-CG-CD2	-5.94	116.64	120.80
1	A	209	ARG	NE-CZ-NH1	5.94	123.27	120.30
1	A	227	TYR	CB-CG-CD1	-5.89	117.47	121.00
2	B	229	TYR	CB-CG-CD1	-5.88	117.47	121.00
2	B	338	ASP	CB-CG-OD2	-5.88	113.01	118.30
2	B	322	ASP	CB-CG-OD1	5.87	123.59	118.30
2	B	97	PHE	CB-CG-CD1	5.84	124.89	120.80
1	A	28	ARG	NE-CZ-NH1	5.84	123.22	120.30
2	D	27	ARG	NE-CZ-NH1	5.82	123.21	120.30
1	C	377	TYR	CA-CB-CG	5.80	124.42	113.40
2	D	163	PHE	CB-CG-CD1	5.78	124.85	120.80
1	C	451	TYR	CB-CG-CD1	5.77	124.46	121.00
1	C	352	TYR	CG-CD2-CE2	5.75	125.90	121.30
1	C	352	TYR	CB-CG-CD1	-5.75	117.55	121.00
1	C	353	MET	CA-CB-CG	5.69	122.97	113.30
2	B	464	ARG	NE-CZ-NH1	5.67	123.14	120.30
2	D	458	PHE	CB-CG-CD2	-5.67	116.83	120.80
2	D	86	TYR	CB-CG-CD2	-5.67	117.60	121.00
2	B	172	PHE	CB-CG-CD2	-5.66	116.84	120.80
1	A	28	ARG	CD-NE-CZ	5.64	131.50	123.60
1	C	113	PHE	CB-CG-CD2	-5.64	116.85	120.80
1	C	95	ARG	NE-CZ-NH2	-5.64	117.48	120.30
1	A	95	ARG	NE-CZ-NH1	5.64	123.12	120.30
2	B	338	ASP	CB-CG-OD1	5.63	123.37	118.30
1	C	98	TYR	CB-CG-CD2	-5.63	117.62	121.00
2	D	478	GLN	CG-CD-OE1	5.62	132.84	121.60
1	C	341	ARG	NE-CZ-NH1	5.62	123.11	120.30
2	D	371	PHE	CB-CG-CD2	-5.59	116.88	120.80
2	D	229	TYR	CB-CG-CD2	-5.54	117.67	121.00
1	A	226	ASP	CB-CG-OD1	-5.53	113.33	118.30
1	C	76	ASP	CB-CG-OD1	5.52	123.27	118.30
1	C	86	GLY	O-C-N	-5.51	113.88	122.70
1	A	50	LYS	C-N-CA	-5.51	107.94	121.70
1	A	319	ARG	NE-CZ-NH1	5.49	123.04	120.30
1	C	282	ARG	NE-CZ-NH1	-5.49	117.56	120.30
1	C	199	ASP	CB-CG-OD2	-5.48	113.36	118.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	359	ARG	NE-CZ-NH1	-5.47	117.56	120.30
1	C	95	ARG	NE-CZ-NH1	5.43	123.02	120.30
1	C	385	TYR	CG-CD1-CE1	-5.38	116.99	121.30
1	C	328	ARG	CD-NE-CZ	5.38	131.13	123.60
1	A	96	ARG	NE-CZ-NH1	5.36	122.98	120.30
1	C	473	LEU	CA-CB-CG	5.36	127.62	115.30
2	D	23	PHE	CB-CG-CD2	-5.36	117.05	120.80
2	D	50	TYR	CB-CG-CD2	5.35	124.21	121.00
2	B	11	TYR	CB-CG-CD2	5.31	124.19	121.00
2	D	97	PHE	CB-CG-CD2	-5.30	117.09	120.80
1	C	365	TYR	CB-CG-CD2	-5.27	117.84	121.00
2	D	519	ARG	NE-CZ-NH2	-5.25	117.67	120.30
1	A	95	ARG	CG-CD-NE	5.25	122.83	111.80
1	A	186	ARG	CD-NE-CZ	5.23	130.92	123.60
1	A	460	ASP	CB-CG-OD2	-5.21	113.61	118.30
1	A	50	LYS	CA-C-N	-5.20	105.76	117.20
2	B	86	TYR	CG-CD2-CE2	-5.20	117.14	121.30
2	D	346	ARG	NE-CZ-NH2	-5.20	117.70	120.30
1	C	295	TYR	CB-CG-CD2	-5.19	117.88	121.00
2	B	405	MET	O-C-N	-5.18	114.41	122.70
2	D	512	ASP	CB-CG-OD1	5.17	122.96	118.30
2	D	515	PHE	CB-CG-CD2	-5.17	117.18	120.80
2	B	478	GLN	CG-CD-OE1	5.16	131.92	121.60
1	C	298	PHE	CB-CG-CD1	-5.15	117.20	120.80
1	C	379	PHE	CB-CG-CD2	-5.15	117.20	120.80
1	A	298	PHE	CB-CG-CD1	-5.12	117.21	120.80
2	D	107	GLU	OE1-CD-OE2	-5.12	117.16	123.30
1	C	103	SER	N-CA-CB	-5.12	102.83	110.50
2	D	174	ASP	CB-CG-OD2	5.11	122.90	118.30
1	C	123	VAL	CA-CB-CG1	5.09	118.54	110.90
2	B	477	ARG	NE-CZ-NH1	5.09	122.84	120.30
2	B	502	ASP	CB-CG-OD1	5.04	122.84	118.30
2	B	418	PHE	CB-CG-CD2	-5.02	117.29	120.80

There are no chirality outliers.

All (9) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	17	GLU	Mainchain
1	A	182	CYS	Mainchain
1	A	231	GLY	Mainchain
1	C	123	VAL	Mainchain

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	C	231	GLY	Mainchain
1	C	240	LEU	Mainchain
1	C	427	TYR	Mainchain
1	C	46	ILE	Mainchain
1	C	95	ARG	Sidechain

## 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	3698	0	3554	16	0
1	C	3706	0	3581	15	0
2	B	4067	0	3912	18	0
2	D	4057	0	3919	8	0
3	B	1	0	0	0	0
3	D	1	0	0	0	0
4	A	1	0	0	0	0
4	B	3	0	0	0	0
4	C	1	0	0	0	0
4	D	2	0	0	0	0
5	A	14	0	6	2	0
5	C	14	0	6	1	0
6	A	17	0	0	0	0
6	C	17	0	0	0	0
7	A	15	0	0	1	0
7	C	15	0	0	1	0
8	A	16	0	24	0	0
8	B	12	0	18	0	0
8	C	4	0	6	0	0
8	D	20	0	30	3	0
9	A	589	0	0	1	0
9	B	737	0	0	6	0
9	C	538	0	0	5	0
9	D	758	0	0	1	0
All	All	18303	0	15056	57	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 (57) 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)
2:B:222:LEU:HB2	2:B:249:LEU:HD23	1.68	0.76
1:C:347:ARG:HG3	1:C:465:LEU:HD21	1.69	0.75
1:C:210:GLU:CG	1:C:210:GLU:OE2	2.37	0.73
1:A:428:ILE:HG13	2:B:108:PRO:HB3	1.79	0.65
1:A:408:GLU:HG2	1:A:412:LYS:HE3	1.79	0.65
2:D:515:PHE:CE1	8:D:2780:EDO:H22	2.36	0.61
2:B:365[A]:LEU:HD13	2:B:375:LEU:HD23	1.83	0.60
2:D:515:PHE:HE1	8:D:2780:EDO:H22	1.66	0.60
5:A:501:HCA:O7	5:A:501:HCA:O2	2.19	0.59
2:B:140:TYR:CD1	2:B:267[B]:MET:HE1	2.40	0.57
1:A:347:ARG:HG3	1:A:465:LEU:HD21	1.85	0.57
2:B:140:TYR:HD1	2:B:267[B]:MET:HE1	1.70	0.56
1:A:84:PRO:HB2	7:A:505:CLF:S2B	2.46	0.56
2:B:68:CYS:HB2	2:B:186:SER:HB2	1.88	0.55
2:B:365[A]:LEU:CD1	2:B:375:LEU:HD23	2.38	0.53
2:B:448:GLN:HG2	9:B:3684:HOH:O	2.09	0.53
1:A:393:LYS:HE3	9:A:3556:HOH:O	2.09	0.52
2:B:400:LYS:HE2	9:B:3538:HOH:O	2.09	0.52
2:D:68:CYS:HB2	2:D:186:SER:HB2	1.91	0.51
1:C:332:GLN:HG3	9:C:2528:HOH:O	2.11	0.51
2:B:500:LYS:HE3	9:B:3511:HOH:O	2.10	0.50
2:B:149:THR:HG23	2:B:160:LEU:HD11	1.93	0.50
1:A:29:LYS:HB3	1:A:46:ILE:HG12	1.93	0.50
1:C:455:ALA:HB2	9:C:1763:HOH:O	2.12	0.50
5:C:501:HCA:O2	5:C:501:HCA:O7	2.30	0.49
1:C:332:GLN:NE2	9:C:1763:HOH:O	2.45	0.49
1:C:435:PRO:HA	1:C:470:TRP:CZ2	2.48	0.48
1:C:46:ILE:HD12	1:C:49:ARG:HD2	1.95	0.48
1:C:99:TYR:HE2	9:C:1653:HOH:O	1.97	0.48
9:C:1571:HOH:O	2:D:118:GLU:HG3	2.14	0.48
2:D:34:ASP:HB3	9:D:3613:HOH:O	2.13	0.48
2:D:68:CYS:HB2	2:D:186:SER:CB	2.44	0.48
2:B:359:HIS:CE1	2:B:411:TYR:OH	2.67	0.48
1:C:428:ILE:HG13	2:D:108:PRO:HB3	1.97	0.47
2:D:288:LEU:HD11	2:D:298:LYS:HA	1.96	0.47
2:B:359:HIS:HE1	2:B:411:TYR:OH	1.97	0.46
1:C:380:ALA:HB1	1:C:384:ASP:HB2	1.97	0.45
2:B:126:ASN:HB2	9:B:3278:HOH:O	2.17	0.45
2:B:288:LEU:HD11	2:B:298:LYS:HA	1.99	0.45
1:A:51[B]:SER:OG	2:B:91:GLN:NE2	2.37	0.45

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:435:PRO:HA	1:A:470:TRP:CZ2	2.51	0.45
1:A:75:LYS:O	1:A:107[B]:SER:HA	2.16	0.45
1:C:38:MET:HB3	1:C:38:MET:HE2	1.69	0.45
1:A:362:ILE:HG12	1:A:372:ILE:HG13	1.99	0.44
9:B:3684:HOH:O	8:D:2780:EDO:H21	2.17	0.43
1:C:228:ASN:HB2	1:C:253:GLY:HA3	2.00	0.42
2:B:68:CYS:HB2	2:B:186:SER:CB	2.49	0.42
1:A:251:TRP:HA	1:A:252:SER:HA	1.79	0.42
2:B:381[B]:GLU:HG3	9:B:3435:HOH:O	2.19	0.42
1:C:46:ILE:HD12	1:C:49:ARG:CD	2.50	0.41
1:A:138:GLU:HG2	1:A:173:LEU:HD22	2.01	0.41
1:A:133:LEU:C	1:A:133:LEU:HD23	2.40	0.41
1:A:228:ASN:HB2	1:A:253:GLY:HA3	2.03	0.41
1:A:440:HIS:HB3	5:A:501:HCA:O5	2.21	0.40
1:C:75:LYS:O	1:C:107[B]:SER:HA	2.20	0.40
1:C:84:PRO:HB2	7:C:505:CLF:S2B	2.61	0.40
1:A:347:ARG:HB2	1:A:370:MET:CE	2.52	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	479/478 (100%)	461 (96%)	18 (4%)	0	100	100
1	C	477/478 (100%)	460 (96%)	17 (4%)	0	100	100
2	B	527/519 (102%)	516 (98%)	9 (2%)	2 (0%)	39	17
2	D	520/519 (100%)	508 (98%)	10 (2%)	2 (0%)	39	17
All	All	2003/1994 (100%)	1945 (97%)	54 (3%)	4 (0%)	52	28

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	D	251	SER
2	B	251	SER
2	B	10	CYS
2	D	10	CYS

### 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	376/397 (95%)	373 (99%)	3 (1%)	86	75
1	C	377/397 (95%)	370 (98%)	7 (2%)	65	39
2	B	430/442 (97%)	428 (100%)	2 (0%)	92	85
2	D	429/442 (97%)	423 (99%)	6 (1%)	74	53
All	All	1612/1678 (96%)	1594 (99%)	18 (1%)	80	63

All (18) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	97	ASN
1	A	360	HIS
1	A	443	ASP
2	B	12	PRO
2	B	86	TYR
1	C	97	ASN
1	C	112	ASN
1	C	123	VAL
1	C	347	ARG
1	C	360	HIS
1	C	443	ASP
1	C	473	LEU
2	D	34	ASP
2	D	56	ARG
2	D	86	TYR
2	D	126	ASN
2	D	253	PRO
2	D	434	GLN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (8) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	118	GLN
2	B	91	GLN
2	B	264	HIS
2	B	359	HIS
1	C	112	ASN
1	C	118	GLN
2	D	91	GLN
2	D	359	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 28 ligands modelled in this entry, 9 are monoatomic - leaving 19 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	$\# Z  > 2$
8	EDO	A	2774	-	3,3,3	0.47	0	2,2,2	0.62	0
8	EDO	A	2775	-	3,3,3	0.48	0	2,2,2	0.52	0
8	EDO	A	2776	-	3,3,3	0.50	0	2,2,2	0.96	0
8	EDO	A	2777	-	3,3,3	0.34	0	2,2,2	0.61	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
5	HCA	A	501	-	4,13,13	3.37	2 (50%)	3,18,18	2.57	2 (66%)
6	CFM	A	503	1	0,24,24	0.00	-	0,45,45	0.00	-
7	CLF	A	505	1,2	0,24,24	0.00	-	0,57,57	0.00	-
8	EDO	B	2778	-	3,3,3	0.45	0	2,2,2	0.48	0
8	EDO	B	2779	-	3,3,3	0.63	0	2,2,2	0.42	0
8	EDO	B	2781	-	3,3,3	0.54	0	2,2,2	0.18	0
8	EDO	C	2782	-	3,3,3	0.33	0	2,2,2	0.58	0
5	HCA	C	501	-	4,13,13	3.36	2 (50%)	3,18,18	2.41	2 (66%)
6	CFM	C	503	1	0,24,24	0.00	-	0,45,45	0.00	-
7	CLF	C	505	1,2	0,24,24	0.00	-	0,57,57	0.00	-
8	EDO	D	2773	-	3,3,3	0.43	0	2,2,2	0.61	0
8	EDO	D	2780	-	3,3,3	0.68	0	2,2,2	0.97	0
8	EDO	D	2783	-	3,3,3	0.67	0	2,2,2	0.33	0
8	EDO	D	2784	-	3,3,3	0.72	0	2,2,2	0.22	0
8	EDO	D	2785	-	3,3,3	0.42	0	2,2,2	0.20	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
8	EDO	A	2774	-	-	0/1/1/1	0/0/0/0
8	EDO	A	2775	-	-	0/1/1/1	0/0/0/0
8	EDO	A	2776	-	-	0/1/1/1	0/0/0/0
8	EDO	A	2777	-	-	0/1/1/1	0/0/0/0
5	HCA	A	501	-	-	0/7/17/17	0/0/0/0
6	CFM	A	503	1	-	0/0/84/84	0/0/8/8
7	CLF	A	505	1,2	-	0/0/132/132	0/12/10/10
8	EDO	B	2778	-	-	0/1/1/1	0/0/0/0
8	EDO	B	2779	-	-	0/1/1/1	0/0/0/0
8	EDO	B	2781	-	-	0/1/1/1	0/0/0/0
8	EDO	C	2782	-	-	0/1/1/1	0/0/0/0
5	HCA	C	501	-	-	0/7/17/17	0/0/0/0
6	CFM	C	503	1	-	0/0/84/84	0/0/8/8
7	CLF	C	505	1,2	-	0/0/132/132	0/12/10/10
8	EDO	D	2773	-	-	0/1/1/1	0/0/0/0
8	EDO	D	2780	-	-	0/1/1/1	0/0/0/0
8	EDO	D	2783	-	-	0/1/1/1	0/0/0/0
8	EDO	D	2784	-	-	0/1/1/1	0/0/0/0
8	EDO	D	2785	-	-	0/1/1/1	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	A	501	HCA	C2-C3	-4.75	1.47	1.54
5	C	501	HCA	C2-C3	-4.17	1.48	1.54
5	A	501	HCA	O7-C3	4.40	1.50	1.43
5	C	501	HCA	O7-C3	5.17	1.51	1.43

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	A	501	HCA	C3-C2-C1	-2.06	111.66	114.96
5	C	501	HCA	C4-C5-C6	2.14	120.95	112.10
5	C	501	HCA	C4-C3-C2	3.16	121.92	111.33
5	A	501	HCA	C4-C3-C2	3.48	123.00	111.33

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5 monomers are involved in 8 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	501	HCA	2	0
7	A	505	CLF	1	0
5	C	501	HCA	1	0
7	C	505	CLF	1	0
8	D	2780	EDO	3	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	478/478 (100%)	-0.56	0 100 100	11, 18, 30, 40	0
1	C	478/478 (100%)	-0.52	1 (0%) 95 94	12, 20, 31, 40	0
2	B	519/519 (100%)	-0.58	1 (0%) 95 94	11, 17, 29, 38	0
2	D	519/519 (100%)	-0.62	2 (0%) 93 93	10, 15, 26, 38	0
All	All	1994/1994 (100%)	-0.57	4 (0%) 95 94	10, 17, 29, 40	0

All (4) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	D	5	ASP	3.2
2	D	4	ILE	2.4
1	C	38	MET	2.3
2	B	122	VAL	2.2

### 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 ⓘ

There are no carbohydrates in this entry.

### 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	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
8	EDO	A	2775	4/4	0.89	0.22	26.65	30,33,34,34	0
8	EDO	D	2773	4/4	0.93	0.13	18.87	21,22,25,25	0
8	EDO	D	2780	4/4	0.92	0.13	12.30	24,25,26,30	0
8	EDO	B	2781	4/4	0.92	0.13	9.20	21,21,22,25	0
8	EDO	D	2784	4/4	0.95	0.10	5.74	15,18,20,20	0
8	EDO	B	2778	4/4	0.94	0.12	4.71	24,24,25,26	0
8	EDO	A	2777	4/4	0.96	0.10	4.38	33,33,34,37	0
8	EDO	B	2779	4/4	0.96	0.11	4.27	17,20,22,22	0
8	EDO	A	2774	4/4	0.97	0.09	2.08	20,23,24,25	0
8	EDO	D	2785	4/4	0.97	0.11	1.59	19,27,27,30	0
5	HCA	C	501	14/14	0.97	0.07	1.09	12,14,16,16	0
8	EDO	A	2776	4/4	0.82	0.11	0.88	30,33,36,36	0
4	MG	D	3006	1/1	0.99	0.06	0.03	12,12,12,12	0
8	EDO	D	2783	4/4	0.96	0.07	-0.05	17,19,19,23	0
3	CL	D	2772	1/1	0.99	0.06	-0.08	17,17,17,17	0
5	HCA	A	501	14/14	0.98	0.06	-0.20	11,14,16,16	0
4	MG	B	3002	1/1	0.99	0.06	-0.39	12,12,12,12	0
7	CLF	A	505	15/15	0.99	0.06	-0.47	12,15,17,17	0
7	CLF	C	505	15/15	0.99	0.06	-0.59	11,13,14,15	0
6	CFM	C	503	17/17	0.99	0.06	-1.03	12,14,15,16	0
6	CFM	A	503	17/17	0.99	0.06	-1.20	11,13,14,15	0
4	MG	A	3001	1/1	0.98	0.05	-1.23	18,18,18,18	0
4	MG	C	3005	1/1	0.99	0.04	-1.84	19,19,19,19	0
3	CL	B	2771	1/1	0.98	0.04	-2.11	20,20,20,20	0
4	MG	B	3003	1/1	0.97	0.08	-	27,27,27,27	0
4	MG	D	3007	1/1	0.97	0.12	-	29,29,29,29	0
4	MG	B	3004	1/1	0.93	0.06	-	25,25,25,25	0
8	EDO	C	2782	4/4	0.96	0.07	-	28,28,29,29	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.