



# Full wwPDB X-ray Structure Validation Report ⓘ

Aug 8, 2020 – 04:31 AM BST

PDB ID : 4JR9  
Title : Crystal structure of nitrate/nitrite exchanger NarK  
Authors : Zheng, H.; Wisedchaisri, G.; Gonen, T.  
Deposited on : 2013-03-21  
Resolution : 2.60 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

<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

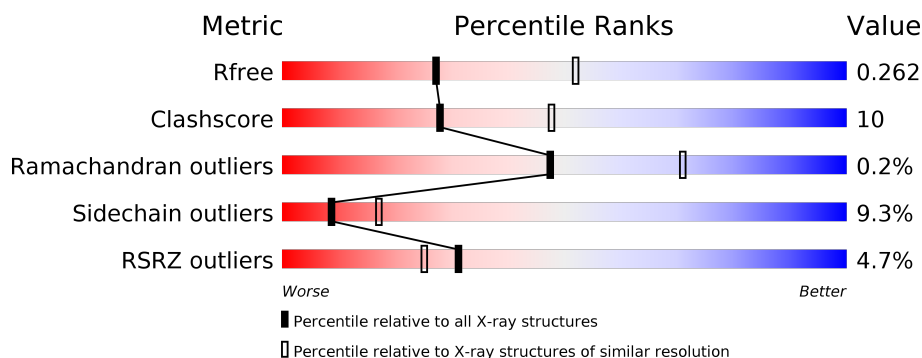
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.60 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))
$R_{free}$	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (2.60-2.60)

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 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	466	<div> <div>6%</div> <div> <div></div> <div>71%</div> <div>15%</div> <div>12%</div> </div> </div>
2	H	217	<div> <div>%</div> <div> <div></div> <div>72%</div> <div>21%</div> <div>5%</div> </div> </div>
3	L	211	<div> <div>3%</div> <div> <div></div> <div>72%</div> <div>23%</div> </div> </div>

## 2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 6337 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 Nitrite extrusion protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	409	Total	C	N	O	S	0	1	0
			3093	2058	493	522	20			

There are 3 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	GLY	-	expression tag	UNP P10903
A	-1	SER	-	expression tag	UNP P10903
A	0	HIS	-	expression tag	UNP P10903

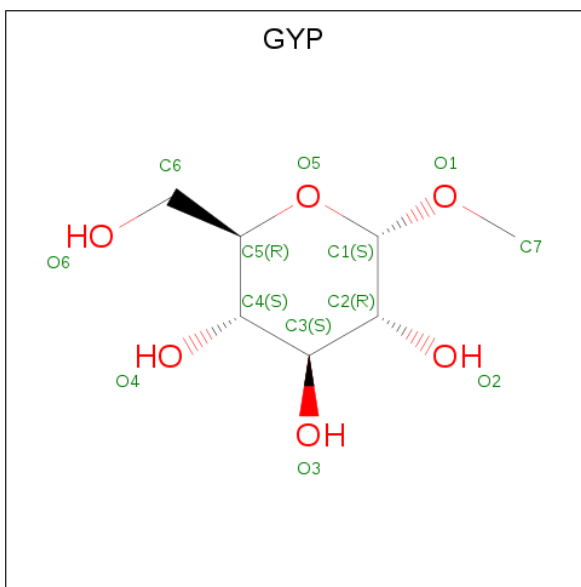
- Molecule 2 is a protein called Immunoglobulin Gamma-2a, Heavy chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	213	Total	C	N	O	S	0	0	0
			1592	1016	261	308	7			

- Molecule 3 is a protein called Immunoglobulin Kappa, Light chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	L	211	Total	C	N	O	S	0	0	0
			1594	1001	261	326	6			

- Molecule 4 is methyl alpha-D-glucopyranoside (three-letter code: GYP) (formula: C<sub>7</sub>H<sub>14</sub>O<sub>6</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			13	7	6		

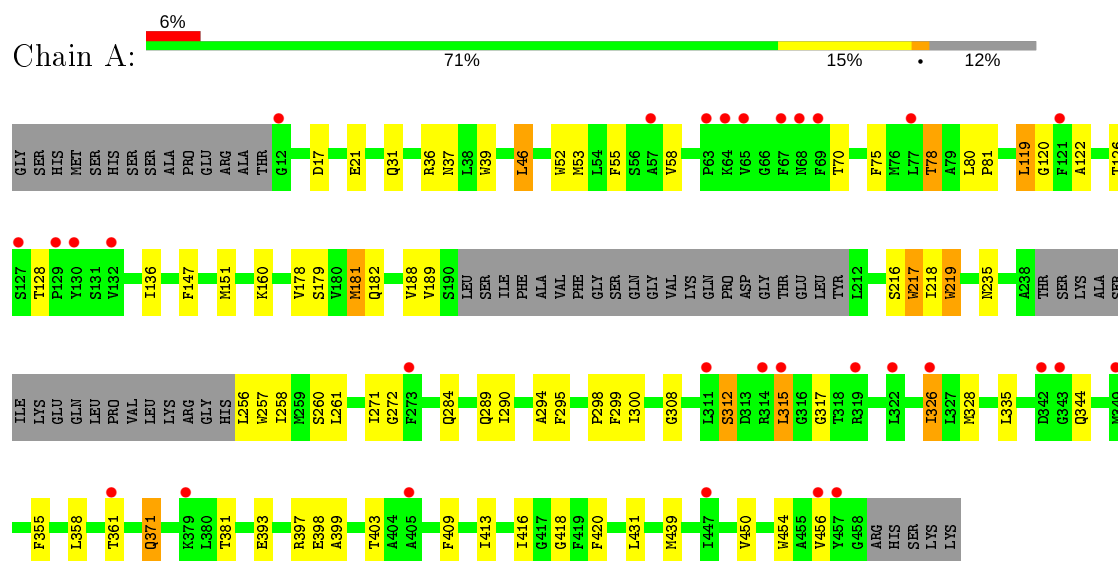
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	5	Total	O	0	0
			5	5		
5	H	22	Total	O	0	0
			22	22		
5	L	18	Total	O	0	0
			18	18		

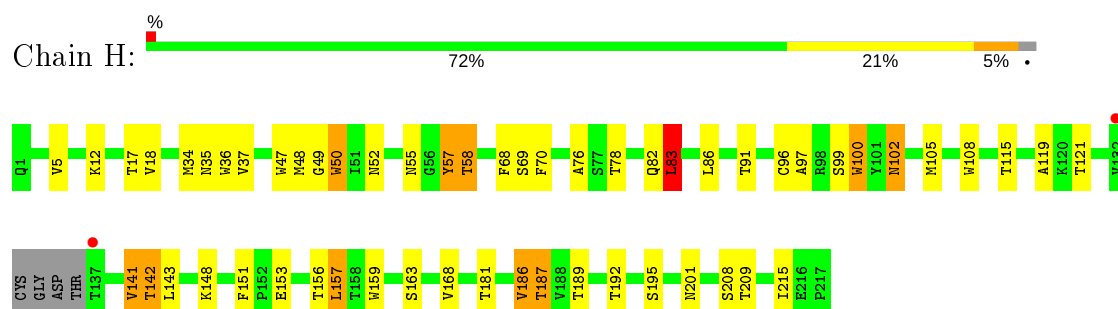
### 3 Residue-property plots [i](#)

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.

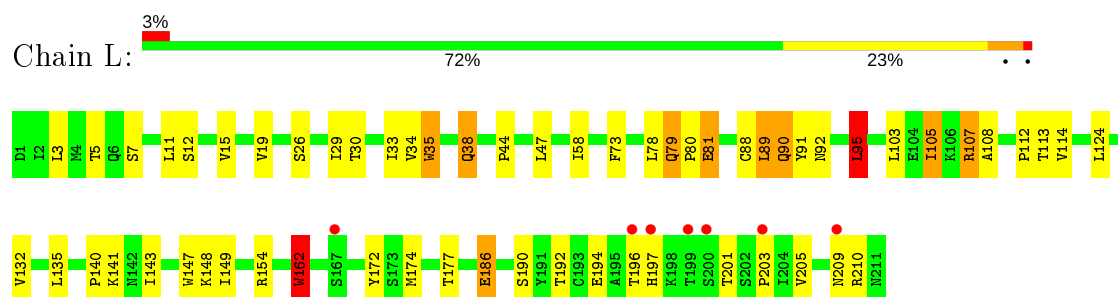
#### • Molecule 1: Nitrite extrusion protein 1



#### • Molecule 2: Immunoglobulin Gamma-2a, Heavy chain



#### • Molecule 3: Immunoglobulin Kappa, Light chain



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	138.27Å 86.55Å 144.79Å 90.00° 115.41° 90.00°	Depositor
Resolution (Å)	50.00 – 2.60 43.27 – 2.60	Depositor EDS
% Data completeness (in resolution range)	99.5 (50.00-2.60) 99.5 (43.27-2.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.96 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.229 , 0.261 0.229 , 0.262	Depositor DCC
$R_{free}$ test set	2406 reflections (5.07%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	56.1	Xtriage
Anisotropy	0.132	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 66.0	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	6337	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	67.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.02% 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: GYP

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	0.52	6/3181 (0.2%)	0.56	2/4322 (0.0%)
2	H	0.65	2/1636 (0.1%)	0.75	2/2234 (0.1%)
3	L	0.56	3/1630 (0.2%)	0.65	1/2226 (0.0%)
All	All	0.57	11/6447 (0.2%)	0.64	5/8782 (0.1%)

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	L	162	TRP	CD2-CE2	6.26	1.48	1.41
1	A	219	TRP	CD2-CE2	5.58	1.48	1.41
2	H	50	TRP	CD2-CE2	5.28	1.47	1.41
3	L	147	TRP	CD2-CE2	5.27	1.47	1.41
1	A	257	TRP	CD2-CE2	5.23	1.47	1.41
3	L	35	TRP	CD2-CE2	5.23	1.47	1.41
1	A	39	TRP	CD2-CE2	5.22	1.47	1.41
2	H	108	TRP	CD2-CE2	5.21	1.47	1.41
1	A	52	TRP	CD2-CE2	5.08	1.47	1.41
1	A	217	TRP	CD2-CE2	5.06	1.47	1.41
1	A	454	TRP	CD2-CE2	5.05	1.47	1.41

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	157	LEU	CA-CB-CG	7.25	131.98	115.30
3	L	95	LEU	CA-CB-CG	6.63	130.56	115.30
2	H	83	LEU	CB-CG-CD2	-6.13	100.58	111.00
1	A	46	LEU	CA-CB-CG	6.05	129.22	115.30
1	A	315	LEU	CA-CB-CG	5.78	128.60	115.30

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	3093	0	3128	40	0
2	H	1592	0	1541	52	0
3	L	1594	0	1494	35	0
4	A	13	0	14	1	0
5	A	5	0	0	0	0
5	H	22	0	0	1	0
5	L	18	0	0	1	0
All	All	6337	0	6177	123	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 10.

All (123) 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:H:100:TRP:O	2:H:100:TRP:HE3	1.21	1.24
2:H:50:TRP:CE2	2:H:100:TRP:HD1	1.72	1.06
2:H:37:VAL:HG21	2:H:105:MET:HE1	1.37	1.04
2:H:100:TRP:O	2:H:100:TRP:CE3	2.12	1.01
2:H:18:VAL:HG23	2:H:86:LEU:HD11	1.39	0.99
2:H:50:TRP:CE2	2:H:100:TRP:CD1	2.51	0.98
2:H:34:MET:CE	2:H:96:CYS:HB2	1.95	0.96
2:H:37:VAL:CG2	2:H:105:MET:HE1	1.96	0.95
2:H:18:VAL:CG2	2:H:86:LEU:HD11	2.01	0.91
2:H:34:MET:HE1	2:H:96:CYS:HB2	1.53	0.90
2:H:35:ASN:HD22	2:H:47:TRP:HE1	1.21	0.87
2:H:50:TRP:CD2	2:H:100:TRP:HD1	1.95	0.83
1:A:36:ARG:HH21	1:A:235:ASN:HD21	1.28	0.82
1:A:312:SER:HB3	1:A:371:GLN:HG2	1.61	0.81
1:A:119:LEU:HB3	1:A:136:ILE:HG21	1.65	0.77
3:L:132:VAL:HG22	3:L:177:THR:HG22	1.66	0.75
3:L:5:THR:HG23	5:L:302:HOH:O	1.92	0.69

*Continued on next page...*



*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:97:ALA:HB3	2:H:105:MET:HE2	1.74	0.69
1:A:260:SER:HB3	1:A:413:ILE:HG21	1.75	0.68
2:H:37:VAL:CG2	2:H:105:MET:CE	2.70	0.68
2:H:35:ASN:ND2	2:H:47:TRP:HE1	1.92	0.68
2:H:55:ASN:HB3	2:H:57:TYR:HB2	1.79	0.64
2:H:100:TRP:C	2:H:100:TRP:HE3	2.00	0.64
1:A:326:ILE:HD12	1:A:450:VAL:HG21	1.79	0.64
1:A:312:SER:CB	1:A:371:GLN:HG2	2.28	0.63
2:H:50:TRP:CZ2	2:H:100:TRP:CD1	2.87	0.63
2:H:50:TRP:CD2	2:H:100:TRP:CD1	2.84	0.63
2:H:34:MET:HE1	2:H:96:CYS:CB	2.27	0.62
3:L:89:LEU:HD21	3:L:95:LEU:HD13	1.83	0.61
3:L:112:PRO:HG3	3:L:143:ILE:HD11	1.82	0.60
1:A:335:LEU:HD23	1:A:439:MET:SD	2.41	0.60
3:L:11:LEU:HD11	3:L:19:VAL:HG13	1.83	0.60
1:A:181:MET:HG3	1:A:219:TRP:CH2	2.37	0.59
2:H:52:ASN:HB3	2:H:55:ASN:HB2	1.84	0.59
2:H:100:TRP:CZ3	3:L:91:TYR:O	2.56	0.59
2:H:100:TRP:C	2:H:100:TRP:CE3	2.76	0.58
2:H:141:VAL:O	2:H:187:THR:HA	2.04	0.58
1:A:182:GLN:HB2	1:A:295:PHE:CB	2.34	0.58
3:L:196:THR:HG22	3:L:203:PRO:HB3	1.85	0.57
3:L:34:VAL:HG13	3:L:91:TYR:CE1	2.38	0.57
2:H:100:TRP:HZ3	3:L:91:TYR:O	1.87	0.57
3:L:148:LYS:HB2	3:L:192:THR:HB	1.86	0.57
1:A:160:LYS:HG3	1:A:397:ARG:HA	1.87	0.56
1:A:381:THR:HG21	1:A:398:GLU:HB2	1.87	0.56
3:L:34:VAL:HG23	3:L:89:LEU:HB3	1.87	0.56
1:A:399:ALA:O	1:A:403:THR:HG23	2.06	0.55
3:L:162:TRP:N	3:L:162:TRP:CD1	2.75	0.55
1:A:178:VAL:O	1:A:182:GLN:HG2	2.06	0.55
1:A:358:LEU:HA	1:A:361:THR:HG22	1.89	0.55
2:H:168:VAL:HG22	2:H:186:VAL:CG1	2.38	0.54
2:H:58:THR:CG2	2:H:70:PHE:HB2	2.38	0.54
1:A:37:ASN:HD21	1:A:235:ASN:HD22	1.55	0.53
3:L:143:ILE:HG22	3:L:162:TRP:CZ3	2.43	0.53
1:A:271:ILE:HG22	1:A:418:GLY:HA3	1.90	0.53
2:H:17:THR:HA	2:H:83:LEU:O	2.08	0.53
3:L:38:GLN:HG2	3:L:44:PRO:HG3	1.90	0.53
2:H:55:ASN:CB	2:H:57:TYR:HB2	2.38	0.53
2:H:159:TRP:HZ3	2:H:215:ILE:HD13	1.73	0.53

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:L:78:LEU:HD11	3:L:103:LEU:HD21	1.89	0.53
1:A:381:THR:CG2	1:A:398:GLU:HB2	2.40	0.51
3:L:143:ILE:HG22	3:L:162:TRP:CH2	2.46	0.51
1:A:75:PHE:O	1:A:78:THR:HG22	2.11	0.51
2:H:91:THR:HG23	2:H:115:THR:HA	1.93	0.51
3:L:140:PRO:O	3:L:197:HIS:HE1	1.94	0.50
3:L:79:GLN:NE2	3:L:81:GLU:OE2	2.38	0.50
3:L:80:PRO:HA	3:L:105:ILE:HD12	1.93	0.49
2:H:55:ASN:CB	2:H:57:TYR:H	2.25	0.49
1:A:312:SER:OG	1:A:317:GLY:HA2	2.12	0.49
2:H:47:TRP:CH2	2:H:49:GLY:HA2	2.49	0.48
2:H:68:PHE:CE2	2:H:83:LEU:HD21	2.48	0.48
1:A:308:GLY:O	1:A:312:SER:HB2	2.13	0.48
2:H:189:THR:O	2:H:192:THR:HG22	2.14	0.48
2:H:143:LEU:HD22	2:H:215:ILE:HG21	1.96	0.48
2:H:97:ALA:CB	2:H:105:MET:HE2	2.41	0.47
3:L:114:VAL:HG22	3:L:135:LEU:HD22	1.95	0.47
3:L:81:GLU:HG3	3:L:81:GLU:H	1.38	0.47
3:L:107:ARG:HD3	3:L:108:ALA:O	2.14	0.47
3:L:35:TRP:CZ3	3:L:88:CYS:HB3	2.50	0.47
2:H:12:LYS:HG3	2:H:18:VAL:HG22	1.96	0.47
1:A:147:PHE:O	1:A:151:MET:HB2	2.15	0.46
1:A:17:ASP:H	2:H:102:ASN:HD21	1.63	0.46
1:A:182:GLN:HE21	1:A:298:PRO:HG2	1.81	0.46
1:A:188:VAL:HG21	1:A:218:ILE:HB	1.97	0.46
2:H:55:ASN:HB3	2:H:57:TYR:H	1.81	0.46
1:A:36:ARG:NH2	1:A:235:ASN:HD21	2.06	0.46
1:A:55:PHE:HA	1:A:58:VAL:HB	1.97	0.46
2:H:153:GLU:HG3	5:H:316:HOH:O	2.14	0.46
1:A:290:ILE:HG13	1:A:294:ALA:HB2	1.98	0.45
1:A:80:LEU:N	1:A:81:PRO:HD2	2.32	0.45
3:L:29:ILE:HG21	3:L:90:GLN:HG3	1.99	0.45
2:H:36:TRP:HB3	2:H:48:MET:CE	2.47	0.45
2:H:34:MET:HE2	2:H:96:CYS:HB2	1.93	0.45
3:L:141:LYS:HD2	3:L:172:TYR:CE1	2.52	0.45
1:A:122:ALA:HB1	1:A:128:THR:HG21	1.98	0.45
1:A:21:GLU:OE2	1:A:160:LYS:HG2	2.16	0.45
1:A:289:GLN:O	1:A:289:GLN:HG2	2.18	0.44
3:L:34:VAL:CG2	3:L:89:LEU:HB3	2.47	0.44
1:A:181:MET:HG3	1:A:219:TRP:HH2	1.81	0.44
1:A:393:GLU:HB3	1:A:397:ARG:NH2	2.32	0.44

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:H:76:ALA:O	2:H:78:THR:HG23	2.17	0.44
2:H:36:TRP:HB3	2:H:48:MET:HE1	1.99	0.43
2:H:119:ALA:HB3	2:H:151:PHE:CE2	2.53	0.43
3:L:35:TRP:CH2	3:L:88:CYS:HB3	2.52	0.43
1:A:258:ILE:HD13	1:A:456:VAL:HG21	2.01	0.43
1:A:271:ILE:HA	1:A:271:ILE:HD13	1.90	0.43
2:H:58:THR:HG22	2:H:70:PHE:HB2	2.01	0.43
3:L:35:TRP:CE2	3:L:73:PHE:HB2	2.54	0.43
3:L:3:LEU:HB2	3:L:26:SER:HB3	2.00	0.43
1:A:420:PHE:HB3	4:A:501:GYP:H2	2.01	0.42
3:L:186:GLU:HA	3:L:210:ARG:NH1	2.34	0.42
1:A:300:ILE:HD13	1:A:300:ILE:HA	1.84	0.41
2:H:142:THR:HA	2:H:186:VAL:O	2.20	0.41
2:H:69:SER:HB2	2:H:82:GLN:HB3	2.02	0.41
2:H:47:TRP:CE2	3:L:95:LEU:HG	2.56	0.41
3:L:149:ILE:HD12	3:L:154:ARG:HD3	2.02	0.41
1:A:120:GLY:HA3	1:A:217:TRP:NE1	2.36	0.41
3:L:47:LEU:HD23	3:L:58:ILE:HD12	2.03	0.41
1:A:272:GLY:HA2	1:A:418:GLY:HA2	2.03	0.40
1:A:179:SER:OG	1:A:299:PHE:HB2	2.21	0.40
2:H:50:TRP:NE1	2:H:100:TRP:CD1	2.89	0.40
2:H:142:THR:HB	2:H:187:THR:HG22	2.02	0.40
3:L:194:GLU:CG	3:L:205:VAL:HG22	2.52	0.40
3:L:190:SER:HB3	3:L:209:ASN:OD1	2.21	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	404/466 (87%)	393 (97%)	11 (3%)	0	<a href="#">100</a> <a href="#">100</a>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	H	209/217 (96%)	206 (99%)	2 (1%)	1 (0%)	29	52
3	L	209/211 (99%)	197 (94%)	11 (5%)	1 (0%)	29	52
All	All	822/894 (92%)	796 (97%)	24 (3%)	2 (0%)	47	71

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	H	99	SER
3	L	92	ASN

### 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	315/364 (86%)	292 (93%)	23 (7%)	14	28
2	H	174/182 (96%)	154 (88%)	20 (12%)	5	10
3	L	177/190 (93%)	158 (89%)	19 (11%)	6	12
All	All	666/736 (90%)	604 (91%)	62 (9%)	9	17

All (62) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	31	GLN
1	A	46	LEU
1	A	53	MET
1	A	70	THR
1	A	78	THR
1	A	119	LEU
1	A	126	THR
1	A	181	MET
1	A	189	VAL
1	A	216	SER
1	A	256	LEU
1	A	261	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	284	GLN
1	A	312	SER
1	A	315	LEU
1	A	326	ILE
1	A	328	MET
1	A	344	GLN
1	A	355	PHE
1	A	371	GLN
1	A	409	PHE
1	A	416	ILE
1	A	431	LEU
2	H	5	VAL
2	H	57	TYR
2	H	58	THR
2	H	83	LEU
2	H	100	TRP
2	H	102	ASN
2	H	121	THR
2	H	141	VAL
2	H	142	THR
2	H	148	LYS
2	H	156	THR
2	H	157	LEU
2	H	163	SER
2	H	181	THR
2	H	186	VAL
2	H	187	THR
2	H	195	SER
2	H	201	ASN
2	H	208	SER
2	H	209	THR
3	L	7	SER
3	L	12	SER
3	L	15	VAL
3	L	30	THR
3	L	33	ILE
3	L	38	GLN
3	L	79	GLN
3	L	81	GLU
3	L	89	LEU
3	L	90	GLN
3	L	95	LEU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
3	L	105	ILE
3	L	107	ARG
3	L	113	THR
3	L	124	LEU
3	L	162	TRP
3	L	174	MET
3	L	186	GLU
3	L	201	THR

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	182	GLN
1	A	235	ASN
1	A	284	GLN
1	A	344	GLN
1	A	371	GLN
2	H	35	ASN
2	H	102	ASN
3	L	27	GLN

### 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](#)

1 ligand is 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
4	GYP	A	501	-	13,13,13	0.72	1 (7%)	18,18,18	1.58	3 (16%)

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	GYP	A	501	-	-	2/4/24/24	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	501	GYP	O1-C1	2.07	1.43	1.40

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
4	A	501	GYP	O1-C1-C2	4.39	113.30	108.15
4	A	501	GYP	C3-C4-C5	-3.09	104.72	110.24
4	A	501	GYP	C1-C2-C3	2.37	114.92	110.00

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	A	501	GYP	C2-C1-O1-C7
4	A	501	GYP	O5-C1-O1-C7

There are no ring outliers.

1 monomer is involved in 1 short contact:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	501	GYP	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	409/466 (87%)	0.34	30 (7%) 15 11	39, 89, 122, 155	0
2	H	213/217 (98%)	-0.25	2 (0%) 84 82	30, 43, 72, 129	0
3	L	211/211 (100%)	-0.06	7 (3%) 46 39	29, 48, 79, 107	0
All	All	833/894 (93%)	0.09	39 (4%) 31 25	29, 61, 116, 155	0

All (39) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	343	GLY	6.2
1	A	132	VAL	6.0
1	A	315	LEU	5.9
3	L	209	ASN	4.7
1	A	68	ASN	4.4
1	A	314	ARG	4.1
1	A	67	PHE	4.1
1	A	65	VAL	3.9
1	A	121	PHE	3.7
1	A	326	ILE	3.6
1	A	319	ARG	3.1
2	H	137	THR	3.1
1	A	130	TYR	3.1
3	L	199	THR	3.1
1	A	457	TYR	3.0
1	A	69	PHE	3.0
1	A	447	ILE	3.0
1	A	77	LEU	3.0
1	A	127	SER	3.0
1	A	63	PRO	2.9
1	A	361	THR	2.7
1	A	342	ASP	2.6
1	A	456	VAL	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	A	129	PRO	2.6
1	A	379	LYS	2.5
1	A	322	LEU	2.4
1	A	57	ALA	2.3
3	L	196	THR	2.3
1	A	273	PHE	2.3
3	L	167	SER	2.2
3	L	197	HIS	2.2
1	A	405	ALA	2.2
1	A	12	GLY	2.2
1	A	311	LEU	2.2
2	H	132	VAL	2.2
3	L	203	PRO	2.2
3	L	200	SER	2.1
1	A	349	MET	2.0
1	A	64	LYS	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
4	GYP	A	501	13/13	0.87	0.17	83,91,98,107	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.