



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 15, 2017 – 05:08 am GMT

PDB ID : 4FM4  
Title : Wild Type Fe-type Nitrile Hydratase from Comamonas testosteroni N1  
Authors : Kuhn, M.L.; Martinez, S.; Gumataotao, N.; Bornscheuer, U.; Liu, D.; Holz, R.C.  
Deposited on : 2012-06-15  
Resolution : 2.38 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

<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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

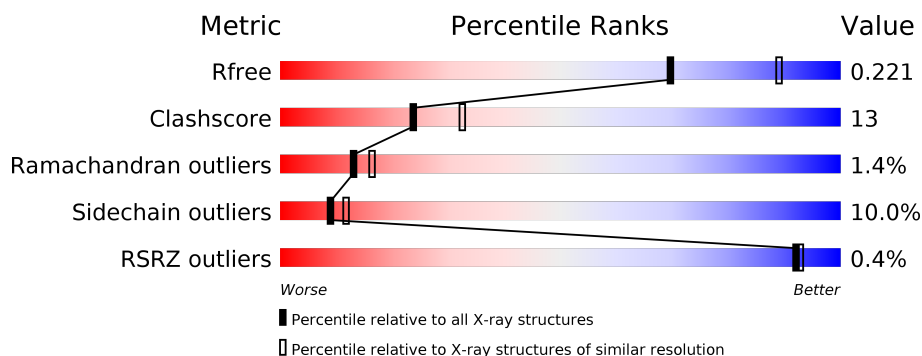
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.38 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))
$R_{free}$	100719	4388 (2.40-2.36)
Clashscore	112137	4984 (2.40-2.36)
Ramachandran outliers	110173	4907 (2.40-2.36)
Sidechain outliers	110143	4909 (2.40-2.36)
RSRZ outliers	101464	4423 (2.40-2.36)

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	209	<div> <div>71%</div> <div>21%</div> <div>7%</div> <div>.</div> </div>
1	C	209	<div> <div>69%</div> <div>24%</div> <div>6%</div> <div>.</div> </div>
1	E	209	<div> <div>72%</div> <div>22%</div> <div>5%</div> <div>.</div> </div>
1	G	209	<div> <div>72%</div> <div>22%</div> <div>5%</div> <div>.</div> </div>
1	I	209	<div> <div>71%</div> <div>23%</div> <div>.</div> <div>.</div> </div>
1	K	209	<div> <div>73%</div> <div>20%</div> <div>5%</div> <div>.</div> </div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
1	M	209	 71% 22% 5% •
1	O	209	 74% 18% 7% •
2	B	206	 79% 18% •
2	D	206	 75% 20% •
2	F	206	 76% 19% •
2	H	206	 81% 14% •
2	J	206	 75% 20% •
2	L	206	 76% 20% •
2	N	206	 77% 19% •
2	P	206	 75% 21% •

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
1	CSD	E	104	-	-	X	-
3	PO4	O	301	-	-	-	X

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 27326 atoms, of which 8 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 Nitrile hydratase alpha subunit.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	206	Total	C	H	N	O	S	0	0	0
			1610	1026	1	268	306	9			
1	C	206	Total	C	H	N	O	S	0	0	0
			1610	1026	1	268	306	9			
1	E	206	Total	C	H	N	O	S	0	0	0
			1610	1026	1	268	306	9			
1	G	206	Total	C	H	N	O	S	0	0	0
			1610	1026	1	268	306	9			
1	I	206	Total	C	H	N	O	S	0	0	0
			1610	1026	1	268	306	9			
1	K	206	Total	C	H	N	O	S	0	0	0
			1610	1026	1	268	306	9			
1	M	206	Total	C	H	N	O	S	0	0	0
			1610	1026	1	268	306	9			
1	O	206	Total	C	H	N	O	S	0	0	0
			1610	1026	1	268	306	9			

- Molecule 2 is a protein called Nitrile hydratase beta subunit.

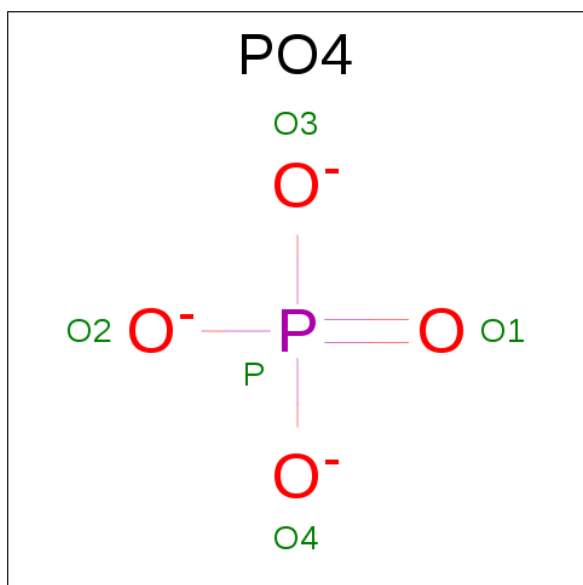
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	206	Total	C	N	O	S	0	0	0
			1589	1017	271	293	8			
2	D	206	Total	C	N	O	S	0	0	0
			1589	1017	271	293	8			
2	F	206	Total	C	N	O	S	0	0	0
			1589	1017	271	293	8			
2	H	206	Total	C	N	O	S	0	0	0
			1589	1017	271	293	8			
2	J	206	Total	C	N	O	S	0	0	0
			1589	1017	271	293	8			
2	L	206	Total	C	N	O	S	0	0	0
			1589	1017	271	293	8			

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	N	206	Total	C	N	O	S	0	0	0
			1589	1017	271	293	8			
2	P	206	Total	C	N	O	S	0	0	0
			1589	1017	271	293	8			

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



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	A	1	Total	O	P	0	0
			5	4	1		
3	C	1	Total	O	P	0	0
			5	4	1		
3	E	1	Total	O	P	0	0
			5	4	1		
3	G	1	Total	O	P	0	0
			5	4	1		
3	I	1	Total	O	P	0	0
			5	4	1		
3	K	1	Total	O	P	0	0
			5	4	1		
3	M	1	Total	O	P	0	0
			5	4	1		
3	O	1	Total	O	P	0	0
			5	4	1		

- Molecule 4 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	G	1	Total Fe 1 1	0	0
4	K	1	Total Fe 1 1	0	0
4	E	1	Total Fe 1 1	0	0
4	I	1	Total Fe 1 1	0	0
4	C	1	Total Fe 1 1	0	0
4	A	1	Total Fe 1 1	0	0
4	O	1	Total Fe 1 1	0	0
4	M	1	Total Fe 1 1	0	0

- Molecule 5 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	98	Total O 98 98	0	0
5	B	113	Total O 113 113	0	0
5	C	91	Total O 91 91	0	0
5	D	106	Total O 106 106	0	0
5	E	107	Total O 107 107	0	0
5	F	112	Total O 112 112	0	0
5	G	93	Total O 93 93	0	0
5	H	110	Total O 110 110	0	0
5	I	108	Total O 108 108	0	0
5	J	92	Total O 92 92	0	0
5	K	113	Total O 113 113	0	0
5	L	122	Total O 122 122	0	0

*Continued on next page...*

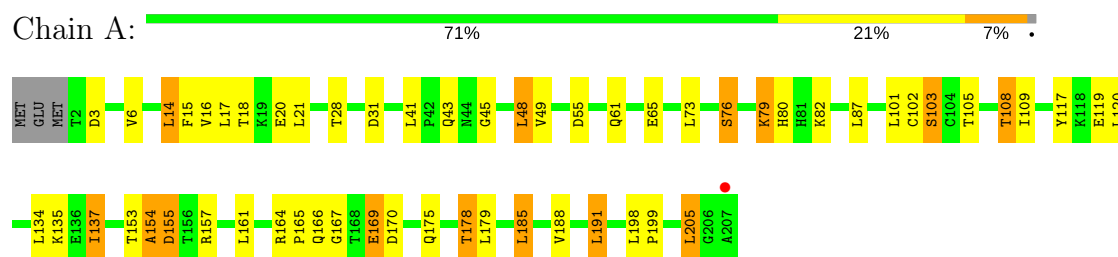
*Continued from previous page...*

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	M	99	Total 99	O 99	0	0
5	N	104	Total 104	O 104	0	0
5	O	115	Total 115	O 115	0	0
5	P	103	Total 103	O 103	0	0

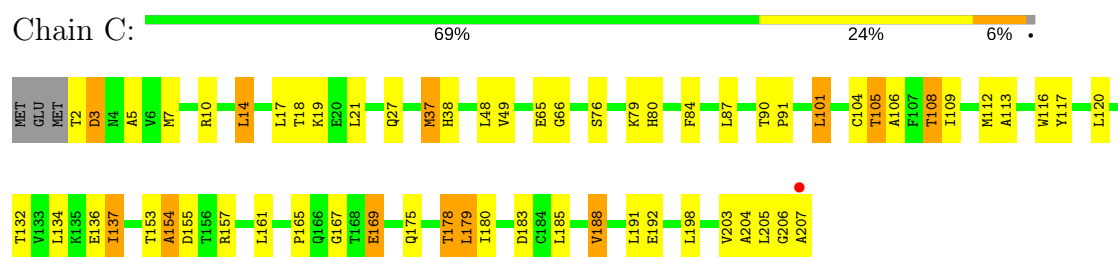
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

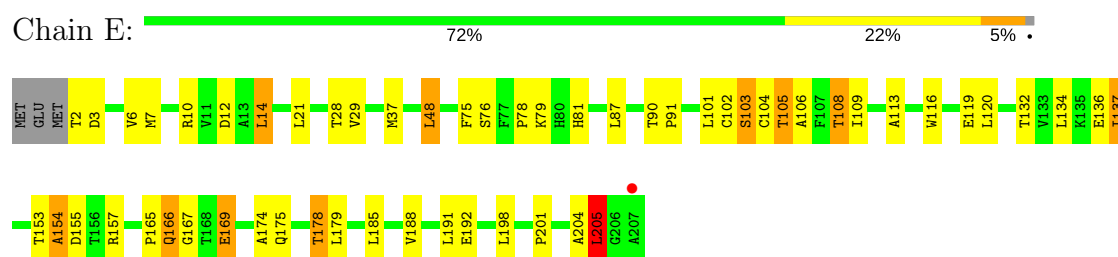
#### • Molecule 1: Nitrile hydratase alpha subunit



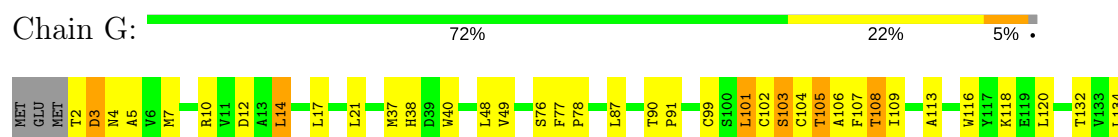
#### • Molecule 1: Nitrile hydratase alpha subunit



#### • Molecule 1: Nitrile hydratase alpha subunit



#### • Molecule 1: Nitrile hydratase alpha subunit

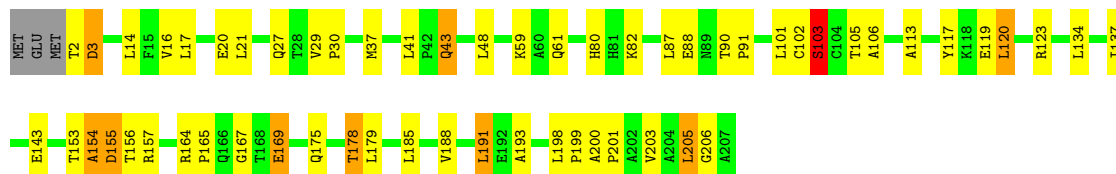






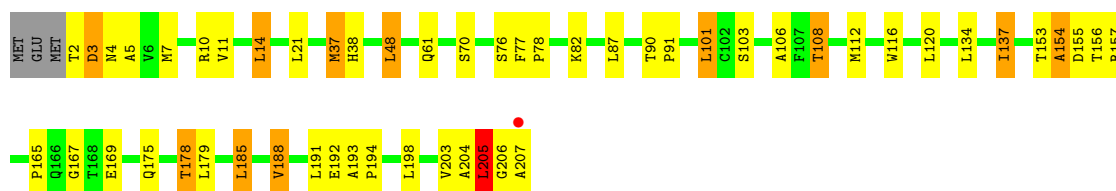
• Molecule 1: Nitrile hydratase alpha subunit

Chain I: 71% 23% . .



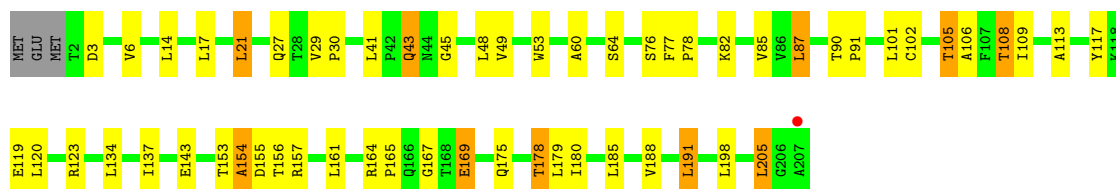
• Molecule 1: Nitrile hydratase alpha subunit

Chain K: 73% 20% 5% .



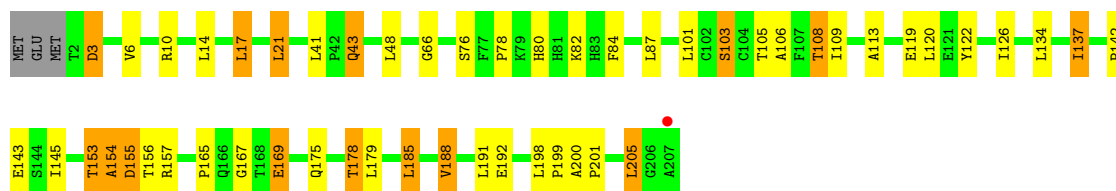
• Molecule 1: Nitrile hydratase alpha subunit

Chain M: 71% 22% 5% .



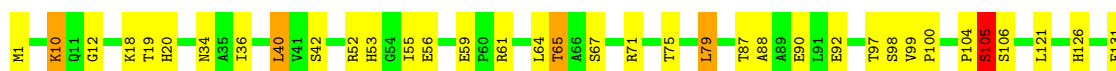
• Molecule 1: Nitrile hydratase alpha subunit

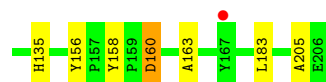
Chain O: 74% 18% 7% .



• Molecule 2: Nitrile hydratase beta subunit

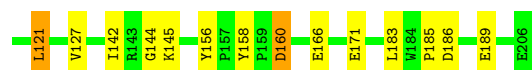
Chain B: 79% 18% .





• Molecule 2: Nitrile hydratase beta subunit

Chain D: 75% 20% •



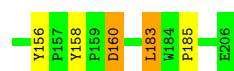
• Molecule 2: Nitrile hydratase beta subunit

Chain F: 76% 19% •



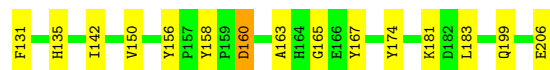
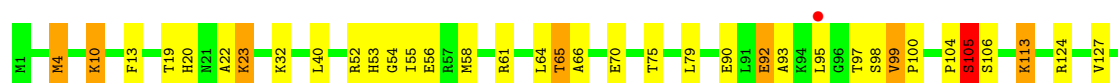
• Molecule 2: Nitrile hydratase beta subunit

Chain H: 81% 14% •



• Molecule 2: Nitrile hydratase beta subunit

Chain J: 75% 20% •



• Molecule 2: Nitrile hydratase beta subunit

Chain L: 76% 20% •





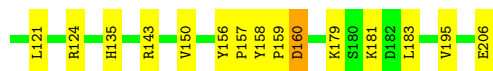
- Molecule 2: Nitrile hydratase beta subunit

Chain N: 77% 19% •



- Molecule 2: Nitrile hydratase beta subunit

Chain P: 75% 21% •



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	111.40Å 111.40Å 475.31Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	36.16 – 2.38 36.16 – 2.38	Depositor EDS
% Data completeness (in resolution range)	99.3 (36.16-2.38) 87.0 (36.16-2.38)	Depositor EDS
$R_{merge}$	0.13	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.14 (at 2.39Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.7.3_928)	Depositor
R, $R_{free}$	0.188 , 0.227 0.177 , 0.221	Depositor DCC
$R_{free}$ test set	13123 reflections (5.02%)	DCC
Wilson B-factor (Å <sup>2</sup> )	33.1	Xtriage
Anisotropy	0.642	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	(Not available) , (Not available)	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.35$	Xtriage
Estimated twinning fraction	0.487 for -h,-k,l 0.487 for h,-h-k,-l 0.487 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	27326	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	47.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.34% 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: CSD, PO4, FE

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	0.54	0/1630	0.66	0/2227
1	C	0.57	0/1630	0.71	0/2227
1	E	0.57	0/1630	0.69	0/2227
1	G	0.57	0/1630	0.70	0/2227
1	I	0.53	0/1630	0.67	0/2227
1	K	0.57	0/1630	0.71	0/2227
1	M	0.55	0/1630	0.68	0/2227
1	O	0.55	0/1630	0.67	0/2227
2	B	0.56	0/1636	0.65	0/2216
2	D	0.58	0/1636	0.65	1/2216 (0.0%)
2	F	0.59	0/1636	0.67	1/2216 (0.0%)
2	H	0.59	0/1636	0.66	1/2216 (0.0%)
2	J	0.56	0/1636	0.66	0/2216
2	L	0.60	0/1636	0.67	1/2216 (0.0%)
2	N	0.57	0/1636	0.65	0/2216
2	P	0.58	0/1636	0.66	0/2216
All	All	0.57	0/26128	0.67	4/35544 (0.0%)

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	1
1	C	0	1
1	E	0	1
1	G	0	1
1	I	0	1
1	K	0	1
1	M	0	1

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	#Chirality outliers	#Planarity outliers
1	O	0	1
All	All	0	8

There are no bond length outliers.

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	H	105	SER	N-CA-C	-6.11	94.51	111.00
2	F	105	SER	N-CA-C	-5.75	95.47	111.00
2	L	105	SER	N-CA-C	-5.52	96.10	111.00
2	D	105	SER	N-CA-C	-5.26	96.79	111.00

There are no chirality outliers.

5 of 8 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	153	THR	Peptide
1	C	153	THR	Peptide
1	E	153	THR	Peptide
1	G	153	THR	Peptide
1	I	153	THR	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	1609	1	1592	43	0
1	C	1609	1	1590	49	0
1	E	1609	1	1591	42	0
1	G	1609	1	1591	42	0
1	I	1609	1	1591	48	0
1	K	1609	1	1590	48	0
1	M	1609	1	1590	52	0
1	O	1609	1	1590	44	0
2	B	1589	0	1538	37	0
2	D	1589	0	1538	43	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	F	1589	0	1538	41	0
2	H	1589	0	1538	38	0
2	J	1589	0	1538	48	0
2	L	1589	0	1538	46	0
2	N	1589	0	1538	33	0
2	P	1589	0	1538	47	0
3	A	5	0	0	0	0
3	C	5	0	0	1	0
3	E	5	0	0	1	0
3	G	5	0	0	0	0
3	I	5	0	0	0	0
3	K	5	0	0	0	0
3	M	5	0	0	0	0
3	O	5	0	0	0	0
4	A	1	0	0	0	0
4	C	1	0	0	0	0
4	E	1	0	0	0	0
4	G	1	0	0	0	0
4	I	1	0	0	0	0
4	K	1	0	0	0	0
4	M	1	0	0	0	0
4	O	1	0	0	0	0
5	A	98	0	0	7	0
5	B	113	0	0	4	0
5	C	91	0	0	5	1
5	D	106	0	0	2	0
5	E	107	0	0	6	0
5	F	112	0	0	4	1
5	G	93	0	0	4	0
5	H	110	0	0	8	0
5	I	108	0	0	5	0
5	J	92	0	0	8	0
5	K	113	0	0	8	0
5	L	122	0	0	5	0
5	M	99	0	0	7	0
5	N	104	0	0	1	0
5	O	115	0	0	4	1
5	P	103	0	0	3	1
All	All	27318	8	25029	640	2

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 640 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:K:207:ALA:CB	2:P:42:SER:HB3	1.71	1.20
2:F:124:ARG:HH11	2:F:124:ARG:HG3	0.99	1.13
1:G:154:ALA:O	5:G:447:HOH:O	1.83	0.96
2:F:97:THR:HG22	2:F:98:SER:H	1.30	0.96
1:K:207:ALA:HB1	2:P:42:SER:HB3	1.46	0.95

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

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
5:C:490:HOH:O	5:O:511:HOH:O[2_555]	1.77	0.43
5:F:376:HOH:O	5:P:374:HOH:O[1_565]	2.12	0.08

## 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	202/209 (97%)	192 (95%)	8 (4%)	2 (1%)	18	24
1	C	202/209 (97%)	195 (96%)	6 (3%)	1 (0%)	32	44
1	E	202/209 (97%)	195 (96%)	4 (2%)	3 (2%)	12	15
1	G	202/209 (97%)	193 (96%)	6 (3%)	3 (2%)	12	15
1	I	202/209 (97%)	196 (97%)	4 (2%)	2 (1%)	18	24
1	K	202/209 (97%)	194 (96%)	5 (2%)	3 (2%)	12	15
1	M	202/209 (97%)	194 (96%)	7 (4%)	1 (0%)	32	44
1	O	202/209 (97%)	194 (96%)	6 (3%)	2 (1%)	18	24
2	B	204/206 (99%)	196 (96%)	6 (3%)	2 (1%)	18	24
2	D	204/206 (99%)	196 (96%)	4 (2%)	4 (2%)	9	9

*Continued on next page...*



Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	F	204/206 (99%)	194 (95%)	6 (3%)	4 (2%)	9	9
2	H	204/206 (99%)	194 (95%)	6 (3%)	4 (2%)	9	9
2	J	204/206 (99%)	190 (93%)	11 (5%)	3 (2%)	12	15
2	L	204/206 (99%)	195 (96%)	6 (3%)	3 (2%)	12	15
2	N	204/206 (99%)	192 (94%)	7 (3%)	5 (2%)	6	6
2	P	204/206 (99%)	191 (94%)	11 (5%)	2 (1%)	18	24
All	All	3248/3320 (98%)	3101 (96%)	103 (3%)	44 (1%)	13	16

5 of 44 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	103	SER
1	A	154	ALA
2	B	106	SER
1	C	154	ALA
2	D	105	SER

### 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	172/175 (98%)	151 (88%)	21 (12%)	6	6
1	C	172/175 (98%)	149 (87%)	23 (13%)	4	5
1	E	172/175 (98%)	152 (88%)	20 (12%)	6	7
1	G	172/175 (98%)	155 (90%)	17 (10%)	9	12
1	I	172/175 (98%)	152 (88%)	20 (12%)	6	7
1	K	172/175 (98%)	153 (89%)	19 (11%)	7	9
1	M	172/175 (98%)	156 (91%)	16 (9%)	10	14
1	O	172/175 (98%)	149 (87%)	23 (13%)	4	5
2	B	164/164 (100%)	152 (93%)	12 (7%)	16	23
2	D	164/164 (100%)	149 (91%)	15 (9%)	11	15

Continued on next page...

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
2	F	164/164 (100%)	151 (92%)	13 (8%)	14	20
2	H	164/164 (100%)	152 (93%)	12 (7%)	16	23
2	J	164/164 (100%)	148 (90%)	16 (10%)	9	12
2	L	164/164 (100%)	149 (91%)	15 (9%)	11	15
2	N	164/164 (100%)	153 (93%)	11 (7%)	19	27
2	P	164/164 (100%)	147 (90%)	17 (10%)	8	10
All	All	2688/2712 (99%)	2418 (90%)	270 (10%)	9	12

5 of 270 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	H	18	LYS
1	I	191	LEU
1	O	185	LEU
2	H	34	ASN
1	I	21	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 8 such sidechains are listed below:

Mol	Chain	Res	Type
1	I	43	GLN
1	O	43	GLN
1	M	43	GLN
1	G	38	HIS
1	M	38	HIS

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

16 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
1	CSD	A	102	1,4	4,7,8	1.45	0	2,8,10	3.23	2 (100%)
1	CSD	A	104	1,4	4,7,8	1.29	1 (25%)	2,8,10	1.97	1 (50%)
1	CSD	C	102	1,4	4,7,8	1.52	0	2,8,10	3.86	2 (100%)
1	CSD	C	104	1,4	4,7,8	1.51	1 (25%)	2,8,10	1.76	1 (50%)
1	CSD	E	102	1,4	4,7,8	1.52	1 (25%)	2,8,10	3.95	2 (100%)
1	CSD	E	104	1,4	4,7,8	1.44	1 (25%)	2,8,10	1.93	1 (50%)
1	CSD	G	102	1,4	4,7,8	1.30	1 (25%)	2,8,10	3.56	2 (100%)
1	CSD	G	104	1,4	4,7,8	1.31	0	2,8,10	1.54	0
1	CSD	I	102	1,4	4,7,8	1.32	0	2,8,10	3.43	2 (100%)
1	CSD	I	104	1,4	4,7,8	1.31	0	2,8,10	1.72	0
1	CSD	K	102	1,4	4,7,8	1.21	1 (25%)	2,8,10	3.67	2 (100%)
1	CSD	K	104	1,4	4,7,8	1.46	1 (25%)	2,8,10	1.55	0
1	CSD	M	102	1,4	4,7,8	1.36	0	2,8,10	3.75	2 (100%)
1	CSD	M	104	1,4	4,7,8	1.45	1 (25%)	2,8,10	1.73	1 (50%)
1	CSD	O	102	1,4	4,7,8	1.20	0	2,8,10	3.88	2 (100%)
1	CSD	O	104	1,4	4,7,8	1.36	1 (25%)	2,8,10	1.76	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
1	CSD	A	102	1,4	-	0/2/6/8	0/0/0/0
1	CSD	A	104	1,4	-	0/2/6/8	0/0/0/0
1	CSD	C	102	1,4	-	0/2/6/8	0/0/0/0
1	CSD	C	104	1,4	-	0/2/6/8	0/0/0/0
1	CSD	E	102	1,4	-	0/2/6/8	0/0/0/0
1	CSD	E	104	1,4	-	0/2/6/8	0/0/0/0
1	CSD	G	102	1,4	-	0/2/6/8	0/0/0/0
1	CSD	G	104	1,4	-	0/2/6/8	0/0/0/0
1	CSD	I	102	1,4	-	0/2/6/8	0/0/0/0
1	CSD	I	104	1,4	-	0/2/6/8	0/0/0/0
1	CSD	K	102	1,4	-	0/2/6/8	0/0/0/0
1	CSD	K	104	1,4	-	0/2/6/8	0/0/0/0
1	CSD	M	102	1,4	-	0/2/6/8	0/0/0/0

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
1	CSD	M	104	1,4	-	0/2/6/8	0/0/0/0
1	CSD	O	102	1,4	-	0/2/6/8	0/0/0/0
1	CSD	O	104	1,4	-	0/2/6/8	0/0/0/0

The worst 5 of 9 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	K	104	CSD	CB-SG	-2.21	1.66	1.79
1	C	104	CSD	CB-SG	-2.19	1.66	1.79
1	E	104	CSD	CB-SG	-2.06	1.67	1.79
1	G	102	CSD	CB-SG	-2.05	1.67	1.79
1	M	104	CSD	CB-SG	-2.03	1.67	1.79

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	O	102	CSD	O-C-CA	-2.91	116.98	125.02
1	A	102	CSD	O-C-CA	-2.79	117.32	125.02
1	G	102	CSD	O-C-CA	-2.64	117.74	125.02
1	I	102	CSD	O-C-CA	-2.64	117.74	125.02
1	E	102	CSD	O-C-CA	-2.56	117.94	125.02

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

8 monomers are involved in 14 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
1	A	102	CSD	1	0
1	C	104	CSD	2	0
1	E	102	CSD	2	0
1	E	104	CSD	4	0
1	G	102	CSD	1	0
1	G	104	CSD	2	0
1	I	102	CSD	1	0
1	M	102	CSD	1	0

## 5.5 Carbohydrates

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

Of 16 ligands modelled in this entry, 8 are monoatomic - leaving 8 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	$\# Z  > 2$
3	PO4	A	301	-	4,4,4	0.75	0	6,6,6	0.41	0
3	PO4	C	301	-	4,4,4	0.69	0	6,6,6	0.43	0
3	PO4	E	301	-	4,4,4	0.62	0	6,6,6	0.82	0
3	PO4	G	301	-	4,4,4	0.67	0	6,6,6	0.93	0
3	PO4	I	301	-	4,4,4	0.73	0	6,6,6	0.47	0
3	PO4	K	301	-	4,4,4	0.70	0	6,6,6	0.41	0
3	PO4	M	301	-	4,4,4	0.75	0	6,6,6	0.40	0
3	PO4	O	301	-	4,4,4	0.68	0	6,6,6	0.48	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
3	PO4	A	301	-	-	0/0/0/0	0/0/0/0
3	PO4	C	301	-	-	0/0/0/0	0/0/0/0
3	PO4	E	301	-	-	0/0/0/0	0/0/0/0
3	PO4	G	301	-	-	0/0/0/0	0/0/0/0
3	PO4	I	301	-	-	0/0/0/0	0/0/0/0
3	PO4	K	301	-	-	0/0/0/0	0/0/0/0
3	PO4	M	301	-	-	0/0/0/0	0/0/0/0
3	PO4	O	301	-	-	0/0/0/0	0/0/0/0

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	C	301	PO4	1	0
3	E	301	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	204/209 (97%)	0.07	1 (0%) 90 91	32, 45, 70, 131	0
1	C	204/209 (97%)	0.06	1 (0%) 90 91	31, 45, 66, 87	1 (0%)
1	E	204/209 (97%)	0.10	1 (0%) 90 91	32, 45, 69, 90	1 (0%)
1	G	204/209 (97%)	0.06	0 100 100	33, 45, 69, 91	1 (0%)
1	I	204/209 (97%)	0.06	0 100 100	33, 44, 68, 102	0
1	K	204/209 (97%)	0.07	1 (0%) 90 91	33, 45, 68, 110	1 (0%)
1	M	204/209 (97%)	0.07	1 (0%) 90 91	33, 45, 69, 125	1 (0%)
1	O	204/209 (97%)	0.09	1 (0%) 90 91	32, 45, 70, 120	0
2	B	206/206 (100%)	0.10	1 (0%) 90 91	31, 44, 77, 100	0
2	D	206/206 (100%)	0.12	0 100 100	30, 43, 75, 98	0
2	F	206/206 (100%)	0.11	1 (0%) 90 91	30, 43, 74, 96	0
2	H	206/206 (100%)	0.07	0 100 100	33, 43, 77, 96	0
2	J	206/206 (100%)	0.08	1 (0%) 90 91	31, 44, 76, 100	0
2	L	206/206 (100%)	0.04	1 (0%) 90 91	31, 44, 75, 96	0
2	N	206/206 (100%)	0.09	1 (0%) 90 91	30, 45, 79, 102	0
2	P	206/206 (100%)	0.10	1 (0%) 90 91	31, 44, 76, 102	0
All	All	3280/3320 (98%)	0.08	12 (0%) 92 93	30, 44, 75, 131	5 (0%)

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	207	ALA	5.9
1	M	207	ALA	5.9
1	O	207	ALA	4.4
1	E	207	ALA	3.0
1	K	207	ALA	3.0

## 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. 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(Å <sup>2</sup> )	Q<0.9
1	CSD	I	104	8/9	0.96	0.13	-	57,60,75,77	8
1	CSD	O	104	8/9	0.97	0.14	-	52,55,76,78	8
1	CSD	M	104	8/9	0.97	0.12	-	46,49,73,77	8
1	CSD	E	102	8/9	0.93	0.19	-	34,49,57,60	8
1	CSD	G	102	8/9	0.95	0.20	-	37,52,61,66	8
1	CSD	A	102	8/9	0.94	0.13	-	39,54,61,64	8
1	CSD	C	102	8/9	0.92	0.19	-	40,55,64,66	8
1	CSD	M	102	8/9	0.95	0.14	-	41,49,63,63	8
1	CSD	O	102	8/9	0.94	0.14	-	47,51,61,70	8
1	CSD	I	102	8/9	0.96	0.14	-	50,62,69,69	0
1	CSD	K	102	8/9	0.96	0.19	-	34,47,59,60	8
1	CSD	C	104	8/9	0.98	0.13	-	51,53,66,69	8
1	CSD	A	104	8/9	0.96	0.13	-	42,50,72,76	8
1	CSD	G	104	8/9	0.98	0.13	-	56,58,64,64	8
1	CSD	E	104	8/9	0.99	0.13	-	42,56,66,67	8
1	CSD	K	104	8/9	0.98	0.13	-	52,59,71,71	8

## 6.3 Carbohydrates [i](#)

There are no carbohydrates 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. 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(Å <sup>2</sup> )	Q<0.9
3	PO4	O	301	5/5	0.99	0.17	2.67	54,63,67,71	0
3	PO4	I	301	5/5	0.99	0.15	0.60	55,64,67,71	0
3	PO4	A	301	5/5	0.98	0.14	0.36	54,67,74,76	0

*Continued on next page...*



*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	PO4	M	301	5/5	0.99	0.15	0.36	55,68,72,76	0
3	PO4	G	301	5/5	0.99	0.14	0.13	64,70,76,81	0
3	PO4	E	301	5/5	0.99	0.12	-0.59	69,76,79,85	0
3	PO4	C	301	5/5	0.98	0.13	-0.81	68,69,81,82	0
3	PO4	K	301	5/5	0.98	0.12	-2.78	70,78,85,86	0
4	FE	K	302	1/1	0.98	0.11	-	59,59,59,59	1
4	FE	A	302	1/1	0.98	0.12	-	65,65,65,65	1
4	FE	I	302	1/1	0.94	0.11	-	74,74,74,74	1
4	FE	M	302	1/1	0.90	0.10	-	66,66,66,66	1
4	FE	G	302	1/1	0.96	0.10	-	60,60,60,60	1
4	FE	C	302	1/1	0.97	0.09	-	65,65,65,65	1
4	FE	E	302	1/1	0.95	0.09	-	70,70,70,70	1
4	FE	O	302	1/1	0.96	0.12	-	68,68,68,68	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.