



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 03:04 PM GMT

PDB ID : 4BET  
Title : Crystal structure of the Legionella pneumophila FIC domain-containing effector AnkX protein (inactive H229A mutant) in complex with cytidine-diphosphate-choline  
Authors : Campanacci, V.; Mukherjee, S.; Roy, C.R.; Cherfils, J.  
Deposited on : 2013-03-12  
Resolution : 2.55 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

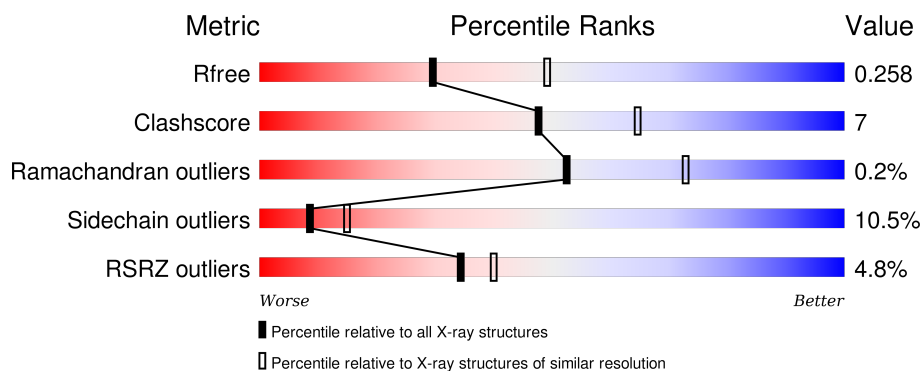
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.55 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))
$R_{free}$	91344	4549 (2.58-2.50)
Clashscore	102246	5292 (2.58-2.50)
Ramachandran outliers	100387	5194 (2.58-2.50)
Sidechain outliers	100360	5196 (2.58-2.50)
RSRZ outliers	91569	4561 (2.58-2.50)

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	512	 6% 73% 19% • 6%
1	B	512	 3% 71% 20% • 6%

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
3	SO4	A	2000	-	-	-	X
3	SO4	A	2001	-	-	-	X
3	SO4	B	2001	-	-	-	X

## 2 Entry composition [i](#)

There are 5 unique types of molecules in this entry. The entry contains 7929 atoms, of which 76 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 PHOSPHOCHOLINE TRANSFERASE ANKX.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	480	Total	C	N	O	S	0	0	0
			3822	2449	636	721	16			
1	B	479	Total	C	N	O	S	0	0	0
			3817	2446	635	720	16			

There are 62 discrepancies between the modelled and reference sequences:

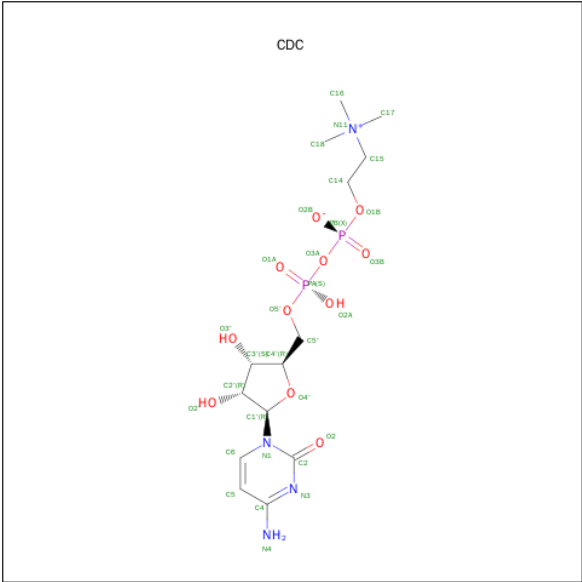
Chain	Residue	Modelled	Actual	Comment	Reference
A	-27	MET	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-26	SER	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-25	TYR	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-24	TYR	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-23	HIS	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-22	HIS	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-21	HIS	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-20	HIS	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-19	HIS	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-18	HIS	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-17	LEU	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-16	GLU	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-15	SER	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-14	THR	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-13	SER	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-12	LEU	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-11	TYR	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-10	LYS	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-9	LYS	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-8	ALA	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-7	GLY	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-6	LEU	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-5	GLU	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-4	ASN	-	EXPRESSION TAG	UNP Q5ZXXN6
A	-3	LEU	-	EXPRESSION TAG	UNP Q5ZXXN6

*Continued on next page...*

*Continued from previous page...*

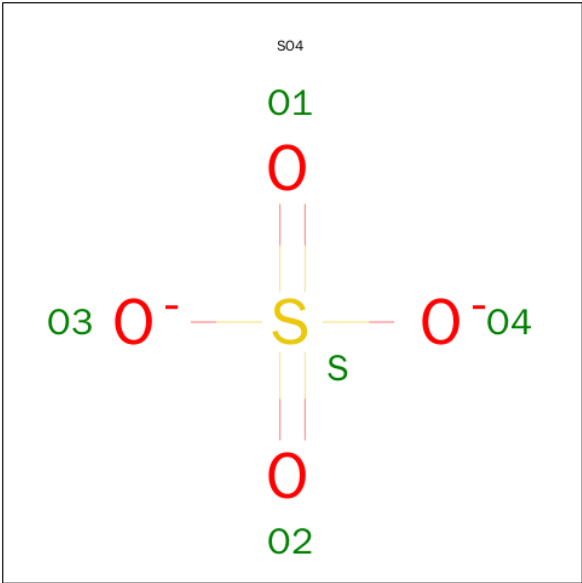
Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	TYR	-	EXPRESSION TAG	UNP Q5ZXX6
A	-1	PHE	-	EXPRESSION TAG	UNP Q5ZXX6
A	0	GLN	-	EXPRESSION TAG	UNP Q5ZXX6
A	1	GLY	-	EXPRESSION TAG	UNP Q5ZXX6
A	229	ALA	HIS	ENGINEERED MUTATION	UNP Q5ZXX6
A	247	PRO	LEU	ENGINEERED MUTATION	UNP Q5ZXX6
B	-27	MET	-	EXPRESSION TAG	UNP Q5ZXX6
B	-26	SER	-	EXPRESSION TAG	UNP Q5ZXX6
B	-25	TYR	-	EXPRESSION TAG	UNP Q5ZXX6
B	-24	TYR	-	EXPRESSION TAG	UNP Q5ZXX6
B	-23	HIS	-	EXPRESSION TAG	UNP Q5ZXX6
B	-22	HIS	-	EXPRESSION TAG	UNP Q5ZXX6
B	-21	HIS	-	EXPRESSION TAG	UNP Q5ZXX6
B	-20	HIS	-	EXPRESSION TAG	UNP Q5ZXX6
B	-19	HIS	-	EXPRESSION TAG	UNP Q5ZXX6
B	-18	HIS	-	EXPRESSION TAG	UNP Q5ZXX6
B	-17	LEU	-	EXPRESSION TAG	UNP Q5ZXX6
B	-16	GLU	-	EXPRESSION TAG	UNP Q5ZXX6
B	-15	SER	-	EXPRESSION TAG	UNP Q5ZXX6
B	-14	THR	-	EXPRESSION TAG	UNP Q5ZXX6
B	-13	SER	-	EXPRESSION TAG	UNP Q5ZXX6
B	-12	LEU	-	EXPRESSION TAG	UNP Q5ZXX6
B	-11	TYR	-	EXPRESSION TAG	UNP Q5ZXX6
B	-10	LYS	-	EXPRESSION TAG	UNP Q5ZXX6
B	-9	LYS	-	EXPRESSION TAG	UNP Q5ZXX6
B	-8	ALA	-	EXPRESSION TAG	UNP Q5ZXX6
B	-7	GLY	-	EXPRESSION TAG	UNP Q5ZXX6
B	-6	LEU	-	EXPRESSION TAG	UNP Q5ZXX6
B	-5	GLU	-	EXPRESSION TAG	UNP Q5ZXX6
B	-4	ASN	-	EXPRESSION TAG	UNP Q5ZXX6
B	-3	LEU	-	EXPRESSION TAG	UNP Q5ZXX6
B	-2	TYR	-	EXPRESSION TAG	UNP Q5ZXX6
B	-1	PHE	-	EXPRESSION TAG	UNP Q5ZXX6
B	0	GLN	-	EXPRESSION TAG	UNP Q5ZXX6
B	1	GLY	-	EXPRESSION TAG	UNP Q5ZXX6
B	229	ALA	HIS	ENGINEERED MUTATION	UNP Q5ZXX6
B	247	PRO	LEU	ENGINEERED MUTATION	UNP Q5ZXX6

- Molecule 2 is [2-CYTIDYLATE-O'-PHOSPHONYLOXYL]-ETHYL-TRIMETHYL-AMMONIUM (three-letter code: CDC) (formula: C<sub>14</sub>H<sub>26</sub>N<sub>4</sub>O<sub>11</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
2	A	1	Total	C	H	N	O	P	0	0
			57	14	26	4	11	2		
2	B	1	Total	C	H	N	O	P	0	0
			57	14	26	4	11	2		

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

Continued on next page...

*Continued from previous page...*

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

- Molecule 4 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
4	A	1	Total	C	H	O	0	0
			14	3	8	3		
4	B	1	Total	C	H	O	0	0
			14	3	8	3		
4	B	1	Total	C	H	O	0	0
			14	3	8	3		

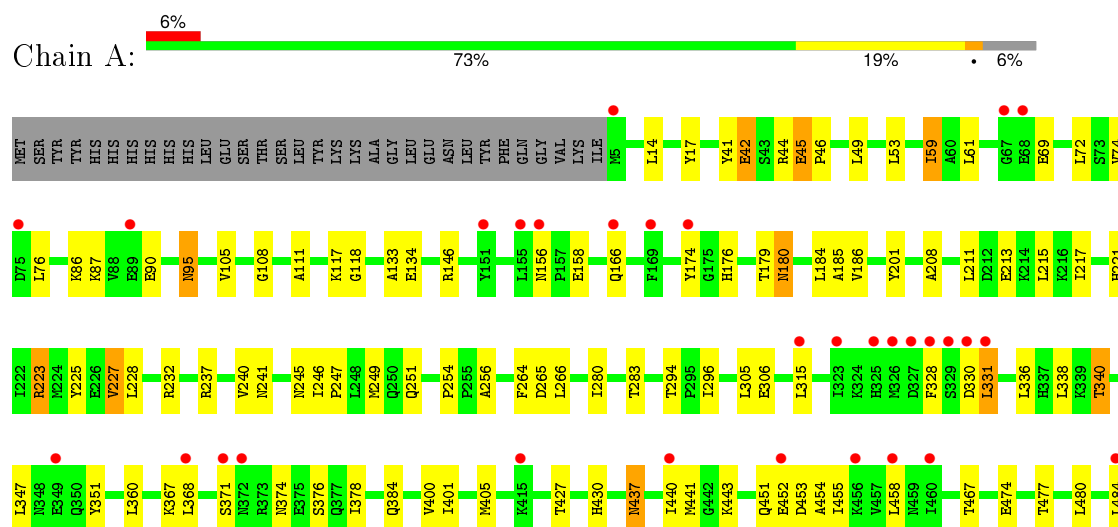
- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	53	Total	O	0	0
			53	53		
5	B	61	Total	O	0	0
			61	61		

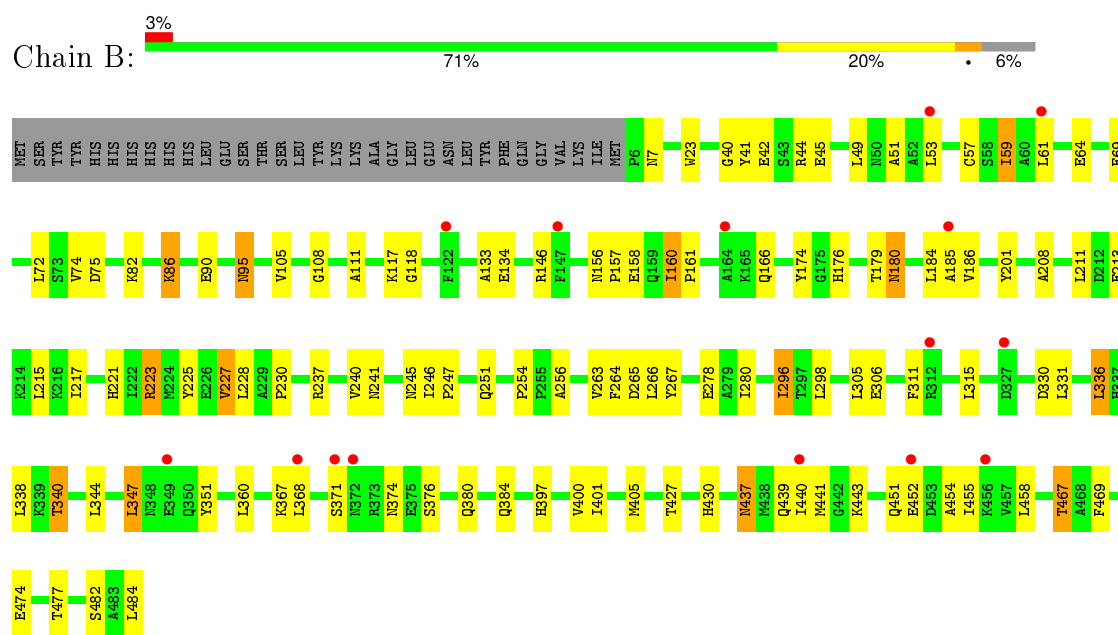
### 3 Residue-property plots

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: PHOSPHOCHOLINE TRANSFERASE ANKX



#### • Molecule 1: PHOSPHOCHOLINE TRANSFERASE ANKX





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	66.10Å 122.15Å 75.70Å 90.00° 107.32° 90.00°	Depositor
Resolution (Å)	43.89 – 2.55 43.89 – 2.55	Depositor EDS
% Data completeness (in resolution range)	98.7 (43.89-2.55) 98.6 (43.89-2.55)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.04 (at 2.54Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, $R_{free}$	0.203 , 0.247 0.210 , 0.258	Depositor DCC
$R_{free}$ test set	1848 reflections (5.27%)	DCC
Wilson B-factor (Å <sup>2</sup> )	49.5	Xtriage
Anisotropy	0.551	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 55.2	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.50$ , $\langle L^2 \rangle = 0.34$	Xtriage
Outliers	0 of 37005 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	7929	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	69.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.97% 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 [i](#)

### 5.1 Standard geometry [i](#)

Bond lengths and bond angles in the following residue types are not validated in this section: GOL, SO4, CDC

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	1/3908 (0.0%)	0.73	1/5282 (0.0%)
1	B	0.51	0/3903	0.72	0/5274
All	All	0.51	1/7811 (0.0%)	0.73	1/10556 (0.0%)

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	46	PRO	N-CD	5.29	1.55	1.47

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	45	GLU	C-N-CD	5.44	139.83	128.40

There are no chirality outliers.

There are no planarity outliers.

### 5.2 Too-close contacts [i](#)

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	3822	0	3802	49	0
1	B	3817	0	3801	65	0
2	A	31	26	25	1	0
2	B	31	26	25	1	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
3	A	10	0	0	0	0
3	B	10	0	0	0	0
4	A	6	8	8	0	0
4	B	12	16	16	0	0
5	A	53	0	0	0	0
5	B	61	0	0	0	0
All	All	7853	76	7677	114	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 7.

All (114) 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:397:HIS:O	1:B:401:ILE:HD12	1.36	1.22
1:A:437:ASN:HB3	1:A:440:ILE:CD1	1.90	1.02
1:B:437:ASN:HB3	1:B:440:ILE:CD1	1.92	0.99
1:A:437:ASN:HB3	1:A:440:ILE:HD11	1.51	0.93
1:A:437:ASN:HB3	1:A:440:ILE:CG1	2.03	0.87
1:B:437:ASN:HB3	1:B:440:ILE:HD11	1.54	0.87
1:B:437:ASN:HB3	1:B:440:ILE:CG1	2.04	0.87
1:B:437:ASN:HB3	1:B:440:ILE:HG12	1.60	0.84
1:B:41:TYR:HA	1:B:44:ARG:NH1	1.93	0.83
1:A:437:ASN:HB3	1:A:440:ILE:HG12	1.59	0.82
1:A:14:LEU:HB2	1:A:249:MET:HE1	1.64	0.80
1:B:42:GLU:HG2	1:B:49:LEU:HD12	1.66	0.78
1:B:397:HIS:C	1:B:401:ILE:HD12	2.07	0.75
1:B:41:TYR:HA	1:B:44:ARG:HH12	1.53	0.73
1:A:14:LEU:HB2	1:A:249:MET:CE	2.20	0.71
1:B:437:ASN:CB	1:B:440:ILE:HG12	2.21	0.71
1:A:437:ASN:CB	1:A:440:ILE:HG12	2.20	0.70
1:B:160:ILE:HG13	1:B:161:PRO:HD3	1.74	0.69
1:B:405:MET:HE1	1:B:443:LYS:HD2	1.76	0.66
1:A:223:ARG:HG2	1:A:264:PHE:O	1.95	0.66
1:B:237:ARG:HA	1:B:241:ASN:HD22	1.61	0.66
1:A:405:MET:HE1	1:A:443:LYS:HD2	1.78	0.65
1:A:237:ARG:HA	1:A:241:ASN:HD22	1.60	0.65
1:B:223:ARG:HG2	1:B:264:PHE:O	1.99	0.63
1:B:40:GLY:O	1:B:44:ARG:HG3	2.02	0.60
1:B:208:ALA:HB1	1:B:213:GLU:HB3	1.84	0.60
1:A:217:ILE:O	1:A:221:HIS:HD2	1.85	0.60

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:245:ASN:ND2	1:A:256:ALA:H	2.00	0.59
1:A:208:ALA:HB1	1:A:213:GLU:HB3	1.84	0.59
1:A:41:TYR:HA	1:A:44:ARG:HH12	1.66	0.59
1:B:217:ILE:O	1:B:221:HIS:HD2	1.85	0.59
1:B:245:ASN:ND2	1:B:256:ALA:H	2.01	0.59
1:A:245:ASN:HD21	1:A:256:ALA:H	1.52	0.57
1:B:245:ASN:HD21	1:B:256:ALA:H	1.53	0.57
1:A:437:ASN:CG	1:A:440:ILE:HG12	2.27	0.56
1:B:228:LEU:O	1:B:230:PRO:HD3	2.06	0.55
1:A:41:TYR:HA	1:A:44:ARG:NH1	2.21	0.55
1:B:439:GLN:HE21	1:B:443:LYS:HZ2	1.55	0.55
1:B:437:ASN:CG	1:B:440:ILE:HG12	2.28	0.54
1:B:397:HIS:O	1:B:401:ILE:CD1	2.31	0.54
1:B:254:PRO:HD2	1:B:280:ILE:HG23	1.90	0.53
1:A:45:GLU:HG2	1:A:90:GLU:HG3	1.90	0.53
1:B:45:GLU:HG2	1:B:90:GLU:HG3	1.89	0.53
1:A:225:TYR:O	1:A:228:LEU:O	2.28	0.52
1:B:157:PRO:HA	1:B:160:ILE:HG23	1.92	0.51
1:A:374:ASN:HD22	1:A:376:SER:H	1.58	0.51
1:B:374:ASN:HD22	1:B:376:SER:H	1.58	0.51
1:B:225:TYR:O	1:B:228:LEU:O	2.28	0.50
1:B:44:ARG:NH1	2:B:1000:CDC:N3	2.59	0.50
1:B:344:LEU:O	1:B:347:LEU:HB2	2.11	0.49
1:B:437:ASN:CB	1:B:440:ILE:CD1	2.80	0.49
1:A:254:PRO:HD2	1:A:280:ILE:HG23	1.94	0.49
1:A:133:ALA:HB2	1:A:186:VAL:HG23	1.95	0.48
1:B:105:VAL:HG13	1:B:185:ALA:HB3	1.95	0.48
1:B:133:ALA:HB2	1:B:186:VAL:HG23	1.96	0.48
1:A:437:ASN:CB	1:A:440:ILE:CD1	2.78	0.48
1:A:111:ALA:HB3	1:A:176:HIS:HA	1.96	0.48
1:B:427:THR:H	1:B:430:HIS:CD2	2.32	0.48
1:B:174:TYR:CD2	1:B:180:ASN:HB2	2.49	0.48
1:A:105:VAL:HG13	1:A:185:ALA:HB3	1.96	0.47
1:B:72:LEU:HD22	1:B:201:TYR:CD2	2.49	0.47
1:A:454:ALA:O	1:A:458:LEU:HG	2.13	0.47
1:A:400:VAL:HG13	1:A:440:ILE:HD12	1.95	0.47
1:A:174:TYR:CD2	1:A:180:ASN:HB2	2.49	0.47
1:A:59:ILE:HD11	1:A:246:ILE:HD11	1.96	0.47
1:A:427:THR:H	1:A:430:HIS:CD2	2.32	0.47
1:A:227:VAL:HG23	1:A:265:ASP:HA	1.97	0.46
1:A:118:GLY:HA3	1:A:266:LEU:O	2.15	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:118:GLY:HA3	1:B:266:LEU:O	2.15	0.46
1:B:227:VAL:HG23	1:B:265:ASP:HA	1.97	0.46
1:A:72:LEU:HD22	1:A:201:TYR:CD2	2.51	0.45
1:B:59:ILE:HD11	1:B:246:ILE:HD11	1.97	0.45
1:B:156:ASN:HD21	1:B:158:GLU:HB2	1.82	0.45
1:B:296:ILE:HD12	1:B:298:LEU:H	1.80	0.45
1:B:134:GLU:HB2	1:B:184:LEU:HB3	1.99	0.45
1:B:440:ILE:HG13	1:B:441:MET:N	2.32	0.45
1:B:400:VAL:HG13	1:B:440:ILE:HD12	1.98	0.45
1:B:57:CYS:O	1:B:61:LEU:HB2	2.17	0.45
1:A:108:GLY:HA3	1:A:179:THR:HG21	2.00	0.44
1:B:454:ALA:O	1:B:458:LEU:HG	2.18	0.44
1:B:397:HIS:C	1:B:401:ILE:CD1	2.82	0.44
1:A:134:GLU:HB2	1:A:184:LEU:HB3	2.00	0.44
1:A:156:ASN:HD21	1:A:158:GLU:HB2	1.82	0.44
1:A:328:PHE:HD1	1:A:331:LEU:HD22	1.81	0.44
1:B:254:PRO:HD3	1:B:340:THR:HG21	2.00	0.43
1:A:440:ILE:HG13	1:A:441:MET:N	2.33	0.43
1:B:263:VAL:O	1:B:267:TYR:HB2	2.18	0.43
1:A:44:ARG:NH1	2:A:1000:CD3:N3	2.66	0.43
1:B:23:TRP:HB3	1:B:311:PHE:CE1	2.54	0.43
1:A:86:LYS:O	1:A:87:LYS:HB2	2.19	0.43
1:B:108:GLY:HA3	1:B:179:THR:HG21	2.00	0.43
1:B:474:GLU:O	1:B:477:THR:HG22	2.19	0.42
1:A:437:ASN:CB	1:A:440:ILE:HD11	2.36	0.42
1:A:427:THR:H	1:A:430:HIS:HD2	1.68	0.42
1:B:59:ILE:HD11	1:B:246:ILE:CD1	2.49	0.42
1:B:247:PRO:O	1:B:251:GLN:HG2	2.19	0.42
1:A:42:GLU:HG3	1:A:49:LEU:HD12	2.01	0.42
1:B:467:THR:HG22	1:B:469:PHE:H	1.83	0.42
1:B:467:THR:HG22	1:B:469:PHE:N	2.35	0.42
1:B:427:THR:H	1:B:430:HIS:HD2	1.68	0.42
1:A:42:GLU:CG	1:A:49:LEU:HD12	2.50	0.42
1:A:336:LEU:HA	1:A:336:LEU:HD12	1.97	0.41
1:A:254:PRO:HD3	1:A:340:THR:HG21	2.01	0.41
1:A:247:PRO:O	1:A:251:GLN:HG2	2.20	0.41
1:B:51:ALA:HA	1:B:86:LYS:HE3	2.02	0.41
1:B:336:LEU:HD12	1:B:336:LEU:HA	1.95	0.41
1:B:437:ASN:CB	1:B:440:ILE:HD11	2.39	0.41
1:A:17:TYR:CD2	1:A:283:THR:HG23	2.56	0.40
1:B:254:PRO:HD2	1:B:280:ILE:CG2	2.51	0.40

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:111:ALA:HB3	1:B:176:HIS:HA	2.03	0.40
1:B:7:ASN:HB2	1:B:64:GLU:HA	2.02	0.40
1:B:439:GLN:HE21	1:B:443:LYS:NZ	2.19	0.40
1:B:108:GLY:HA3	1:B:179:THR:CG2	2.52	0.40
1:A:474:GLU:O	1:A:477:THR:HG22	2.20	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	478/512 (93%)	461 (96%)	16 (3%)	1 (0%)	52	73
1	B	477/512 (93%)	459 (96%)	17 (4%)	1 (0%)	52	73
All	All	955/1024 (93%)	920 (96%)	33 (4%)	2 (0%)	52	73

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	95	ASN
1	A	95	ASN

### 5.3.2 Protein sidechains [i](#)

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

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

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	A	416/447 (93%)	372 (89%)	44 (11%)	8	15
1	B	416/447 (93%)	373 (90%)	43 (10%)	9	15
All	All	832/894 (93%)	745 (90%)	87 (10%)	8	15

All (87) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	42	GLU
1	A	53	LEU
1	A	59	ILE
1	A	61	LEU
1	A	69	GLU
1	A	74	VAL
1	A	76	LEU
1	A	95	ASN
1	A	117	LYS
1	A	146	ARG
1	A	166	GLN
1	A	180	ASN
1	A	211	LEU
1	A	215	LEU
1	A	223	ARG
1	A	227	VAL
1	A	232	ARG
1	A	240	VAL
1	A	294	THR
1	A	296	ILE
1	A	305	LEU
1	A	306	GLU
1	A	315	LEU
1	A	330	ASP
1	A	331	LEU
1	A	338	LEU
1	A	340	THR
1	A	347	LEU
1	A	351	TYR
1	A	360	LEU
1	A	367	LYS
1	A	368	LEU
1	A	371	SER
1	A	378	ILE
1	A	384	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	401	ILE
1	A	437	ASN
1	A	451	GLN
1	A	452	GLU
1	A	453	ASP
1	A	455	ILE
1	A	467	THR
1	A	480	LEU
1	A	484	LEU
1	B	53	LEU
1	B	59	ILE
1	B	69	GLU
1	B	74	VAL
1	B	75	ASP
1	B	82	LYS
1	B	86	LYS
1	B	95	ASN
1	B	117	LYS
1	B	146	ARG
1	B	160	ILE
1	B	166	GLN
1	B	180	ASN
1	B	211	LEU
1	B	215	LEU
1	B	223	ARG
1	B	227	VAL
1	B	240	VAL
1	B	278	GLU
1	B	296	ILE
1	B	305	LEU
1	B	306	GLU
1	B	315	LEU
1	B	330	ASP
1	B	331	LEU
1	B	336	LEU
1	B	338	LEU
1	B	340	THR
1	B	347	LEU
1	B	351	TYR
1	B	360	LEU
1	B	367	LYS
1	B	368	LEU

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	B	371	SER
1	B	380	GLN
1	B	384	GLN
1	B	437	ASN
1	B	451	GLN
1	B	452	GLU
1	B	455	ILE
1	B	467	THR
1	B	482	SER
1	B	484	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (24) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	156	ASN
1	A	221	HIS
1	A	241	ASN
1	A	245	ASN
1	A	251	GLN
1	A	332	ASN
1	A	350	GLN
1	A	374	ASN
1	A	380	GLN
1	A	430	HIS
1	A	437	ASN
1	A	439	GLN
1	A	463	ASN
1	B	156	ASN
1	B	221	HIS
1	B	241	ASN
1	B	245	ASN
1	B	251	GLN
1	B	348	ASN
1	B	374	ASN
1	B	384	GLN
1	B	430	HIS
1	B	439	GLN
1	B	463	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

9 ligands are modelled in this entry.

In the following table, the Counts columns list the number of bonds (or angles) for which Mogul statistics could be retrieved, the number of bonds (or angles) that are observed in the model and the number of bonds (or angles) that are defined in the chemical component dictionary. The Link column lists molecule types, if any, to which the group is linked. The Z score for a bond length (or angle) is the number of standard deviations the observed value is removed from the expected value. A bond length (or angle) with  $|Z| > 2$  is considered an outlier worth inspection. RMSZ is the root-mean-square of all Z scores of the bond lengths (or angles).

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\# Z  > 2$	Counts	RMSZ	$\# Z  > 2$
2	CDC	A	1000	-	25,32,32	0.74	1 (4%)	35,49,49	1.48	2 (5%)
3	SO4	A	2000	-	4,4,4	0.16	0	6,6,6	0.10	0
3	SO4	A	2001	-	4,4,4	0.19	0	6,6,6	0.09	0
4	GOL	A	3000	-	5,5,5	0.16	0	5,5,5	0.31	0
2	CDC	B	1000	-	25,32,32	0.76	1 (4%)	35,49,49	1.44	2 (5%)
3	SO4	B	2000	-	4,4,4	0.13	0	6,6,6	0.08	0
3	SO4	B	2001	-	4,4,4	0.16	0	6,6,6	0.09	0
4	GOL	B	3000	-	5,5,5	0.19	0	5,5,5	0.20	0
4	GOL	B	3001	-	5,5,5	0.10	0	5,5,5	0.41	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
2	CDC	A	1000	-	-	0/20/40/40	0/2/2/2

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
3	SO4	A	2000	-	-	0/0/0/0	0/0/0/0
3	SO4	A	2001	-	-	0/0/0/0	0/0/0/0
4	GOL	A	3000	-	-	0/4/4/4	0/0/0/0
2	CDC	B	1000	-	-	0/20/40/40	0/2/2/2
3	SO4	B	2000	-	-	0/0/0/0	0/0/0/0
3	SO4	B	2001	-	-	0/0/0/0	0/0/0/0
4	GOL	B	3000	-	-	0/4/4/4	0/0/0/0
4	GOL	B	3001	-	-	0/4/4/4	0/0/0/0

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	1000	CDC	C6-N1	2.23	1.38	1.35
2	A	1000	CDC	C6-N1	2.28	1.39	1.35

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	1000	CDC	C2-N3-C4	3.68	120.80	115.61
2	B	1000	CDC	C2-N3-C4	3.74	120.89	115.61
2	B	1000	CDC	PB-O3A-PA	6.79	151.79	132.73
2	A	1000	CDC	PB-O3A-PA	7.46	153.68	132.73

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
2	A	1000	CDC	1	0
2	B	1000	CDC	1	0

## 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 ⓘ

### 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	480/512 (93%)	0.54	31 (6%) 22 25	48, 68, 101, 125	1 (0%)
1	B	479/512 (93%)	0.44	15 (3%) 52 58	44, 64, 96, 125	1 (0%)
All	All	959/1024 (93%)	0.49	46 (4%) 34 40	44, 66, 99, 125	2 (0%)

All (46) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	330	ASP	6.4
1	A	349	GLU	4.5
1	A	328	PHE	4.1
1	A	372	ASN	4.0
1	B	372	ASN	3.9
1	A	323	ILE	3.8
1	A	368	LEU	3.3
1	B	349	GLU	3.3
1	A	331	LEU	3.2
1	A	329	SER	3.2
1	A	325	HIS	3.1
1	A	155	LEU	3.0
1	A	440	ILE	2.8
1	A	169	PHE	2.7
1	A	166	GLN	2.7
1	A	371	SER	2.7
1	A	326	MET	2.6
1	A	5	MET	2.5
1	A	460	ILE	2.5
1	B	327	ASP	2.4
1	A	452	GLU	2.4
1	B	371	SER	2.4
1	B	122	PHE	2.4
1	A	327	ASP	2.3

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	456	LYS	2.2
1	A	156	ASN	2.2
1	A	67	GLY	2.2
1	B	61	LEU	2.2
1	B	452	GLU	2.1
1	A	315	LEU	2.1
1	A	456	LYS	2.1
1	B	312	ARG	2.1
1	B	440	ILE	2.1
1	B	147	PHE	2.1
1	B	164	ALA	2.1
1	B	185	ALA	2.1
1	A	75	ASP	2.1
1	A	458	LEU	2.1
1	A	415	LYS	2.1
1	A	151	TYR	2.1
1	A	174	TYR	2.0
1	B	368	LEU	2.0
1	A	89	GLU	2.0
1	A	484	LEU	2.0
1	B	53	LEU	2.0
1	A	68	GLU	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 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( $\text{\AA}^2$ )	Q<0.9
3	SO4	A	2000	5/5	0.81	0.27	5.08	146,147,147,147	0
3	SO4	A	2001	5/5	0.73	0.33	4.21	154,154,154,155	0
3	SO4	B	2001	5/5	0.82	0.26	2.71	140,140,141,141	0
3	SO4	B	2000	5/5	0.90	0.19	0.96	122,122,122,122	0
2	CDC	A	1000	31/31	0.95	0.18	0.20	54,56,64,64	0
2	CDC	B	1000	31/31	0.97	0.17	-0.13	47,54,59,63	0
4	GOL	A	3000	6/6	0.92	0.17	-0.22	60,61,65,65	0
4	GOL	B	3000	6/6	0.95	0.16	-1.31	64,64,64,64	0
4	GOL	B	3001	6/6	0.76	0.30	-	77,78,79,79	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.