



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 10:32 PM GMT

PDB ID : 4Y73  
Title : Crystal structure of IRAK4 kinase domain with inhibitor  
Authors : Lesburg, C.A.  
Deposited on : 2015-02-13  
Resolution : 2.14 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

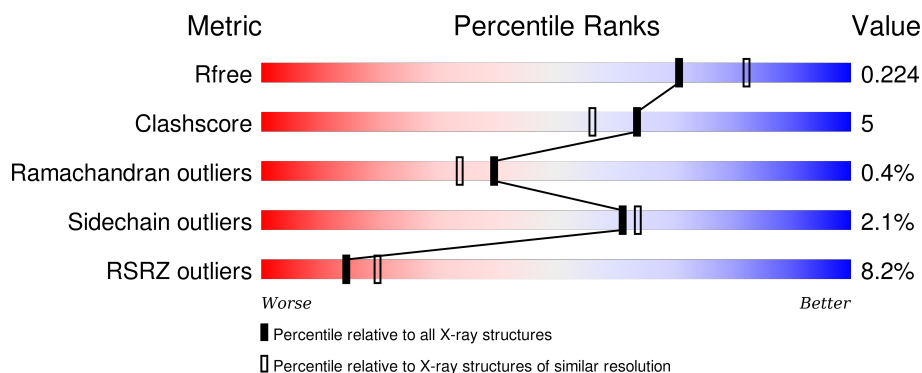
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.14 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	1693 (2.16-2.12)
Clashscore	102246	1824 (2.16-2.12)
Ramachandran outliers	100387	1798 (2.16-2.12)
Sidechain outliers	100360	1798 (2.16-2.12)
RSRZ outliers	91569	1699 (2.16-2.12)

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	302	<div> <div>6%</div> <div>89%</div> <div>6%</div> <div>• •</div> </div>
1	B	302	<div> <div>11%</div> <div>87%</div> <div>9%</div> <div>•</div> </div>
1	C	302	<div> <div>6%</div> <div>86%</div> <div>10%</div> <div>• •</div> </div>
1	D	302	<div> <div>9%</div> <div>86%</div> <div>9%</div> <div>• •</div> </div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 9744 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.

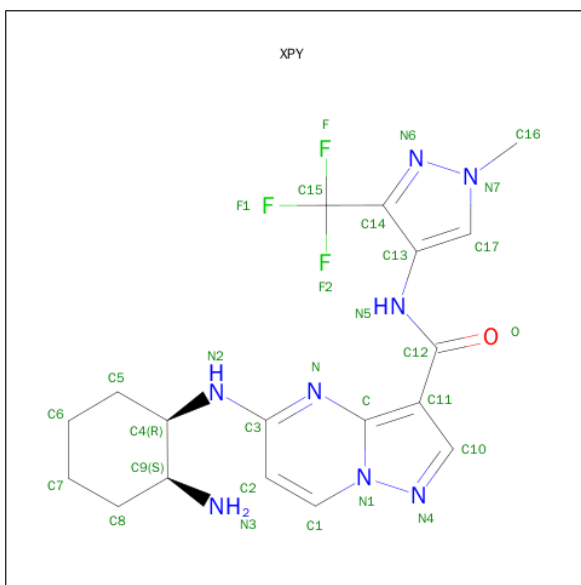
- Molecule 1 is a protein called Interleukin-1 receptor-associated kinase 4.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	292	Total	C	N	O	P	S	0	1	0
			2310	1446	387	458	3	16			
1	B	292	Total	C	N	O	P	S	0	0	0
			2305	1442	386	459	3	15			
1	C	293	Total	C	N	O	P	S	0	0	0
			2310	1445	387	460	3	15			
1	D	291	Total	C	N	O	P	S	0	1	0
			2304	1442	386	458	3	15			

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	159	GLY	-	expression tag	UNP Q9NWZ3
B	159	GLY	-	expression tag	UNP Q9NWZ3
C	159	GLY	-	expression tag	UNP Q9NWZ3
D	159	GLY	-	expression tag	UNP Q9NWZ3

- Molecule 2 is 5-[[[(1R,2S)-2-aminocyclohexyl]amino]-N-[1-methyl-3-(trifluoromethyl)-1H-pyrazol-4-yl]pyrazolo[1,5-a]pyrimidine-3-carboxamide (three-letter code: XPY) (formula: C<sub>18</sub>H<sub>21</sub>F<sub>3</sub>N<sub>8</sub>O).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total	C	F	N	O	0	0
			30	18	3	8	1		
2	B	1	Total	C	F	N	O	0	0
			30	18	3	8	1		
2	C	1	Total	C	F	N	O	0	0
			30	18	3	8	1		
2	D	1	Total	C	F	N	O	0	0
			30	18	3	8	1		

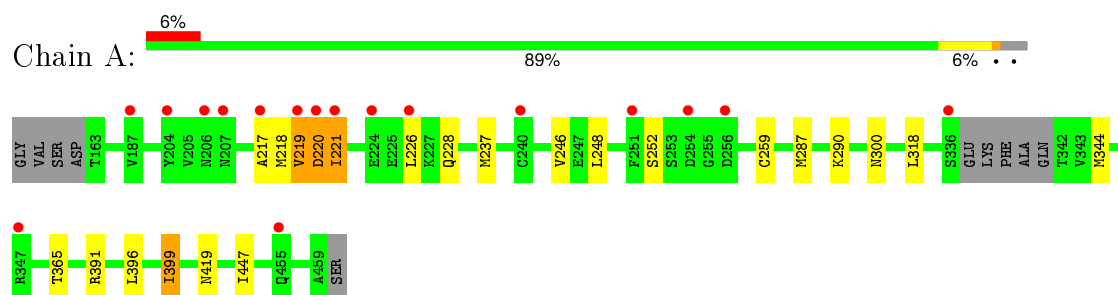
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	98	Total	O	0	0
			98	98		
3	B	86	Total	O	0	0
			86	86		
3	C	105	Total	O	0	0
			105	105		
3	D	106	Total	O	0	0
			106	106		

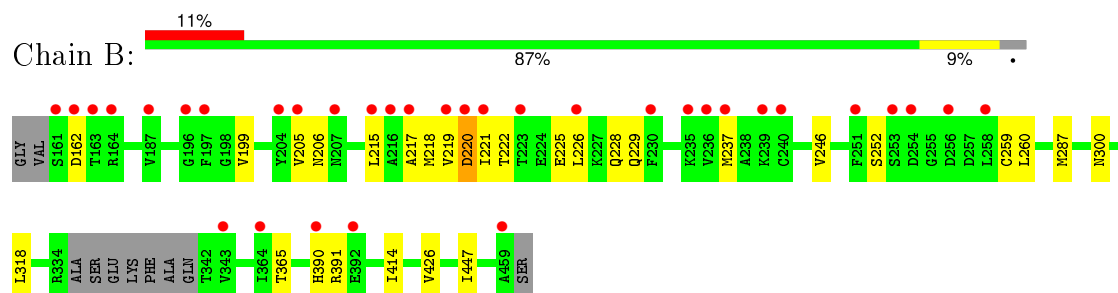
### 3 Residue-property plots [i](#)

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 ( $\text{RSRZ} > 2$ ). 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.

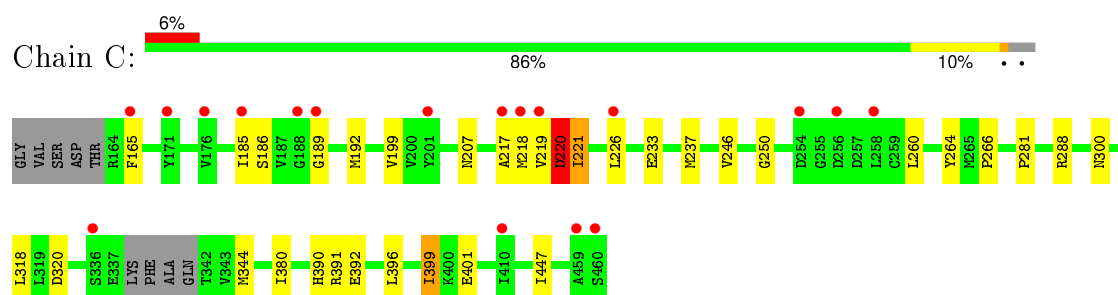
- Molecule 1: Interleukin-1 receptor-associated kinase 4



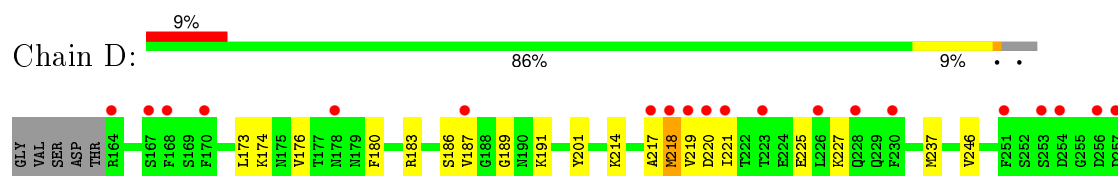
- Molecule 1: Interleukin-1 receptor-associated kinase 4

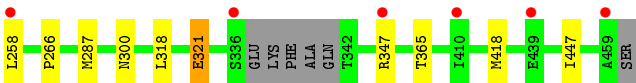


- Molecule 1: Interleukin-1 receptor-associated kinase 4



- Molecule 1: Interleukin-1 receptor-associated kinase 4





## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	142.86 Å   138.75 Å   88.02 Å 90.00°   124.59°   90.00°	Depositor
Resolution (Å)	74.24 – 2.14 74.23 – 2.14	Depositor EDS
% Data completeness (in resolution range)	99.2 (74.24-2.14) 99.4 (74.23-2.14)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.10 (at 2.14 Å)	Xtriage
Refinement program	BUSTER 2.11.4	Depositor
R, $R_{free}$	0.195   ,   0.217 0.201   ,   0.224	Depositor DCC
$R_{free}$ test set	3869 reflections (5.29%)	DCC
Wilson B-factor (Å <sup>2</sup> )	38.7	Xtriage
Anisotropy	0.504	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 52.3	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 76998 reflections	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	9744	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	55.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 17.93% 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: TPO, XPY, SEP

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	0.49	0/2315	0.66	1/3120 (0.0%)
1	B	0.50	0/2310	0.66	1/3114 (0.0%)
1	C	0.49	0/2315	0.67	0/3120
1	D	0.50	0/2309	0.67	0/3112
All	All	0.49	0/9249	0.66	2/12466 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	220	ASP	C-N-CA	5.12	134.51	121.70
1	A	220	ASP	C-N-CA	5.04	134.30	121.70

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	2310	0	2274	19	0
1	B	2305	0	2265	29	0
1	C	2310	0	2270	28	0
1	D	2304	0	2264	30	0
2	A	30	0	21	0	0

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	30	0	21	0	0
2	C	30	0	21	0	0
2	D	30	0	21	0	0
3	A	98	0	0	1	0
3	B	86	0	0	1	0
3	C	105	0	0	0	0
3	D	106	0	0	0	0
All	All	9744	0	9157	97	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 5.

All (97) 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:221:ILE:HG21	1:B:226:LEU:HD11	1.34	1.08
1:D:217:ALA:HA	1:D:218:MET:HB3	1.33	1.06
1:B:222:THR:HB	1:B:225:GLU:HG3	1.44	0.95
1:B:217:ALA:HA	1:B:218:MET:CB	1.95	0.95
1:C:217:ALA:HA	1:C:218:MET:HB2	1.49	0.95
1:C:220:ASP:HB3	1:C:221:ILE:HA	1.50	0.93
1:A:217:ALA:HA	1:A:218:MET:CB	1.98	0.93
1:D:266:PRO:HG2	1:D:321[A]:GLU:HG3	1.55	0.89
1:B:221:ILE:HG21	1:B:226:LEU:CD1	2.02	0.89
1:B:221:ILE:CG2	1:B:226:LEU:CD1	2.56	0.83
1:B:217:ALA:HA	1:B:218:MET:HB3	1.61	0.82
1:D:180:PHE:CD2	1:D:214:LYS:HD2	2.15	0.80
1:A:217:ALA:HA	1:A:218:MET:HB3	1.63	0.80
1:B:221:ILE:CG2	1:B:226:LEU:HD11	2.12	0.80
1:B:217:ALA:HA	1:B:218:MET:HB2	1.75	0.69
1:B:390:HIS:O	1:C:391:ARG:HA	1.94	0.67
1:D:217:ALA:HA	1:D:218:MET:CB	2.15	0.66
1:B:287:MET:HG3	3:B:1116:HOH:O	1.96	0.65
1:D:237:MET:HE1	1:D:246:VAL:HG23	1.78	0.65
1:A:217:ALA:HA	1:A:218:MET:HB2	1.76	0.65
1:C:220:ASP:HB2	1:C:221:ILE:HG13	1.79	0.65
1:B:391:ARG:HA	1:C:390:HIS:O	1.98	0.64
1:D:287:MET:HA	1:D:287:MET:HE3	1.78	0.63
1:A:217:ALA:HB3	1:A:221:ILE:CG2	2.28	0.62
1:D:174:LYS:HG2	1:D:180:PHE:HE1	1.63	0.62
1:B:217:ALA:HB3	1:B:221:ILE:HG21	1.82	0.61

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:220:ASP:HB2	1:A:221:ILE:HG13	1.81	0.61
1:D:217:ALA:HB3	1:D:221:ILE:HG21	1.81	0.60
1:C:217:ALA:HA	1:C:218:MET:CB	2.30	0.60
1:B:246:VAL:HG11	1:B:318:LEU:HD12	1.84	0.60
1:B:217:ALA:HB3	1:B:221:ILE:CG2	2.33	0.59
1:A:217:ALA:HB3	1:A:221:ILE:HG21	1.84	0.58
1:C:233:GLU:HG2	1:C:260:LEU:HD13	1.85	0.58
1:C:217:ALA:HB3	1:C:221:ILE:HG21	1.85	0.58
1:C:220:ASP:CB	1:C:221:ILE:HA	2.30	0.57
1:B:221:ILE:HG23	1:B:226:LEU:CD1	2.35	0.56
1:B:220:ASP:HB2	1:B:221:ILE:HG13	1.88	0.56
1:C:281:PRO:HD3	1:D:321[B]:GLU:HG2	1.88	0.56
1:B:220:ASP:HB2	1:B:221:ILE:HA	1.89	0.55
1:D:220:ASP:OD2	1:D:225:GLU:OE2	2.24	0.55
1:A:419:ASN:HB2	1:D:186:SER:CB	2.36	0.55
1:C:207:ASN:HA	1:D:187:VAL:HG23	1.90	0.54
1:C:220:ASP:HB3	1:C:221:ILE:CA	2.32	0.53
1:C:396:LEU:HD12	1:C:399:ILE:HD13	1.91	0.52
1:D:246:VAL:HG11	1:D:318:LEU:HD12	1.92	0.52
1:A:396:LEU:HD12	1:A:399:ILE:HD13	1.92	0.52
1:A:220:ASP:HB2	1:A:221:ILE:HA	1.92	0.51
1:D:173:LEU:HA	1:D:176:VAL:HG22	1.92	0.51
1:C:237:MET:HE2	1:C:246:VAL:HG23	1.93	0.51
1:B:221:ILE:CG2	1:B:226:LEU:HD13	2.40	0.51
1:A:419:ASN:HB2	1:D:186:SER:HB3	1.92	0.50
1:D:217:ALA:HB3	1:D:221:ILE:CG2	2.41	0.49
1:B:221:ILE:HG23	1:B:226:LEU:HD13	1.94	0.48
1:D:174:LYS:HG2	1:D:180:PHE:CE1	2.46	0.48
1:B:225:GLU:O	1:B:228:GLN:HB3	2.14	0.48
1:C:217:ALA:HB3	1:C:221:ILE:CG2	2.43	0.48
1:A:246:VAL:HG11	1:A:318:LEU:HD12	1.96	0.48
1:D:183:ARG:HB2	1:D:189:GLY:HA3	1.95	0.48
1:A:237:MET:HG2	1:A:248:LEU:CB	2.44	0.47
1:C:220:ASP:CB	1:C:221:ILE:HG13	2.44	0.47
1:A:391:ARG:HD3	3:A:1162:HOH:O	2.14	0.47
1:D:227:LYS:HG2	1:D:258:LEU:HD11	1.97	0.47
1:D:217:ALA:CA	1:D:218:MET:HB3	2.24	0.46
1:B:237:MET:HE2	1:B:246:VAL:HG23	1.96	0.46
1:B:237:MET:CE	1:B:246:VAL:HG23	2.45	0.46
1:B:215:LEU:HD11	1:B:260:LEU:HD12	1.97	0.46
1:D:237:MET:CE	1:D:246:VAL:HG23	2.47	0.45

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:300:ASN:HA	1:B:447:ILE:HG21	1.99	0.44
1:D:183:ARG:HB3	1:D:187:VAL:CG1	2.48	0.44
1:A:300:ASN:HA	1:A:447:ILE:HG21	1.99	0.44
1:B:390:HIS:HB3	1:C:392:GLU:H	1.83	0.44
1:C:185:ILE:HD13	1:C:189:GLY:O	2.17	0.43
1:D:287:MET:HA	1:D:287:MET:CE	2.46	0.43
1:A:237:MET:HG2	1:A:248:LEU:HB2	2.00	0.43
1:D:217:ALA:HB1	1:D:219:VAL:H	1.83	0.43
1:C:192:MET:SD	1:C:264:TYR:HE1	2.41	0.42
1:A:287[B]:MET:CE	1:A:290:LYS:HD3	2.49	0.42
1:C:281:PRO:CD	1:D:321[B]:GLU:HG2	2.49	0.42
1:C:300:ASN:HA	1:C:447:ILE:HG21	2.01	0.42
1:B:217:ALA:CA	1:B:218:MET:CB	2.82	0.42
1:B:414:ILE:HG12	1:B:426:VAL:HG11	2.01	0.42
1:C:246:VAL:HG11	1:C:318:LEU:HD12	2.02	0.42
1:D:191:LYS:HA	1:D:201:TYR:HD1	1.85	0.42
1:A:252:SER:HB3	1:A:259:CYS:HB2	2.02	0.42
1:A:287[B]:MET:HE1	1:A:290:LYS:HD3	2.01	0.42
1:C:165:PHE:HB3	1:C:250:GLY:HA2	2.02	0.42
1:C:221:ILE:HG21	1:C:226:LEU:HD13	2.02	0.41
1:B:252:SER:HB3	1:B:259:CYS:HB2	2.00	0.41
1:C:281:PRO:HD3	1:D:321[B]:GLU:CG	2.49	0.41
1:D:300:ASN:HA	1:D:447:ILE:HG21	2.01	0.41
1:C:237:MET:CE	1:C:246:VAL:HG23	2.51	0.41
1:B:205:VAL:HG23	1:B:206:ASN:H	1.86	0.41
1:C:266:PRO:HD2	1:C:320:ASP:HA	2.04	0.41
1:D:217:ALA:CA	1:D:218:MET:CB	2.94	0.40
1:D:321[A]:GLU:H	1:D:321[A]:GLU:CD	2.25	0.40
1:C:288:ARG:HB3	1:C:380:ILE:HG23	2.03	0.40
1:A:221:ILE:HG21	1:A:226:LEU:HD13	2.03	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	287/302 (95%)	277 (96%)	9 (3%)	1 (0%)	46	41
1	B	286/302 (95%)	272 (95%)	13 (4%)	1 (0%)	46	41
1	C	287/302 (95%)	275 (96%)	10 (4%)	2 (1%)	26	17
1	D	286/302 (95%)	272 (95%)	13 (4%)	1 (0%)	46	41
All	All	1146/1208 (95%)	1096 (96%)	45 (4%)	5 (0%)	39	33

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	219	VAL
1	B	219	VAL
1	C	219	VAL
1	C	220	ASP
1	D	218	MET

### 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	252/259 (97%)	246 (98%)	6 (2%)	57	58
1	B	252/259 (97%)	248 (98%)	4 (2%)	70	74
1	C	252/259 (97%)	245 (97%)	7 (3%)	51	50
1	D	251/259 (97%)	246 (98%)	5 (2%)	63	66
All	All	1007/1036 (97%)	985 (98%)	22 (2%)	61	62

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	219	VAL
1	A	221	ILE
1	A	228	GLN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	A	344	MET
1	A	365	THR
1	A	399	ILE
1	B	162	ASP
1	B	199	VAL
1	B	229	GLN
1	B	365	THR
1	C	186	SER
1	C	199	VAL
1	C	220	ASP
1	C	221	ILE
1	C	344	MET
1	C	399	ILE
1	C	401	GLU
1	D	321[A]	GLU
1	D	321[B]	GLU
1	D	347	ARG
1	D	365	THR
1	D	418	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	D	232	GLN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

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

12 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	TPO	A	342	1	8,10,11	1.03	0	7,14,16	1.24	1 (14%)
1	TPO	A	345	1	8,10,11	1.18	1 (12%)	7,14,16	1.32	1 (14%)
1	SEP	A	346	1	8,9,10	0.61	0	8,12,14	1.39	2 (25%)
1	TPO	B	342	1	8,10,11	0.93	0	7,14,16	0.90	0
1	TPO	B	345	1	8,10,11	0.99	0	7,14,16	1.39	1 (14%)
1	SEP	B	346	1	8,9,10	0.99	0	8,12,14	1.09	0
1	TPO	C	342	1	8,10,11	0.91	0	7,14,16	1.05	0
1	TPO	C	345	1	8,10,11	1.32	2 (25%)	7,14,16	1.76	2 (28%)
1	SEP	C	346	1	8,9,10	0.83	0	8,12,14	1.37	2 (25%)
1	TPO	D	342	1	8,10,11	0.98	0	7,14,16	1.07	0
1	TPO	D	345	1	8,10,11	1.12	1 (12%)	7,14,16	1.23	1 (14%)
1	SEP	D	346	1	8,9,10	0.87	0	8,12,14	1.35	1 (12%)

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	TPO	A	342	1	-	0/8/11/13	0/0/0/0
1	TPO	A	345	1	-	1/8/11/13	0/0/0/0
1	SEP	A	346	1	-	0/6/8/10	0/0/0/0
1	TPO	B	342	1	-	0/8/11/13	0/0/0/0
1	TPO	B	345	1	-	1/8/11/13	0/0/0/0
1	SEP	B	346	1	-	0/6/8/10	0/0/0/0
1	TPO	C	342	1	-	0/8/11/13	0/0/0/0
1	TPO	C	345	1	-	1/8/11/13	0/0/0/0
1	SEP	C	346	1	-	0/6/8/10	0/0/0/0
1	TPO	D	342	1	-	0/8/11/13	0/0/0/0
1	TPO	D	345	1	-	1/8/11/13	0/0/0/0
1	SEP	D	346	1	-	0/6/8/10	0/0/0/0

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	345	TPO	CG2-CB	2.01	1.56	1.51
1	D	345	TPO	CB-CA	2.02	1.57	1.54
1	A	345	TPO	CB-CA	2.33	1.58	1.54
1	C	345	TPO	CB-CA	2.63	1.58	1.54

All (11) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	346	SEP	O-C-CA	-2.45	119.10	125.49
1	C	346	SEP	O-C-CA	-2.36	119.33	125.49
1	A	345	TPO	O-C-CA	-2.18	119.67	125.44
1	A	346	SEP	O-C-CA	-2.17	119.84	125.49
1	C	345	TPO	O-C-CA	-2.13	119.82	125.44
1	D	345	TPO	O-C-CA	-2.11	119.87	125.44
1	B	345	TPO	O-C-CA	-2.08	119.94	125.44
1	A	346	SEP	OG-CB-CA	2.21	110.16	108.27
1	C	346	SEP	OG-CB-CA	2.35	110.28	108.27
1	A	342	TPO	OG1-P-O1P	2.37	113.04	107.11
1	C	345	TPO	OG1-P-O1P	3.73	116.42	107.11

There are no chirality outliers.

All (4) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	B	345	TPO	OG1-CB-CA-N
1	C	345	TPO	OG1-CB-CA-N
1	D	345	TPO	OG1-CB-CA-N
1	A	345	TPO	OG1-CB-CA-N

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

4 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	XPY	A	1000	-	25,33,33	1.85	4 (16%)	26,49,49	1.44	5 (19%)
2	XPY	B	1000	-	25,33,33	1.86	5 (20%)	26,49,49	1.59	6 (23%)
2	XPY	C	1000	-	25,33,33	1.84	3 (12%)	26,49,49	1.35	3 (11%)
2	XPY	D	1000	-	25,33,33	1.96	5 (20%)	26,49,49	1.53	5 (19%)

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	XPY	A	1000	-	-	0/14/29/29	0/4/4/4
2	XPY	B	1000	-	-	0/14/29/29	0/4/4/4
2	XPY	C	1000	-	-	0/14/29/29	0/4/4/4
2	XPY	D	1000	-	-	0/14/29/29	0/4/4/4

All (17) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	1000	XPY	C17-C13	2.05	1.41	1.37
2	A	1000	XPY	C2-C3	2.10	1.44	1.39
2	A	1000	XPY	C3-N	2.30	1.36	1.32
2	B	1000	XPY	C3-N	2.36	1.36	1.32
2	B	1000	XPY	C17-C13	2.51	1.42	1.37
2	D	1000	XPY	C2-C3	2.55	1.45	1.39
2	C	1000	XPY	C3-N	2.58	1.37	1.32
2	B	1000	XPY	C2-C3	2.68	1.45	1.39
2	D	1000	XPY	C3-N	2.76	1.37	1.32
2	B	1000	XPY	C4-N2	4.08	1.50	1.45
2	D	1000	XPY	C4-N2	4.14	1.50	1.45
2	A	1000	XPY	C4-N2	4.25	1.51	1.45
2	C	1000	XPY	C4-N2	4.31	1.51	1.45
2	B	1000	XPY	C3-N2	6.11	1.48	1.36
2	C	1000	XPY	C3-N2	6.50	1.49	1.36
2	A	1000	XPY	C3-N2	6.75	1.50	1.36
2	D	1000	XPY	C3-N2	7.07	1.50	1.36

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	1000	XPY	C2-C3-N2	-3.82	114.24	121.18

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1000	XPY	C2-C3-N2	-3.79	114.29	121.18
2	A	1000	XPY	C2-C3-N2	-3.38	115.03	121.18
2	C	1000	XPY	C2-C3-N2	-3.36	115.08	121.18
2	B	1000	XPY	C15-C14-C13	-2.93	126.16	129.12
2	D	1000	XPY	C15-C14-C13	-2.83	126.27	129.12
2	B	1000	XPY	C5-C4-N2	-2.81	106.18	111.19
2	C	1000	XPY	C5-C4-N2	-2.79	106.22	111.19
2	A	1000	XPY	C5-C4-N2	-2.69	106.40	111.19
2	A	1000	XPY	C15-C14-C13	-2.57	126.53	129.12
2	D	1000	XPY	C3-N2-C4	-2.08	121.45	124.15
2	B	1000	XPY	C2-C3-N	2.01	126.50	123.55
2	B	1000	XPY	C15-C14-N6	2.43	122.66	119.68
2	A	1000	XPY	C15-C14-N6	2.45	122.69	119.68
2	D	1000	XPY	C15-C14-N6	2.46	122.69	119.68
2	A	1000	XPY	C10-N4-N1	2.55	105.61	103.70
2	D	1000	XPY	C10-N4-N1	3.15	106.06	103.70
2	C	1000	XPY	C10-N4-N1	3.28	106.16	103.70
2	B	1000	XPY	C10-N4-N1	3.66	106.45	103.70

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

## 5.7 Other polymers [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	289/302 (95%)	0.54	17 (5%) 26 34	31, 52, 91, 108	0
1	B	289/302 (95%)	0.83	34 (11%) 6 9	28, 51, 94, 120	0
1	C	290/302 (96%)	0.59	18 (6%) 24 31	27, 48, 95, 112	0
1	D	288/302 (95%)	0.69	26 (9%) 12 17	28, 50, 95, 121	0
All	All	1156/1208 (95%)	0.67	95 (8%) 14 20	27, 50, 94, 121	0

All (95) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	217	ALA	11.0
1	B	221	ILE	7.6
1	B	220	ASP	6.4
1	D	221	ILE	6.1
1	A	254	ASP	6.1
1	D	253	SER	5.8
1	C	256	ASP	5.7
1	B	163	THR	5.5
1	C	217	ALA	5.3
1	D	217	ALA	5.0
1	B	240	CYS	4.9
1	C	218	MET	4.8
1	B	216	ALA	4.5
1	D	254	ASP	4.5
1	B	205	VAL	4.4
1	B	239	LYS	4.4
1	A	240	CYS	4.1
1	B	226	LEU	4.1
1	A	221	ILE	4.1
1	A	207	ASN	4.0
1	D	220	ASP	3.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	251	PHE	3.7
1	A	219	VAL	3.7
1	D	164	ARG	3.6
1	D	459	ALA	3.6
1	C	410	ILE	3.6
1	A	220	ASP	3.4
1	C	460	SER	3.4
1	B	219	VAL	3.4
1	D	219	VAL	3.4
1	D	226	LEU	3.4
1	B	254	ASP	3.4
1	B	237	MET	3.4
1	A	217	ALA	3.3
1	B	204	TYR	3.3
1	B	207	ASN	3.1
1	C	185	ILE	3.1
1	B	343	VAL	3.0
1	B	161	SER	3.0
1	A	187	VAL	3.0
1	D	167	SER	2.9
1	A	204	TYR	2.9
1	B	230	PHE	2.9
1	D	223	THR	2.9
1	D	347	ARG	2.8
1	D	168	PHE	2.8
1	B	187	VAL	2.8
1	C	254	ASP	2.8
1	A	347	ARG	2.8
1	A	206	ASN	2.8
1	D	258	LEU	2.8
1	B	251	PHE	2.7
1	B	196	GLY	2.7
1	B	258	LEU	2.7
1	D	170	PHE	2.7
1	D	439	GLU	2.7
1	B	215	LEU	2.6
1	D	256	ASP	2.6
1	D	410	ILE	2.6
1	C	219	VAL	2.6
1	C	459	ALA	2.6
1	D	218	MET	2.6
1	B	253	SER	2.6

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	B	236	VAL	2.5
1	B	459	ALA	2.5
1	C	171	TYR	2.5
1	B	256	ASP	2.4
1	C	165	PHE	2.4
1	B	235	LYS	2.4
1	A	336	SER	2.4
1	C	201	TYR	2.4
1	B	390	HIS	2.4
1	B	164	ARG	2.3
1	C	258	LEU	2.3
1	B	223	THR	2.3
1	D	228	GLN	2.3
1	C	226	LEU	2.2
1	B	197	PHE	2.2
1	C	176	VAL	2.2
1	A	251	PHE	2.2
1	C	336	SER	2.2
1	C	188	GLY	2.1
1	D	178	ASN	2.1
1	D	230	PHE	2.1
1	D	257	ASP	2.1
1	B	392	GLU	2.1
1	B	364	ILE	2.1
1	A	224	GLU	2.1
1	C	189	GLY	2.1
1	A	455	GLN	2.1
1	A	226	LEU	2.1
1	D	336	SER	2.1
1	A	256	ASP	2.1
1	D	187	VAL	2.1
1	B	162	ASP	2.0

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

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
1	TPO	C	345	11/12	0.96	0.11	-	57,60,63,65	0
1	TPO	A	345	11/12	0.91	0.13	-	80,82,85,86	0
1	SEP	D	346	10/11	0.75	0.18	-	73,80,88,89	0
1	SEP	B	346	10/11	0.72	0.14	-	79,83,90,90	0
1	SEP	C	346	10/11	0.89	0.11	-	66,75,84,85	0
1	SEP	A	346	10/11	0.84	0.15	-	83,87,93,93	0
1	TPO	B	342	11/12	0.89	0.12	-	88,89,92,92	0
1	TPO	B	345	11/12	0.91	0.13	-	75,77,81,82	0
1	TPO	D	345	11/12	0.97	0.12	-	67,69,71,73	0
1	TPO	D	342	11/12	0.80	0.14	-	76,77,84,85	0
1	TPO	A	342	11/12	0.70	0.18	-	95,96,101,101	0
1	TPO	C	342	11/12	0.69	0.20	-	79,85,94,95	0

### 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
2	XPY	C	1000	30/30	0.97	0.14	0.70	27,35,42,42	0
2	XPY	A	1000	30/30	0.95	0.13	0.65	29,40,48,50	0
2	XPY	B	1000	30/30	0.97	0.13	-0.18	25,33,39,42	0
2	XPY	D	1000	30/30	0.98	0.13	-0.33	26,39,44,49	0

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

There are no such residues in this entry.