



# wwPDB X-ray Structure Validation Summary Report ⓘ

Mar 9, 2018 – 11:41 am GMT

PDB ID : 4JD3  
Title : Crystal Structure of Mycobacterium tuberculosis PKS11 Reveals Intermediates in the Synthesis of Methyl-branched Alkylpyrones  
Authors : Gokulan, K.; Sacchettini, J.C.; Mycobacterium Tuberculosis Structural Proteomics Project (XMTB)  
Deposited on : 2013-02-22  
Resolution : 2.25 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

<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 : 1.7.3 (157068), CSD as539be (2018)  
Xtriage (Phenix) : 1.13  
EDS : trunk30967  
Percentile statistics : 20171227.v01 (using entries in the PDB archive December 27th 2017)  
Refmac : 5.8.0158  
CCP4 : 7.0 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : trunk30967

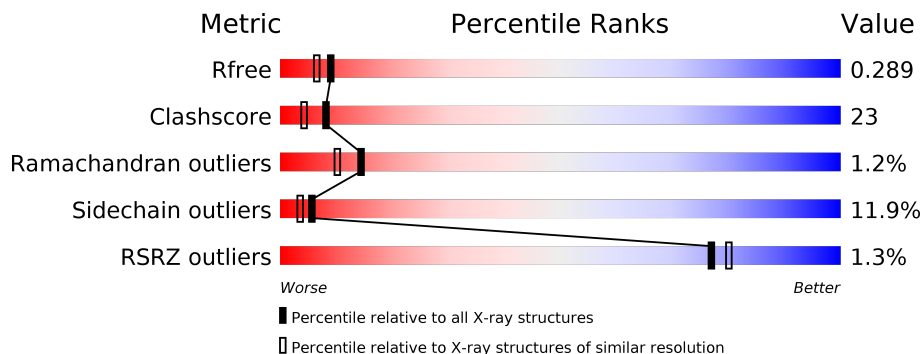
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.25 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))
$R_{free}$	111664	1178 (2.26-2.26)
Clashscore	122126	1286 (2.26-2.26)
Ramachandran outliers	120053	1253 (2.26-2.26)
Sidechain outliers	120020	1254 (2.26-2.26)
RSRZ outliers	108989	1158 (2.26-2.26)

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	353	
1	B	353	
1	C	353	
1	D	353	

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 crite-

ria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	PLM	A	400	-	-	X	-
2	PLM	B	401	-	-	X	-

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10842 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 Alpha-pyrone synthesis polyketide synthase-like Pks11.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	352	Total	C	N	O	S	0	0	0
			2642	1671	469	495	7			
1	C	352	Total	C	N	O	S	0	0	0
			2642	1671	469	495	7			
1	B	352	Total	C	N	O	S	0	1	0
			2652	1677	472	496	7			
1	D	352	Total	C	N	O	S	0	0	0
			2642	1671	469	495	7			

There are 4 discrepancies between the modelled and reference sequences:

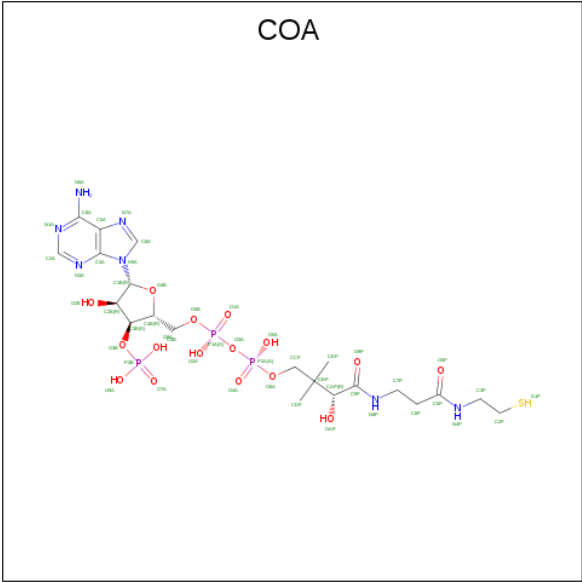
Chain	Residue	Modelled	Actual	Comment	Reference
A	138	SER	CYS	ENGINEERED MUTATION	UNP O06587
C	138	SER	CYS	ENGINEERED MUTATION	UNP O06587
B	138	SER	CYS	ENGINEERED MUTATION	UNP O06587
D	138	SER	CYS	ENGINEERED MUTATION	UNP O06587

- Molecule 2 is PALMITIC ACID (three-letter code: PLM) (formula: C<sub>16</sub>H<sub>32</sub>O<sub>2</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	A	1	Total	C	O	0	0
			18	16	2		
2	C	1	Total	C	O	0	0
			18	16	2		
2	B	1	Total	C	O	0	0
			18	16	2		
2	D	1	Total	C	O	0	0
			18	16	2		

- Molecule 3 is COENZYME A (three-letter code: COA) (formula: C<sub>21</sub>H<sub>36</sub>N<sub>7</sub>O<sub>16</sub>P<sub>3</sub>S).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
3	C	1	Total 48	C 21	N 7	O 16	P 3	S 1	0	0
3	B	1	Total 48	C 21	N 7	O 16	P 3	S 1	0	0

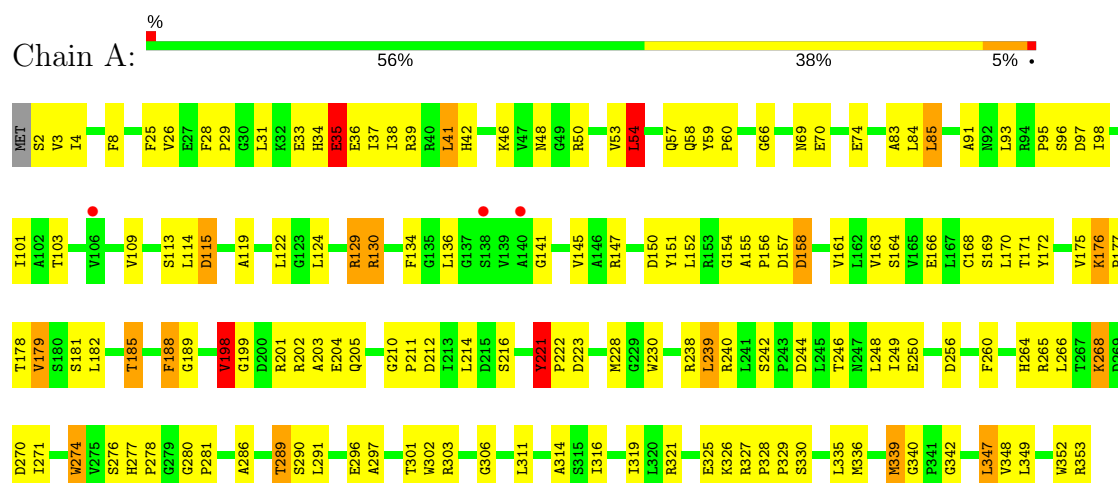
- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	20	Total 20	O 20	0	0
4	C	25	Total 25	O 25	0	0
4	B	30	Total 30	O 30	0	0
4	D	21	Total 21	O 21	0	0

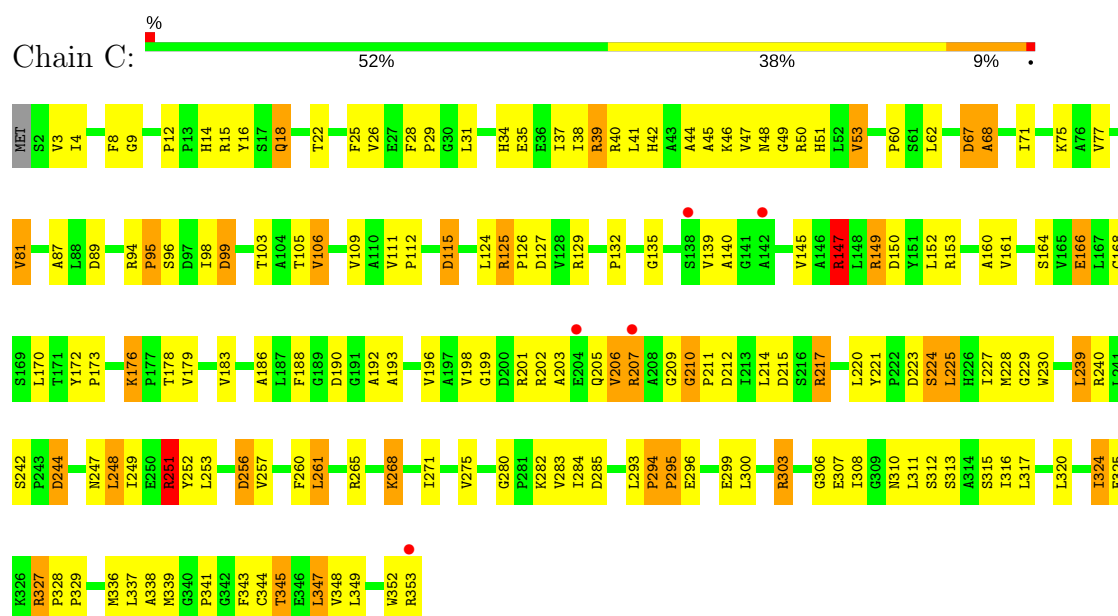
### 3 Residue-property plots

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: Alpha-pyrone synthesis polyketide synthase-like Pks11

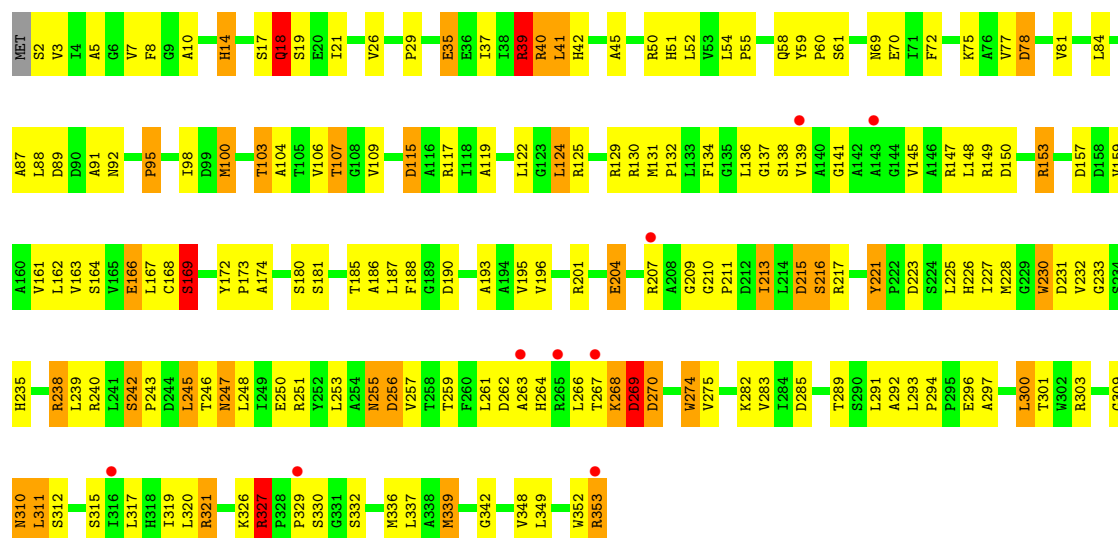


- Molecule 1: Alpha-pyrone synthesis polyketide synthase-like Pks11



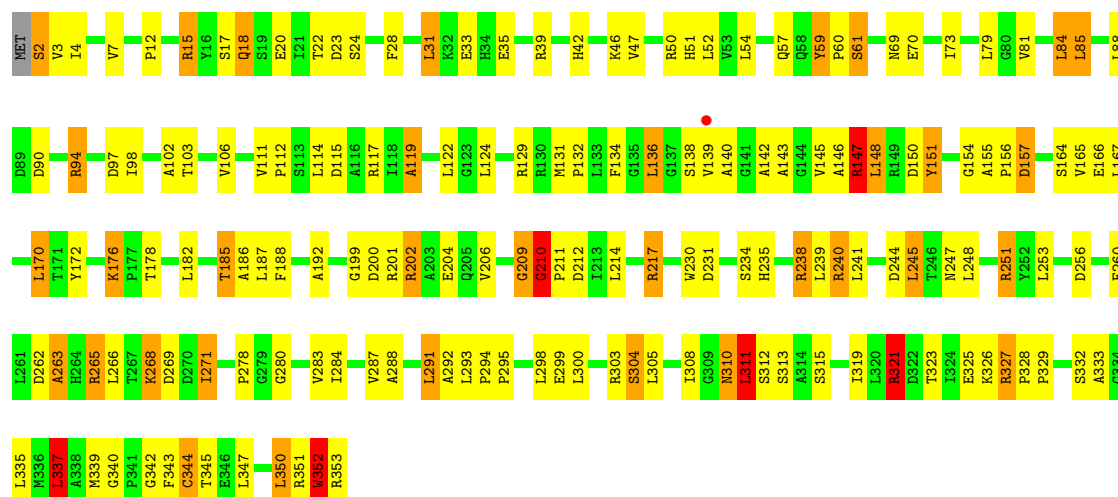
- Molecule 1: Alpha-pyrone synthesis polyketide synthase-like Pks11





● Molecule 1: Alpha-pyrone synthesis polyketide synthase-like Pks11

Chain D: 53% 36% 10% ●





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	72.17Å 48.86Å 194.47Å 90.00° 97.82° 90.00°	Depositor
Resolution (Å)	32.11 – 2.25 32.12 – 2.25	Depositor EDS
% Data completeness (in resolution range)	(Not available) (32.11-2.25) 68.6 (32.12-2.25)	Depositor EDS
$R_{merge}$	0.16	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.01 (at 2.24Å)	Xtriage
Refinement program	PHENIX	Depositor
R, $R_{free}$	0.194 , 0.294 0.201 , 0.289	Depositor DCC
$R_{free}$ test set	2255 reflections (5.08%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	38.5	Xtriage
Anisotropy	0.039	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 25.4	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.011 for h,-k,-h-l	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	10842	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.50% 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: COA, PLM

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	1.43	15/2695 (0.6%)	1.41	21/3671 (0.6%)
1	B	1.44	7/2706 (0.3%)	1.44	38/3686 (1.0%)
1	C	1.51	19/2695 (0.7%)	1.47	32/3671 (0.9%)
1	D	1.45	12/2695 (0.4%)	1.39	34/3671 (0.9%)
All	All	1.46	53/10791 (0.5%)	1.43	125/14699 (0.9%)

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	D	0	1
All	All	0	3

The worst 5 of 53 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	166	GLU	CD-OE2	-6.28	1.18	1.25
1	C	244	ASP	CB-CG	5.97	1.64	1.51
1	C	16	TYR	CE1-CZ	5.96	1.46	1.38
1	A	35	GLU	CG-CD	5.80	1.60	1.51
1	B	221	TYR	CE1-CZ	5.80	1.46	1.38

The worst 5 of 125 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	149	ARG	NE-CZ-NH1	-15.71	112.45	120.30

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	67	ASP	CB-CG-OD1	13.27	130.24	118.30
1	D	202	ARG	NE-CZ-NH1	-11.67	114.47	120.30
1	C	251	ARG	NE-CZ-NH1	-11.14	114.73	120.30
1	B	238	ARG	NE-CZ-NH1	10.65	125.62	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	326	LYS	Peptide
1	C	239	LEU	Peptide
1	D	352	TRP	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	2642	0	2675	112	0
1	B	2652	0	2681	142	0
1	C	2642	0	2675	134	0
1	D	2642	0	2673	126	1
2	A	18	0	31	11	0
2	B	18	0	31	9	0
2	C	18	0	31	3	0
2	D	18	0	31	5	0
3	B	48	0	32	4	0
3	C	48	0	32	5	1
4	A	20	0	0	0	0
4	B	30	0	0	5	0
4	C	25	0	0	3	0
4	D	21	0	0	2	0
All	All	10842	0	10892	490	1

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 23.

The worst 5 of 490 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:C:224:SER:HB3	1:C:343:PHE:H	1.10	1.11
1:B:353:ARG:HH11	1:B:353:ARG:HG2	1.12	1.07
1:B:107:THR:HG21	1:B:167:LEU:H	1.00	1.06
1:C:106:VAL:HG11	2:C:401:PLM:H21	1.36	1.04
1:C:34:HIS:HB3	1:C:37:ILE:HD11	1.37	1.03

All (1) 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)
1:D:176:LYS:NZ	3:C:402:COA:O8A[2_555]	1.52	0.68

## 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	350/353 (99%)	320 (91%)	27 (8%)	3 (1%)	19	16
1	B	351/353 (99%)	317 (90%)	29 (8%)	5 (1%)	12	8
1	C	350/353 (99%)	313 (89%)	33 (9%)	4 (1%)	16	11
1	D	350/353 (99%)	314 (90%)	31 (9%)	5 (1%)	12	8
All	All	1401/1412 (99%)	1264 (90%)	120 (9%)	17 (1%)	14	10

5 of 17 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	327	ARG
1	B	169	SER
1	A	210	GLY
1	C	96	SER
1	C	210	GLY

### 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	280/281 (100%)	253 (90%)	27 (10%)	9	7
1	B	281/281 (100%)	241 (86%)	40 (14%)	3	2
1	C	280/281 (100%)	257 (92%)	23 (8%)	12	11
1	D	280/281 (100%)	237 (85%)	43 (15%)	3	1
All	All	1121/1124 (100%)	988 (88%)	133 (12%)	6	3

5 of 133 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	103	THR
1	B	242	SER
1	D	300	LEU
1	B	107	THR
1	B	180	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 23 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	18	GLN
1	B	51	HIS
1	D	247	ASN
1	B	34	HIS
1	B	58	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 carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

6 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	PLM	A	400	-	14,17,17	0.44	0	13,17,17	0.65	0
2	PLM	B	401	-	14,17,17	0.78	0	13,17,17	1.67	3 (23%)
3	COA	B	402	-	41,50,50	1.19	3 (7%)	51,75,75	1.90	14 (27%)
2	PLM	C	401	-	14,17,17	0.53	0	13,17,17	1.83	3 (23%)
3	COA	C	402	-	41,50,50	1.21	3 (7%)	51,75,75	1.70	11 (21%)
2	PLM	D	400	-	14,17,17	0.64	0	13,17,17	1.82	3 (23%)

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	PLM	A	400	-	-	0/13/15/15	0/0/0/0
2	PLM	B	401	-	-	0/13/15/15	0/0/0/0
3	COA	B	402	-	-	0/44/64/64	0/3/3/3
2	PLM	C	401	-	-	0/13/15/15	0/0/0/0
3	COA	C	402	-	-	0/44/64/64	0/3/3/3
2	PLM	D	400	-	-	0/13/15/15	0/0/0/0

The worst 5 of 6 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	C	402	COA	C8A-N9A	-3.32	1.32	1.36
3	B	402	COA	C8A-N9A	-2.95	1.33	1.36

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	402	COA	O5P-C5P	-2.84	1.17	1.23
3	C	402	COA	C5A-N7A	-2.36	1.31	1.39
3	B	402	COA	P3B-O8A	-2.19	1.46	1.54

The worst 5 of 34 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	402	COA	N3A-C2A-N1A	-7.58	122.37	128.86
3	C	402	COA	N3A-C2A-N1A	-5.04	124.55	128.86
2	D	400	PLM	C9-C8-C7	-3.97	92.85	114.41
2	C	401	PLM	C6-C5-C4	-3.75	94.08	114.41
3	C	402	COA	O3B-C3B-C4B	-3.44	97.39	110.05

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 38 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	400	PLM	11	0
2	B	401	PLM	9	0
3	B	402	COA	4	0
2	C	401	PLM	3	0
3	C	402	COA	5	1
2	D	400	PLM	5	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 [i](#)

### 6.1 Protein, DNA and RNA chains [i](#)

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	352/353 (99%)	-0.15	3 (0%) 84 86	21, 36, 48, 61	0
1	B	352/353 (99%)	-0.14	9 (2%) 56 60	21, 35, 50, 66	0
1	C	352/353 (99%)	-0.20	5 (1%) 75 78	21, 34, 49, 70	0
1	D	352/353 (99%)	-0.18	1 (0%) 93 94	21, 35, 50, 65	0
All	All	1408/1412 (99%)	-0.17	18 (1%) 77 80	21, 35, 49, 70	0

The worst 5 of 18 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	267	THR	3.1
1	A	138	SER	3.0
1	B	329	PRO	3.0
1	C	207	ARG	3.0
1	B	265	ARG	2.9

### 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. 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
3	COA	C	402	48/48	0.89	0.24	24,37,54,63	0
3	COA	B	402	48/48	0.90	0.26	24,37,50,56	0
2	PLM	A	400	18/18	0.92	0.25	29,36,46,47	0
2	PLM	D	400	18/18	0.92	0.20	25,33,39,41	0
2	PLM	B	401	18/18	0.93	0.16	28,37,50,53	0
2	PLM	C	401	18/18	0.93	0.17	29,31,36,39	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.