Full wwPDB X-ray Structure Validation Report (i)

Jun 7, 2025 – 01:05 pm BST

PDB ID	:	9FKU / pdb_00009fku
Title	:	Crystal Structure of AimR from Katmira phage
Authors	:	Gallego del Sol, F.; Marina, A.
Deposited on	:	2024-06-04
Resolution	:	1.90 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The types of validation reports are described at
http://www.wwpdb.org/validation/2017/FAQs#types.

The following versions of software and data (see references (1)) were used in the production of this report:

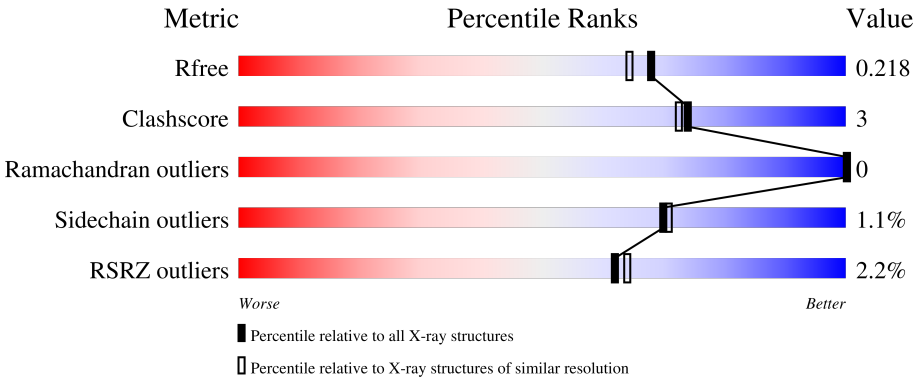
MolProbity	:	4-5-2 with Phenix2.0rc1
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	2.0rc1
EDS	:	3.0
Percentile statistics	:	20231227.v01 (using entries in the PDB archive December 27th 2023)
CCP4	:	9.0.003 (Gargrove)
Density-Fitness	:	1.0.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.43.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 1.90 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}	164625	7293 (1.90-1.90)
Clashscore	180529	8090 (1.90-1.90)
Ramachandran outliers	177936	8022 (1.90-1.90)
Sidechain outliers	177891	8022 (1.90-1.90)
RSRZ outliers	164620	7292 (1.90-1.90)

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 of the lower bar indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria respectively. A grey segment represents the fraction of residues that are not modelled. The numeric value for each fraction is indicated below the corresponding segment, with a dot representing fractions <=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	386	3% 91% 8%
1	B	386	% 93% 7%
2	C	6	50% 33% 17%
2	D	6	83% 17%

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
4	ACT	A	402	-	-	X	-

2 Entry composition

There are 5 unique types of molecules in this entry. The entry contains 6875 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.

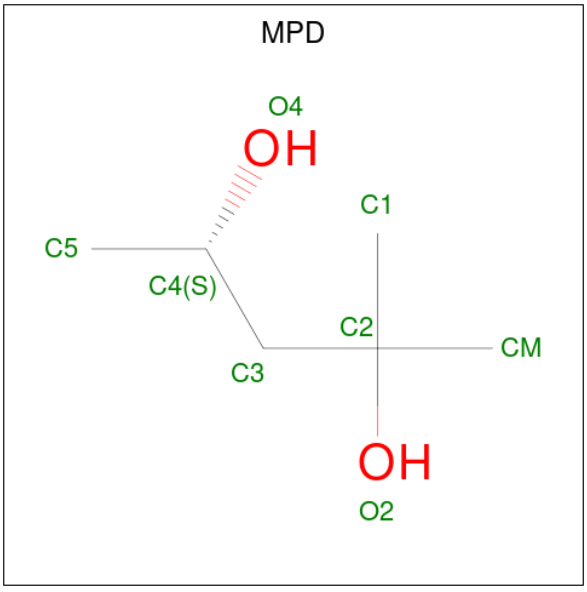
- Molecule 1 is a protein called Arbitrium receptor.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	386	Total	C	N	O	S	0	2	0
			3188	2043	520	605	20			
1	B	386	Total	C	N	O	S	0	0	0
			3177	2036	519	602	20			

- Molecule 2 is a protein called GLY-VAL-VAL-ARG-GLY-ALA.

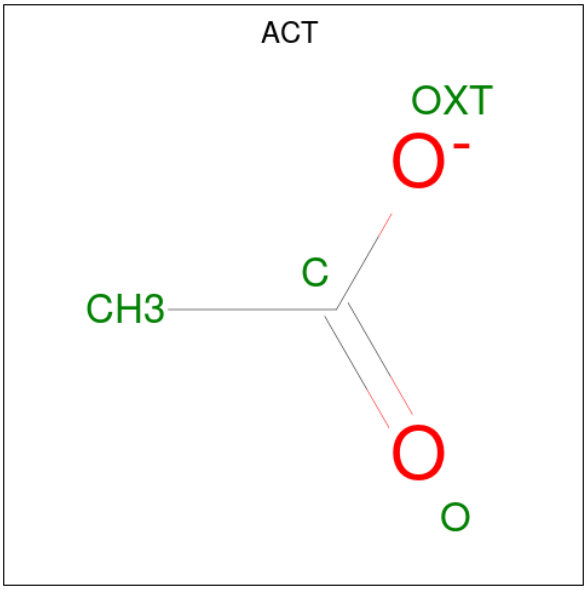
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	C	6	Total	C	N	O	0	0	0
			39	23	9	7			
2	D	6	Total	C	N	O	0	0	0
			39	23	9	7			

- Molecule 3 is (4S)-2-METHYL-2,4-PENTANEDIOL (CCD ID: MPD) (formula: $C_6H_{14}O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C O 8 6 2	0	0
3	B	1	Total C O 8 6 2	0	0
3	B	1	Total C O 8 6 2	0	0

- Molecule 4 is ACETATE ION (CCD ID: ACT) (formula: $C_2H_3O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 4 2 2	0	0
4	A	1	Total C O 4 2 2	0	0
4	B	1	Total C O 4 2 2	0	0

- Molecule 5 is water.

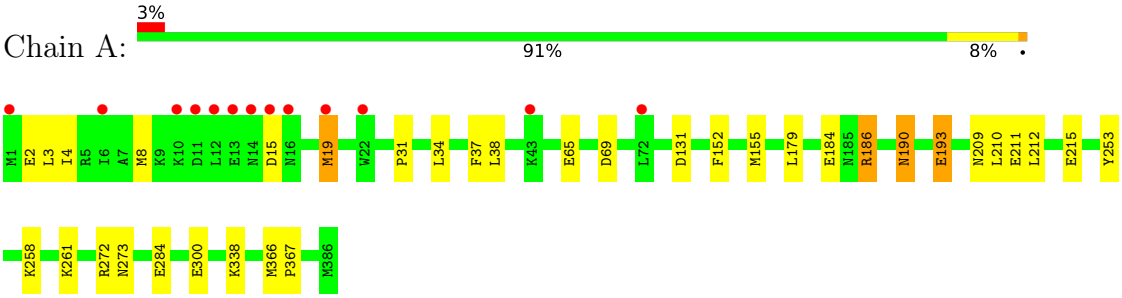
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	191	Total O 191 191	0	0
5	B	196	Total O 196 196	0	0
5	C	5	Total O 5 5	0	0
5	D	4	Total O 4 4	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

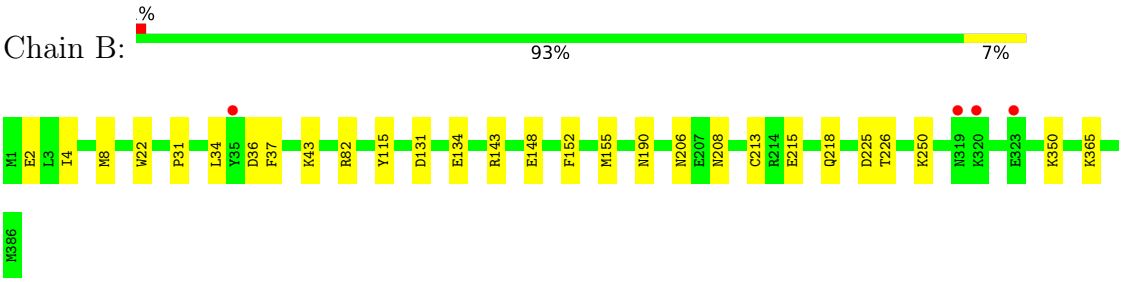
• Molecule 1: Arbitrium receptor

Chain A:



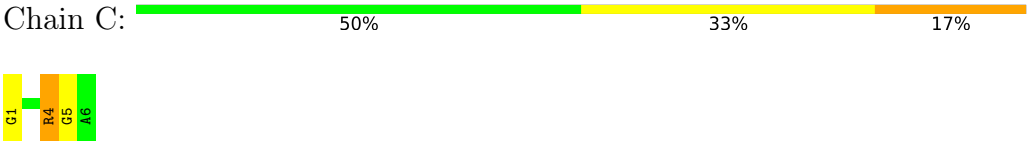
• Molecule 1: Arbitrium receptor

Chain B:



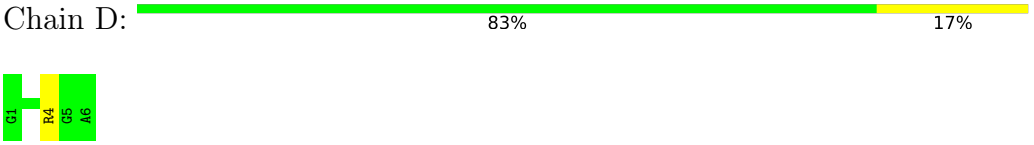
• Molecule 2: GLY-VAL-VAL-ARG-GLY-ALA

Chain C:



• Molecule 2: GLY-VAL-VAL-ARG-GLY-ALA

Chain D:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	69.77Å 203.58Å 143.07Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	48.69 – 1.90 48.69 – 1.90	Depositor EDS
% Data completeness (in resolution range)	100.0 (48.69-1.90) 100.0 (48.69-1.90)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.56 (at 1.90Å)	Xtriage
Refinement program	REFMAC 5.8.0425	Depositor
R, R_{free}	0.178 , 0.212 0.188 , 0.218	Depositor DCC
R_{free} test set	4067 reflections (5.06%)	wwPDB-VP
Wilson B-factor (Å^2)	33.6	Xtriage
Anisotropy	0.176	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 32.3	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6875	wwPDB-VP
Average B, all atoms (Å^2)	40.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.07% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: ACT, MPD

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	0.77	2/3253 (0.1%)	1.18	9/4369 (0.2%)
1	B	0.70	1/3236 (0.0%)	1.16	4/4346 (0.1%)
2	C	2.03	2/38 (5.3%)	1.31	0/48
2	D	1.61	0/38	1.09	0/48
All	All	0.76	5/6565 (0.1%)	1.17	13/8811 (0.1%)

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	2

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	C	5	GLY	C-O	-6.29	1.17	1.24
2	C	4	ARG	C-O	-5.83	1.16	1.23
1	A	215	GLU	C-O	-5.78	1.17	1.24
1	A	211	GLU	C-O	-5.18	1.18	1.24
1	B	213	CYS	C-O	-5.14	1.18	1.24

All (13) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	31	PRO	N-CA-C	6.97	122.61	114.03
1	B	31	PRO	N-CA-C	6.68	122.22	114.20
1	A	210	LEU	N-CA-C	6.26	118.91	111.33
1	B	225	ASP	CA-CB-CG	6.13	118.73	112.60

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	19	MET	CG-SD-CE	-5.94	87.83	100.90
1	A	15	ASP	CA-CB-CG	5.91	118.51	112.60
1	A	193	GLU	CB-CG-CD	5.69	122.28	112.60
1	A	284	GLU	CB-CA-C	5.66	118.73	109.90
1	A	131	ASP	CA-CB-CG	5.65	118.25	112.60
1	A	261	LYS	CB-CA-C	-5.57	101.97	109.71
1	A	190	ASN	OD1-CG-ND2	5.25	127.85	122.60
1	B	131	ASP	CA-CB-CG	5.10	117.70	112.60
1	B	226	THR	CA-CB-OG1	-5.07	102.00	109.60

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	186	ARG	Sidechain
1	A	272	ARG	Sidechain

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	3188	0	3177	18	0
1	B	3177	0	3165	18	0
2	C	39	0	44	2	0
2	D	39	0	44	3	0
3	A	8	0	14	0	0
3	B	16	0	28	2	0
4	A	8	0	6	2	0
4	B	4	0	3	0	0
5	A	191	0	0	0	0
5	B	196	0	0	4	0
5	C	5	0	0	1	0
5	D	4	0	0	1	0
All	All	6875	0	6481	39	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 3.

All (39) close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:C:4:ARG:HD2	5:C:101:HOH:O	1.77	0.82
1:B:8:MET:HE1	1:B:37:PHE:HE2	1.47	0.80
1:B:250:LYS:HD2	3:B:402:MPD:HM3	1.67	0.75
1:A:4:ILE:O	1:A:8:MET:HG3	1.85	0.75
3:B:402:MPD:HM1	5:B:521:HOH:O	1.90	0.71
1:B:215:GLU:OE1	5:B:501:HOH:O	2.10	0.70
2:D:4:ARG:HD2	5:D:103:HOH:O	1.91	0.69
1:B:134:GLU:OE1	5:B:502:HOH:O	2.14	0.64
1:A:8:MET:HE3	1:A:38:LEU:HD21	1.81	0.62
1:B:206:ASN:HD21	2:D:4:ARG:HE	1.48	0.61
1:B:152:PHE:CD1	1:B:155:MET:HE3	2.37	0.60
1:A:338:LYS:HE2	4:A:402:ACT:H2	1.85	0.58
1:A:338:LYS:CE	4:A:402:ACT:H2	2.33	0.57
1:A:3:LEU:HD13	1:A:69:ASP:CG	2.29	0.56
1:A:152:PHE:CD1	1:A:155:MET:HE3	2.41	0.56
1:A:209[A]:ASN:ND2	1:A:212:LEU:HD12	2.23	0.53
1:B:208:ASN:HD21	1:B:365:LYS:NZ	2.07	0.52
1:A:19:MET:HE1	1:A:34:LEU:HB3	1.92	0.52
1:B:152:PHE:HD1	1:B:155:MET:HE3	1.74	0.52
1:B:218:GLN:NE2	5:B:503:HOH:O	2.42	0.52
1:B:8:MET:HE1	1:B:37:PHE:CE2	2.37	0.49
1:B:115:TYR:OH	1:B:148:GLU:HG3	2.13	0.48
1:B:4:ILE:HG13	1:B:8:MET:HE2	1.95	0.48
1:A:8:MET:HE2	1:A:37:PHE:HE2	1.79	0.48
1:A:152:PHE:HD1	1:A:155:MET:HE3	1.80	0.47
1:B:208:ASN:HD21	1:B:365:LYS:HZ3	1.62	0.47
1:A:253:TYR:CZ	1:A:273:ASN:HB3	2.51	0.46
1:A:8:MET:CE	1:A:37:PHE:HE2	2.29	0.46
1:A:179:LEU:HD13	1:A:193:GLU:HG3	1.98	0.46
1:B:206:ASN:HD21	2:D:4:ARG:NE	2.14	0.45
1:A:19:MET:CE	1:A:34:LEU:HD23	2.47	0.44
1:B:143:ARG:HG2	1:B:143:ARG:HH11	1.83	0.43
1:A:8:MET:HE3	1:A:38:LEU:CD2	2.48	0.43
1:B:36:ASP:HB3	1:B:43:LYS:HG2	2.01	0.42
1:B:22:TRP:HB2	1:B:34:LEU:HD21	2.01	0.42
1:A:3:LEU:HD23	1:A:65:GLU:OE2	2.20	0.41
1:B:82:ARG:HB3	1:B:115:TYR:CE2	2.55	0.41
1:A:366:MET:HB2	1:A:367:PRO:CD	2.50	0.41
1:A:300:GLU:OE2	2:C:1:GLY:N	2.50	0.41

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	386/386 (100%)	379 (98%)	7 (2%)	0	100	100
1	B	384/386 (100%)	382 (100%)	2 (0%)	0	100	100
2	C	4/6 (67%)	4 (100%)	0	0	100	100
2	D	4/6 (67%)	4 (100%)	0	0	100	100
All	All	778/784 (99%)	769 (99%)	9 (1%)	0	100	100

There are no Ramachandran outliers to report.

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	355/353 (101%)	350 (99%)	5 (1%)	62	62
1	B	353/353 (100%)	350 (99%)	3 (1%)	79	80
2	C	3/3 (100%)	3 (100%)	0	100	100
2	D	3/3 (100%)	3 (100%)	0	100	100
All	All	714/712 (100%)	706 (99%)	8 (1%)	70	71

All (8) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	2	GLU
1	A	184	GLU
1	A	186	ARG
1	A	190	ASN
1	A	258	LYS
1	B	2	GLU
1	B	190	ASN
1	B	350	LYS

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (9) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	150	ASN
1	A	252	ASN
1	B	142	GLN
1	B	206	ASN
1	B	208	ASN
1	B	252	ASN
1	B	282	ASN
1	B	326	GLN
1	B	359	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 oligosaccharides 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	$\#\|Z\| > 2$
3	MPD	B	401	-	7,7,7	0.43	0	9,10,10	0.94	0
4	ACT	B	403	-	3,3,3	1.05	0	3,3,3	0.80	0
3	MPD	A	401	-	7,7,7	0.30	0	9,10,10	0.71	0
4	ACT	A	403	-	3,3,3	0.88	0	3,3,3	0.80	0
4	ACT	A	402	-	3,3,3	0.93	0	3,3,3	0.41	0
3	MPD	B	402	-	7,7,7	0.61	0	9,10,10	1.06	1 (11%)

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	MPD	B	401	-	-	1/5/5/5	-
3	MPD	A	401	-	-	0/5/5/5	-
3	MPD	B	402	-	-	2/5/5/5	-

There are no bond length outliers.

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	402	MPD	O2-C2-CM	-2.32	100.63	108.08

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
3	B	402	MPD	O2-C2-C3-C4
3	B	401	MPD	C2-C3-C4-C5
3	B	402	MPD	C2-C3-C4-O4

There are no ring outliers.

2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	402	ACT	2	0
3	B	402	MPD	2	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^{th} 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^2)	Q<0.9
1	A	386/386 (100%)	0.06	13 (3%) 48 50	16, 36, 70, 86	2 (0%)
1	B	386/386 (100%)	-0.04	4 (1%) 79 81	24, 37, 73, 98	0
2	C	6/6 (100%)	-0.42	0 100 100	23, 26, 29, 30	0
2	D	6/6 (100%)	-0.15	0 100 100	25, 29, 32, 32	0
All	All	784/784 (100%)	0.00	17 (2%) 62 64	16, 36, 71, 98	2 (0%)

All (17) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	12	LEU	4.7
1	A	6	ILE	3.4
1	A	10	LYS	3.4
1	A	72	LEU	3.2
1	A	13	GLU	3.1
1	A	14	ASN	3.0
1	A	16	ASN	2.7
1	B	35	TYR	2.6
1	A	19	MET	2.6
1	A	43	LYS	2.5
1	B	323	GLU	2.5
1	A	1	MET	2.4
1	A	15	ASP	2.3
1	A	22	TRP	2.2
1	B	319	ASN	2.2
1	B	320	LYS	2.0
1	A	11	ASP	2.0

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

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

6.3 Carbohydrates (i)

There are no monosaccharides in this entry.

6.4 Ligands (i)

In the following table, the Atoms column lists the number of modelled atoms in the group and the number defined in the chemical component dictionary. The B-factors column lists the minimum, median, 95^{th} 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(Å^2)	Q<0.9
4	ACT	A	403	4/4	0.81	0.16	47,53,56,68	0
4	ACT	A	402	4/4	0.82	0.17	38,41,57,86	0
3	MPD	A	401	8/8	0.82	0.21	74,81,89,91	0
3	MPD	B	401	8/8	0.83	0.18	52,64,65,76	0
3	MPD	B	402	8/8	0.84	0.19	32,42,64,80	0
4	ACT	B	403	4/4	0.89	0.14	58,60,69,70	0

6.5 Other polymers (i)

There are no such residues in this entry.