



# wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 09:48 AM GMT

PDB ID : 3KTJ  
Title : Structure of ClpP in complex with ADEP2 in monoclinic crystal form  
Authors : Lee, B.-G.; Brotz-Oesterhelt, H.; Song, H.K.  
Deposited on : 2009-11-25  
Resolution : 2.60 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at [validation@mail.wwpdb.org](mailto:validation@mail.wwpdb.org)

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

---

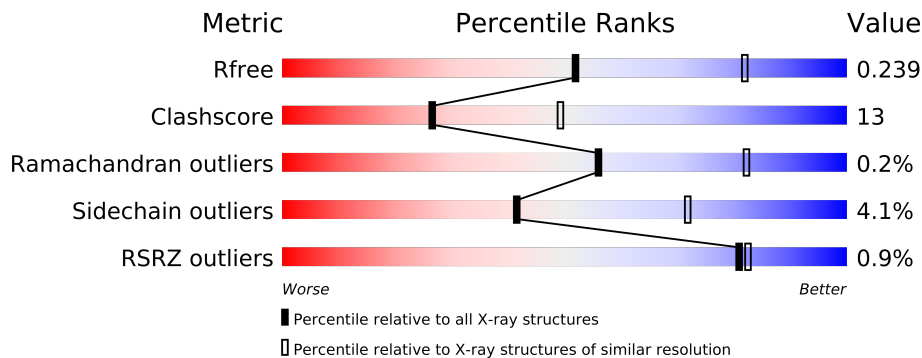
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity	:	4.02b-467
Mogul	:	1.15 2013
Xtriage (Phenix)	:	dev-1323
EDS	:	stable22639
Percentile statistics	:	21963
Refmac	:	5.8.0049
CCP4	:	6.3.0 (Settle)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP)	:	stable22683

# 1 Overall quality at a glance

The reported resolution of this entry is 2.60 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))
$R_{free}$	66092	1718 (2.60-2.60)
Clashscore	79885	2154 (2.60-2.60)
Ramachandran outliers	78287	2113 (2.60-2.60)
Sidechain outliers	78261	2113 (2.60-2.60)
RSRZ outliers	66119	1718 (2.60-2.60)

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. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	199	
1	B	199	
1	C	199	
1	D	199	
1	E	199	
1	F	199	
1	G	199	
2	H	7	
2	I	7	
2	J	7	
2	K	7	
2	L	7	
2	M	7	
2	N	7	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
3	NHE	B	502	-	X
3	NHE	D	504	-	X

## 2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 10173 atoms, of which 0 are hydrogen and 0 are deuterium.

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 ATP-dependent Clp protease proteolytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	174	Total	C	N	O	S	0	0	0
			1326	839	223	257	7			
1	B	174	Total	C	N	O	S	0	0	0
			1326	839	223	257	7			
1	C	174	Total	C	N	O	S	0	0	0
			1326	839	223	257	7			
1	D	174	Total	C	N	O	S	0	0	0
			1326	839	223	257	7			
1	E	174	Total	C	N	O	S	0	0	0
			1326	839	223	257	7			
1	F	174	Total	C	N	O	S	0	0	0
			1326	839	223	257	7			
1	G	174	Total	C	N	O	S	0	0	0
			1326	839	223	257	7			

There are 21 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	197	HIS	-	EXPRESSION TAG	UNP P80244
A	198	HIS	-	EXPRESSION TAG	UNP P80244
A	199	HIS	-	EXPRESSION TAG	UNP P80244
B	197	HIS	-	EXPRESSION TAG	UNP P80244
B	198	HIS	-	EXPRESSION TAG	UNP P80244
B	199	HIS	-	EXPRESSION TAG	UNP P80244
C	197	HIS	-	EXPRESSION TAG	UNP P80244
C	198	HIS	-	EXPRESSION TAG	UNP P80244
C	199	HIS	-	EXPRESSION TAG	UNP P80244
D	197	HIS	-	EXPRESSION TAG	UNP P80244
D	198	HIS	-	EXPRESSION TAG	UNP P80244
D	199	HIS	-	EXPRESSION TAG	UNP P80244
E	197	HIS	-	EXPRESSION TAG	UNP P80244
E	198	HIS	-	EXPRESSION TAG	UNP P80244
E	199	HIS	-	EXPRESSION TAG	UNP P80244

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
F	197	HIS	-	EXPRESSION TAG	UNP P80244
F	198	HIS	-	EXPRESSION TAG	UNP P80244
F	199	HIS	-	EXPRESSION TAG	UNP P80244
G	197	HIS	-	EXPRESSION TAG	UNP P80244
G	198	HIS	-	EXPRESSION TAG	UNP P80244
G	199	HIS	-	EXPRESSION TAG	UNP P80244

- Molecule 2 is a protein called Acyldepsipeptide 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	H	7	Total	C	F	N	O	0	0	0
			57	41	2	6	8			
2	I	7	Total	C	F	N	O	0	0	0
			57	41	2	6	8			
2	J	7	Total	C	F	N	O	0	0	0
			57	41	2	6	8			
2	K	7	Total	C	F	N	O	0	0	0
			57	41	2	6	8			
2	L	7	Total	C	F	N	O	0	0	0
			57	41	2	6	8			
2	M	7	Total	C	F	N	O	0	0	0
			57	41	2	6	8			
2	N	7	Total	C	F	N	O	0	0	0
			57	41	2	6	8			

There are 21 discrepancies between the modelled and reference sequences:

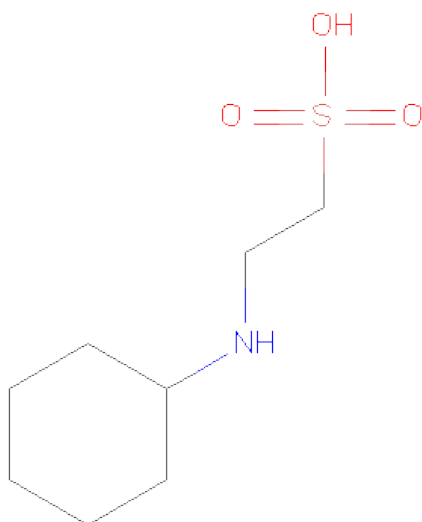
Chain	Residue	Modelled	Actual	Comment	Reference
H	1	CXP	OTT	SEE REMARK 999	NOR NOR01131
H	2	WFP	PHE	SEE REMARK 999	NOR NOR01131
H	5	YCP	MAA	SEE REMARK 999	NOR NOR01131
I	1	CXP	OTT	SEE REMARK 999	NOR NOR01131
I	2	WFP	PHE	SEE REMARK 999	NOR NOR01131
I	5	YCP	MAA	SEE REMARK 999	NOR NOR01131
J	1	CXP	OTT	SEE REMARK 999	NOR NOR01131
J	2	WFP	PHE	SEE REMARK 999	NOR NOR01131
J	5	YCP	MAA	SEE REMARK 999	NOR NOR01131
K	1	CXP	OTT	SEE REMARK 999	NOR NOR01131
K	2	WFP	PHE	SEE REMARK 999	NOR NOR01131
K	5	YCP	MAA	SEE REMARK 999	NOR NOR01131
L	1	CXP	OTT	SEE REMARK 999	NOR NOR01131
L	2	WFP	PHE	SEE REMARK 999	NOR NOR01131

*Continued on next page...*

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
L	5	YCP	MAA	SEE REMARK 999	NOR NOR01131
M	1	CXP	OTT	SEE REMARK 999	NOR NOR01131
M	2	WFP	PHE	SEE REMARK 999	NOR NOR01131
M	5	YCP	MAA	SEE REMARK 999	NOR NOR01131
N	1	CXP	OTT	SEE REMARK 999	NOR NOR01131
N	2	WFP	PHE	SEE REMARK 999	NOR NOR01131
N	5	YCP	MAA	SEE REMARK 999	NOR NOR01131

- Molecule 3 is 2-[N-CYCLOHEXYLAMINO]ETHANESULFONIC ACID (three-letter code: NHE) (formula: C<sub>8</sub>H<sub>17</sub>NO<sub>3</sub>S).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
3	A	1	Total	C	N	O	S	0	0
			13	8	1	3	1		
3	B	1	Total	C	N	O	S	0	0
			13	8	1	3	1		
3	C	1	Total	C	N	O	S	0	0
			13	8	1	3	1		
3	D	1	Total	C	N	O	S	0	0
			13	8	1	3	1		
3	E	1	Total	C	N	O	S	0	0
			13	8	1	3	1		
3	F	1	Total	C	N	O	S	0	0
			13	8	1	3	1		
3	G	1	Total	C	N	O	S	0	0
			13	8	1	3	1		

- Molecule 4 is water.

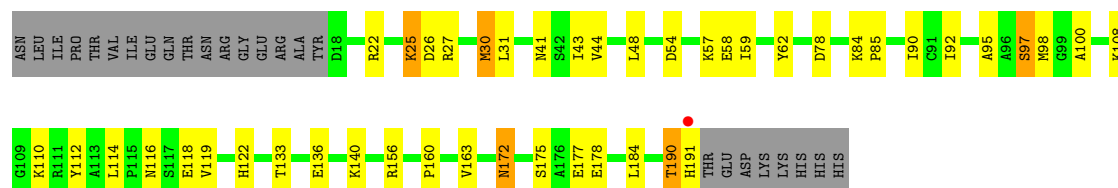
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	57	Total 57	O 57	0	0
4	B	61	Total 61	O 61	0	0
4	C	52	Total 52	O 52	0	0
4	D	58	Total 58	O 58	0	0
4	E	51	Total 51	O 51	0	0
4	F	56	Total 56	O 56	0	0
4	G	58	Total 58	O 58	0	0
4	H	1	Total 1	O 1	0	0
4	I	1	Total 1	O 1	0	0
4	J	1	Total 1	O 1	0	0
4	K	3	Total 3	O 3	0	0
4	L	1	Total 1	O 1	0	0
4	M	1	Total 1	O 1	0	0





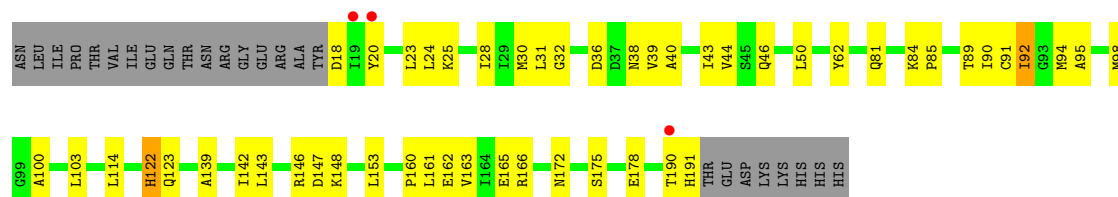
- Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain E:



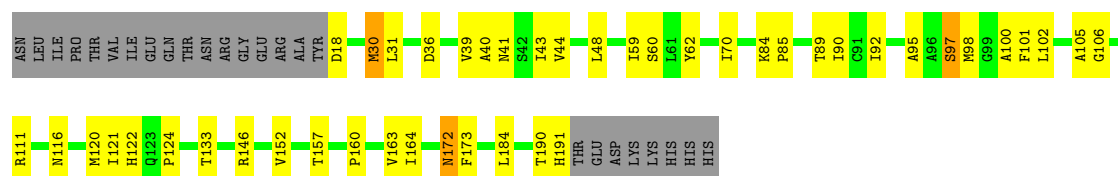
- Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain F:



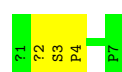
- Molecule 1: ATP-dependent Clp protease proteolytic subunit

Chain G:



- Molecule 2: Acyldepsipeptide 2

Chain H:



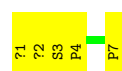
- Molecule 2: Acyldepsipeptide 2

Chain I:



- Molecule 2: Acyldepsipeptide 2

Chain J:



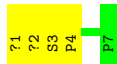
- Molecule 2: Acyldepsipeptide 2

Chain K:



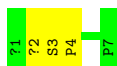
- Molecule 2: Acyldepsipeptide 2

Chain L: 



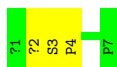
- Molecule 2: Acyldepsipeptide 2

Chain M: 



- Molecule 2: Acyldepsipeptide 2

Chain N: 



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 1 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	121.43Å 152.24Å 100.38Å 90.00° 118.98° 90.00°	Depositor
Resolution (Å)	50.00 – 2.60 39.83 – 2.60	Depositor EDS
% Data completeness (in resolution range)	88.8 (50.00-2.60) 94.4 (39.83-2.60)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.24 (at 2.61Å)	Xtriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.226 , 0.296 0.237 , 0.239	Depositor DCC
$R_{free}$ test set	2368 reflections (5.37%)	DCC
Wilson B-factor (Å <sup>2</sup> )	35.1	Xtriage
Anisotropy	0.305	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 20.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 46486 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.91	EDS
Total number of atoms	10173	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	31.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.08% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

## 5 Model quality

### 5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NHE, CXP, WFP, YCP, MP8

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	0.41	0/1341	0.62	0/1807
1	B	0.40	0/1341	0.61	0/1807
1	C	0.40	0/1341	0.60	0/1807
1	D	0.39	0/1341	0.63	0/1807
1	E	0.40	0/1341	0.60	0/1807
1	F	0.41	0/1341	0.60	0/1807
1	G	0.41	0/1341	0.60	0/1807
2	H	2.81	3/17 (17.6%)	1.91	0/21
2	I	2.51	0/17	1.79	0/21
2	J	2.75	3/17 (17.6%)	1.86	0/21
2	K	2.87	3/17 (17.6%)	1.93	0/21
2	L	2.88	3/17 (17.6%)	1.94	0/21
2	M	2.81	3/17 (17.6%)	1.99	0/21
2	N	2.76	3/17 (17.6%)	2.00	0/21
All	All	0.51	18/9506 (0.2%)	0.64	0/12796

The worst 5 of 18 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	L	3	SER	C-N	6.04	1.45	1.34
2	L	4	PRO	N-CA	5.91	1.57	1.47
2	N	3	SER	C-N	5.84	1.45	1.34
2	M	4	PRO	N-CA	5.62	1.56	1.47
2	K	3	SER	C-N	5.57	1.44	1.34

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

## 5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	1326	0	1366	41	0
1	B	1326	0	1366	30	0
1	C	1326	0	1366	47	1
1	D	1326	0	1366	43	0
1	E	1326	0	1366	39	0
1	F	1326	0	1366	49	0
1	G	1326	0	1366	41	0
2	H	57	0	55	2	0
2	I	57	0	55	2	0
2	J	57	0	55	4	0
2	K	57	0	55	2	0
2	L	57	0	55	3	0
2	M	57	0	55	1	0
2	N	57	0	55	1	0
3	A	13	0	16	0	0
3	B	13	0	16	0	0
3	C	13	0	16	0	0
3	D	13	0	16	0	0
3	E	13	0	16	0	0
3	F	13	0	16	0	0
3	G	13	0	16	0	0
4	A	57	0	0	2	0
4	B	61	0	0	5	0
4	C	52	0	0	4	0
4	D	58	0	0	3	0
4	E	51	0	0	3	0
4	F	56	0	0	6	0
4	G	58	0	0	3	0
4	H	1	0	0	1	0
4	I	1	0	0	0	0
4	J	1	0	0	0	0
4	K	3	0	0	0	0
4	L	1	0	0	0	0
4	M	1	0	0	0	0
All	All	10173	0	10059	262	1

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 13.

The worst 5 of 262 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:F:98:MET:HE2	1:F:98:MET:HA	1.41	0.97
1:B:74:MET:HG3	4:B:239:HOH:O	1.76	0.84
1:A:41:ASN:HB3	1:B:30:MET:HE2	1.64	0.79
1:E:133:THR:HG21	4:F:292:HOH:O	1.85	0.77
1:D:148:LYS:HD2	1:E:116:ASN:HD22	1.49	0.76

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	Distance(Å)	Clash(Å)
1:C:86:LYS:NZ	1:C:86:LYS:NZ[2.455]	1.82	0.38

## 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	172/199 (86%)	161 (94%)	11 (6%)	0	100	100
1	B	172/199 (86%)	158 (92%)	13 (8%)	1 (1%)	33	63
1	C	172/199 (86%)	157 (91%)	15 (9%)	0	100	100
1	D	172/199 (86%)	163 (95%)	8 (5%)	1 (1%)	33	63
1	E	172/199 (86%)	162 (94%)	10 (6%)	0	100	100
1	F	172/199 (86%)	160 (93%)	12 (7%)	0	100	100
1	G	172/199 (86%)	161 (94%)	11 (6%)	0	100	100
2	H	3/7 (43%)	2 (67%)	1 (33%)	0	100	100
2	I	3/7 (43%)	2 (67%)	1 (33%)	0	100	100
2	J	3/7 (43%)	2 (67%)	1 (33%)	0	100	100
2	K	3/7 (43%)	2 (67%)	1 (33%)	0	100	100
2	L	3/7 (43%)	2 (67%)	1 (33%)	0	100	100

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
2	M	3/7 (43%)	2 (67%)	1 (33%)	0	100	100
2	N	3/7 (43%)	2 (67%)	1 (33%)	0	100	100
All	All	1225/1442 (85%)	1136 (93%)	87 (7%)	2 (0%)	56	82

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	190	THR
1	D	54	ASP

### 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	141/165 (86%)	135 (96%)	6 (4%)	40	69
1	B	141/165 (86%)	134 (95%)	7 (5%)	34	61
1	C	141/165 (86%)	138 (98%)	3 (2%)	66	90
1	D	141/165 (86%)	135 (96%)	6 (4%)	40	69
1	E	141/165 (86%)	132 (94%)	9 (6%)	25	47
1	F	141/165 (86%)	135 (96%)	6 (4%)	40	69
1	G	141/165 (86%)	137 (97%)	4 (3%)	56	84
2	H	2/2 (100%)	2 (100%)	0	100	100
2	I	2/2 (100%)	2 (100%)	0	100	100
2	J	2/2 (100%)	2 (100%)	0	100	100
2	K	2/2 (100%)	2 (100%)	0	100	100
2	L	2/2 (100%)	2 (100%)	0	100	100
2	M	2/2 (100%)	2 (100%)	0	100	100
2	N	2/2 (100%)	2 (100%)	0	100	100
All	All	1001/1169 (86%)	960 (96%)	41 (4%)	41	72

5 of 41 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	146	ARG
1	E	25	LYS
1	G	30	MET
1	D	162	GLU
1	D	170	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 26 such sidechains are listed below:

Mol	Chain	Res	Type
1	C	123	GLN
1	D	123	GLN
1	G	116	ASN
1	C	172	ASN
1	D	116	ASN

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

21 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	$\# Z  > 2$
2	WFP	H	2	2	13,13,14	7.54	9 (69%)	15,17,19	3.97	6 (40%)
2	YCP	H	5	2	8,8,9	8.36	4 (50%)	7,9,11	2.34	2 (28%)
2	MP8	H	7	2	8,8,9	6.60	2 (25%)	8,10,12	1.15	1 (12%)
2	WFP	I	2	2	13,13,14	7.08	8 (61%)	15,17,19	3.54	5 (33%)
2	YCP	I	5	2	8,8,9	8.83	3 (37%)	7,9,11	2.27	2 (28%)
2	MP8	I	7	2	8,8,9	6.42	2 (25%)	8,10,12	1.13	1 (12%)
2	WFP	J	2	2	13,13,14	7.50	9 (69%)	15,17,19	4.27	6 (40%)



Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	YCP	J	5	2	8,8,9	8.45	2 (25%)	7,9,11	2.35	2 (28%)
2	MP8	J	7	2	8,8,9	6.65	2 (25%)	8,10,12	0.96	0
2	WFP	K	2	2	13,13,14	7.23	9 (69%)	15,17,19	3.40	5 (33%)
2	YCP	K	5	2	8,8,9	8.35	3 (37%)	7,9,11	2.39	2 (28%)
2	MP8	K	7	2	8,8,9	6.25	2 (25%)	8,10,12	1.23	2 (25%)
2	WFP	L	2	2	13,13,14	7.52	9 (69%)	15,17,19	3.49	6 (40%)
2	YCP	L	5	2	8,8,9	8.70	4 (50%)	7,9,11	2.40	2 (28%)
2	MP8	L	7	2	8,8,9	6.53	2 (25%)	8,10,12	1.04	0
2	WFP	M	2	2	13,13,14	7.35	9 (69%)	15,17,19	3.67	6 (40%)
2	YCP	M	5	2	8,8,9	8.46	3 (37%)	7,9,11	2.44	2 (28%)
2	MP8	M	7	2	8,8,9	6.20	2 (25%)	8,10,12	1.10	1 (12%)
2	WFP	N	2	2	13,13,14	7.29	9 (69%)	15,17,19	2.65	7 (46%)
2	YCP	N	5	2	8,8,9	8.49	3 (37%)	7,9,11	2.39	2 (28%)
2	MP8	N	7	2	8,8,9	6.74	2 (25%)	8,10,12	1.04	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	WFP	H	2	2	-	0/4/6/8	0/1/1/1
2	YCP	H	5	2	-	0/0/10/12	0/1/1/1
2	MP8	H	7	2	-	0/0/11/13	0/1/1/1
2	WFP	I	2	2	-	0/4/6/8	0/1/1/1
2	YCP	I	5	2	-	0/0/10/12	0/1/1/1
2	MP8	I	7	2	-	0/0/11/13	0/1/1/1
2	WFP	J	2	2	-	0/4/6/8	0/1/1/1
2	YCP	J	5	2	-	0/0/10/12	0/1/1/1
2	MP8	J	7	2	-	0/0/11/13	0/1/1/1
2	WFP	K	2	2	-	0/4/6/8	0/1/1/1
2	YCP	K	5	2	-	0/0/10/12	0/1/1/1
2	MP8	K	7	2	-	0/0/11/13	0/1/1/1
2	WFP	L	2	2	-	0/4/6/8	0/1/1/1
2	YCP	L	5	2	-	0/0/10/12	0/1/1/1
2	MP8	L	7	2	-	0/0/11/13	0/1/1/1
2	WFP	M	2	2	-	0/4/6/8	0/1/1/1
2	YCP	M	5	2	-	0/0/10/12	0/1/1/1
2	MP8	M	7	2	-	0/0/11/13	0/1/1/1

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	WFP	N	2	2	-	0/4/6/8	0/1/1/1
2	YCP	N	5	2	-	0/0/10/12	0/1/1/1
2	MP8	N	7	2	-	0/0/11/13	0/1/1/1

The worst 5 of 98 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	I	5	YCP	O-C	23.89	1.27	1.11
2	L	5	YCP	O-C	23.37	1.27	1.11
2	J	5	YCP	O-C	22.92	1.27	1.11
2	N	5	YCP	O-C	22.79	1.27	1.11
2	M	5	YCP	O-C	22.76	1.27	1.11

The worst 5 of 60 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	J	2	WFP	C-CA-N	-14.95	98.89	113.83
2	H	2	WFP	C-CA-N	-13.76	100.08	113.83
2	M	2	WFP	C-CA-N	-12.41	101.44	113.83
2	I	2	WFP	C-CA-N	-12.11	101.73	113.83
2	L	2	WFP	C-CA-N	-11.59	102.26	113.83

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 5.5 Carbohydrates

There are no carbohydrates in this entry.

## 5.6 Ligand geometry

7 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	NHE	A	501	-	13,13,13	1.84	1 (7%)	17,17,17	8.17	9 (52%)
3	NHE	B	502	-	13,13,13	2.11	2 (15%)	17,17,17	8.42	10 (58%)
3	NHE	C	503	-	13,13,13	1.98	2 (15%)	17,17,17	8.24	10 (58%)
3	NHE	D	504	-	13,13,13	1.98	2 (15%)	17,17,17	8.43	10 (58%)
3	NHE	E	505	-	13,13,13	1.93	1 (7%)	17,17,17	8.41	11 (64%)
3	NHE	F	506	-	13,13,13	2.09	2 (15%)	17,17,17	8.40	10 (58%)
3	NHE	G	507	-	13,13,13	2.16	2 (15%)	17,17,17	8.34	9 (52%)

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	NHE	A	501	-	-	0/7/15/15	0/1/1/1
3	NHE	B	502	-	-	0/7/15/15	0/1/1/1
3	NHE	C	503	-	-	0/7/15/15	0/1/1/1
3	NHE	D	504	-	-	0/7/15/15	0/1/1/1
3	NHE	E	505	-	-	0/7/15/15	0/1/1/1
3	NHE	F	506	-	-	0/7/15/15	0/1/1/1
3	NHE	G	507	-	-	0/7/15/15	0/1/1/1

The worst 5 of 12 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	G	507	NHE	C2-S	-6.77	1.67	1.77
3	F	506	NHE	C2-S	-6.60	1.67	1.77
3	B	502	NHE	C2-S	-6.55	1.67	1.77
3	D	504	NHE	C2-S	-6.12	1.68	1.77
3	C	503	NHE	C2-S	-6.07	1.68	1.77

The worst 5 of 69 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	F	506	NHE	O1-S-C2	21.02	124.82	106.81
3	G	507	NHE	O1-S-C2	21.00	124.80	106.81
3	D	504	NHE	O1-S-C2	20.80	124.63	106.81
3	B	502	NHE	O1-S-C2	20.79	124.63	106.81

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	E	505	NHE	O1-S-C2	20.73	124.57	106.81

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

## 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	174/199 (87%)	-0.29	3 (1%) 67 66	15, 26, 52, 79	0
1	B	174/199 (87%)	-0.44	1 (0%) 86 89	16, 27, 51, 76	0
1	C	174/199 (87%)	-0.21	2 (1%) 77 79	18, 30, 52, 74	0
1	D	174/199 (87%)	-0.34	2 (1%) 77 79	15, 29, 54, 75	0
1	E	174/199 (87%)	-0.34	1 (0%) 86 89	17, 29, 52, 74	0
1	F	174/199 (87%)	-0.32	3 (1%) 67 66	17, 28, 51, 78	0
1	G	174/199 (87%)	-0.36	0 100 100	16, 30, 52, 75	0
2	H	7/7 (100%)	-0.24	0 100 100	32, 35, 39, 40	0
2	I	7/7 (100%)	-0.45	0 100 100	27, 36, 40, 41	0
2	J	7/7 (100%)	0.00	0 100 100	30, 37, 39, 41	0
2	K	7/7 (100%)	-0.08	0 100 100	43, 43, 45, 45	0
2	L	7/7 (100%)	0.31	0 100 100	36, 47, 52, 53	0
2	M	7/7 (100%)	-0.88	0 100 100	34, 40, 44, 46	0
2	N	7/7 (100%)	0.42	0 100 100	34, 46, 50, 52	0
All	All	1267/1442 (87%)	-0.32	12 (0%) 81 82	15, 29, 52, 79	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	191	HIS	5.5
1	A	19	ILE	4.6
1	E	191	HIS	4.3
1	A	20	TYR	3.4
1	F	20	TYR	2.9

## 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	RSR	LLDF	B-factors(Å <sup>2</sup> )	Q<0.9
2	MP8	N	7	8/9	0.26	2.28	48,49,49,50	0
2	YCP	N	5	8/9	0.28	1.82	51,51,52,53	0
2	MP8	I	7	8/9	0.19	1.38	39,40,41,41	0
2	YCP	I	5	8/9	0.19	1.27	39,40,41,42	0
2	MP8	J	7	8/9	0.21	1.11	37,38,40,40	0
2	WFP	H	2	13/14	0.17	1.01	30,31,33,33	0
2	MP8	L	7	8/9	0.23	0.99	49,50,50,51	0
2	YCP	H	5	8/9	0.19	0.77	37,38,39,39	0
2	WFP	K	2	13/14	0.18	0.61	43,44,46,47	0
2	YCP	L	5	8/9	0.21	0.46	52,53,53,54	0
2	MP8	H	7	8/9	0.15	0.41	36,37,38,38	0
2	MP8	M	7	8/9	0.15	0.11	38,39,40,41	0
2	WFP	I	2	13/14	0.15	0.09	23,25,31,32	0
2	YCP	J	5	8/9	0.19	0.05	35,38,40,43	0
2	YCP	M	5	8/9	0.15	-0.00	42,44,44,45	0
2	MP8	K	7	8/9	0.15	-0.11	40,43,44,46	0
2	WFP	L	2	13/14	0.17	-0.22	38,40,43,43	0
2	YCP	K	5	8/9	0.17	-0.40	43,44,45,47	0
2	WFP	J	2	13/14	0.13	-0.51	25,27,34,35	0
2	WFP	M	2	13/14	0.13	-0.60	31,34,41,41	0
2	WFP	N	2	13/14	0.14	-0.98	29,32,40,40	0

## 6.3 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 6.4 Ligands ⓘ

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	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
3	NHE	B	502	13/13	0.19	4.46	54,56,56,57	0
3	NHE	D	504	13/13	0.21	3.17	58,59,60,61	0
3	NHE	C	503	13/13	0.19	1.62	42,43,44,44	0
3	NHE	F	506	13/13	0.17	1.46	49,53,55,55	0
3	NHE	A	501	13/13	0.14	0.16	28,31,36,36	0
3	NHE	E	505	13/13	0.13	-0.06	40,41,42,42	0
3	NHE	G	507	13/13	0.12	-0.61	29,30,32,33	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.