



# Full wwPDB X-ray Structure Validation Report ⓘ

Jun 15, 2020 – 11:28 pm BST

PDB ID : 3WNH  
Title : Cyclic hexapeptide PKZDNv in complex with HIV-1 integrase  
Authors : Wielens, J.; Chalmers, D.K.; Parker, M.W.  
Deposited on : 2013-12-10  
Resolution : 1.50 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

<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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
Percentile statistics : 20191225.v01 (using entries in the PDB archive December 25th 2019)  
Refmac : 5.8.0158  
CCP4 : 7.0.044 (Gargrove)  
Ideal geometry (proteins) : Engh & Huber (2001)  
Ideal geometry (DNA, RNA) : Parkinson et al. (1996)  
Validation Pipeline (wwPDB-VP) : 2.11

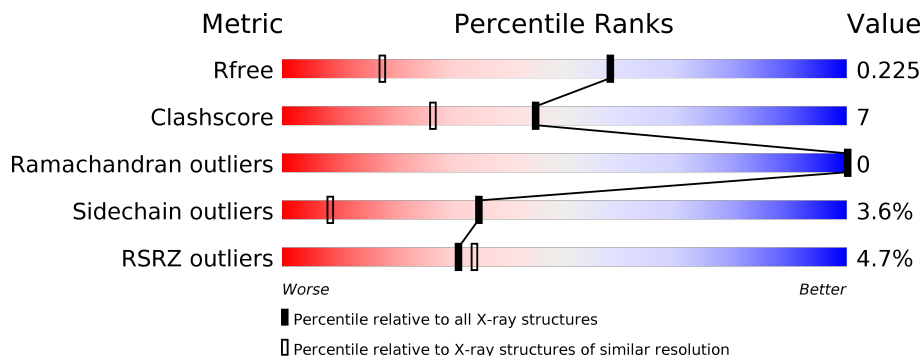
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.50 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))
$R_{free}$	130704	2936 (1.50-1.50)
Clashscore	141614	3144 (1.50-1.50)
Ramachandran outliers	138981	3066 (1.50-1.50)
Sidechain outliers	138945	3064 (1.50-1.50)
RSRZ outliers	127900	2884 (1.50-1.50)

The table below summarises the geometric issues observed across the polymeric chains and their fit to the electron density. The red, orange, yellow and green segments on the lower bar indicate the fraction of residues that contain outliers for  $\geq 3$ , 2, 1 and 0 types of geometric quality criteria 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  $\leq 5\%$ . 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	157	<div style="display: flex; align-items: center;"> <div style="width: 4%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 73%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 13%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 10%; height: 10px; background-color: orange; margin-right: 5px;"></div> <div style="width: 13%; height: 10px; background-color: grey;"></div> </div> <p style="font-size: small; margin-top: 5px;">4%      73%      13%      •      13%</p>
1	B	157	<div style="display: flex; align-items: center;"> <div style="width: 3%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 68%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 18%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 11%; height: 10px; background-color: grey;"></div> </div> <p style="font-size: small; margin-top: 5px;">3%      68%      18%      15%</p>
2	C	6	<div style="display: flex; align-items: center;"> <div style="width: 17%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 67%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 17%; height: 10px; background-color: yellow; margin-right: 5px;"></div> <div style="width: 17%; height: 10px; background-color: orange;"></div> </div> <p style="font-size: small; margin-top: 5px;">17%      67%      17%      17%</p>
2	D	6	<div style="display: flex; align-items: center;"> <div style="width: 17%; height: 10px; background-color: red; margin-right: 5px;"></div> <div style="width: 67%; height: 10px; background-color: green; margin-right: 5px;"></div> <div style="width: 33%; height: 10px; background-color: yellow;"></div> </div> <p style="font-size: small; margin-top: 5px;">17%      67%      33%</p>

## 2 Entry composition [i](#)

There are 6 unique types of molecules in this entry. The entry contains 2480 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 Gag-Pol polyprotein.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	137	1064	673	189	198	4	0	2	0
1	B	134	1045	661	185	195	4	0	1	0

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	56	SER	CYS	ENGINEERED MUTATION	UNP P12497
A	123	GLY	SER	ENGINEERED MUTATION	UNP P12497
A	124	ALA	THR	ENGINEERED MUTATION	UNP P12497
A	127	ARG	LYS	ENGINEERED MUTATION	UNP P12497
A	131	ASP	TRP	ENGINEERED MUTATION	UNP P12497
A	139	ASP	PHE	ENGINEERED MUTATION	UNP P12497
A	185	HIS	PHE	ENGINEERED MUTATION	UNP P12497
B	56	SER	CYS	ENGINEERED MUTATION	UNP P12497
B	123	GLY	SER	ENGINEERED MUTATION	UNP P12497
B	124	ALA	THR	ENGINEERED MUTATION	UNP P12497
B	127	ARG	LYS	ENGINEERED MUTATION	UNP P12497
B	131	ASP	TRP	ENGINEERED MUTATION	UNP P12497
B	139	ASP	PHE	ENGINEERED MUTATION	UNP P12497
B	185	HIS	PHE	ENGINEERED MUTATION	UNP P12497

- Molecule 2 is a protein called PK(NLE)DN(DVA) peptide.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	C	6	47	30	8	9	0	0	0
2	D	6	47	30	8	9	0	0	0

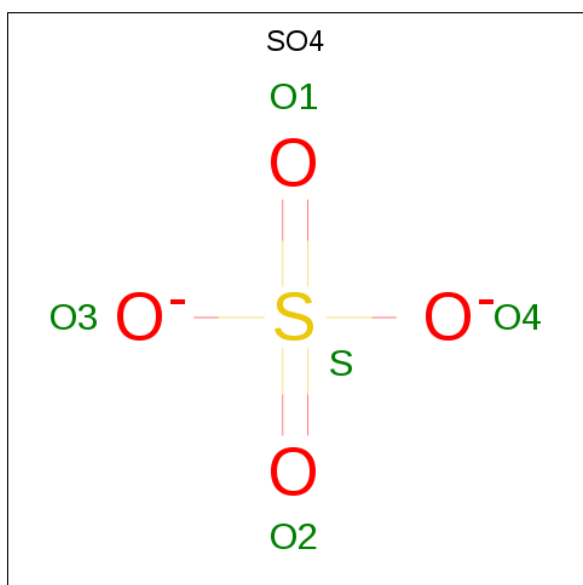
- Molecule 3 is CADMIUM ION (three-letter code: CD) (formula: Cd).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	B	2	Total Cd 2 2	0	0
3	A	2	Total Cd 2 2	0	0

- Molecule 4 is CHLORIDE ION (three-letter code: CL) (formula: Cl).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	B	1	Total Cl 1 1	0	0
4	A	1	Total Cl 1 1	0	0

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

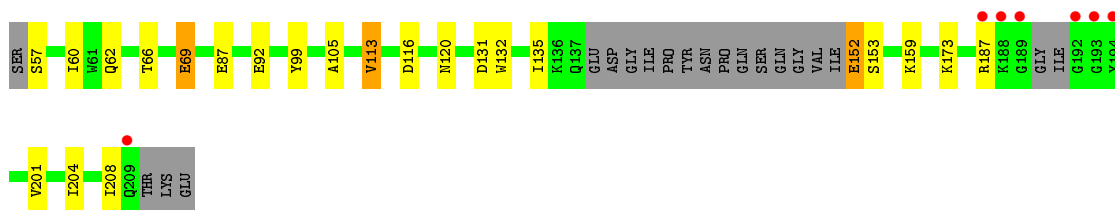
- Molecule 6 is water.

<b>Mol</b>	<b>Chain</b>	<b>Residues</b>	<b>Atoms</b>		<b>ZeroOcc</b>	<b>AltConf</b>
6	A	113	Total 113	O 113	0	0
6	B	121	Total 121	O 121	0	0
6	C	4	Total 4	O 4	0	0
6	D	3	Total 3	O 3	0	0

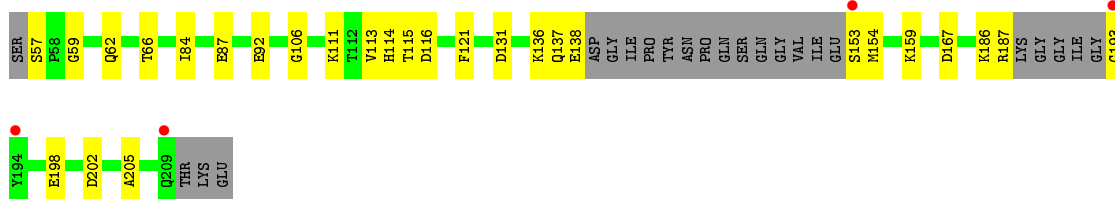
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: Gag-Pol polyprotein



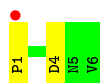
- Molecule 1: Gag-Pol polyprotein



- Molecule 2: PK(NLE)DN(DVA) peptide



- Molecule 2: PK(NLE)DN(DVA) peptide



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 32	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	49.44Å 49.44Å 101.33Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	42.82 – 1.50 42.82 – 1.50	Depositor EDS
% Data completeness (in resolution range)	99.0 (42.82-1.50) 98.9 (42.82-1.50)	Depositor EDS
$R_{merge}$	0.04	Depositor
$R_{sym}$	0.04	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.99 (at 1.50Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, $R_{free}$	0.190 , 0.227 0.190 , 0.225	Depositor DCC
$R_{free}$ test set	2196 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	15.5	Xtriage
Anisotropy	0.021	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.38 , 29.5	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.51$ , $\langle L^2 \rangle = 0.34$	Xtriage
Estimated twinning fraction	0.028 for -h,-k,l 0.489 for h,-h-k,-l 0.028 for -k,-h,-l	Xtriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	2480	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	22.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 9.11% 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 i

### 5.1 Standard geometry i

Bond lengths and bond angles in the following residue types are not validated in this section: NLE, CD, DVA, SO4, CL

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	1.36	4/1082 (0.4%)	1.46	8/1463 (0.5%)
1	B	1.42	5/1063 (0.5%)	1.36	9/1437 (0.6%)
2	C	1.89	1/31 (3.2%)	2.09	2/38 (5.3%)
2	D	1.89	1/31 (3.2%)	2.07	1/38 (2.6%)
All	All	1.40	11/2207 (0.5%)	1.43	20/2976 (0.7%)

All (11) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	92	GLU	CD-OE2	-6.75	1.18	1.25
1	B	131	ASP	CG-OD2	-6.64	1.10	1.25
1	A	132	TRP	CD2-CE2	6.16	1.48	1.41
2	D	1	PRO	C-N	5.98	1.47	1.34
1	A	99	TYR	CE1-CZ	-5.96	1.30	1.38
2	C	1	PRO	C-N	5.67	1.47	1.34
1	A	92	GLU	CD-OE2	-5.47	1.19	1.25
1	B	106	GLY	CA-C	-5.39	1.43	1.51
1	B	87[A]	GLU	CD-OE2	5.21	1.31	1.25
1	B	87[B]	GLU	CD-OE2	5.21	1.31	1.25
1	A	87	GLU	CD-OE2	5.04	1.31	1.25

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	131	ASP	CB-CG-OD1	-13.67	106.00	118.30
1	A	116	ASP	CB-CG-OD2	-12.06	107.44	118.30
1	A	131	ASP	CB-CG-OD2	-12.02	107.48	118.30
1	B	116	ASP	CB-CG-OD2	-11.84	107.64	118.30
1	A	131	ASP	OD1-CG-OD2	11.70	145.53	123.30
1	B	131	ASP	CB-CG-OD2	-10.50	108.85	118.30

*Continued on next page...*



Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	116	ASP	CB-CG-OD1	8.42	125.88	118.30
1	B	116	ASP	CB-CG-OD1	8.29	125.76	118.30
1	B	131	ASP	OD1-CG-OD2	7.86	138.23	123.30
1	B	113	VAL	CG1-CB-CG2	-7.80	98.41	110.90
2	D	4	ASP	CB-CG-OD1	7.06	124.65	118.30
1	A	69	GLU	OE1-CD-OE2	6.50	131.10	123.30
1	B	131	ASP	CB-CG-OD1	-6.21	112.71	118.30
1	A	113[A]	VAL	CG1-CB-CG2	-5.70	101.79	110.90
1	A	113[B]	VAL	CG1-CB-CG2	-5.70	101.79	110.90
1	B	87[A]	GLU	OE1-CD-OE2	5.42	129.81	123.30
1	B	87[B]	GLU	OE1-CD-OE2	5.42	129.81	123.30
2	C	1	PRO	N-CA-CB	5.27	109.62	103.30
2	C	5	ASN	CA-C-O	-5.11	109.38	120.10
1	B	202	ASP	CB-CG-OD2	-5.07	113.73	118.30

There are no chirality outliers.

There are no planarity outliers.

## 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	1064	0	1059	16	0
1	B	1045	0	1040	15	0
2	C	47	0	49	0	0
2	D	47	0	49	0	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
4	A	1	0	0	0	0
4	B	1	0	0	0	0
5	A	15	0	0	0	0
5	B	15	0	0	0	0
6	A	113	0	0	3	0
6	B	121	0	0	3	0
6	C	4	0	0	0	0
6	D	3	0	0	0	0
All	All	2480	0	2197	29	0

The all-atom clashscore is defined as the number of clashes found per 1000 atoms (including hydrogen atoms). The all-atom clashscore for this structure is 7.

All (29) 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:A:113[B]:VAL:HG23	1:A:113[B]:VAL:O	1.97	0.65
1:B:167:ASP:OD2	6:B:471:HOH:O	2.15	0.65
1:B:111:LYS:HE2	6:B:479:HOH:O	1.98	0.63
1:B:59:GLY:HA3	1:B:111:LYS:HE2	1.82	0.61
1:A:57:SER:N	6:A:1163:HOH:O	2.36	0.59
1:A:173:LYS:NZ	6:A:1204:HOH:O	2.36	0.57
1:A:152:GLU:CD	1:A:153:SER:H	2.09	0.55
1:A:208:ILE:HD13	1:B:198:GLU:HG2	1.91	0.53
1:B:66:THR:OG1	1:B:159:LYS:HE3	2.09	0.53
1:A:187:ARG:HG2	1:A:187:ARG:NH1	2.28	0.49
1:A:201:VAL:HG12	1:B:205:ALA:HB2	1.95	0.48
1:A:66:THR:OG1	1:A:159:LYS:HE3	2.13	0.48
1:A:105:ALA:HB2	1:A:135:ILE:HD11	1.96	0.48
1:A:120:ASN:N	1:A:120:ASN:OD1	2.47	0.47
1:A:57:SER:HA	1:A:60:ILE:HD12	1.95	0.47
1:B:137:GLN:O	1:B:138:GLU:CG	2.63	0.46
1:B:154:MET:HB3	6:B:456:HOH:O	2.16	0.46
1:B:114:HIS:CE1	1:B:138:GLU:OE2	2.68	0.46
1:A:113[B]:VAL:CG2	1:A:113[B]:VAL:O	2.64	0.44
1:A:152:GLU:OE2	1:A:153:SER:N	2.51	0.44
1:B:114:HIS:ND1	1:B:138:GLU:HB2	2.33	0.44
1:B:115:THR:HG21	1:B:121:PHE:CE1	2.54	0.42
1:B:84:ILE:HD13	1:B:84:ILE:HG21	1.76	0.42
1:B:137:GLN:O	1:B:138:GLU:HG3	2.20	0.42
1:A:113[A]:VAL:HG23	1:A:135:ILE:CG2	2.50	0.42
1:A:69:GLU:O	6:A:1205:HOH:O	2.22	0.41
1:B:114:HIS:HE1	1:B:138:GLU:OE2	2.03	0.41
1:A:204:ILE:O	1:A:208:ILE:HG13	2.21	0.41
1:B:187:ARG:HD3	1:B:193:GLY:O	2.21	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles [i](#)

### 5.3.1 Protein backbone [i](#)

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

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

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	133/157 (85%)	132 (99%)	1 (1%)	0	100	100
1	B	129/157 (82%)	128 (99%)	1 (1%)	0	100	100
2	C	3/6 (50%)	3 (100%)	0	0	100	100
2	D	3/6 (50%)	3 (100%)	0	0	100	100
All	All	268/326 (82%)	266 (99%)	2 (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	110/126 (87%)	108 (98%)	2 (2%)	59	30
1	B	109/126 (86%)	104 (95%)	5 (5%)	27	5
2	C	4/4 (100%)	3 (75%)	1 (25%)	0	0
2	D	4/4 (100%)	4 (100%)	0	100	100
All	All	227/260 (87%)	219 (96%)	8 (4%)	35	9

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

Mol	Chain	Res	Type
1	A	62	GLN
1	A	152	GLU

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	B	57	SER
1	B	62	GLN
1	B	136	LYS
1	B	153	SER
1	B	186	LYS
2	C	5	ASN

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (7) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	95	GLN
1	A	155	ASN
1	A	164	GLN
1	B	95	GLN
1	B	114	HIS
1	B	137	GLN
2	D	5	ASN

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

4 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	NLE	D	3	2	6,7,8	0.41	0	2,7,9	1.32	0
2	NLE	C	3	2	6,7,8	0.54	0	2,7,9	1.21	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	NLE	D	3	2	-	0/5/6/8	-
2	NLE	C	3	2	-	0/5/6/8	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

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](#)

Of 12 ligands modelled in this entry, 6 are monoatomic - leaving 6 for Mogul analysis.

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	$\# Z  > 2$
5	SO4	A	1004	-	4,4,4	0.42	0	6,6,6	0.67	0
5	SO4	B	306	-	4,4,4	0.52	0	6,6,6	1.02	1 (16%)
5	SO4	A	1005	-	4,4,4	0.72	0	6,6,6	0.40	0
5	SO4	B	305	-	4,4,4	0.77	0	6,6,6	0.74	0
5	SO4	A	1006	-	4,4,4	0.43	0	6,6,6	1.07	1 (16%)
5	SO4	B	304	-	4,4,4	0.78	0	6,6,6	0.95	0

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
5	A	1006	SO4	O4-S-O2	-2.20	97.82	109.31
5	B	306	SO4	O4-S-O3	2.03	117.71	109.06

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 [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	137/157 (87%)	0.16	7 (5%) 28 30	10, 17, 48, 72	0
1	B	134/157 (85%)	0.13	4 (2%) 50 55	10, 17, 40, 68	0
2	C	4/6 (66%)	1.00	1 (25%) 0 0	22, 32, 39, 39	0
2	D	4/6 (66%)	0.72	1 (25%) 0 0	22, 32, 35, 39	0
All	All	279/326 (85%)	0.16	13 (4%) 31 34	10, 17, 44, 72	0

All (13) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	193	GLY	4.8
1	A	193	GLY	4.5
1	B	194	TYR	3.7
1	A	209	GLN	3.3
1	A	194	TYR	2.8
2	C	1	PRO	2.6
1	A	192	GLY	2.6
1	A	189	GLY	2.5
1	B	153	SER	2.5
1	A	188	LYS	2.3
1	B	209	GLN	2.3
2	D	1	PRO	2.1
1	A	187	ARG	2.0

### 6.2 Non-standard residues in protein, DNA, RNA chains [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
2	DVA	C	6	7/8	0.78	0.20	35,39,44,44	0
2	DVA	D	6	7/8	0.79	0.21	36,39,45,45	0
2	NLE	D	3	8/9	0.93	0.10	21,24,26,26	0
2	NLE	C	3	8/9	0.94	0.10	21,24,25,26	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. 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
5	SO4	A	1005	5/5	0.88	0.18	36,44,50,55	0
5	SO4	B	305	5/5	0.92	0.19	29,42,45,54	0
5	SO4	B	304	5/5	0.95	0.12	20,28,30,33	0
5	SO4	A	1004	5/5	0.96	0.09	19,30,32,33	0
4	CL	A	1003	1/1	0.97	0.08	14,14,14,14	0
5	SO4	A	1006	5/5	0.98	0.10	28,31,39,40	0
5	SO4	B	306	5/5	0.98	0.08	32,32,43,47	0
3	CD	A	1001	1/1	0.99	0.06	14,14,14,14	0
4	CL	B	303	1/1	0.99	0.06	14,14,14,14	0
3	CD	B	302	1/1	0.99	0.06	13,13,13,13	0
3	CD	A	1002	1/1	0.99	0.06	13,13,13,13	0
3	CD	B	301	1/1	0.99	0.07	14,14,14,14	0

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

There are no such residues in this entry.