



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 07:48 PM GMT

PDB ID : 1H8T  
Title : Echovirus 11  
Authors : Stuart, A.; Mckee, T.; Williams, P.A.; Harley, C.; Stuart, D.I.; Brown, T.D.K.;  
Lea, S.M.  
Deposited on : 2001-02-15  
Resolution : 2.90 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  
<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

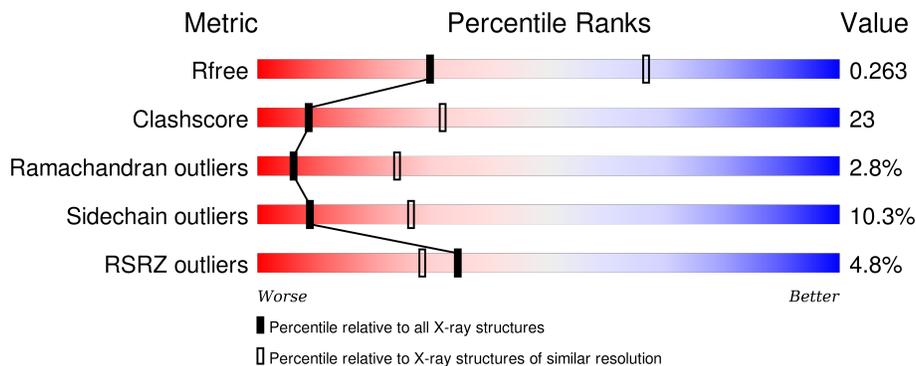
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.90 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))
$R_{free}$	91344	1451 (2.90-2.90)
Clashscore	102246	1668 (2.90-2.90)
Ramachandran outliers	100387	1630 (2.90-2.90)
Sidechain outliers	100360	1632 (2.90-2.90)
RSRZ outliers	91569	1456 (2.90-2.90)

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 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	292	
2	B	262	
3	C	238	
4	D	68	

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:

<b>Mol</b>	<b>Type</b>	<b>Chain</b>	<b>Res</b>	<b>Chirality</b>	<b>Geometry</b>	<b>Clashes</b>	<b>Electron density</b>
5	DOA	A	1000	-	-	-	X

## 2 Entry composition [i](#)

There are 7 unique types of molecules in this entry. The entry contains 6575 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 ECHOVIRUS 11 COAT PROTEIN VP1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	290	2284	1439	399	435	11	0	0	1

There are 14 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	48	VAL	MET	CONFLICT	UNP P29813
A	78	GLU	GLY	CONFLICT	UNP P29813
A	84	SER	THR	CONFLICT	UNP P29813
A	131	THR	SER	CONFLICT	UNP P29813
A	132	GLN	ARG	CONFLICT	UNP P29813
A	161	THR	ALA	CONFLICT	UNP P29813
A	267	SER	THR	CONFLICT	UNP P29813
A	270	ASP	ASN	CONFLICT	UNP P29813
A	271	ILE	VAL	CONFLICT	UNP P29813
A	276	ASN	THR	CONFLICT	UNP P29813
A	279	THR	ASN	CONFLICT	UNP P29813
A	283	ASP	GLU	CONFLICT	UNP P29813
A	289	VAL	LEU	CONFLICT	UNP P29813
A	292	HIS	TYR	CONFLICT	UNP P29813

- Molecule 2 is a protein called ECHOVIRUS 11 COAT PROTEIN VP2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	253	1964	1240	332	376	16	0	0	1

There are 12 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	1043	ARG	LYS	CONFLICT	UNP P29813
B	1045	ASP	ASN	CONFLICT	UNP P29813

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	1073	LYS	ARG	CONFLICT	UNP P29813
B	1108	ILE	LEU	CONFLICT	UNP P29813
B	1136	THR	GLN	CONFLICT	UNP P29813
B	1157	ALA	SER	CONFLICT	UNP P29813
B	1159	GLY	SER	CONFLICT	UNP P29813
B	1168	SER	THR	CONFLICT	UNP P29813
B	1185	PHE	TYR	CONFLICT	UNP P29813
B	1230	ASN	ASP	CONFLICT	UNP P29813
B	1235	PHE	SER	CONFLICT	UNP P29813
B	1260	ALA	SER	CONFLICT	UNP P29813

- Molecule 3 is a protein called ECHOVIRUS 11 COAT PROTEIN VP3.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
3	C	238	1819	1161	299	346	13	0	0	0

There are 13 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	2005	ILE	MET	CONFLICT	UNP P29813
C	2059	ALA	GLU	CONFLICT	UNP P29813
C	2061	ASN	LYS	CONFLICT	UNP P29813
C	2063	GLU	ASP	CONFLICT	UNP P29813
C	2066	ASP	GLU	CONFLICT	UNP P29813
C	2067	ILE	VAL	CONFLICT	UNP P29813
C	2080	SER	ASP	CONFLICT	UNP P29813
C	2093	GLY	SER	CONFLICT	UNP P29813
C	2107	TYR	PHE	CONFLICT	UNP P29813
C	2144	SER	ASN	CONFLICT	UNP P29813
C	2168	ILE	VAL	CONFLICT	UNP P29813
C	2232	GLN	GLU	CONFLICT	UNP P29813
C	2234	ALA	THR	CONFLICT	UNP P29813

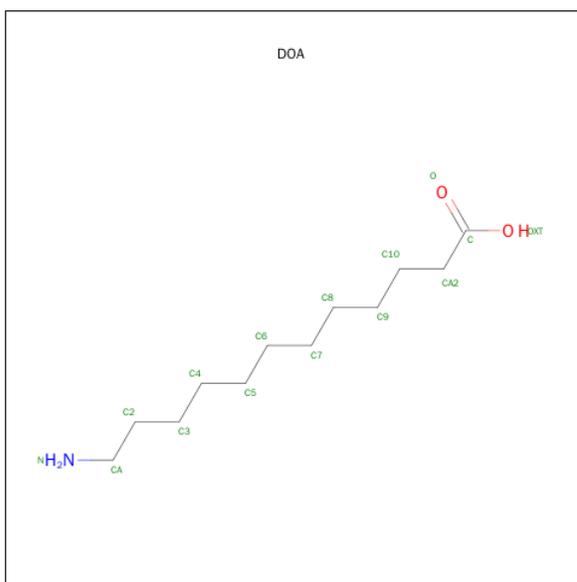
- Molecule 4 is a protein called ECHOVIRUS 11 COAT PROTEIN VP4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
4	D	61	467	289	82	95	1	0	0	1

There are 4 discrepancies between the modelled and reference sequences:

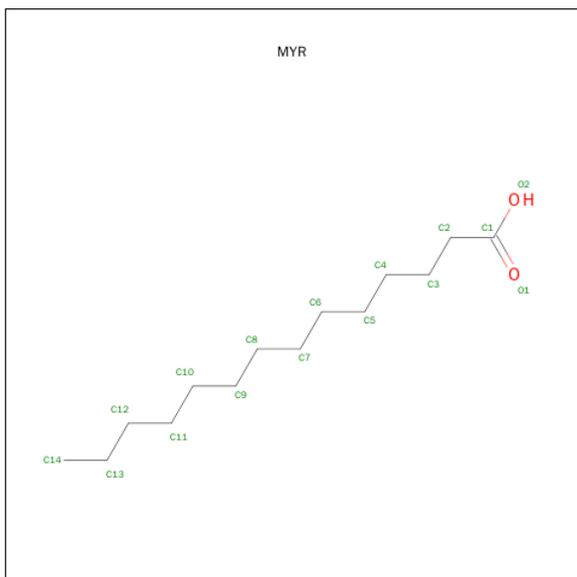
Chain	Residue	Modelled	Actual	Comment	Reference
D	3018	ARG	ASN	CONFLICT	UNP P29813
D	3022	ASN	SER	CONFLICT	UNP P29813
D	3045	ASP	GLU	CONFLICT	UNP P29813
D	3047	THR	SER	CONFLICT	UNP P29813

- Molecule 5 is 12-AMINO-DODECANOIC ACID (three-letter code: DOA) (formula:  $C_{12}H_{25}NO_2$ ).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
5	A	1	15	12	1	2	0	0

- Molecule 6 is MYRISTIC ACID (three-letter code: MYR) (formula:  $C_{14}H_{28}O_2$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	D	1	Total	C	O	0	0
			15	14	1		

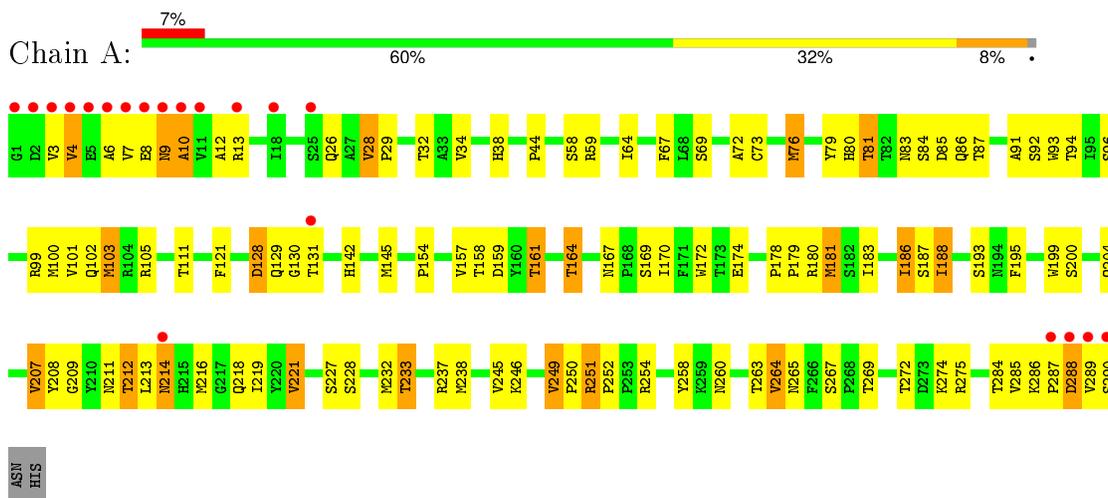
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	2	Total	O	0	0
			2	2		
7	B	4	Total	O	0	0
			4	4		
7	C	5	Total	O	0	0
			5	5		

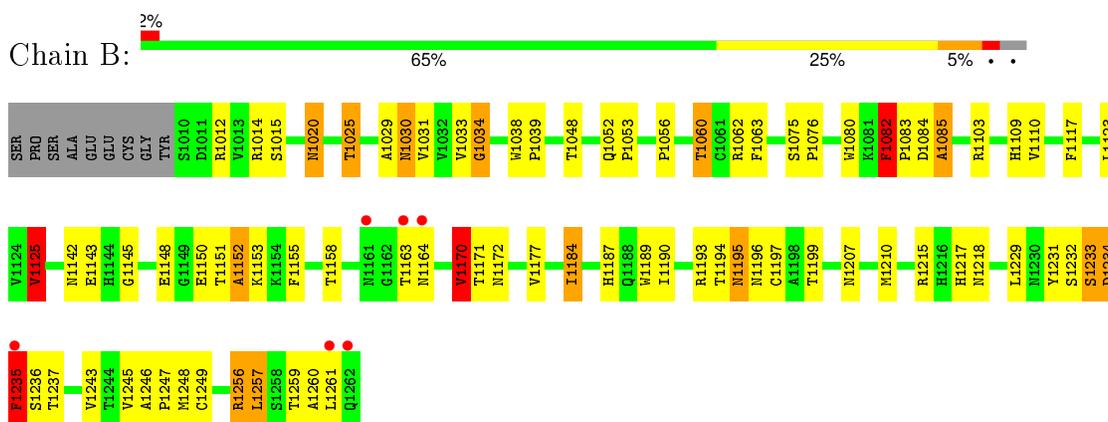
### 3 Residue-property plots [\(i\)](#)

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 ( $RSRZ > 2$ ). 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.

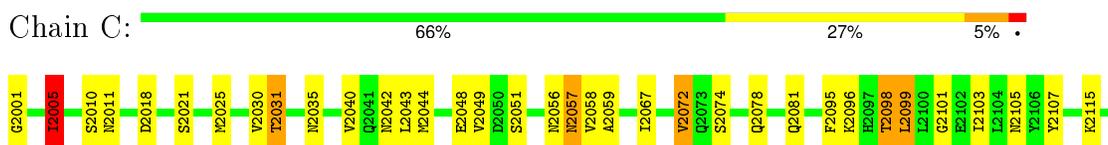
- Molecule 1: ECHOVIRUS 11 COAT PROTEIN VP1

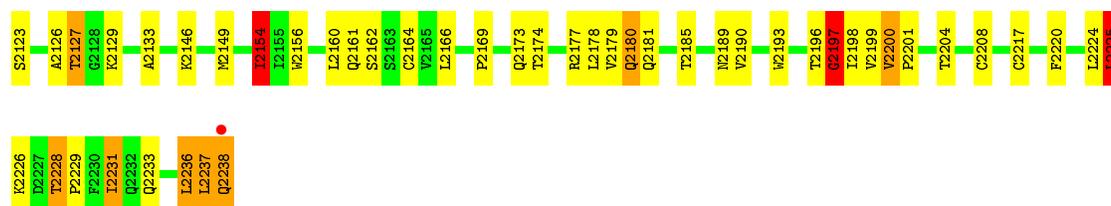


- Molecule 2: ECHOVIRUS 11 COAT PROTEIN VP2



- Molecule 3: ECHOVIRUS 11 COAT PROTEIN VP3





● Molecule 4: ECHOVIRUS 11 COAT PROTEIN VP4



## 4 Data and refinement statistics

Property	Value	Source
Space group	H 3 2	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	300.85Å 300.85Å 1476.62Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	100.00 – 2.90 39.78 – 2.86	Depositor EDS
% Data completeness (in resolution range)	71.3 (100.00-2.90) 69.7 (39.78-2.86)	Depositor EDS
$R_{merge}$	0.22	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.77 (at 2.86Å)	Xtriage
Refinement program	CNS 1.0	Depositor
R, $R_{free}$	0.246 , 0.255 0.265 , 0.263	Depositor DCC
$R_{free}$ test set	2014 reflections (0.50%)	DCC
Wilson B-factor (Å <sup>2</sup> )	22.7	Xtriage
Anisotropy	0.015	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.95 , 36.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.43$ , $\langle L^2 \rangle = 0.25$	Xtriage
Outliers	0 of 411034 reflections	Xtriage
$F_o, F_c$ correlation	0.81	EDS
Total number of atoms	6575	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	17.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.47% 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: DOA, MYR

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	1/2349 (0.0%)	0.95	3/3208 (0.1%)
2	B	0.77	0/2016	1.05	8/2755 (0.3%)
3	C	0.76	1/1866 (0.1%)	0.99	6/2549 (0.2%)
4	D	0.89	0/475	1.04	3/641 (0.5%)
All	All	0.78	2/6706 (0.0%)	1.00	20/9153 (0.2%)

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	76	MET	SD-CE	6.53	2.14	1.77
3	C	2001	GLY	N-CA	6.04	1.55	1.46

The worst 5 of 20 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	1195	ASN	N-CA-CB	-8.60	95.12	110.60
2	B	1082	PHE	N-CA-C	8.59	134.20	111.00
3	C	2237	LEU	CB-CG-CD1	-7.95	97.48	111.00
3	C	2197	GLY	N-CA-C	7.88	132.79	113.10
1	A	237	ARG	NE-CZ-NH2	-7.44	116.58	120.30

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	2284	0	2212	150	0
2	B	1964	0	1874	65	0
3	C	1819	0	1772	111	0
4	D	467	0	447	47	0
5	A	15	0	24	2	0
6	D	15	0	27	1	0
7	A	2	0	0	1	0
7	B	4	0	0	0	0
7	C	5	0	0	0	0
All	All	6575	0	6356	303	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 23.

The worst 5 of 303 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:76:MET:SD	1:A:76:MET:CE	2.14	1.34
1:A:260:ASN:OD1	1:A:263:THR:HG22	1.49	1.09
3:C:2098:THR:HG22	3:C:2101:GLY:H	1.20	1.04
4:D:3054:THR:O	4:D:3055:GLU:HB2	1.51	1.04
3:C:2233:GLN:NE2	3:C:2238:GLN:HG2	1.73	1.04

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	288/292 (99%)	259 (90%)	20 (7%)	9 (3%)	<b>5</b>   <b>21</b>
2	B	251/262 (96%)	220 (88%)	23 (9%)	8 (3%)	<b>5</b>   <b>20</b>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
3	C	236/238 (99%)	224 (95%)	10 (4%)	2 (1%)	24	60
4	D	57/68 (84%)	44 (77%)	9 (16%)	4 (7%)	1	4
All	All	832/860 (97%)	747 (90%)	62 (8%)	23 (3%)	6	24

5 of 23 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	8	GLU
1	A	9	ASN
1	A	214	ASN
1	A	228	SER
2	B	1152	ALA

### 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	259/262 (99%)	236 (91%)	23 (9%)	12	35
2	B	217/225 (96%)	199 (92%)	18 (8%)	14	38
3	C	204/204 (100%)	176 (86%)	28 (14%)	4	13
4	D	51/56 (91%)	45 (88%)	6 (12%)	6	19
All	All	731/747 (98%)	656 (90%)	75 (10%)	9	26

5 of 75 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	B	1171	THR
3	C	2005	ILE
4	D	3010	THR
2	B	1184	ILE
2	B	1245	VAL

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

Mol	Chain	Res	Type
2	B	1109	HIS
2	B	1195	ASN
4	D	3044	GLN
2	B	1119	GLN
2	B	1172	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](#)

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](#)

2 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	$\# Z  > 2$
5	DOA	A	1000	-	11,14,14	0.50	0	11,14,14	1.58	4 (36%)
6	MYR	D	3500	-	14,14,15	0.63	0	12,13,15	1.13	1 (8%)

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
5	DOA	A	1000	-	-	0/10/12/12	0/0/0/0
6	MYR	D	3500	-	-	0/11/12/13	0/0/0/0

There are no bond length outliers.

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed( $^{\circ}$ )	Ideal( $^{\circ}$ )
5	A	1000	DOA	C5-C4-C3	-2.85	99.83	114.53
5	A	1000	DOA	C7-C6-C5	-2.55	101.36	114.53
6	D	3500	MYR	C7-C6-C5	-2.45	101.89	114.53
5	A	1000	DOA	C3-C2-CA	-2.35	102.94	114.01
5	A	1000	DOA	C10-C9-C8	-2.27	102.80	114.53

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	A	1000	DOA	2	0
6	D	3500	MYR	1	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	290/292 (99%)	0.33	20 (6%) 20 14	1, 12, 71, 119	0
2	B	253/262 (96%)	0.00	6 (2%) 62 57	2, 9, 37, 85	0
3	C	238/238 (100%)	-0.18	1 (0%) 93 92	1, 8, 29, 75	0
4	D	61/68 (89%)	0.88	13 (21%) 1 0	6, 40, 83, 106	0
All	All	842/860 (97%)	0.13	40 (4%) 34 28	1, 10, 60, 119	0

The worst 5 of 40 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	1	GLY	13.3
1	A	289	VAL	12.4
1	A	4	VAL	10.3
2	B	1262	GLN	10.1
1	A	2	ASP	9.7

### 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. 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
5	DOA	A	1000	15/15	0.92	0.24	3.08	7,19,41,52	0
6	MYR	D	3500	15/16	0.71	0.43	-	24,51,73,75	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.