



# wwPDB X-ray Structure Validation Summary Report i

Feb 28, 2014 – 11:25 PM GMT

PDB ID : 2B5L  
Title : Crystal Structure of DDB1 In Complex with Simian Virus 5 V Protein  
Authors : Li, T.; Chen, X.; Garbutt, K.C.; Zhou, P.; Zheng, N.  
Deposited on : 2005-09-28  
Resolution : 2.85 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>

---

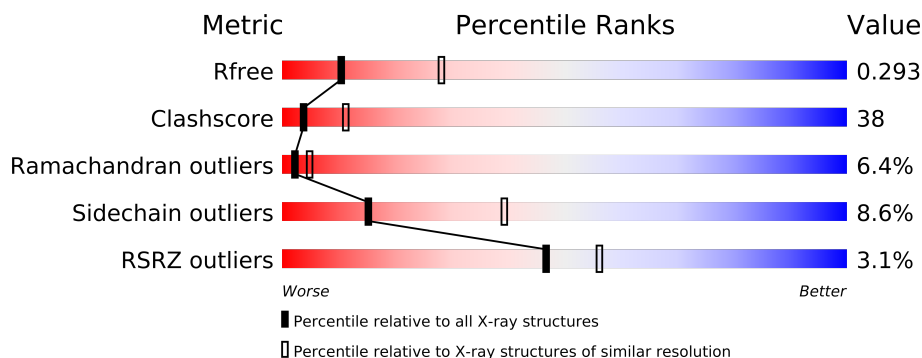
The following versions of software and data (see [references](#)) were used in the production of this report:

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 2.85 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	1524 (2.90-2.82)
Clashscore	79885	1879 (2.90-2.82)
Ramachandran outliers	78287	1824 (2.90-2.82)
Sidechain outliers	78261	1827 (2.90-2.82)
RSRZ outliers	66119	1526 (2.90-2.82)

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	1140	
1	B	1140	
2	C	222	
2	D	222	

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 20399 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 damage-specific DNA binding protein 1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	1132	Total	C	N	O	S	0	0	0
			8860	5610	1493	1708	49			
1	B	1134	Total	C	N	O	S	0	0	0
			8876	5619	1495	1713	49			

- Molecule 2 is a protein called Nonstructural protein V.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	C	174	Total	C	N	O	S	0	0	0
			1324	826	229	261	8			
2	D	175	Total	C	N	O	S	0	0	0
			1335	834	233	260	8			

- Molecule 3 is ZINC ION (three-letter code: ZN) (formula: Zn).

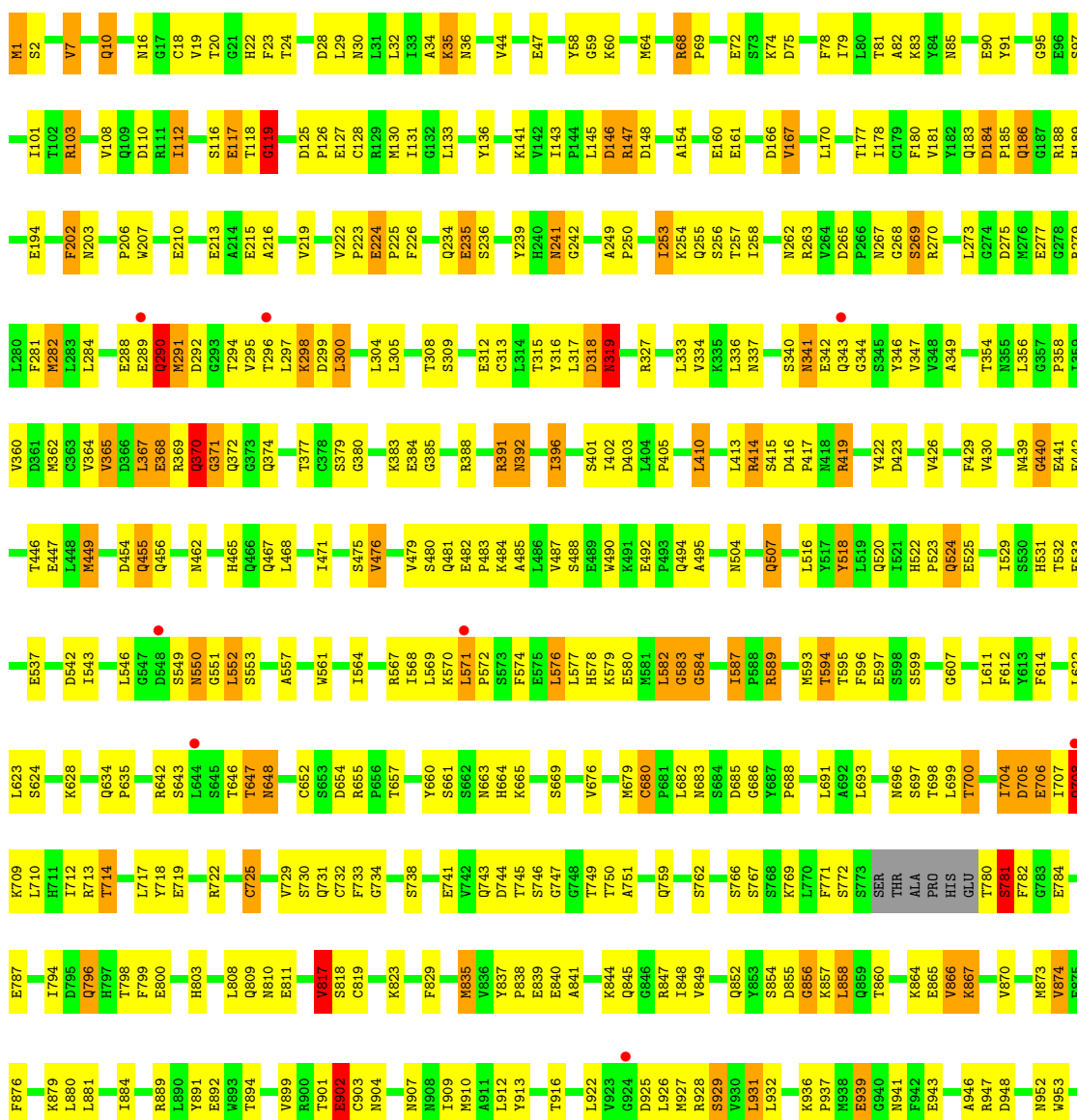
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	D	2	Total	Zn	0	0
			2	2		
3	C	2	Total	Zn	0	0
			2	2		

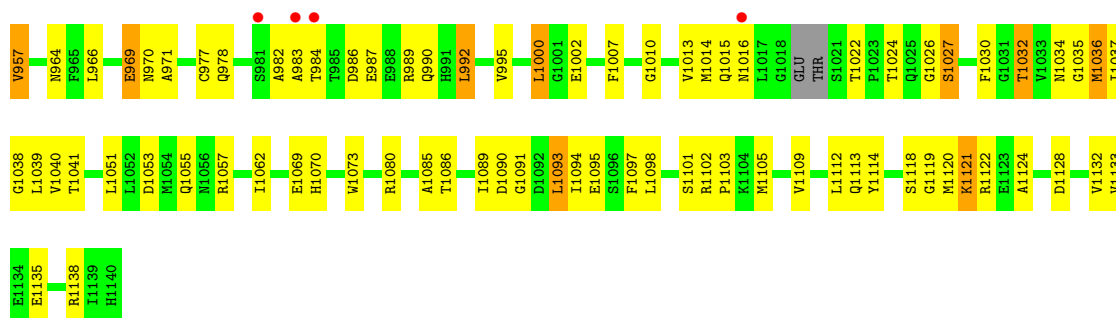
### 3 Residue-property plots

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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: damage-specific DNA binding protein 1

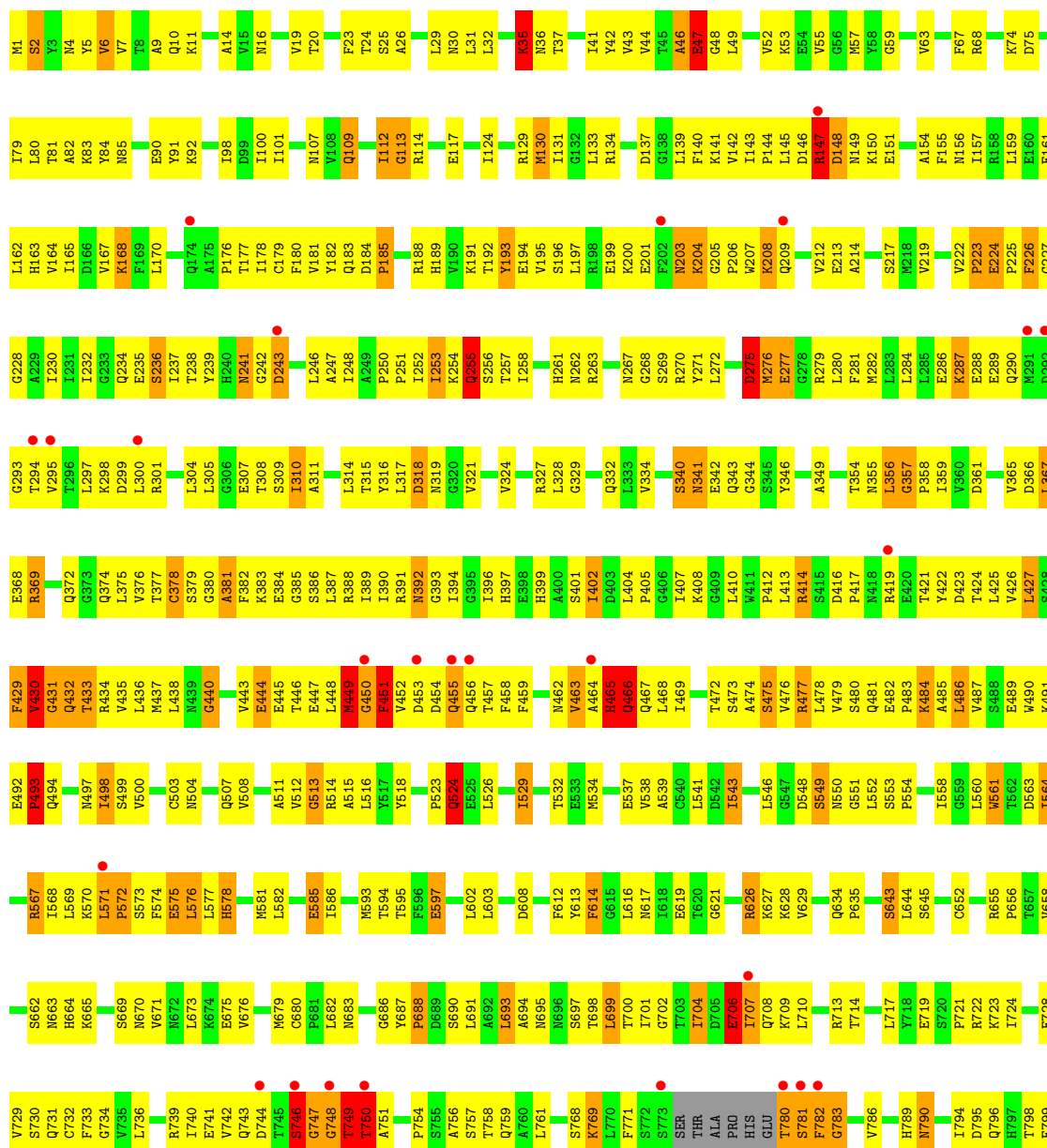
Chain A: 

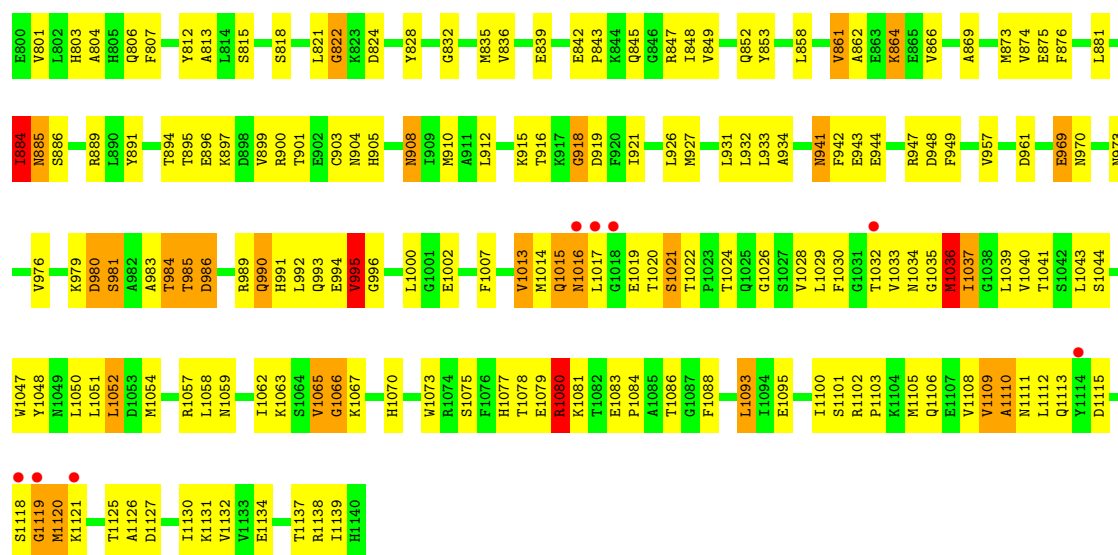




• Molecule 1: damage-specific DNA binding protein 1

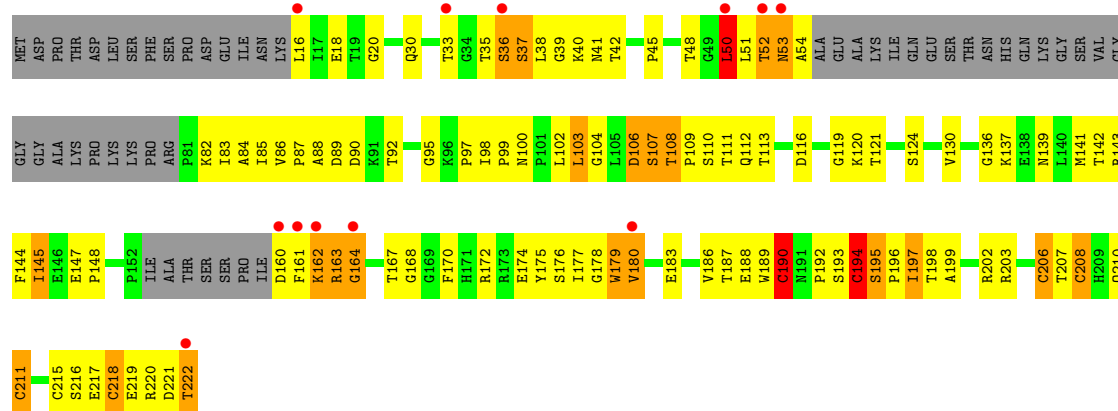
Chain B:





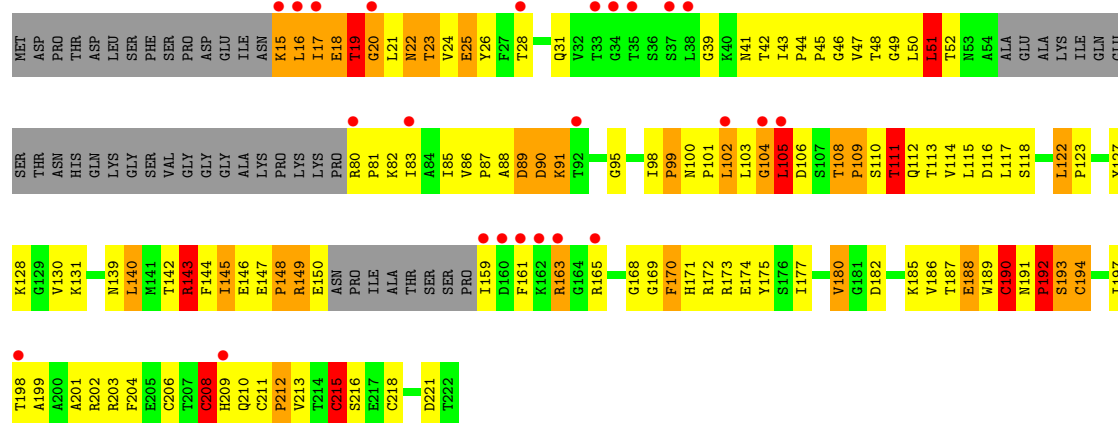
• Molecule 2: Nonstructural protein V

Chain C:



• Molecule 2: Nonstructural protein V

Chain D:



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	62.76Å 240.79Å 117.18Å 90.00° 101.79° 90.00°	Depositor
Resolution (Å)	47.80 – 2.85 47.77 – 2.63	Depositor EDS
% Data completeness (in resolution range)	(Not available) (47.80-2.85) 86.5 (47.77-2.63)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.84 (at 2.65Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.229 , 0.299 0.226 , 0.293	Depositor DCC
$R_{free}$ test set	3805 reflections (5.02%)	DCC
Wilson B-factor (Å <sup>2</sup> )	46.9	Xtriage
Anisotropy	0.604	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 35.4	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	0 of 92110 reflections	Xtriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	20399	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	64.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.54% 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:  
ZN

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	0.45	1/9022 (0.0%)	0.77	7/12219 (0.1%)
1	B	0.45	3/9039 (0.0%)	0.91	41/12244 (0.3%)
2	C	0.56	0/1351	0.95	9/1832 (0.5%)
2	D	0.71	2/1361 (0.1%)	1.24	19/1843 (1.0%)
All	All	0.48	6/20773 (0.0%)	0.88	76/28138 (0.3%)

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	B	1	1
2	D	0	1
All	All	1	2

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

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	189	TRP	C-N	-15.06	0.99	1.34
2	D	190	CYS	C-N	11.09	1.59	1.34
1	B	748	GLY	N-CA	-7.48	1.34	1.46
1	B	431	GLY	N-CA	-6.93	1.35	1.46
1	A	725	CYS	CB-SG	-5.43	1.73	1.81

The worst 5 of 76 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	749	THR	N-CA-C	-19.59	58.10	111.00
1	B	35	LYS	CB-CA-C	18.15	146.70	110.40

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	15	LYS	N-CA-C	-16.98	65.16	111.00
1	B	431	GLY	N-CA-C	-16.54	71.75	113.10
1	B	749	THR	N-CA-CB	14.29	137.44	110.30

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
1	B	1	MET	CA

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	B	430	VAL	Peptide
2	D	190	CYS	Mainchain

## 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	8860	0	8828	466	0
1	B	8876	0	8842	793	0
2	C	1324	0	1306	128	0
2	D	1335	0	1323	177	0
3	C	2	0	0	0	0
3	D	2	0	0	0	0
All	All	20399	0	20299	1549	0

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 38.

The worst 5 of 1549 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:367:LEU:HD21	1:B:374:GLN:CG	1.51	1.38
1:B:367:LEU:CD2	1:B:374:GLN:HB2	1.53	1.37
1:B:367:LEU:HD21	1:B:374:GLN:CB	1.50	1.37

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:B:742:VAL:CG2	1:B:750:THR:HG21	1.57	1.35
1:B:367:LEU:CD2	1:B:374:GLN:CG	2.06	1.34

There are no symmetry-related clashes.

## 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	1126/1140 (99%)	960 (85%)	128 (11%)	38 (3%)	6	21
1	B	1130/1140 (99%)	892 (79%)	145 (13%)	93 (8%)	1	3
2	C	168/222 (76%)	123 (73%)	33 (20%)	12 (7%)	2	4
2	D	169/222 (76%)	115 (68%)	32 (19%)	22 (13%)	0	1
All	All	2593/2724 (95%)	2090 (81%)	338 (13%)	165 (6%)	2	5

5 of 165 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	254	LYS
1	A	269	SER
1	A	290	GLN
1	A	341	ASN
1	A	370	GLN

### 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	992/999 (99%)	900 (91%)	92 (9%)	13	35

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	B	994/999 (100%)	925 (93%)	69 (7%)	22	52
2	C	151/191 (79%)	134 (89%)	17 (11%)	9	23
2	D	151/191 (79%)	133 (88%)	18 (12%)	8	19
All	All	2288/2380 (96%)	2092 (91%)	196 (9%)	15	39

5 of 196 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	1069	GLU
1	B	402	ILE
2	D	51	LEU
1	A	1135	GLU
1	B	147	ARG

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 82 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	1056	ASN
1	B	255	GLN
2	C	210	GLN
1	A	1070	HIS
1	B	22	HIS

### 5.3.3 RNA ⓘ

There are no RNA chains in this entry.

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

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

## 5.5 Carbohydrates ⓘ

There are no carbohydrates in this entry.

## 5.6 Ligand geometry ⓘ

Of 4 ligands modelled in this entry, 4 are monoatomic - leaving 0 for Mogul analysis.

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.

## 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	1132/1140 (99%)	-0.35	12 (1%) 77 84	18, 47, 99, 160	0
1	B	1134/1140 (99%)	-0.03	34 (2%) 48 57	29, 68, 122, 173	0
2	C	174/222 (78%)	0.19	12 (6%) 17 19	32, 62, 118, 183	0
2	D	175/222 (78%)	0.54	24 (13%) 4 4	39, 81, 138, 168	0
All	All	2615/2724 (95%)	-0.11	82 (3%) 47 56	18, 59, 117, 183	0

The worst 5 of 82 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	1018	GLY	6.3
1	B	750	THR	5.5
1	B	780	THR	5.5
2	D	16	LEU	4.4
2	D	161	PHE	4.3

### 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 ⓘ

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	ZN	C	3001	1/1	0.09	-1.48	35,35,35,35	0
3	ZN	D	3003	1/1	0.05	-1.86	35,35,35,35	0
3	ZN	C	3002	1/1	0.07	-2.51	35,35,35,35	0
3	ZN	D	3004	1/1	0.08	-2.68	35,35,35,35	0

## 6.5 Other polymers ⓘ

There are no such residues in this entry.