



# wwPDB X-ray Structure Validation Summary Report ⓘ

Feb 1, 2016 – 12:30 PM GMT

PDB ID : 3REB  
Title : HIV-1 Nef protein in complex with engineered Hck-SH3 domain  
Authors : Schulte, A.; Blankenfeldt, W.; Geyer, M.  
Deposited on : 2011-04-04  
Resolution : 3.45 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

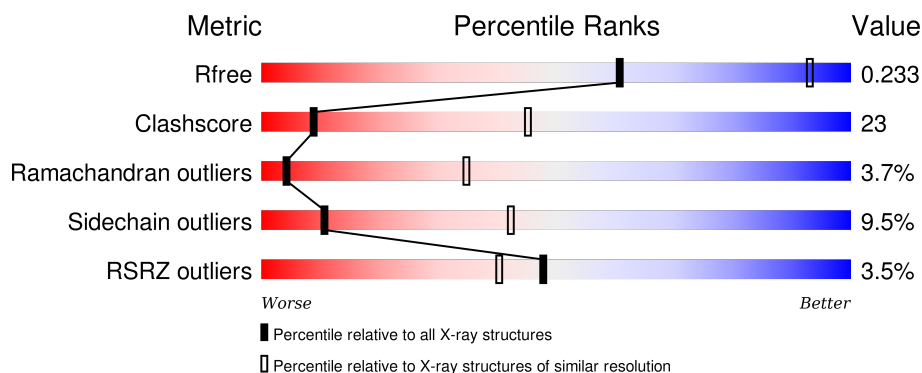
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 3.45 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	1000 (3.56-3.36)
Clashscore	102246	1090 (3.56-3.36)
Ramachandran outliers	100387	1057 (3.56-3.36)
Sidechain outliers	100360	1058 (3.56-3.36)
RSRZ outliers	91569	1005 (3.56-3.36)

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	166	<div> <div>4%</div> <div>32% 28% • • 36%</div> </div>
1	C	166	<div> <div>2%</div> <div>33% 30% 5% • 31%</div> </div>
2	B	90	<div> <div>2%</div> <div>46% 19% • 34%</div> </div>
2	D	90	<div> <div></div> <div>44% 20% • 33%</div> </div>

## 2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 2761 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 Protein Nef.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	106	Total	C	N	O	S	0	0	0
			886	587	153	143	3			
1	C	114	Total	C	N	O	S	0	0	0
			914	602	156	153	3			

There are 8 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	47	MET	ILE	ENGINEERED MUTATION	UNP P03407
A	48	ALA	THR	ENGINEERED MUTATION	UNP P03407
A	59	SER	CYS	ENGINEERED MUTATION	UNP P03407
A	210	ALA	CYS	ENGINEERED MUTATION	UNP P03407
C	47	MET	ILE	ENGINEERED MUTATION	UNP P03407
C	48	ALA	THR	ENGINEERED MUTATION	UNP P03407
C	59	SER	CYS	ENGINEERED MUTATION	UNP P03407
C	210	ALA	CYS	ENGINEERED MUTATION	UNP P03407

- Molecule 2 is a protein called Tyrosine-protein kinase HCK.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	59	Total	C	N	O	S	0	0	0
			479	307	77	94	1			
2	D	60	Total	C	N	O	S	0	0	0
			482	309	75	97	1			

There are 72 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	78	MET	-	INITIATING METHIONINE	UNP P08631
B	90	VAL	GLU	ENGINEERED MUTATION	UNP P08631
B	91	SER	ALA	ENGINEERED MUTATION	UNP P08631
B	92	TRP	ILE	ENGINEERED MUTATION	UNP P08631

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	93	SER	HIS	ENGINEERED MUTATION	UNP P08631
B	94	PRO	HIS	ENGINEERED MUTATION	UNP P08631
B	95	ASP	GLU	ENGINEERED MUTATION	UNP P08631
B	139	GLY	-	EXPRESSION TAG	UNP P08631
B	140	GLY	-	EXPRESSION TAG	UNP P08631
B	141	GLY	-	EXPRESSION TAG	UNP P08631
B	142	THR	-	EXPRESSION TAG	UNP P08631
B	143	SER	-	EXPRESSION TAG	UNP P08631
B	144	GLY	-	EXPRESSION TAG	UNP P08631
B	145	GLY	-	EXPRESSION TAG	UNP P08631
B	146	GLY	-	EXPRESSION TAG	UNP P08631
B	147	ARG	-	EXPRESSION TAG	UNP P08631
B	148	HIS	-	EXPRESSION TAG	UNP P08631
B	149	ARG	-	EXPRESSION TAG	UNP P08631
B	150	ARG	-	EXPRESSION TAG	UNP P08631
B	151	ARG	-	EXPRESSION TAG	UNP P08631
B	152	GLN	-	EXPRESSION TAG	UNP P08631
B	153	ALA	-	EXPRESSION TAG	UNP P08631
B	154	GLU	-	EXPRESSION TAG	UNP P08631
B	155	ARG	-	EXPRESSION TAG	UNP P08631
B	156	MET	-	EXPRESSION TAG	UNP P08631
B	157	SER	-	EXPRESSION TAG	UNP P08631
B	158	GLN	-	EXPRESSION TAG	UNP P08631
B	159	ILE	-	EXPRESSION TAG	UNP P08631
B	160	LYS	-	EXPRESSION TAG	UNP P08631
B	161	ARG	-	EXPRESSION TAG	UNP P08631
B	162	LEU	-	EXPRESSION TAG	UNP P08631
B	163	LEU	-	EXPRESSION TAG	UNP P08631
B	164	SER	-	EXPRESSION TAG	UNP P08631
B	165	GLU	-	EXPRESSION TAG	UNP P08631
B	166	LYS	-	EXPRESSION TAG	UNP P08631
B	167	LYS	-	EXPRESSION TAG	UNP P08631
D	78	MET	-	INITIATING METHIONINE	UNP P08631
D	90	VAL	GLU	ENGINEERED MUTATION	UNP P08631
D	91	SER	ALA	ENGINEERED MUTATION	UNP P08631
D	92	TRP	ILE	ENGINEERED MUTATION	UNP P08631
D	93	SER	HIS	ENGINEERED MUTATION	UNP P08631
D	94	PRO	HIS	ENGINEERED MUTATION	UNP P08631
D	95	ASP	GLU	ENGINEERED MUTATION	UNP P08631
D	139	GLY	-	EXPRESSION TAG	UNP P08631
D	140	GLY	-	EXPRESSION TAG	UNP P08631
D	141	GLY	-	EXPRESSION TAG	UNP P08631

*Continued on next page...*

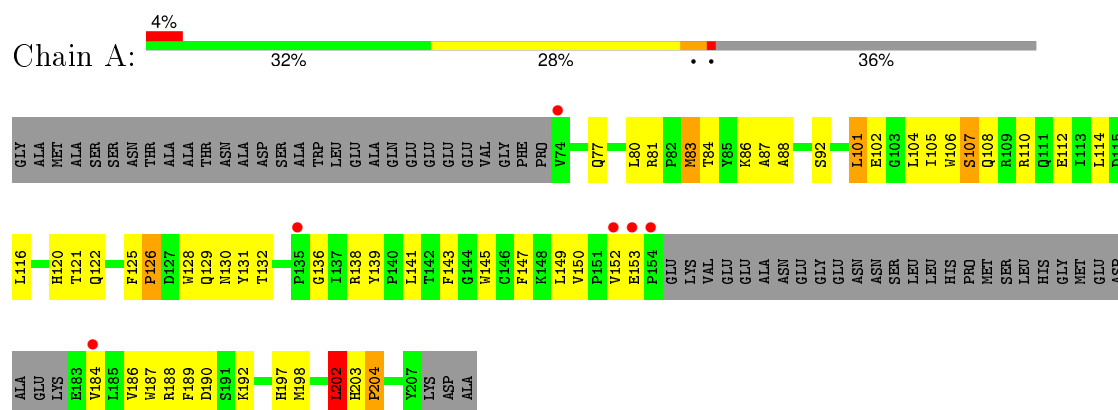
*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
D	142	THR	-	EXPRESSION TAG	UNP P08631
D	143	SER	-	EXPRESSION TAG	UNP P08631
D	144	GLY	-	EXPRESSION TAG	UNP P08631
D	145	GLY	-	EXPRESSION TAG	UNP P08631
D	146	GLY	-	EXPRESSION TAG	UNP P08631
D	147	ARG	-	EXPRESSION TAG	UNP P08631
D	148	HIS	-	EXPRESSION TAG	UNP P08631
D	149	ARG	-	EXPRESSION TAG	UNP P08631
D	150	ARG	-	EXPRESSION TAG	UNP P08631
D	151	ARG	-	EXPRESSION TAG	UNP P08631
D	152	GLN	-	EXPRESSION TAG	UNP P08631
D	153	ALA	-	EXPRESSION TAG	UNP P08631
D	154	GLU	-	EXPRESSION TAG	UNP P08631
D	155	ARG	-	EXPRESSION TAG	UNP P08631
D	156	MET	-	EXPRESSION TAG	UNP P08631
D	157	SER	-	EXPRESSION TAG	UNP P08631
D	158	GLN	-	EXPRESSION TAG	UNP P08631
D	159	ILE	-	EXPRESSION TAG	UNP P08631
D	160	LYS	-	EXPRESSION TAG	UNP P08631
D	161	ARG	-	EXPRESSION TAG	UNP P08631
D	162	LEU	-	EXPRESSION TAG	UNP P08631
D	163	LEU	-	EXPRESSION TAG	UNP P08631
D	164	SER	-	EXPRESSION TAG	UNP P08631
D	165	GLU	-	EXPRESSION TAG	UNP P08631
D	166	LYS	-	EXPRESSION TAG	UNP P08631
D	167	LYS	-	EXPRESSION TAG	UNP P08631

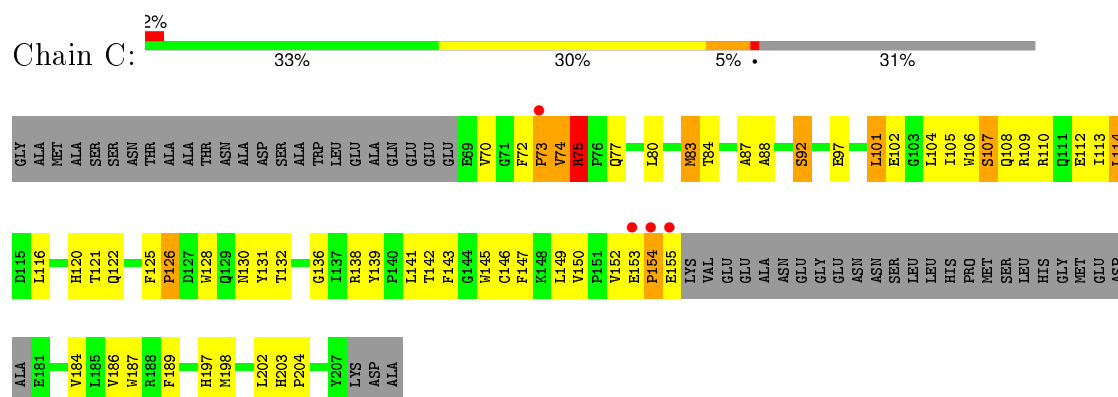
### 3 Residue-property plots [i](#)

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.

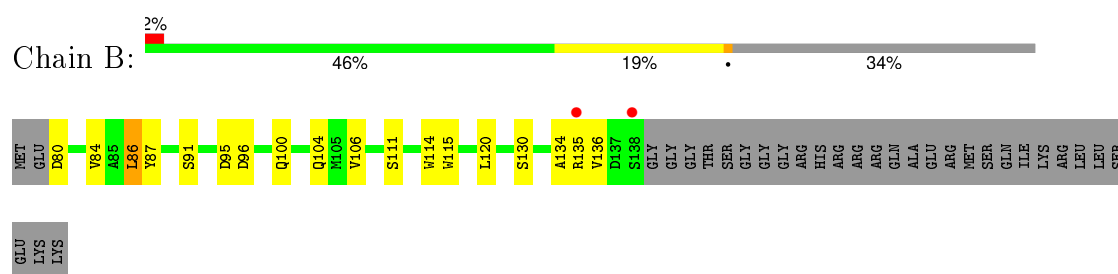
#### • Molecule 1: Protein Nef



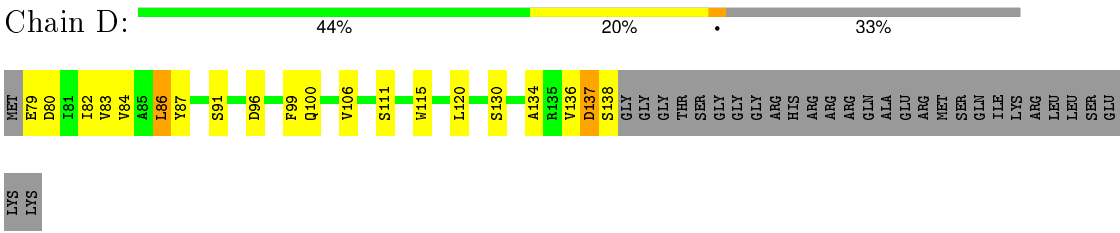
#### • Molecule 1: Protein Nef



#### • Molecule 2: Tyrosine-protein kinase HCK



#### • Molecule 2: Tyrosine-protein kinase HCK



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 61	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	112.46Å 112.46Å 130.03Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	19.86 – 3.45 19.86 – 3.45	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.86-3.45) 100.0 (19.86-3.45)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.37 (at 3.44Å)	Xtriage
Refinement program	REFMAC 5.5.0102	Depositor
R, $R_{free}$	0.216 , 0.243 0.212 , 0.233	Depositor DCC
$R_{free}$ test set	634 reflections (5.44%)	DCC
Wilson B-factor (Å <sup>2</sup> )	114.2	Xtriage
Anisotropy	0.043	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.26 , 86.9	EDS
Estimated twinning fraction	0.076 for h,-h-k,-l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.45$ , $\langle L^2 \rangle = 0.28$	Xtriage
Outliers	0 of 12285 reflections	Xtriage
$F_o, F_c$ correlation	0.93	EDS
Total number of atoms	2761	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	131.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.31% 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

### 5.1 Standard geometry

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	0.68	0/921	0.69	0/1255
1	C	0.84	1/948 (0.1%)	0.93	2/1291 (0.2%)
2	B	0.82	0/491	0.74	0/668
2	D	0.86	0/494	0.76	0/673
All	All	0.79	1/2854 (0.0%)	0.80	2/3887 (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	1
1	C	0	1
All	All	0	2

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	146	CYS	CB-SG	-6.70	1.70	1.82

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	75	ARG	N-CA-CB	18.67	144.21	110.60
1	C	75	ARG	N-CA-C	-9.77	84.63	111.00

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	202	LEU	Peptide

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Group
1	C	202	LEU	Peptide

## 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	886	0	854	51	0
1	C	914	0	854	49	0
2	B	479	0	456	16	0
2	D	482	0	451	15	0
All	All	2761	0	2615	126	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 126 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:C:74:VAL:HG22	1:C:74:VAL:O	1.55	1.06
1:C:152:VAL:HG23	1:C:153:GLU:O	1.68	0.91
1:C:130:ASN:HB2	1:C:141:LEU:HD12	1.58	0.85
1:A:130:ASN:HB2	1:A:141:LEU:HD12	1.60	0.83
1:C:72:PHE:H	2:D:87:TYR:HE1	1.27	0.82

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	102/166 (61%)	88 (86%)	11 (11%)	3 (3%)	6	40
1	C	110/166 (66%)	91 (83%)	11 (10%)	8 (7%)	1	15
2	B	57/90 (63%)	54 (95%)	3 (5%)	0	100	100
2	D	58/90 (64%)	56 (97%)	1 (2%)	1 (2%)	11	52
All	All	327/512 (64%)	289 (88%)	26 (8%)	12 (4%)	4	35

5 of 12 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	70	VAL
1	C	73	PRO
1	C	75	ARG
1	C	203	HIS
2	D	137	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	91/141 (64%)	84 (92%)	7 (8%)	16	53
1	C	90/141 (64%)	82 (91%)	8 (9%)	12	46
2	B	52/76 (68%)	47 (90%)	5 (10%)	10	42
2	D	52/76 (68%)	45 (86%)	7 (14%)	5	25
All	All	285/434 (66%)	258 (90%)	27 (10%)	11	42

5 of 27 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	74	VAL
1	C	92	SER
2	D	111	SER
1	C	80	LEU
1	A	184	VAL

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	129	GLN
1	C	120	HIS

### 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](#)

There are no ligands in this entry.

## 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	106/166 (63%)	0.07	6 (5%) 27 24	74, 147, 226, 348	0
1	C	114/166 (68%)	-0.25	4 (3%) 48 40	63, 115, 185, 286	0
2	B	59/90 (65%)	-0.32	2 (3%) 49 42	91, 118, 176, 219	0
2	D	60/90 (66%)	-0.41	0 100 100	78, 112, 169, 205	0
All	All	339/512 (66%)	-0.19	12 (3%) 48 40	63, 124, 216, 348	0

The worst 5 of 12 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	154	PRO	21.4
1	A	153	GLU	6.1
1	C	73	PRO	3.6
1	C	154	PRO	3.1
2	B	138	SER	2.9

### 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](#)

There are no ligands in this entry.

## 6.5 Other polymers [i](#)

There are no such residues in this entry.