



# wwPDB X-ray Structure Validation Summary Report ⓘ

May 13, 2020 – 07:22 am BST

PDB ID : 3PBP  
Title : Structure of the yeast heterotrimeric Nup82-Nup159-Nup116 nucleoporin complex  
Authors : Debler, E.W.; Hoelz, A.  
Deposited on : 2010-10-20  
Resolution : 2.60 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

<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

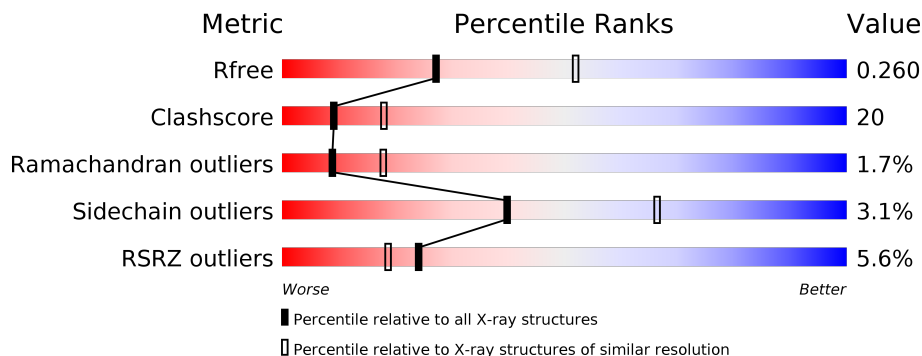
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

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))
$R_{free}$	130704	3163 (2.60-2.60)
Clashscore	141614	3518 (2.60-2.60)
Ramachandran outliers	138981	3455 (2.60-2.60)
Sidechain outliers	138945	3455 (2.60-2.60)
RSRZ outliers	127900	3104 (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 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	452	
1	D	452	
1	G	452	
1	J	452	
2	B	148	
2	E	148	

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	H	148	
2	K	148	
3	C	36	
3	F	36	
3	I	36	
3	L	36	

## 2 Entry composition [i](#)

There are 4 unique types of molecules in this entry. The entry contains 19653 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 Nucleoporin NUP82.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
1	A	438	3539	2273	569	686	4	7	0	0	0
1	D	437	3533	2270	568	684	4	7	0	0	0
1	G	438	3539	2273	569	686	4	7	0	0	0
1	J	439	3550	2279	573	687	4	7	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	396	SER	CYS	ENGINEERED MUTATION	UNP P40368
D	396	SER	CYS	ENGINEERED MUTATION	UNP P40368
G	396	SER	CYS	ENGINEERED MUTATION	UNP P40368
J	396	SER	CYS	ENGINEERED MUTATION	UNP P40368

- Molecule 2 is a protein called Nucleoporin NUP116/NSP116.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	S	Se			
2	B	146	1165	745	204	212	3	1	0	0	0
2	E	146	1165	745	204	212	3	1	0	0	0
2	H	146	1165	745	204	212	3	1	0	0	0
2	K	146	1165	745	204	212	3	1	0	0	0

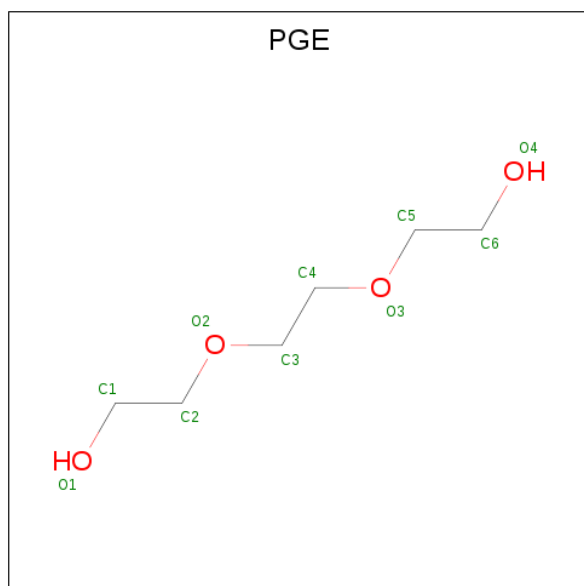
There are 4 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	966	MSE	-	INITIATING METHIONINE	UNP Q02630
E	966	MSE	-	INITIATING METHIONINE	UNP Q02630
H	966	MSE	-	EXPRESSION TAG	UNP Q02630
K	966	MSE	-	INITIATING METHIONINE	UNP Q02630

- Molecule 3 is a protein called Nucleoporin NUP159.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	C	26	Total	C	N	O	Se	0	0	0
			204	133	34	35	2			
3	F	26	Total	C	N	O	Se	0	0	0
			204	133	34	35	2			
3	I	28	Total	C	N	O	Se	0	0	0
			222	145	37	38	2			
3	L	24	Total	C	N	O	Se	0	0	0
			192	126	32	33	1			

- Molecule 4 is TRIETHYLENE GLYCOL (three-letter code: PGE) (formula: C<sub>6</sub>H<sub>14</sub>O<sub>4</sub>).

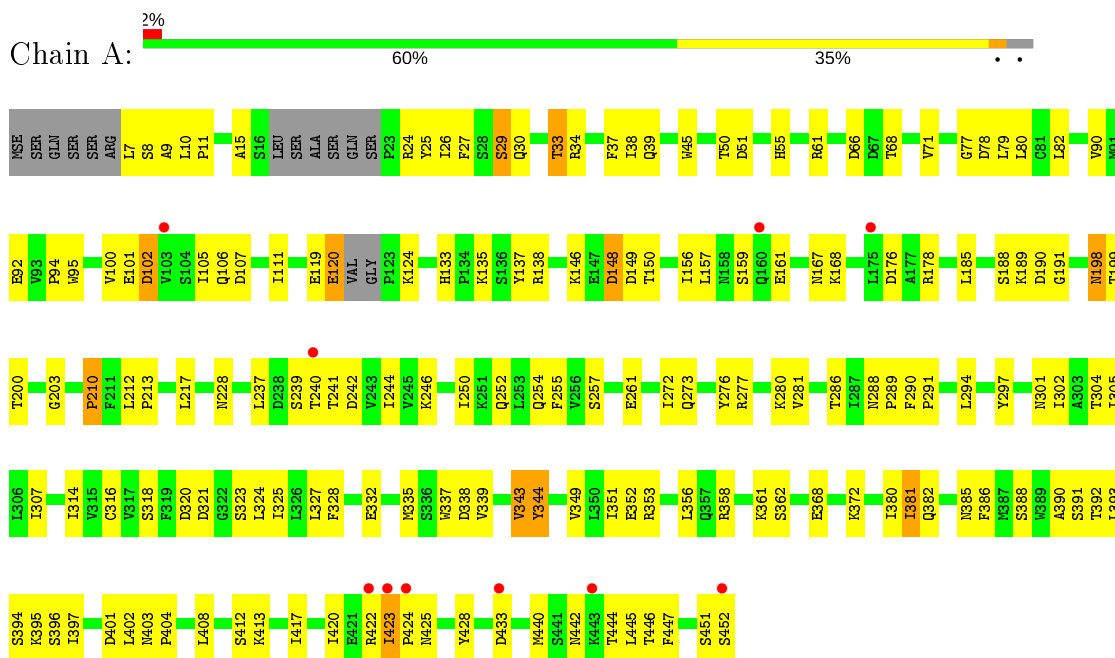


Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	D	1	Total	C	O	0	0
			10	6	4		

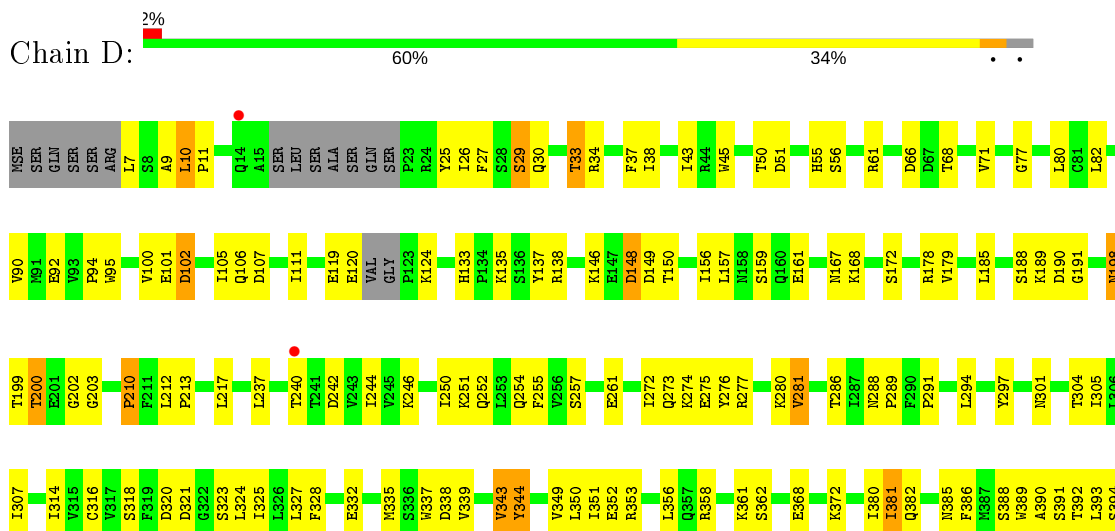
### 3 Residue-property plots

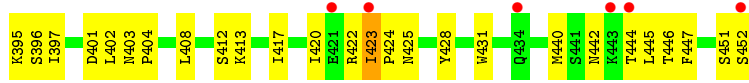
These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are 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: Nucleoporin NUP82

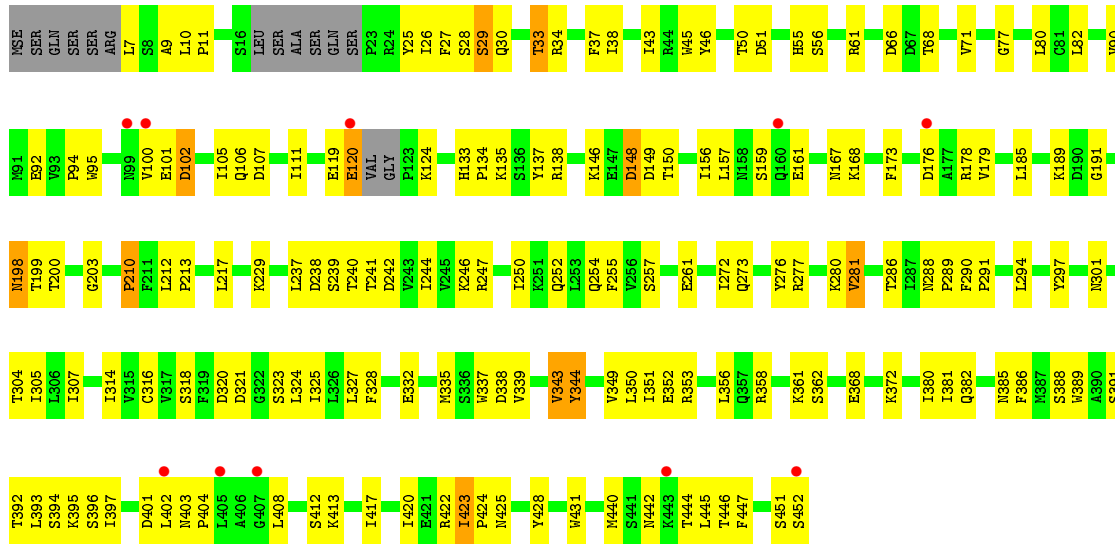


- Molecule 1: Nucleoporin NUP82

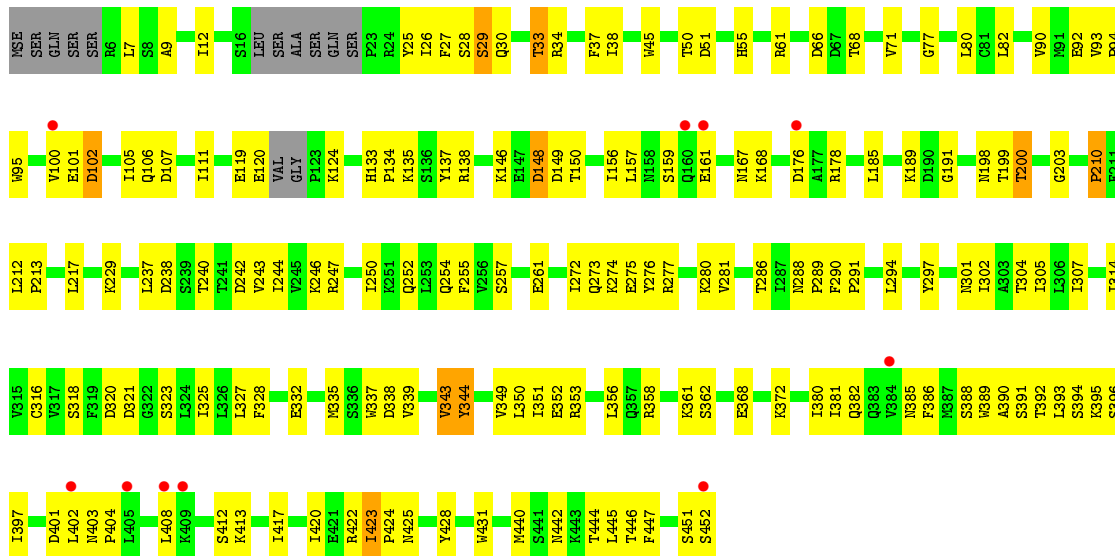




• Molecule 1: Nucleoporin NUP82

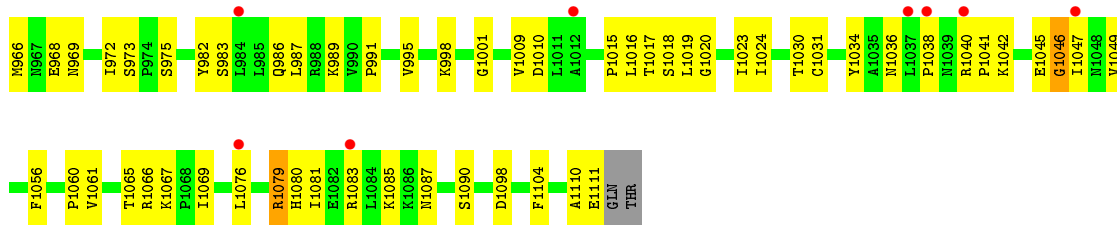


• Molecule 1: Nucleoporin NUP82

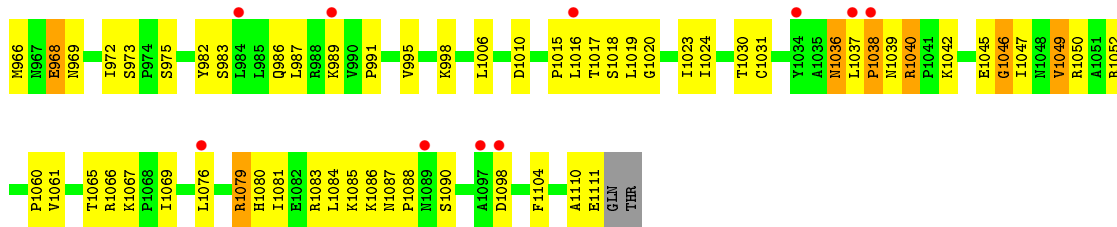


• Molecule 2: Nucleoporin NUP116/NSP116

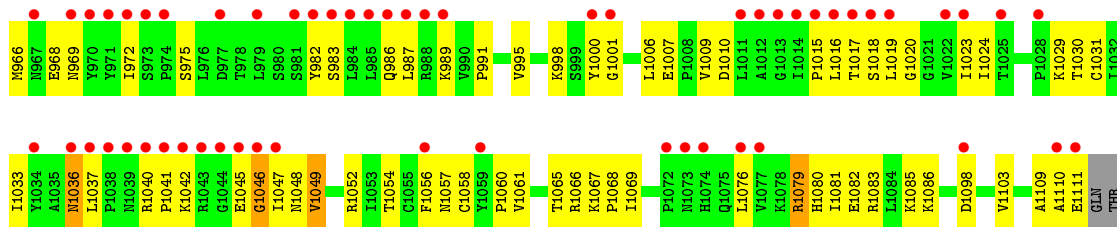




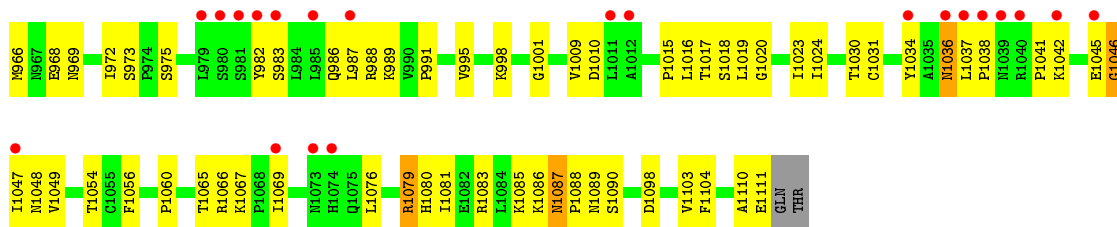
• Molecule 2: Nucleoporin NUP116/NSP116



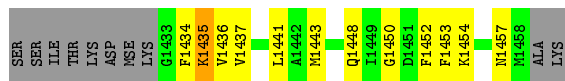
• Molecule 2: Nucleoporin NUP116/NSP116



• Molecule 2: Nucleoporin NUP116/NSP116



• Molecule 3: Nucleoporin NUP159



• Molecule 3: Nucleoporin NUP159

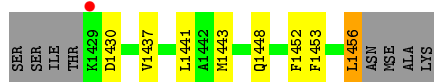


Chain F:  44% 25% 28%



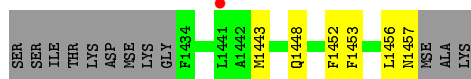
- Molecule 3: Nucleoporin NUP159

Chain I:  3% 56% 19% 22%



- Molecule 3: Nucleoporin NUP159

Chain L:  3% 50% 17% 33%



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	61.50Å 96.77Å 144.28Å 105.98° 93.97° 108.24°	Depositor
Resolution (Å)	50.00 – 2.60 47.11 – 2.81	Depositor EDS
% Data completeness (in resolution range)	(Not available) (50.00-2.60) 93.3 (47.11-2.81)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.10	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.40 (at 2.81Å)	Xtrriage
Refinement program	CNS 1.2	Depositor
R, $R_{free}$	0.257 , 0.272 0.233 , 0.260	Depositor DCC
$R_{free}$ test set	6890 reflections (9.80%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	53.4	Xtrriage
Anisotropy	0.889	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.31 , 40.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	0.013 for -h,-k,h+k+l	Xtrriage
$F_o, F_c$ correlation	0.92	EDS
Total number of atoms	19653	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	74.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.64% 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: PGE

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	0.43	0/3603	0.68	2/4877 (0.0%)
1	D	0.42	0/3597	0.68	2/4869 (0.0%)
1	G	0.42	0/3603	0.68	1/4877 (0.0%)
1	J	0.43	0/3614	0.68	1/4891 (0.0%)
2	B	0.35	0/1193	0.64	0/1617
2	E	0.35	0/1193	0.62	0/1617
2	H	0.41	0/1193	0.63	0/1617
2	K	0.37	0/1193	0.63	0/1617
3	C	0.53	0/204	0.61	0/266
3	F	0.48	0/204	0.56	0/266
3	I	0.44	0/222	0.64	0/288
3	L	0.49	0/193	0.52	0/254
All	All	0.41	0/20012	0.66	6/27056 (0.0%)

There are no bond length outliers.

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

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	J	343	VAL	N-CA-C	-5.42	96.38	111.00
1	G	343	VAL	N-CA-C	-5.15	97.09	111.00
1	D	343	VAL	N-CA-C	-5.10	97.23	111.00
1	A	343	VAL	N-CA-C	-5.05	97.36	111.00
1	A	381	ILE	N-CA-C	-5.04	97.38	111.00

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	3539	0	3507	141	0
1	D	3533	0	3502	146	0
1	G	3539	0	3507	154	0
1	J	3550	0	3520	157	0
2	B	1165	0	1183	39	0
2	E	1165	0	1183	46	0
2	H	1165	0	1183	54	0
2	K	1165	0	1183	46	0
3	C	204	0	213	11	0
3	F	204	0	213	12	0
3	I	222	0	237	13	0
3	L	192	0	201	10	0
4	D	10	0	14	5	0
All	All	19653	0	19646	784	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 20.

The worst 5 of 784 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)
2:H:1010:ASP:HB3	2:H:1046:GLY:HA2	1.41	0.99
1:G:10:LEU:HD12	1:G:11:PRO:HD2	1.51	0.92
1:J:444:THR:HG22	1:J:445:LEU:H	1.40	0.86
3:I:1437:VAL:HG11	1:J:243:VAL:HG12	1.58	0.85
1:G:444:THR:HG22	1:G:445:LEU:H	1.40	0.84

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	432/452 (96%)	396 (92%)	32 (7%)	4 (1%)	17	35
1	D	431/452 (95%)	397 (92%)	29 (7%)	5 (1%)	13	27
1	G	432/452 (96%)	396 (92%)	32 (7%)	4 (1%)	17	35
1	J	433/452 (96%)	399 (92%)	29 (7%)	5 (1%)	13	27
2	B	144/148 (97%)	120 (83%)	19 (13%)	5 (4%)	3	5
2	E	144/148 (97%)	117 (81%)	20 (14%)	7 (5%)	2	2
2	H	144/148 (97%)	121 (84%)	19 (13%)	4 (3%)	5	7
2	K	144/148 (97%)	120 (83%)	19 (13%)	5 (4%)	3	5
3	C	24/36 (67%)	21 (88%)	2 (8%)	1 (4%)	3	3
3	F	24/36 (67%)	21 (88%)	3 (12%)	0	100	100
3	I	26/36 (72%)	25 (96%)	1 (4%)	0	100	100
3	L	22/36 (61%)	21 (96%)	1 (4%)	0	100	100
All	All	2400/2544 (94%)	2154 (90%)	206 (9%)	40 (2%)	9	18

5 of 40 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	344	TYR
2	B	1036	ASN
3	C	1435	LYS
1	D	344	TYR
1	G	344	TYR

### 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	411/415 (99%)	396 (96%)	15 (4%)	35	61

*Continued on next page...*

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	D	410/415 (99%)	398 (97%)	12 (3%)	42	68
1	G	411/415 (99%)	398 (97%)	13 (3%)	39	65
1	J	412/415 (99%)	400 (97%)	12 (3%)	42	68
2	B	132/133 (99%)	128 (97%)	4 (3%)	41	67
2	E	132/133 (99%)	128 (97%)	4 (3%)	41	67
2	H	132/133 (99%)	128 (97%)	4 (3%)	41	67
2	K	132/133 (99%)	128 (97%)	4 (3%)	41	67
3	C	22/28 (79%)	22 (100%)	0	100	100
3	F	22/28 (79%)	21 (96%)	1 (4%)	27	52
3	I	24/28 (86%)	22 (92%)	2 (8%)	11	22
3	L	21/28 (75%)	21 (100%)	0	100	100
All	All	2261/2304 (98%)	2190 (97%)	71 (3%)	40	66

5 of 71 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
2	E	1049	VAL
1	G	102	ASP
1	J	281	VAL
2	E	1079	ARG
1	G	29	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
2	E	969	ASN
2	E	1075	GLN
1	J	262	ASN
1	D	430	ASN
2	E	967	ASN

### 5.3.3 RNA ⓘ

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](#)

1 ligand is 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	$\# Z  > 2$
4	PGE	D	6119	-	9,9,9	0.92	1 (11%)	8,8,8	0.98	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
4	PGE	D	6119	-	-	1/7/7/7	-

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	D	6119	PGE	C5-C6	2.23	1.61	1.49

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	D	6119	PGE	O1-C1-C2-O2

There are no ring outliers.

1 monomer is involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	D	6119	PGE	5	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	431/452 (95%)	0.15	10 (2%) 60 54	44, 65, 97, 118	0
1	D	430/452 (95%)	0.13	8 (1%) 66 62	41, 65, 98, 118	0
1	G	431/452 (95%)	0.12	10 (2%) 60 54	43, 65, 97, 117	0
1	J	432/452 (95%)	0.14	10 (2%) 60 54	41, 64, 99, 117	0
2	B	145/148 (97%)	0.45	8 (5%) 25 19	53, 90, 116, 129	0
2	E	145/148 (97%)	0.41	10 (6%) 16 12	52, 90, 116, 129	0
2	H	145/148 (97%)	1.80	56 (38%) 0 0	58, 94, 119, 144	0
2	K	145/148 (97%)	0.73	21 (14%) 2 1	56, 92, 117, 137	0
3	C	24/36 (66%)	0.09	0 100 100	59, 69, 91, 108	0
3	F	24/36 (66%)	0.29	0 100 100	58, 69, 89, 109	0
3	I	26/36 (72%)	0.25	1 (3%) 40 33	57, 71, 118, 123	0
3	L	23/36 (63%)	0.10	1 (4%) 35 28	62, 70, 89, 96	0
All	All	2401/2544 (94%)	0.31	135 (5%) 24 19	41, 70, 106, 144	0

The worst 5 of 135 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	H	1038	PRO	7.5
2	H	1037	LEU	7.4
2	H	1036	ASN	6.9
2	H	1039	ASN	6.8
2	K	1037	LEU	6.6

### 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. 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
4	PGE	D	6119	10/10	0.82	0.26	70,72,77,80	0

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

There are no such residues in this entry.