



# wwPDB X-ray Structure Validation Summary Report ⓘ

Jan 31, 2016 – 10:59 PM GMT

PDB ID : 1VYJ  
Title : STRUCTURAL AND BIOCHEMICAL STUDIES OF HUMAN PCNA COMPLEXES PROVIDE THE BASIS FOR ASSOCIATION WITH CDK/CYCLIN AND RATIONALE FOR INHIBITOR DESIGN  
Authors : Kontopidis, G.; Wu, S.; Zheleva, D.; Taylor, P.; Mcinnes, C.; Lane, D.; Fischer, P.; Walkinshaw, M.D.  
Deposited on : 2004-04-30  
Resolution : 2.80 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

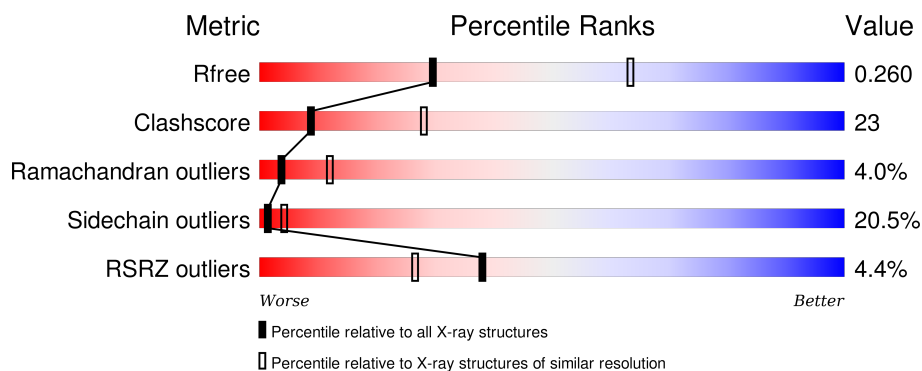
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.80 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	2393 (2.80-2.80)
Clashscore	102246	2827 (2.80-2.80)
Ramachandran outliers	100387	2782 (2.80-2.80)
Sidechain outliers	100360	2784 (2.80-2.80)
RSRZ outliers	91569	2404 (2.80-2.80)

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	261	<div> <div>4%</div> <div>56%</div> <div>33%</div> <div>7%</div> <div>..</div> </div>
1	C	261	<div> <div>2%</div> <div>58%</div> <div>29%</div> <div>9%</div> <div>..</div> </div>
1	E	261	<div> <div>5%</div> <div>57%</div> <div>31%</div> <div>9%</div> <div>..</div> </div>
1	G	261	<div> <div>4%</div> <div>51%</div> <div>35%</div> <div>13%</div> <div>.</div> </div>
1	I	261	<div> <div>3%</div> <div>52%</div> <div>37%</div> <div>7%</div> <div>..</div> </div>

*Continued on next page...*

Continued from previous page...

Mol	Chain	Length	Quality of chain
1	K	261	<div><div></div><div>4%</div><div>48%</div><div>36%</div><div>13%</div><div></div><div></div></div>
2	B	16	<div><div></div><div>25%</div><div>38%</div><div>38%</div><div></div><div></div></div>
2	D	16	<div><div></div><div>19%</div><div>50%</div><div>25%</div><div>25%</div><div></div></div>
2	F	16	<div><div></div><div>19%</div><div>44%</div><div>38%</div><div>19%</div><div></div></div>
2	H	16	<div><div></div><div>13%</div><div>50%</div><div>25%</div><div>25%</div><div></div></div>
2	J	16	<div><div></div><div>13%</div><div>50%</div><div>38%</div><div>13%</div><div></div></div>
2	L	16	<div><div></div><div>19%</div><div>38%</div><div>44%</div><div>19%</div><div></div></div>

## 2 Entry composition

There are 3 unique types of molecules in this entry. The entry contains 13057 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 PROLIFERATING CELL NUCLEAR ANTIGEN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	257	Total	C	N	O	S	0	0	0
			1980	1242	324	398	16			
1	C	257	Total	C	N	O	S	0	3	0
			2005	1257	328	404	16			
1	E	257	Total	C	N	O	S	0	0	0
			1980	1242	324	398	16			
1	G	260	Total	C	N	O	S	0	0	0
			2002	1254	327	405	16			
1	I	257	Total	C	N	O	S	0	0	0
			1980	1242	324	398	16			
1	K	257	Total	C	N	O	S	0	0	0
			1980	1242	324	398	16			

- Molecule 2 is a protein called SMALL PEPTIDE SAVLQKKITDYFHPKK.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
2	B	16	Total	C	N	O	0	0	0
			135	89	23	23			
2	D	16	Total	C	N	O	0	0	0
			135	89	23	23			
2	F	16	Total	C	N	O	0	0	0
			135	89	23	23			
2	H	16	Total	C	N	O	0	0	0
			135	89	23	23			
2	J	16	Total	C	N	O	0	0	0
			135	89	23	23			
2	L	16	Total	C	N	O	0	0	0
			135	89	23	23			

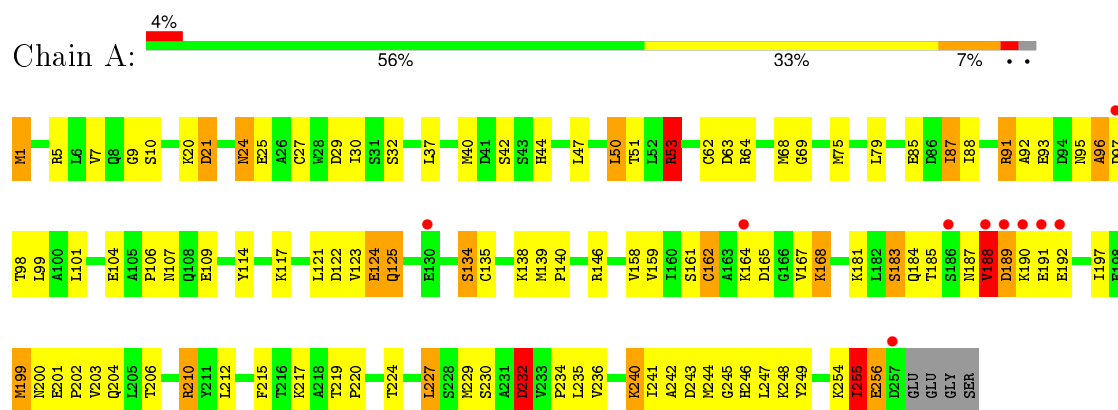
- Molecule 3 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	41	Total 41	O 41	0	0
3	B	3	Total 3	O 3	0	0
3	C	50	Total 50	O 50	0	0
3	D	3	Total 3	O 3	0	0
3	E	59	Total 59	O 59	0	0
3	G	51	Total 51	O 51	0	0
3	H	5	Total 5	O 5	0	0
3	I	56	Total 56	O 56	0	0
3	J	5	Total 5	O 5	0	0
3	K	45	Total 45	O 45	0	0
3	L	2	Total 2	O 2	0	0

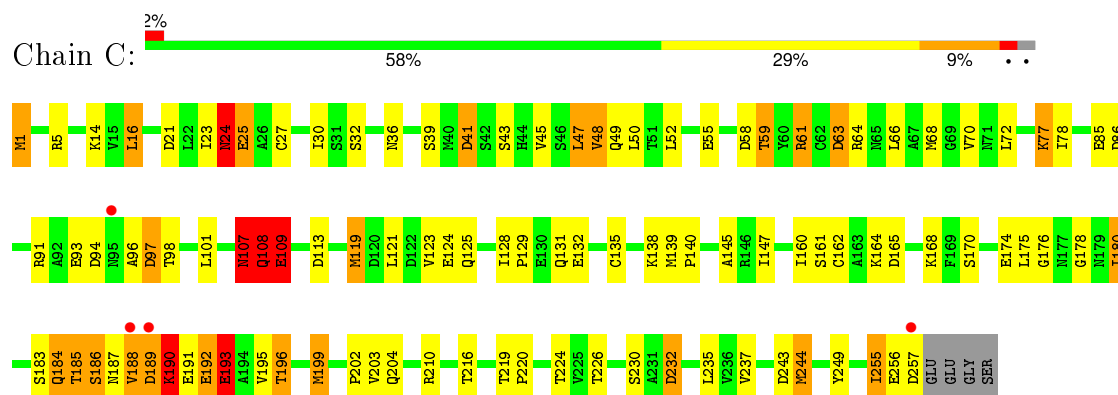
### 3 Residue-property plots [i](#)

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

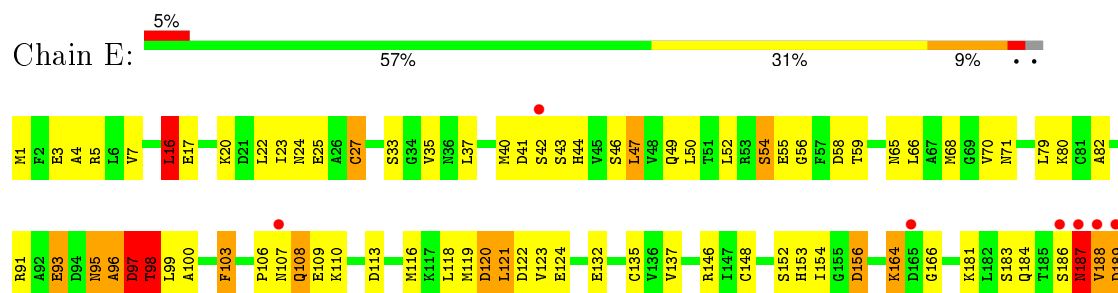
#### • Molecule 1: PROLIFERATING CELL NUCLEAR ANTIGEN



#### • Molecule 1: PROLIFERATING CELL NUCLEAR ANTIGEN

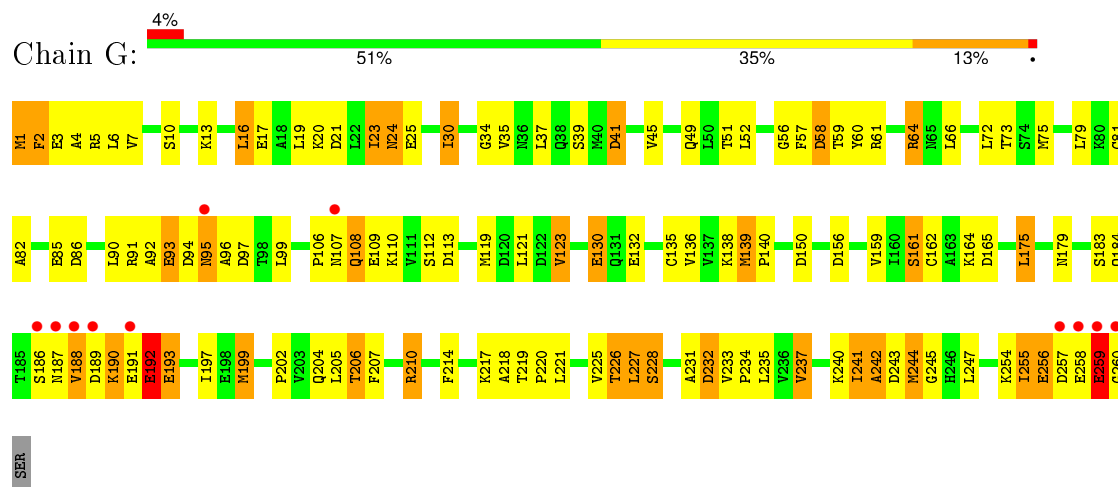


#### • Molecule 1: PROLIFERATING CELL NUCLEAR ANTIGEN

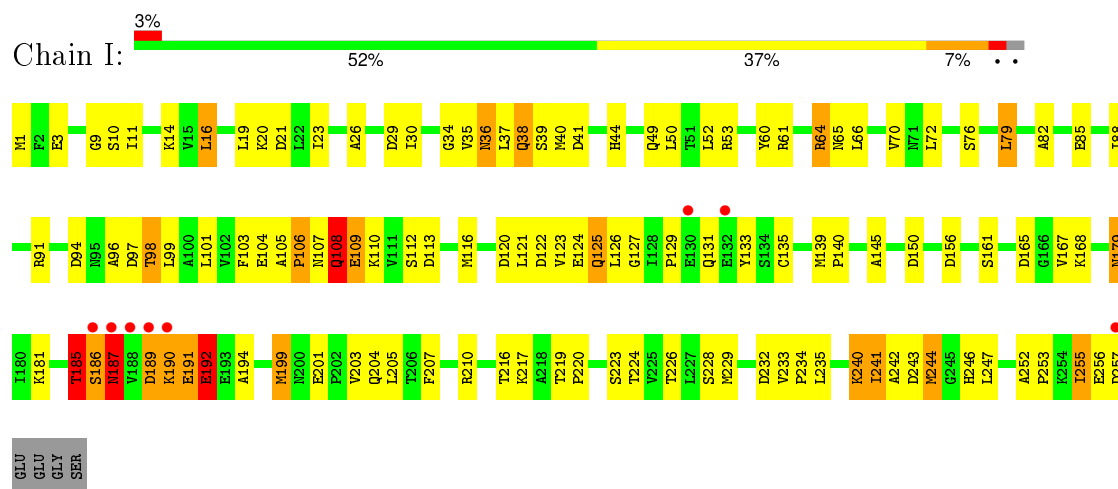




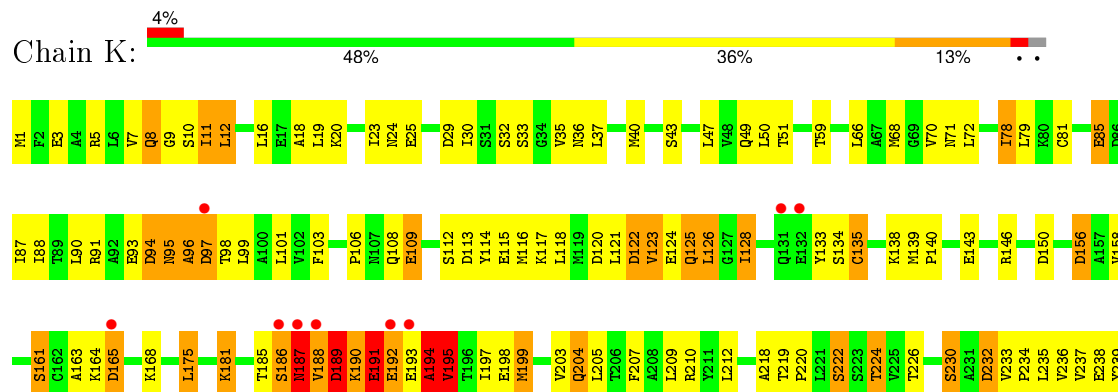
● Molecule 1: PROLIFERATING CELL NUCLEAR ANTIGEN



● Molecule 1: PROLIFERATING CELL NUCLEAR ANTIGEN

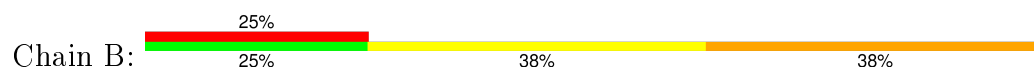


● Molecule 1: PROLIFERATING CELL NUCLEAR ANTIGEN

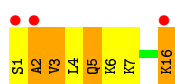




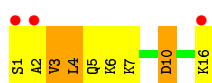
- Molecule 2: SMALL PEPTIDE SAVLQKKITDYFHPKK



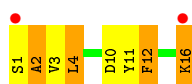
- Molecule 2: SMALL PEPTIDE SAVLQKKITDYFHPKK



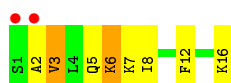
- Molecule 2: SMALL PEPTIDE SAVLQKKITDYFHPKK



- Molecule 2: SMALL PEPTIDE SAVLQKKITDYFHPKK



- Molecule 2: SMALL PEPTIDE SAVLQKKITDYFHPKK



- Molecule 2: SMALL PEPTIDE SAVLQKKITDYFHPKK





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 31 2 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	119.10Å 119.10Å 305.82Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	14.00 – 2.80 13.99 – 2.80	Depositor EDS
% Data completeness (in resolution range)	99.5 (14.00-2.80) 99.5 (13.99-2.80)	Depositor EDS
$R_{merge}$	0.09	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.74 (at 2.80Å)	Xtriage
Refinement program	REFMAC	Depositor
R, $R_{free}$	0.176 , 0.256 0.191 , 0.260	Depositor DCC
$R_{free}$ test set	1917 reflections (3.19%)	DCC
Wilson B-factor (Å <sup>2</sup> )	61.1	Xtriage
Anisotropy	0.027	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.36 , 66.2	EDS
Estimated twinning fraction	0.019 for -h,-k,l	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	1 of 61989 reflections (0.002%)	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	13057	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	58.0	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.

<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	0.63	1/2006 (0.0%)	1.11	11/2710 (0.4%)
1	C	0.84	1/2031 (0.0%)	1.15	11/2744 (0.4%)
1	E	0.66	1/2006 (0.0%)	1.11	9/2710 (0.3%)
1	G	0.65	1/2028 (0.0%)	1.12	7/2739 (0.3%)
1	I	0.67	1/2006 (0.0%)	1.13	12/2710 (0.4%)
1	K	0.65	0/2006	1.09	8/2710 (0.3%)
2	B	0.71	0/138	1.12	1/182 (0.5%)
2	D	0.55	0/138	0.87	0/182
2	F	0.66	0/138	1.14	1/182 (0.5%)
2	H	0.53	0/138	1.06	1/182 (0.5%)
2	J	0.56	0/138	0.95	0/182
2	L	0.69	0/138	1.04	0/182
All	All	0.68	5/12911 (0.0%)	1.11	61/17415 (0.4%)

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	C	0	3
1	E	0	4
1	G	0	1
1	I	0	2
1	K	0	4
All	All	0	14

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	255	ILE	C-N	20.87	1.82	1.34
1	G	255	ILE	C-N	6.81	1.49	1.34
1	A	255	ILE	C-N	5.59	1.47	1.34

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	I	252	ALA	C-N	-5.28	1.24	1.34
1	E	253	PRO	C-N	5.15	1.45	1.34

The worst 5 of 61 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	180	ILE	O-C-N	-11.38	104.50	122.70
1	I	192	GLU	O-C-N	10.51	139.51	122.70
1	A	245	GLY	O-C-N	-8.53	109.06	122.70
1	E	16	LEU	CA-CB-CG	-8.46	95.85	115.30
1	I	192	GLU	CA-C-N	-8.36	98.82	117.20

There are no chirality outliers.

5 of 14 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	107	ASN	Peptide
1	C	180	ILE	Mainchain
1	C	193	GLU	Mainchain
1	E	110	LYS	Mainchain
1	E	98	THR	Mainchain

## 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	1980	0	1983	90	0
1	C	2005	0	2001	71	2
1	E	1980	0	1981	91	1
1	G	2002	0	1996	106	3
1	I	1980	0	1981	88	0
1	K	1980	0	1981	120	0
2	B	135	0	146	22	0
2	D	135	0	146	10	0
2	F	135	0	146	10	0
2	H	135	0	146	10	0
2	J	135	0	146	3	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	L	135	0	146	8	0
3	A	41	0	0	9	0
3	B	3	0	0	0	0
3	C	50	0	0	5	0
3	D	3	0	0	0	0
3	E	59	0	0	5	0
3	G	51	0	0	3	0
3	H	5	0	0	0	0
3	I	56	0	0	2	0
3	J	5	0	0	0	0
3	K	45	0	0	5	0
3	L	2	0	0	0	0
All	All	13057	0	12799	599	3

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 599 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:255:ILE:C	1:C:256:GLU:N	1.82	1.33
1:G:108:GLN:O	1:G:109:GLU:CG	1.82	1.24
1:E:189:ASP:OD1	1:E:189:ASP:O	1.55	1.22
2:F:2:ALA:O	2:F:3:VAL:HG13	1.08	1.20
1:G:108:GLN:O	1:G:109:GLU:HG2	1.01	1.16

All (3) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:77:LYS:NZ	1:G:130:GLU:OE2[6_555]	1.42	0.78
1:E:42:SER:OG	1:G:243:ASP:OD1[6_555]	1.94	0.26
1:C:77:LYS:CE	1:G:130:GLU:OE2[6_555]	2.08	0.12

## 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	255/261 (98%)	236 (92%)	15 (6%)	4 (2%)	12	38
1	C	258/261 (99%)	233 (90%)	14 (5%)	11 (4%)	3	10
1	E	255/261 (98%)	229 (90%)	16 (6%)	10 (4%)	4	12
1	G	258/261 (99%)	237 (92%)	12 (5%)	9 (4%)	4	15
1	I	255/261 (98%)	236 (92%)	13 (5%)	6 (2%)	7	25
1	K	255/261 (98%)	224 (88%)	17 (7%)	14 (6%)	2	6
2	B	14/16 (88%)	9 (64%)	2 (14%)	3 (21%)	0	0
2	D	14/16 (88%)	11 (79%)	1 (7%)	2 (14%)	0	1
2	F	14/16 (88%)	10 (71%)	3 (21%)	1 (7%)	1	3
2	H	14/16 (88%)	11 (79%)	2 (14%)	1 (7%)	1	3
2	J	14/16 (88%)	11 (79%)	1 (7%)	2 (14%)	0	1
2	L	14/16 (88%)	12 (86%)	0	2 (14%)	0	1
All	All	1620/1662 (98%)	1459 (90%)	96 (6%)	65 (4%)	4	12

5 of 65 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	96	ALA
1	A	188	VAL
2	B	2	ALA
1	C	24[A]	ASN
1	C	24[B]	ASN

### 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	225/228 (99%)	179 (80%)	46 (20%)	1	4
1	C	228/228 (100%)	175 (77%)	53 (23%)	1	3
1	E	225/228 (99%)	184 (82%)	41 (18%)	2	6
1	G	227/228 (100%)	185 (82%)	42 (18%)	2	6
1	I	225/228 (99%)	180 (80%)	45 (20%)	1	5
1	K	225/228 (99%)	180 (80%)	45 (20%)	1	5
2	B	15/15 (100%)	11 (73%)	4 (27%)	0	1
2	D	15/15 (100%)	11 (73%)	4 (27%)	0	1
2	F	15/15 (100%)	11 (73%)	4 (27%)	0	1
2	H	15/15 (100%)	11 (73%)	4 (27%)	0	1
2	J	15/15 (100%)	11 (73%)	4 (27%)	0	1
2	L	15/15 (100%)	10 (67%)	5 (33%)	0	1
All	All	1445/1458 (99%)	1148 (79%)	297 (21%)	1	4

5 of 297 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	E	181	LYS
1	G	110	LYS
1	K	156	ASP
1	E	190	LYS
2	F	7	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 33 such sidechains are listed below:

Mol	Chain	Res	Type
1	E	107	ASN
1	G	49	GLN
1	K	65	ASN
1	E	184	GLN
1	E	187	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](#)

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	257/261 (98%)	-0.27	10 (3%) 43 31	32, 52, 102, 162	0
1	C	257/261 (98%)	-0.49	4 (1%) 74 66	28, 44, 88, 146	0
1	E	257/261 (98%)	-0.31	12 (4%) 35 24	28, 49, 95, 151	0
1	G	260/261 (99%)	-0.31	11 (4%) 40 28	33, 54, 105, 158	0
1	I	257/261 (98%)	-0.37	8 (3%) 52 40	29, 49, 95, 156	0
1	K	257/261 (98%)	-0.32	11 (4%) 39 27	38, 55, 96, 132	0
2	B	16/16 (100%)	1.41	4 (25%) 1 0	55, 68, 151, 153	0
2	D	16/16 (100%)	0.69	3 (18%) 2 1	44, 56, 119, 120	0
2	F	16/16 (100%)	0.56	3 (18%) 2 1	50, 64, 121, 124	0
2	H	16/16 (100%)	0.45	2 (12%) 5 2	44, 63, 114, 128	0
2	J	16/16 (100%)	0.51	2 (12%) 5 2	50, 67, 125, 129	0
2	L	16/16 (100%)	0.34	3 (18%) 2 1	47, 57, 114, 121	0
All	All	1641/1662 (98%)	-0.29	73 (4%) 38 26	28, 52, 104, 162	0

The worst 5 of 73 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	189	ASP	14.2
1	G	189	ASP	10.5
2	B	1	SER	10.4
1	E	189	ASP	9.8
1	E	188	VAL	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](#)

There are no ligands in this entry.

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

There are no such residues in this entry.