



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 14, 2017 – 09:51 am GMT

PDB ID : 3FKB  
Title : Structure of NDPK H122G and tenofovir-diphosphate  
Authors : Morera, S.; Chen, Y.X.  
Deposited on : 2008-12-16  
Resolution : 1.65 Å(reported)

This is a Full wwPDB X-ray Structure Validation 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 : 1.7.2 (RC1), CSD as538be (2017)  
Xtriage (Phenix) : 1.9-1692  
EDS : trunk28620  
Percentile statistics : 20161228.v01 (using entries in the PDB archive December 28th 2016)  
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) : recalc28949

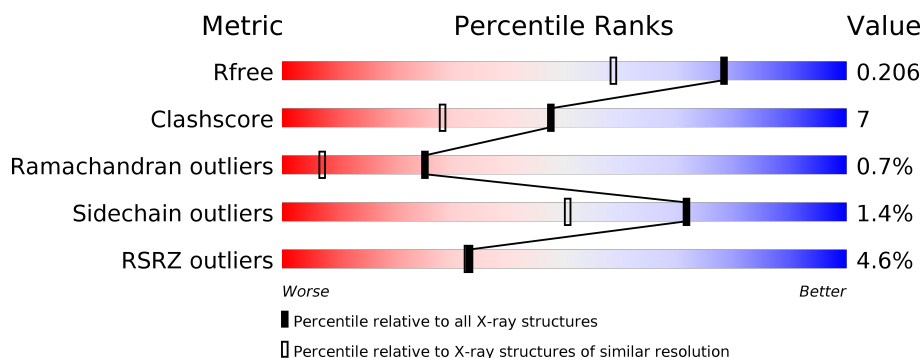
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.65 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))
$R_{free}$	100719	1368 (1.66-1.66)
Clashscore	112137	1468 (1.66-1.66)
Ramachandran outliers	110173	1438 (1.66-1.66)
Sidechain outliers	110143	1438 (1.66-1.66)
RSRZ outliers	101464	1371 (1.66-1.66)

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	155	<div> <div>3%</div> <div> <div></div> <div>82%</div> <div>14%</div> <div>••</div> </div> </div>
1	B	155	<div> <div>5%</div> <div> <div></div> <div>81%</div> <div>15%</div> <div>••</div> </div> </div>
1	C	155	<div> <div>8%</div> <div> <div></div> <div>85%</div> <div>12%</div> <div>••</div> </div> </div>
1	D	155	<div> <div>3%</div> <div> <div></div> <div>81%</div> <div>13%</div> <div>••</div> </div> </div>
1	E	155	<div> <div>3%</div> <div> <div></div> <div>87%</div> <div>8%</div> <div>••</div> </div> </div>
1	F	155	<div> <div>4%</div> <div> <div></div> <div>83%</div> <div>14%</div> <div>•</div> </div> </div>

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 criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
4	EDO	F	157	-	-	-	X
5	GOL	C	157	-	X	-	X
5	GOL	D	157	-	X	-	-

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 8284 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.

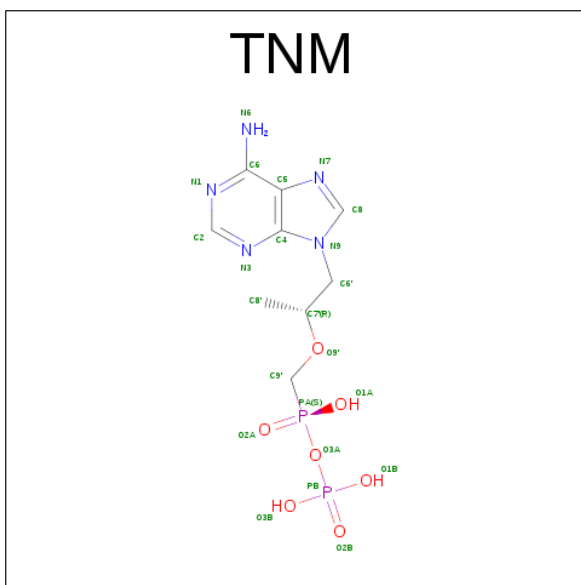
- Molecule 1 is a protein called Nucleoside diphosphate kinase, cytosolic.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	150	Total	C	N	O	S	0	0	0
			1140	732	196	210	2			
1	B	150	Total	C	N	O	S	0	0	0
			1140	732	196	210	2			
1	C	150	Total	C	N	O	S	0	0	0
			1140	732	196	210	2			
1	D	149	Total	C	N	O	S	0	0	0
			1133	727	195	209	2			
1	E	149	Total	C	N	O	S	0	0	0
			1133	727	195	209	2			
1	F	150	Total	C	N	O	S	0	0	0
			1140	732	196	210	2			

There are 6 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	122	GLY	HIS	ENGINEERED	UNP P22887
B	122	GLY	HIS	ENGINEERED	UNP P22887
C	122	GLY	HIS	ENGINEERED	UNP P22887
D	122	GLY	HIS	ENGINEERED	UNP P22887
E	122	GLY	HIS	ENGINEERED	UNP P22887
F	122	GLY	HIS	ENGINEERED	UNP P22887

- Molecule 2 is [(2R)-1-(6-AMINOPURIN-9-YL)PROPAN-2-YL]OXYMETHYL-PHOSPHO NOOXY-PHOSPHINIC ACID (three-letter code: TNM) (formula: C<sub>9</sub>H<sub>15</sub>N<sub>5</sub>O<sub>7</sub>P<sub>2</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	A	1	Total 23	C 9	N 5	O 7	P 2	0	0
2	B	1	Total 23	C 9	N 5	O 7	P 2	0	0
2	C	1	Total 23	C 9	N 5	O 7	P 2	0	0
2	E	1	Total 23	C 9	N 5	O 7	P 2	0	0
2	F	1	Total 23	C 9	N 5	O 7	P 2	0	0

- Molecule 3 is MAGNESIUM ION (three-letter code: MG) (formula: Mg).

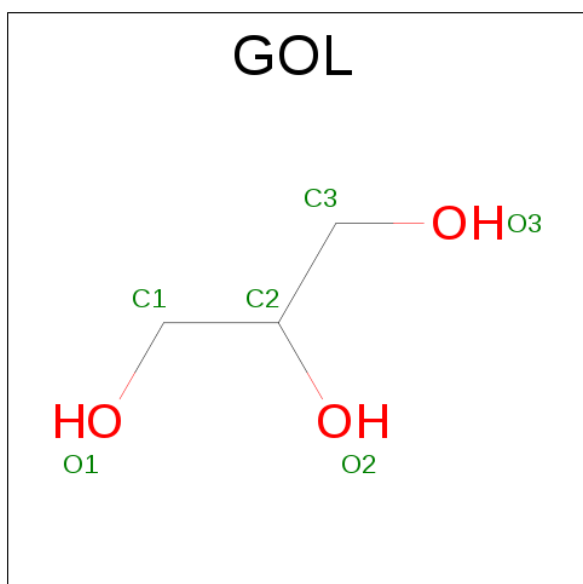
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	D	1	Total Mg 1 1	0	0
3	E	1	Total Mg 1 1	0	0
3	B	1	Total Mg 1 1	0	0
3	C	1	Total Mg 1 1	0	0
3	A	1	Total Mg 1 1	0	0
3	F	1	Total Mg 1 1	0	0

- Molecule 4 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula:  $\text{C}_2\text{H}_6\text{O}_2$ ).



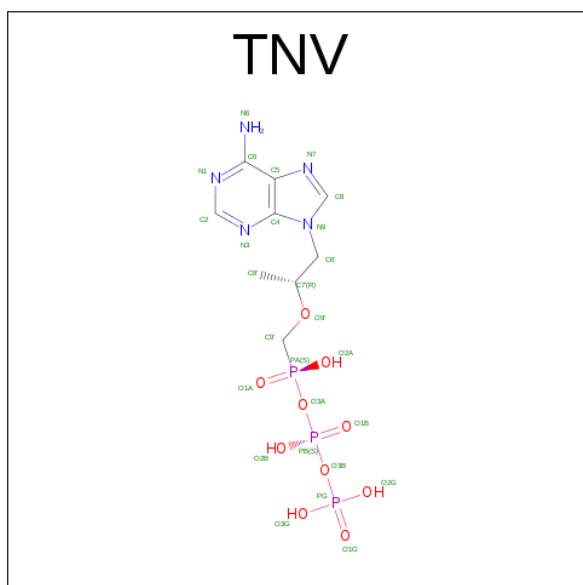
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
4	A	1	Total	C	O	0	0
			4	2	2		
4	B	1	Total	C	O	0	0
			4	2	2		
4	D	1	Total	C	O	0	0
			4	2	2		
4	F	1	Total	C	O	0	0
			4	2	2		
4	F	1	Total	C	O	0	0
			4	2	2		

- Molecule 5 is GLYCEROL (three-letter code: GOL) (formula:  $C_3H_8O_3$ ).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
5	C	1	Total	C	O	0	0
			6	3	3		
5	D	1	Total	C	O	0	0
			6	3	3		

- Molecule 6 is [2-(6-AMINO-9H-PURIN-9-YL)-1-METHYLETHOXY]METHYL-TRIPHOSPHATE (three-letter code: TNV) (formula: C<sub>9</sub>H<sub>16</sub>N<sub>5</sub>O<sub>10</sub>P<sub>3</sub>).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
6	D	1	Total	C	N	O	P	0	0
			27	9	5	10	3		

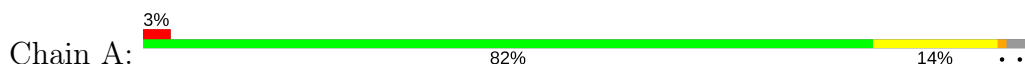
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	203	Total	O	0	0
			203	203		
7	B	200	Total	O	0	0
			200	200		
7	C	212	Total	O	0	0
			212	212		
7	D	210	Total	O	0	0
			210	210		
7	E	239	Total	O	0	0
			239	239		
7	F	214	Total	O	0	0
			214	214		

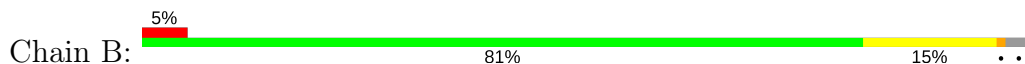
### 3 Residue-property plots [i](#)

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

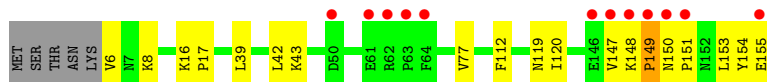
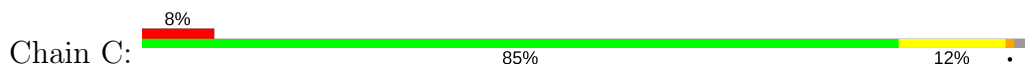
- Molecule 1: Nucleoside diphosphate kinase, cytosolic



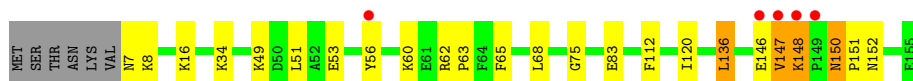
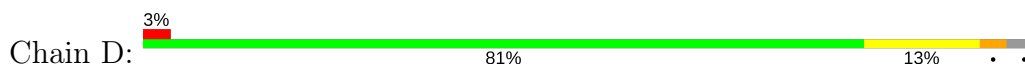
- Molecule 1: Nucleoside diphosphate kinase, cytosolic



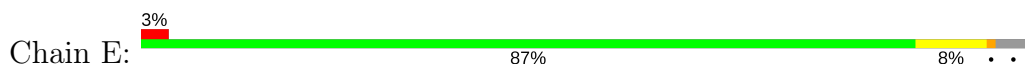
- Molecule 1: Nucleoside diphosphate kinase, cytosolic



- Molecule 1: Nucleoside diphosphate kinase, cytosolic

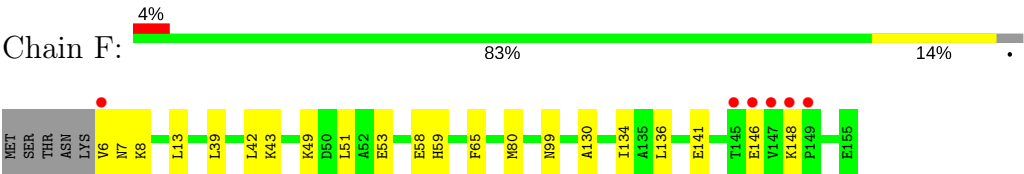


- Molecule 1: Nucleoside diphosphate kinase, cytosolic



- Molecule 1: Nucleoside diphosphate kinase, cytosolic





## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	69.42Å 104.57Å 69.21Å 90.00° 117.96° 90.00°	Depositor
Resolution (Å)	20.00 – 1.65 19.89 – 1.65	Depositor EDS
% Data completeness (in resolution range)	99.8 (20.00-1.65) 99.8 (19.89-1.65)	Depositor EDS
$R_{merge}$	(Not available)	Depositor
$R_{sym}$	0.07	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	4.45 (at 1.65Å)	Xtriage
Refinement program	CNS 1.1	Depositor
R, $R_{free}$	0.181 , 0.205 0.181 , 0.206	Depositor DCC
$R_{free}$ test set	5276 reflections (5.05%)	DCC
Wilson B-factor (Å <sup>2</sup> )	12.0	Xtriage
Anisotropy	0.283	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.34 , 49.2	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtriage
Estimated twinning fraction	0.004 for l,k,-h-l 0.004 for -h-l,k,h 0.018 for -h-l,-k,l 0.019 for h,-k,-h-l 0.021 for l,-k,h	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	8284	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.78% 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: TNM, GOL, MG, TNV, EDO

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	0.29	0/1163	0.57	0/1573
1	B	0.28	0/1163	0.57	0/1573
1	C	0.28	0/1163	0.56	0/1573
1	D	0.29	0/1156	0.58	0/1563
1	E	0.29	0/1156	0.59	0/1563
1	F	0.29	0/1163	0.57	0/1573
All	All	0.29	0/6964	0.57	0/9418

There are no bond length outliers.

There are no bond angle outliers.

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	1140	0	1163	16	0
1	B	1140	0	1163	18	0
1	C	1140	0	1163	19	0
1	D	1133	0	1154	21	0
1	E	1133	0	1154	13	0
1	F	1140	0	1163	14	0
2	A	23	0	12	0	0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
2	B	23	0	12	0	0
2	C	23	0	12	0	0
2	E	23	0	12	0	0
2	F	23	0	12	1	0
3	A	1	0	0	0	0
3	B	1	0	0	0	0
3	C	1	0	0	0	0
3	D	1	0	0	0	0
3	E	1	0	0	0	0
3	F	1	0	0	0	0
4	A	4	0	6	1	0
4	B	4	0	6	1	0
4	D	4	0	6	0	0
4	F	8	0	12	0	0
5	C	6	0	3	0	0
5	D	6	0	3	0	0
6	D	27	0	12	0	0
7	A	203	0	0	3	0
7	B	200	0	0	2	0
7	C	212	0	0	3	0
7	D	210	0	0	1	0
7	E	239	0	0	1	0
7	F	214	0	0	2	0
All	All	8284	0	7068	98	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 7.

All (98) 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:D:146:GLU:HG3	1:D:147:VAL:H	1.35	0.92
1:F:6:VAL:HG13	1:F:8:LYS:H	1.41	0.86
1:A:151:PRO:HD2	7:A:2265:HOH:O	1.84	0.78
1:D:56:TYR:CD2	1:D:68:LEU:HD21	2.20	0.76
1:E:148:LYS:HD2	1:E:149:PRO:HD2	1.70	0.73
1:A:6:VAL:HG13	1:A:8:LYS:H	1.53	0.72
1:B:150:ASN:HD21	1:B:152:ASN:HD22	1.41	0.69
1:B:62:ARG:HA	1:B:62:ARG:HE	1.59	0.68
1:D:68:LEU:C	1:D:68:LEU:HD23	2.17	0.66
1:C:150:ASN:HB2	1:C:153:LEU:HD23	1.81	0.63

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:8:LYS:HD3	1:D:83:GLU:OE2	2.00	0.62
1:C:16:LYS:HE3	7:C:1855:HOH:O	1.99	0.61
1:C:149:PRO:C	1:C:150:ASN:HD22	2.04	0.60
1:C:149:PRO:HG3	1:C:154:TYR:HE2	1.69	0.58
1:A:49:LYS:O	1:A:53:GLU:HG3	2.04	0.57
1:D:146:GLU:HG3	1:D:147:VAL:N	2.13	0.57
1:F:6:VAL:HG22	1:F:7:ASN:H	1.70	0.57
1:A:6:VAL:HG22	1:A:7:ASN:N	2.20	0.56
1:A:6:VAL:HG22	1:A:7:ASN:H	1.70	0.56
1:C:151:PRO:HD2	7:C:1910:HOH:O	2.06	0.56
1:E:147:VAL:O	1:E:148:LYS:HB3	2.06	0.55
1:B:38:VAL:HG13	4:B:157:EDO:H21	1.89	0.55
1:A:101:LEU:HD21	1:A:114:VAL:HG13	1.89	0.54
1:E:150:ASN:ND2	1:E:152:ASN:H	2.04	0.54
1:E:150:ASN:HD22	1:E:150:ASN:C	2.11	0.54
1:D:150:ASN:C	1:D:150:ASN:HD22	2.12	0.53
1:B:150:ASN:ND2	1:B:152:ASN:HB2	2.24	0.53
1:D:146:GLU:O	1:D:147:VAL:HB	2.10	0.52
1:D:150:ASN:ND2	1:D:152:ASN:H	2.08	0.51
1:A:141:GLU:CD	1:A:141:GLU:H	2.12	0.51
1:C:17:PRO:HD3	1:C:77:VAL:HG22	1.92	0.51
1:B:8:LYS:HE2	7:B:2156:HOH:O	2.10	0.51
1:E:150:ASN:HD22	1:E:151:PRO:N	2.09	0.50
1:A:54:SER:HB3	1:A:136:LEU:HD11	1.93	0.50
1:D:49:LYS:O	1:D:53:GLU:HG3	2.11	0.50
1:F:141:GLU:H	1:F:141:GLU:CD	2.16	0.50
1:B:51:LEU:CD1	1:B:136:LEU:HG	2.43	0.49
1:E:8:LYS:HD3	1:E:8:LYS:O	2.10	0.49
1:F:6:VAL:HG22	1:F:7:ASN:N	2.27	0.49
1:F:39:LEU:HD21	1:F:42:LEU:HD22	1.93	0.49
1:A:39:LEU:HD21	1:A:42:LEU:HD22	1.93	0.49
1:D:7:ASN:N	7:D:1879:HOH:O	2.45	0.49
1:F:51:LEU:CD1	1:F:136:LEU:HG	2.42	0.49
1:B:59:HIS:O	1:B:65:PHE:HB2	2.13	0.48
1:D:51:LEU:CD1	1:D:136:LEU:HD22	2.44	0.48
1:E:49:LYS:O	1:E:53:GLU:HG3	2.14	0.47
7:A:1068:HOH:O	1:D:34:LYS:HE3	2.15	0.46
1:A:112:PHE:HB2	1:A:120:ILE:HD13	1.98	0.46
1:C:149:PRO:HG3	1:C:154:TYR:CE2	2.48	0.46
1:D:150:ASN:HD22	1:D:151:PRO:N	2.13	0.46
1:B:39:LEU:HD21	1:B:42:LEU:HD22	1.98	0.46

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:17:PRO:HD3	1:B:77:VAL:HG22	1.97	0.46
1:B:8:LYS:HD2	7:B:2229:HOH:O	2.16	0.46
1:E:39:LEU:HD21	1:E:42:LEU:HD22	1.98	0.46
1:F:58:GLU:HG2	7:F:1899:HOH:O	2.15	0.45
1:B:62:ARG:NE	1:B:62:ARG:HA	2.27	0.45
1:D:112:PHE:HB2	1:D:120:ILE:HD13	1.98	0.45
1:A:51:LEU:HD12	1:A:136:LEU:HD22	1.99	0.45
1:F:59:HIS:O	1:F:65:PHE:HB2	2.18	0.44
1:D:16:LYS:HE2	1:D:56:TYR:OH	2.17	0.44
1:B:118:ARG:NH1	1:C:155:GLU:OE2	2.50	0.44
1:C:39:LEU:HD21	1:C:42:LEU:HD22	1.99	0.44
1:C:43:LYS:HA	1:F:42:LEU:O	2.18	0.43
1:A:150:ASN:OD1	1:A:152:ASN:HB2	2.18	0.43
1:A:51:LEU:CD1	1:A:136:LEU:HD22	2.48	0.43
1:D:146:GLU:CG	1:D:147:VAL:N	2.78	0.43
1:C:112:PHE:HB2	1:C:120:ILE:HD13	2.00	0.43
1:F:130:ALA:O	1:F:134:ILE:HG13	2.19	0.43
1:B:49:LYS:O	1:B:53:GLU:HG3	2.19	0.42
1:D:60:LYS:HA	1:D:65:PHE:CD1	2.54	0.42
1:C:149:PRO:C	1:C:150:ASN:ND2	2.71	0.42
1:D:147:VAL:O	1:D:148:LYS:HB2	2.19	0.42
1:A:60:LYS:HA	1:A:65:PHE:CD1	2.55	0.42
1:A:60:LYS:HE2	7:A:1275:HOH:O	2.18	0.42
1:C:16:LYS:HB3	1:C:17:PRO:HD2	2.02	0.42
1:D:60:LYS:HA	1:D:65:PHE:CG	2.54	0.42
1:C:6:VAL:N	7:C:2046:HOH:O	2.53	0.41
1:E:150:ASN:HD22	1:E:151:PRO:CD	2.33	0.41
1:A:150:ASN:OD1	1:A:152:ASN:ND2	2.52	0.41
1:B:43:LYS:HA	1:E:42:LEU:O	2.20	0.41
1:B:7:ASN:OD1	1:B:85:LYS:HE3	2.20	0.41
1:F:49:LYS:O	1:F:53:GLU:HG3	2.20	0.41
1:C:16:LYS:CE	1:C:119:ASN:OD1	2.69	0.41
1:F:13:LEU:HB2	1:F:80:MET:HG3	2.02	0.41
1:C:150:ASN:N	1:C:150:ASN:ND2	2.68	0.41
1:E:150:ASN:HD22	1:E:152:ASN:H	1.68	0.41
1:E:34:LYS:HE2	7:E:1697:HOH:O	2.19	0.41
1:B:112:PHE:HB2	1:B:120:ILE:HD13	2.03	0.41
1:B:16:LYS:HB3	1:B:17:PRO:HD2	2.03	0.40
1:C:147:VAL:HG12	1:C:148:LYS:N	2.36	0.40
2:F:160:TNM:H8	7:F:1038:HOH:O	2.21	0.40
1:F:6:VAL:HG13	1:F:8:LYS:N	2.22	0.40

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:8:LYS:N	1:C:8:LYS:HD2	2.36	0.40
1:D:62:ARG:HA	1:D:63:PRO:HD3	1.95	0.40
4:A:157:EDO:H11	1:D:75:GLY:HA3	2.02	0.40
1:B:99:ASN:ND2	1:B:101:LEU:H	2.19	0.40
1:C:42:LEU:O	1:F:43:LYS:HA	2.20	0.40
1:E:17:PRO:HD3	1:E:77:VAL:HG22	2.03	0.40

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	148/155 (96%)	141 (95%)	6 (4%)	1 (1%)	25	7
1	B	148/155 (96%)	144 (97%)	4 (3%)	0	100	100
1	C	148/155 (96%)	141 (95%)	6 (4%)	1 (1%)	25	7
1	D	147/155 (95%)	141 (96%)	4 (3%)	2 (1%)	13	1
1	E	147/155 (95%)	142 (97%)	4 (3%)	1 (1%)	25	7
1	F	148/155 (96%)	143 (97%)	4 (3%)	1 (1%)	25	7
All	All	886/930 (95%)	852 (96%)	28 (3%)	6 (1%)	25	7

All (6) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	C	149	PRO
1	A	149	PRO
1	D	148	LYS
1	E	148	LYS
1	F	148	LYS
1	D	147	VAL

### 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	121/126 (96%)	118 (98%)	3 (2%)	53	25
1	B	121/126 (96%)	119 (98%)	2 (2%)	66	43
1	C	121/126 (96%)	121 (100%)	0	100	100
1	D	120/126 (95%)	118 (98%)	2 (2%)	66	43
1	E	120/126 (95%)	119 (99%)	1 (1%)	85	73
1	F	121/126 (96%)	119 (98%)	2 (2%)	66	43
All	All	724/756 (96%)	714 (99%)	10 (1%)	71	50

All (10) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	99	ASN
1	A	115	ASP
1	A	136	LEU
1	B	80	MET
1	B	99	ASN
1	D	136	LEU
1	D	150	ASN
1	E	150	ASN
1	F	99	ASN
1	F	146	GLU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (10) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	7	ASN
1	A	99	ASN
1	B	59	HIS
1	B	99	ASN
1	B	150	ASN
1	D	7	ASN
1	D	150	ASN

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Res	Type
1	E	150	ASN
1	F	7	ASN
1	F	99	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](#)

Of 19 ligands modelled in this entry, 6 are monoatomic - leaving 13 for Mogul analysis.

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	EDO	A	157	-	3,3,3	0.59	0	2,2,2	0.01	0
2	TNM	A	160	3	22,24,24	1.94	6 (27%)	18,36,36	2.77	4 (22%)
4	EDO	B	157	-	3,3,3	0.55	0	2,2,2	0.06	0
2	TNM	B	160	3	22,24,24	1.93	6 (27%)	18,36,36	2.38	3 (16%)
5	GOL	C	157	-	5,5,5	4.67	5 (100%)	5,5,5	2.01	2 (40%)
2	TNM	C	160	3	22,24,24	1.93	6 (27%)	18,36,36	1.95	3 (16%)
5	GOL	D	157	-	5,5,5	4.64	5 (100%)	5,5,5	2.04	2 (40%)
4	EDO	D	158	-	3,3,3	0.55	0	2,2,2	0.05	0
6	TNV	D	160	3	24,28,28	1.90	6 (25%)	19,43,43	2.33	3 (15%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
2	TNM	E	160	3	22,24,24	1.98	6 (27%)	18,36,36	2.51	2 (11%)
4	EDO	F	157	-	3,3,3	0.57	0	2,2,2	0.04	0
4	EDO	F	158	-	3,3,3	0.56	0	2,2,2	0.05	0
2	TNM	F	160	3	22,24,24	1.93	6 (27%)	18,36,36	2.87	2 (11%)

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	EDO	A	157	-	-	0/1/1/1	0/0/0/0
2	TNM	A	160	3	-	0/11/16/16	0/2/2/2
4	EDO	B	157	-	-	0/1/1/1	0/0/0/0
2	TNM	B	160	3	-	0/11/16/16	0/2/2/2
5	GOL	C	157	-	-	0/4/4/4	0/0/0/0
2	TNM	C	160	3	-	0/11/16/16	0/2/2/2
5	GOL	D	157	-	-	0/4/4/4	0/0/0/0
4	EDO	D	158	-	-	0/1/1/1	0/0/0/0
6	TNV	D	160	3	-	0/17/22/22	0/2/2/2
2	TNM	E	160	3	-	0/11/16/16	0/2/2/2
4	EDO	F	157	-	-	0/1/1/1	0/0/0/0
4	EDO	F	158	-	-	0/1/1/1	0/0/0/0
2	TNM	F	160	3	-	0/11/16/16	0/2/2/2

All (46) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	157	GOL	C3-C2	-5.89	1.30	1.52
5	D	157	GOL	C3-C2	-5.89	1.30	1.52
5	C	157	GOL	C1-C2	-5.89	1.30	1.52
5	D	157	GOL	C1-C2	-5.82	1.30	1.52
2	E	160	TNM	O9'-C9'	-5.52	1.37	1.42
6	D	160	TNV	O9'-C9'	-5.15	1.38	1.42
2	A	160	TNM	O9'-C9'	-5.13	1.38	1.42
2	C	160	TNM	O9'-C9'	-4.98	1.38	1.42
2	B	160	TNM	O9'-C9'	-4.98	1.38	1.42
2	F	160	TNM	O9'-C9'	-4.82	1.38	1.42
6	D	160	TNV	PA-O2A	-3.11	1.48	1.56
2	B	160	TNM	PA-O1A	-3.06	1.48	1.56
2	F	160	TNM	PA-O1A	-2.95	1.49	1.56
2	C	160	TNM	PA-O1A	-2.95	1.49	1.56

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	C	157	GOL	O2-C2	-2.92	1.34	1.43
2	A	160	TNM	PA-O1A	-2.89	1.49	1.56
5	D	157	GOL	O2-C2	-2.88	1.34	1.43
2	E	160	TNM	PA-O1A	-2.80	1.49	1.56
2	A	160	TNM	C6'-N9	-2.62	1.45	1.48
6	D	160	TNV	C6'-N9	-2.45	1.45	1.48
2	F	160	TNM	C6'-N9	-2.42	1.46	1.48
2	C	160	TNM	C6'-N9	-2.32	1.46	1.48
2	E	160	TNM	C6'-N9	-2.25	1.46	1.48
2	B	160	TNM	C6'-N9	-2.10	1.46	1.48
2	A	160	TNM	C4-N3	2.39	1.39	1.35
2	B	160	TNM	C4-N3	2.39	1.39	1.35
2	C	160	TNM	C4-N3	2.42	1.39	1.35
2	F	160	TNM	C4-N3	2.44	1.39	1.35
6	D	160	TNV	C4-N3	2.53	1.39	1.35
2	E	160	TNM	C4-N3	2.54	1.39	1.35
2	B	160	TNM	C2-N1	2.81	1.39	1.33
2	A	160	TNM	C2-N1	2.84	1.39	1.33
6	D	160	TNV	C2-N1	2.84	1.39	1.33
2	C	160	TNM	C2-N1	2.85	1.39	1.33
2	E	160	TNM	C2-N1	2.90	1.39	1.33
2	F	160	TNM	C2-N1	2.97	1.39	1.33
2	A	160	TNM	C2-N3	3.23	1.37	1.32
6	D	160	TNV	C2-N3	3.24	1.37	1.32
2	C	160	TNM	C2-N3	3.27	1.37	1.32
2	E	160	TNM	C2-N3	3.31	1.37	1.32
2	F	160	TNM	C2-N3	3.31	1.37	1.32
2	B	160	TNM	C2-N3	3.48	1.38	1.32
5	D	157	GOL	O3-C3	3.92	1.58	1.42
5	C	157	GOL	O1-C1	3.93	1.59	1.42
5	C	157	GOL	O3-C3	3.95	1.59	1.42
5	D	157	GOL	O1-C1	3.96	1.59	1.42

All (21) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
6	D	160	TNV	PB-O3A-PA	-3.15	122.24	132.39
2	A	160	TNM	PB-O3A-PA	-3.00	121.80	132.38
2	C	160	TNM	PB-O3A-PA	-2.93	122.05	132.38
2	B	160	TNM	PB-O3A-PA	-2.92	122.08	132.38
2	F	160	TNM	PB-O3A-PA	-2.88	122.23	132.38
2	E	160	TNM	PB-O3A-PA	-2.73	122.75	132.38

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	A	160	TNM	O1A-PA-O2A	2.01	116.81	110.09
2	A	160	TNM	C5-C6-N6	2.02	124.58	120.47
2	B	160	TNM	C5-C6-N6	2.04	124.63	120.47
2	C	160	TNM	C5-C6-N6	2.05	124.64	120.47
6	D	160	TNV	C5-C6-N6	2.06	124.67	120.47
5	C	157	GOL	O3-C3-C2	3.10	125.70	110.07
5	D	157	GOL	O3-C3-C2	3.12	125.79	110.07
5	C	157	GOL	O1-C1-C2	3.13	125.82	110.07
5	D	157	GOL	O1-C1-C2	3.19	126.13	110.07
2	C	160	TNM	C9'-O9'-C7'	5.68	118.91	114.33
2	B	160	TNM	C9'-O9'-C7'	8.16	120.91	114.33
6	D	160	TNV	C9'-O9'-C7'	8.49	121.17	114.33
2	E	160	TNM	C9'-O9'-C7'	8.88	121.48	114.33
2	A	160	TNM	C9'-O9'-C7'	10.01	122.39	114.33
2	F	160	TNM	C9'-O9'-C7'	10.59	122.86	114.33

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

3 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
4	A	157	EDO	1	0
4	B	157	EDO	1	0
2	F	160	TNM	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

### 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	150/155 (96%)	0.02	5 (3%) 47 49	6, 11, 28, 44	0
1	B	150/155 (96%)	0.03	8 (5%) 27 26	7, 12, 28, 43	0
1	C	150/155 (96%)	0.15	12 (8%) 13 12	6, 13, 37, 45	0
1	D	149/155 (96%)	-0.04	5 (3%) 46 48	7, 12, 24, 41	0
1	E	149/155 (96%)	-0.14	5 (3%) 46 48	6, 10, 24, 39	0
1	F	150/155 (96%)	-0.06	6 (4%) 39 40	7, 11, 27, 42	0
All	All	898/930 (96%)	-0.01	41 (4%) 33 33	6, 11, 28, 45	0

All (41) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	149	PRO	8.2
1	D	147	VAL	7.6
1	A	149	PRO	6.6
1	C	148	LYS	6.0
1	A	6	VAL	6.0
1	E	149	PRO	5.8
1	F	146	GLU	5.8
1	A	151	PRO	5.7
1	F	149	PRO	5.7
1	E	146	GLU	5.5
1	D	149	PRO	5.0
1	F	148	LYS	5.0
1	A	148	LYS	4.7
1	C	64	PHE	4.7
1	F	6	VAL	4.5
1	C	63	PRO	4.4
1	B	149	PRO	4.4
1	D	148	LYS	4.0
1	E	147	VAL	3.9

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	F	147	VAL	3.8
1	E	148	LYS	3.7
1	D	146	GLU	3.6
1	B	62	ARG	3.4
1	A	150	ASN	3.4
1	B	148	LYS	3.2
1	B	60	LYS	3.2
1	C	150	ASN	3.2
1	C	151	PRO	3.1
1	B	61	GLU	3.1
1	C	146	GLU	3.1
1	C	61	GLU	2.9
1	C	155	GLU	2.9
1	B	63	PRO	2.7
1	B	151	PRO	2.6
1	D	56	TYR	2.6
1	C	62	ARG	2.4
1	C	147	VAL	2.4
1	E	7	ASN	2.4
1	C	50	ASP	2.3
1	F	145	THR	2.2
1	B	6	VAL	2.0

## 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
4	EDO	F	157	4/4	0.84	0.25	3.15	31,32,33,34	0
5	GOL	C	157	6/6	0.75	0.18	2.90	60,60,61,61	0
6	TNV	D	160	27/27	0.85	0.15	1.84	24,28,34,36	0
2	TNM	C	160	23/23	0.90	0.17	1.31	22,33,40,41	0
4	EDO	B	157	4/4	0.63	0.32	1.23	40,41,41,43	0
4	EDO	A	157	4/4	0.80	0.27	0.80	42,43,43,44	0
2	TNM	B	160	23/23	0.91	0.15	0.54	18,27,40,40	0
2	TNM	A	160	23/23	0.94	0.10	0.26	15,18,21,22	0
2	TNM	F	160	23/23	0.96	0.10	0.06	15,18,25,25	0
4	EDO	D	158	4/4	0.74	0.19	0.06	35,37,37,38	0
2	TNM	E	160	23/23	0.95	0.08	-0.08	11,14,20,23	0
3	MG	F	156	1/1	0.99	0.12	-	11,11,11,11	0
3	MG	E	156	1/1	0.99	0.04	-	11,11,11,11	0
4	EDO	F	158	4/4	0.38	0.41	-	51,51,51,52	0
3	MG	D	156	1/1	0.95	0.08	-	27,27,27,27	0
5	GOL	D	157	6/6	0.32	0.32	-	76,77,77,77	0
3	MG	B	156	1/1	0.98	0.04	-	15,15,15,15	0
3	MG	A	156	1/1	0.99	0.06	-	12,12,12,12	0
3	MG	C	156	1/1	0.97	0.10	-	20,20,20,20	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.