



# Full wwPDB X-ray Structure Validation Report ⓘ

Sep 3, 2023 – 01:05 PM EDT

PDB ID : 3SQ8  
Title : Crystal Structure Analysis of the Yeast Tyrosyl-DNA Phosphodiesterase 1 H432R Mutant (SCAN1 Mutant)  
Authors : Gajewski, S.; White, S.W.  
Deposited on : 2011-07-05  
Resolution : 2.10 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

<https://www.wwpdb.org/validation/2017/XrayValidationReportHelp>

with specific help available everywhere you see the ⓘ symbol.

The types of validation reports are described at

<http://www.wwpdb.org/validation/2017/FAQs#types>.

---

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 : 2.35  
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) : 2.35

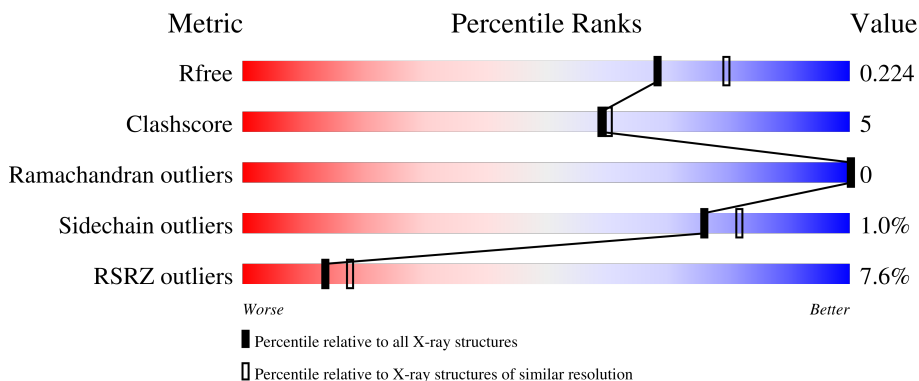
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*



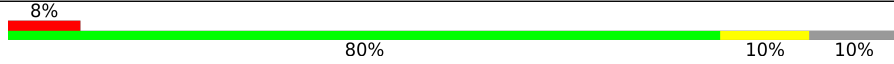
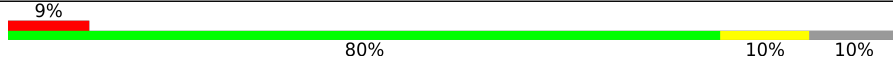
The reported resolution of this entry is 2.10 Å.

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	5197 (2.10-2.10)
Clashscore	141614	5710 (2.10-2.10)
Ramachandran outliers	138981	5647 (2.10-2.10)
Sidechain outliers	138945	5648 (2.10-2.10)
RSRZ outliers	127900	5083 (2.10-2.10)

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 of 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	470	
1	B	470	
1	C	470	
1	D	470	

## 2 Entry composition i

There are 2 unique types of molecules in this entry. The entry contains 14246 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 Tyrosyl-DNA phosphodiesterase 1.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
			Total	C	N	O	P	S			
1	A	422	Total 3430	C 2223	N 564	O 622	P 1	S 20	0	0	0
1	B	423	Total 3450	C 2235	N 570	O 624	P 1	S 20	0	1	0
1	C	423	Total 3437	C 2227	N 565	O 624	P 1	S 20	0	0	0
1	D	425	Total 3457	C 2241	N 569	O 626	P 1	S 20	0	1	0

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	78	MET	-	initiating methionine	UNP P38319
A	432	ARG	HIS	engineered mutation	UNP P38319
A	540	LEU	-	expression tag	UNP P38319
A	541	HIS	-	expression tag	UNP P38319
A	542	HIS	-	expression tag	UNP P38319
A	543	HIS	-	expression tag	UNP P38319
A	544	HIS	-	expression tag	UNP P38319
A	545	HIS	-	expression tag	UNP P38319
A	546	HIS	-	expression tag	UNP P38319
A	547	HIS	-	expression tag	UNP P38319
B	78	MET	-	initiating methionine	UNP P38319
B	432	ARG	HIS	engineered mutation	UNP P38319
B	540	LEU	-	expression tag	UNP P38319
B	541	HIS	-	expression tag	UNP P38319
B	542	HIS	-	expression tag	UNP P38319
B	543	HIS	-	expression tag	UNP P38319
B	544	HIS	-	expression tag	UNP P38319
B	545	HIS	-	expression tag	UNP P38319
B	546	HIS	-	expression tag	UNP P38319
B	547	HIS	-	expression tag	UNP P38319
C	78	MET	-	initiating methionine	UNP P38319

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
C	432	ARG	HIS	engineered mutation	UNP P38319
C	540	LEU	-	expression tag	UNP P38319
C	541	HIS	-	expression tag	UNP P38319
C	542	HIS	-	expression tag	UNP P38319
C	543	HIS	-	expression tag	UNP P38319
C	544	HIS	-	expression tag	UNP P38319
C	545	HIS	-	expression tag	UNP P38319
C	546	HIS	-	expression tag	UNP P38319
C	547	HIS	-	expression tag	UNP P38319
D	78	MET	-	initiating methionine	UNP P38319
D	432	ARG	HIS	engineered mutation	UNP P38319
D	540	LEU	-	expression tag	UNP P38319
D	541	HIS	-	expression tag	UNP P38319
D	542	HIS	-	expression tag	UNP P38319
D	543	HIS	-	expression tag	UNP P38319
D	544	HIS	-	expression tag	UNP P38319
D	545	HIS	-	expression tag	UNP P38319
D	546	HIS	-	expression tag	UNP P38319
D	547	HIS	-	expression tag	UNP P38319

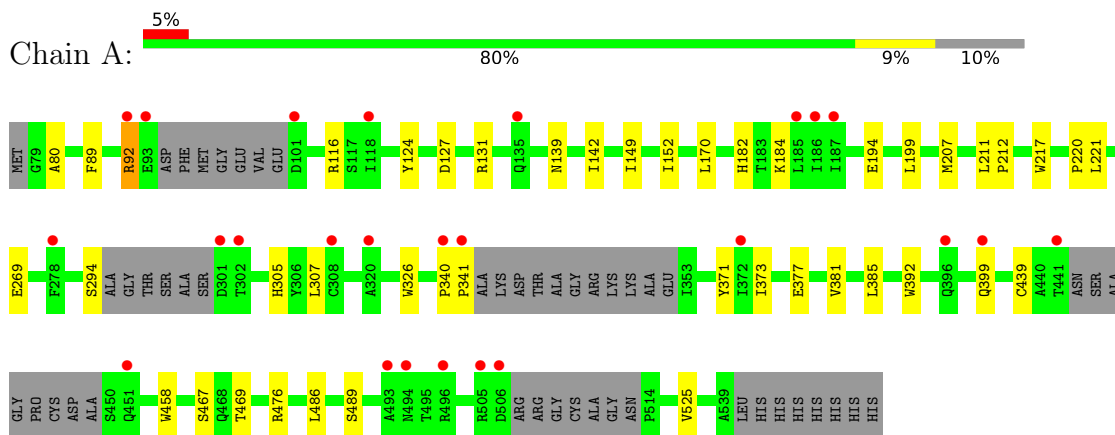
- Molecule 2 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	168	Total O 168 168	0	0
2	B	146	Total O 146 146	0	0
2	C	95	Total O 95 95	0	0
2	D	63	Total O 63 63	0	0

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

These plots are drawn for all protein, RNA, DNA and oligosaccharide 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: Tyrosyl-DNA phosphodiesterase 1





## 4 Data and refinement statistics i

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	63.87Å 81.49Å 96.79Å 89.52° 84.90° 67.82°	Depositor
Resolution (Å)	24.79 – 2.10 24.79 – 2.10	Depositor EDS
% Data completeness (in resolution range)	93.4 (24.79-2.10) 93.4 (24.79-2.10)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	3.38 (at 2.10Å)	Xtrriage
Refinement program	REFMAC 5.5.0109	Depositor
R, $R_{free}$	0.187 , 0.226 0.183 , 0.224	Depositor DCC
$R_{free}$ test set	4912 reflections (5.01%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	34.7	Xtrriage
Anisotropy	0.259	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.32 , 42.8	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.49$ , $\langle L^2 \rangle = 0.32$	Xtrriage
Estimated twinning fraction	0.003 for -h,-h+k,-l	Xtrriage
$F_o, F_c$ correlation	0.96	EDS
Total number of atoms	14246	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	49.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.82% 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: NEP

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	0.64	4/3501 (0.1%)	0.69	0/4734
1	B	0.65	2/3525 (0.1%)	0.68	0/4765
1	C	0.58	3/3508 (0.1%)	0.63	1/4743 (0.0%)
1	D	0.55	1/3532 (0.0%)	0.61	0/4776
All	All	0.60	10/14066 (0.1%)	0.66	1/19018 (0.0%)

All (10) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	326	TRP	CD2-CE2	5.68	1.48	1.41
1	A	458	TRP	CD2-CE2	5.60	1.48	1.41
1	B	458	TRP	CD2-CE2	5.43	1.47	1.41
1	A	217	TRP	CD2-CE2	5.38	1.47	1.41
1	A	392	TRP	CD2-CE2	5.36	1.47	1.41
1	C	217	TRP	CD2-CE2	5.35	1.47	1.41
1	C	387	TRP	CD2-CE2	5.34	1.47	1.41
1	B	217	TRP	CD2-CE2	5.32	1.47	1.41
1	C	326	TRP	CD2-CE2	5.28	1.47	1.41
1	D	217	TRP	CD2-CE2	5.13	1.47	1.41

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	432	ARG	NE-CZ-NH2	-5.51	117.54	120.30

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	3430	0	3419	28	0
1	B	3450	0	3436	46	0
1	C	3437	0	3418	29	0
1	D	3457	0	3438	41	0
2	A	168	0	0	2	0
2	B	146	0	0	2	0
2	C	95	0	0	3	0
2	D	63	0	0	1	0
All	All	14246	0	13711	128	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 5.

All (128) 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:B:207:MET:HE2	1:D:541:HIS:HE1	1.26	0.99
1:B:149:ILE:HD13	1:B:170:LEU:HD11	1.44	0.96
1:D:149:ILE:HD13	1:D:170:LEU:HD11	1.48	0.96
1:C:149:ILE:HD13	1:C:170:LEU:HD11	1.52	0.89
1:C:514:PRO:HD3	1:D:177:PRO:HG3	1.57	0.85
1:A:149:ILE:HD13	1:A:170:LEU:HD11	1.61	0.82
1:B:541:HIS:HE1	1:D:207:MET:HE2	1.45	0.80
1:B:207:MET:CE	1:D:541:HIS:HE1	1.96	0.78
1:B:149:ILE:CD1	1:B:170:LEU:HD11	2.15	0.76
1:B:207:MET:HE2	1:D:541:HIS:CE1	2.18	0.75
1:B:207:MET:HE1	1:B:211:LEU:HD11	1.72	0.72
1:B:207:MET:CE	1:B:211:LEU:HD11	2.22	0.69
1:C:302:THR:HG22	1:C:442:ASN:HB2	1.74	0.69
1:A:149:ILE:CD1	1:A:170:LEU:HD11	2.25	0.67
1:B:541:HIS:HE1	1:D:207:MET:CE	2.08	0.66
1:B:207:MET:CE	1:D:541:HIS:CE1	2.76	0.66
1:D:441:THR:O	1:D:442:ASN:HB2	1.96	0.65
1:C:149:ILE:CD1	1:C:170:LEU:HD11	2.27	0.63
1:B:149:ILE:HD13	1:B:170:LEU:CD1	2.23	0.63

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:207:MET:HE1	1:A:211:LEU:HD11	1.81	0.62
1:C:135:GLN:NE2	1:C:160:THR:HG23	2.13	0.62
1:B:275:PRO:HG2	1:C:229:LEU:HD11	1.80	0.62
1:B:400:LYS:HG2	2:B:615:HOH:O	2.00	0.62
1:D:149:ILE:CD1	1:D:170:LEU:HD11	2.25	0.61
1:B:321:ARG:NH1	1:B:323:GLU:OE2	2.34	0.60
1:C:357:ASN:HB2	2:C:617:HOH:O	2.01	0.60
1:B:116:ARG:HE	1:B:139:ASN:ND2	1.99	0.59
1:B:394[B]:HIS:CE1	1:B:396:GLN:HG2	2.38	0.58
1:D:230:PRO:HA	2:D:560:HOH:O	2.01	0.58
1:D:207:MET:HE1	1:D:211:LEU:HD11	1.87	0.57
1:C:116:ARG:HE	1:C:139:ASN:ND2	2.02	0.57
1:D:207:MET:CE	1:D:211:LEU:HD11	2.35	0.57
1:B:124:TYR:CZ	1:B:142:ILE:HG23	2.40	0.57
1:D:116:ARG:HE	1:D:139:ASN:ND2	2.03	0.57
1:A:207:MET:CE	1:A:211:LEU:HD11	2.36	0.56
1:A:116:ARG:HE	1:A:139:ASN:ND2	2.03	0.56
1:C:127:ASP:HB3	1:C:152:ILE:HG12	1.89	0.55
1:D:269:GLU:HB2	1:D:486:LEU:HB3	1.89	0.55
1:B:245:HIS:HD2	1:C:229:LEU:CD2	2.20	0.54
1:B:541:HIS:CE1	1:D:207:MET:CE	2.88	0.54
1:C:307:LEU:HD11	1:C:373:ILE:HD11	1.90	0.54
1:C:192:ASN:OD1	1:C:192:ASN:N	2.41	0.53
1:C:490:ARG:HB3	2:C:615:HOH:O	2.09	0.53
1:A:127:ASP:HB3	1:A:152:ILE:HG12	1.90	0.53
1:D:149:ILE:HD13	1:D:170:LEU:CD1	2.32	0.53
1:A:467:SER:OG	1:A:469:THR:HG22	2.08	0.53
1:B:360:ILE:HD11	1:B:410:LYS:HB3	1.91	0.52
1:B:86:LYS:HB3	1:B:104:THR:HG22	1.92	0.52
1:C:401:LYS:O	1:C:405:GLU:HG2	2.09	0.52
1:C:207:MET:CE	1:C:211:LEU:HD11	2.40	0.51
1:B:116:ARG:HG3	1:B:139:ASN:HD22	1.76	0.51
1:A:89:PHE:O	1:A:92:ARG:NH1	2.44	0.50
1:D:124:TYR:CZ	1:D:142:ILE:HG23	2.46	0.50
1:B:307:LEU:HD11	1:B:373:ILE:HD11	1.93	0.50
1:C:124:TYR:CZ	1:C:142:ILE:HG23	2.46	0.50
1:C:456:LEU:HB2	1:C:498:VAL:HB	1.92	0.50
1:C:207:MET:HE1	1:C:211:LEU:HD11	1.94	0.49
1:B:289:LYS:HE3	2:B:655:HOH:O	2.12	0.49
1:D:302:THR:HB	1:D:441:THR:O	2.12	0.49
1:B:376:THR:HG22	1:B:528:TYR:CZ	2.48	0.49

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:441:THR:O	1:B:442:ASN:HB2	2.13	0.49
1:C:490:ARG:CB	2:C:615:HOH:O	2.61	0.48
1:A:149:ILE:HD13	1:A:170:LEU:CD1	2.39	0.48
1:D:392:TRP:HB3	1:D:432:ARG:HD3	1.94	0.48
1:D:293:LEU:O	1:D:490:ARG:HD3	2.13	0.48
1:D:397:TYR:HB2	1:D:404:TYR:HB2	1.95	0.47
1:D:230:PRO:HB3	1:D:235:ARG:HH21	1.80	0.47
1:A:373:ILE:CD1	1:A:525:VAL:HG11	2.45	0.47
1:B:192:ASN:OD1	1:B:192:ASN:N	2.47	0.47
1:B:541:HIS:NE2	1:D:207:MET:HG3	2.30	0.47
1:C:434:LYS:HB2	1:C:462:THR:O	2.14	0.47
1:D:536:PHE:CE1	1:D:541:HIS:CD2	3.02	0.47
1:B:199:LEU:HD12	1:B:217:TRP:CE3	2.50	0.46
1:B:127:ASP:HB3	1:B:152:ILE:HG12	1.97	0.46
1:D:441:THR:O	1:D:442:ASN:CB	2.62	0.46
1:B:421:MET:CE	1:B:522:THR:HB	2.46	0.45
1:D:135:GLN:NE2	1:D:160:THR:HG23	2.32	0.45
1:A:381:VAL:CG1	1:B:230:PRO:HD3	2.47	0.45
1:A:211:LEU:CD2	1:A:385:LEU:CD1	2.94	0.45
1:D:194:GLU:HB3	1:D:221:LEU:HD11	1.98	0.45
1:B:394[B]:HIS:CE1	1:B:396:GLN:CG	3.00	0.45
1:D:401:LYS:O	1:D:405:GLU:HG2	2.16	0.45
1:D:229:LEU:HA	1:D:230:PRO:HD2	1.79	0.44
1:D:307:LEU:HD11	1:D:373:ILE:HD11	1.99	0.44
1:C:536:PHE:HE1	1:C:541:HIS:CE1	2.36	0.44
1:A:307:LEU:HD11	1:A:373:ILE:HD11	1.98	0.44
1:B:230:PRO:HB3	1:B:235:ARG:HH21	1.83	0.43
1:A:124:TYR:CZ	1:A:142:ILE:HG23	2.53	0.43
1:A:194:GLU:CD	1:A:221:LEU:HD21	2.38	0.43
1:A:381:VAL:HG12	1:B:230:PRO:HD3	1.99	0.43
1:B:207:MET:HG3	1:D:541:HIS:CE1	2.54	0.43
1:B:393:PHE:CZ	1:B:430:PRO:HB2	2.54	0.43
1:A:80:ALA:HB2	1:A:220:PRO:HD3	2.00	0.43
1:A:194:GLU:OE2	1:A:221:LEU:HD21	2.18	0.43
1:C:396:GLN:HA	1:C:396:GLN:OE1	2.19	0.43
1:D:80:ALA:HB2	1:D:220:PRO:HD3	2.00	0.43
1:C:277:LYS:HB2	1:C:277:LYS:HE3	1.65	0.43
1:A:127:ASP:O	1:A:131:ARG:HG3	2.19	0.43
1:B:207:MET:HE3	1:B:211:LEU:HD11	2.00	0.43
1:D:340:PRO:HA	1:D:341:PRO:HD2	1.84	0.43
1:D:211:LEU:HB2	1:D:212:PRO:HD3	2.01	0.42

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:207:MET:HE3	1:B:211:LEU:CD1	2.49	0.42
1:A:116:ARG:HE	1:A:139:ASN:HD22	1.67	0.42
1:A:377:GLU:HG2	2:A:596:HOH:O	2.18	0.42
1:B:541:HIS:HA	1:D:392:TRP:CZ2	2.54	0.42
1:D:277:LYS:HB2	1:D:277:LYS:HE3	1.73	0.42
1:A:184:LYS:O	1:A:199:LEU:HD23	2.19	0.42
1:A:211:LEU:HB2	1:A:212:PRO:HD3	2.02	0.42
1:D:154:ALA:O	1:D:157:MET:HB2	2.20	0.41
1:A:211:LEU:CD2	1:A:385:LEU:HD12	2.50	0.41
1:C:154:ALA:O	1:C:157:MET:HB2	2.20	0.41
1:A:294:SER:HA	2:A:656:HOH:O	2.20	0.41
1:A:305:HIS:HB2	1:A:439:CYS:HB3	2.03	0.41
1:C:149:ILE:HD13	1:C:170:LEU:CD1	2.37	0.41
1:C:230:PRO:HB3	1:C:235:ARG:HH21	1.86	0.41
1:D:230:PRO:HB3	1:D:235:ARG:NH2	2.35	0.41
1:B:541:HIS:CE1	1:D:207:MET:HE2	2.36	0.41
1:B:207:MET:CE	1:B:211:LEU:CD1	2.97	0.40
1:B:207:MET:HE3	1:B:207:MET:HB3	1.95	0.40
1:B:373:ILE:HD13	1:B:525:VAL:HG11	2.02	0.40
1:C:421:MET:SD	1:C:522:THR:HB	2.62	0.40
1:A:269:GLU:HB2	1:A:486:LEU:HB3	2.03	0.40
1:D:152:ILE:HB	1:D:157:MET:HE3	2.03	0.40
1:A:340:PRO:HA	1:A:341:PRO:HD2	1.78	0.40
1:B:238:ILE:HD13	1:B:258:GLU:OE2	2.21	0.40
1:C:86:LYS:HB3	1:C:104:THR:HG22	2.04	0.40
1:C:211:LEU:HB2	1:C:212:PRO:CD	2.51	0.40
1:B:157:MET:CE	1:B:161:LEU:HB3	2.52	0.40

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	409/470 (87%)	398 (97%)	11 (3%)	0	100	100
1	B	409/470 (87%)	401 (98%)	8 (2%)	0	100	100
1	C	408/470 (87%)	398 (98%)	10 (2%)	0	100	100
1	D	411/470 (87%)	403 (98%)	8 (2%)	0	100	100
All	All	1637/1880 (87%)	1600 (98%)	37 (2%)	0	100	100

There are no Ramachandran outliers to report.

### 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	383/422 (91%)	378 (99%)	5 (1%)	69	75
1	B	386/422 (92%)	383 (99%)	3 (1%)	81	86
1	C	384/422 (91%)	380 (99%)	4 (1%)	76	82
1	D	386/422 (92%)	382 (99%)	4 (1%)	76	82
All	All	1539/1688 (91%)	1523 (99%)	16 (1%)	76	82

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	92	ARG
1	A	371	TYR
1	A	399	GLN
1	A	476	ARG
1	A	489	SER
1	B	117	SER
1	B	371	TYR
1	B	399	GLN
1	C	106	LYS
1	C	111	THR
1	C	208	GLU
1	C	371	TYR
1	D	111	THR

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	196	LYS
1	D	371	TYR
1	D	399	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (18) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	135	GLN
1	A	139	ASN
1	A	242	ASN
1	A	288	ASN
1	B	139	ASN
1	B	242	ASN
1	B	288	ASN
1	B	541	HIS
1	C	135	GLN
1	C	139	ASN
1	C	242	ASN
1	C	288	ASN
1	C	309	GLN
1	D	135	GLN
1	D	139	ASN
1	D	242	ASN
1	D	288	ASN
1	D	541	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](#)

4 non-standard protein/DNA/RNA residues are 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
1	NEP	D	182	1	10,14,15	2.20	4 (40%)	5,20,22	1.34	0
1	NEP	A	182	1	10,14,15	2.18	5 (50%)	5,20,22	2.24	2 (40%)
1	NEP	B	182	1	10,14,15	2.38	4 (40%)	5,20,22	2.58	3 (60%)
1	NEP	C	182	1	10,14,15	2.41	5 (50%)	5,20,22	1.83	3 (60%)

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
1	NEP	D	182	1	-	0/5/12/14	0/1/1/1
1	NEP	A	182	1	-	1/5/12/14	0/1/1/1
1	NEP	B	182	1	-	0/5/12/14	0/1/1/1
1	NEP	C	182	1	-	0/5/12/14	0/1/1/1

All (18) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	B	182	NEP	CD2-NE2	-4.03	1.32	1.39
1	B	182	NEP	CD2-CG	3.85	1.41	1.36
1	C	182	NEP	P-O1P	3.81	1.62	1.54
1	D	182	NEP	CD2-NE2	-3.61	1.32	1.39
1	A	182	NEP	P-O3P	3.50	1.50	1.47
1	B	182	NEP	P-O3P	3.49	1.50	1.47
1	C	182	NEP	P-O3P	3.47	1.50	1.47
1	D	182	NEP	P-O2P	3.32	1.61	1.54
1	A	182	NEP	P-O1P	3.21	1.61	1.54
1	D	182	NEP	P-O3P	3.21	1.50	1.47
1	C	182	NEP	CD2-CG	3.20	1.41	1.36
1	C	182	NEP	P-O2P	3.08	1.61	1.54
1	C	182	NEP	CD2-NE2	-2.99	1.33	1.39
1	D	182	NEP	P-O1P	2.97	1.61	1.54
1	B	182	NEP	P-O2P	2.93	1.60	1.54
1	A	182	NEP	P-O2P	2.90	1.60	1.54
1	A	182	NEP	CD2-NE2	-2.79	1.34	1.39
1	A	182	NEP	CD2-CG	2.67	1.40	1.36

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	A	182	NEP	CB-CA-C	3.69	118.39	111.47
1	B	182	NEP	O1P-P-O3P	-3.63	105.60	113.44
1	A	182	NEP	O1P-P-O3P	-3.27	106.38	113.44
1	B	182	NEP	O2P-P-O3P	-3.07	106.81	113.44
1	B	182	NEP	CB-CA-C	2.75	116.62	111.47
1	C	182	NEP	O1P-P-O3P	-2.68	107.65	113.44
1	C	182	NEP	O2P-P-O3P	-2.03	109.05	113.44
1	C	182	NEP	CB-CA-C	2.01	115.23	111.47

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
1	A	182	NEP	CA-CB-CG-CD2

There are no ring outliers.

No monomer is involved in short contacts.

## 5.5 Carbohydrates [i](#)

There are no monosaccharides 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

### 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	421/470 (89%)	0.02	25 (5%) 22 27	25, 39, 68, 104	0
1	B	422/470 (89%)	0.12	26 (6%) 20 25	25, 40, 72, 103	0
1	C	422/470 (89%)	0.38	38 (9%) 9 12	32, 51, 81, 132	0
1	D	424/470 (90%)	0.43	40 (9%) 8 11	33, 54, 89, 142	0
All	All	1689/1880 (89%)	0.24	129 (7%) 13 18	25, 46, 80, 142	0

All (129) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	341	PRO	10.7
1	C	341	PRO	10.4
1	D	442	ASN	8.8
1	D	178	PHE	7.0
1	A	341	PRO	6.5
1	D	441	THR	5.9
1	B	341	PRO	5.8
1	C	442	ASN	5.7
1	C	178	PHE	5.5
1	D	278	PHE	5.3
1	B	442	ASN	5.1
1	A	340	PRO	4.8
1	D	450	SER	4.8
1	D	506	ASP	4.8
1	C	441	THR	4.7
1	C	326	TRP	4.6
1	D	177	PRO	4.3
1	D	277	LYS	4.1
1	A	494	ASN	4.1
1	C	230	PRO	4.0
1	C	331	ILE	4.0

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	C	307	LEU	3.9
1	C	372	ILE	3.9
1	B	278	PHE	3.7
1	D	308	CYS	3.7
1	D	396	GLN	3.5
1	D	320	ALA	3.5
1	B	92	ARG	3.5
1	C	506	ASP	3.5
1	C	308	CYS	3.4
1	A	441	THR	3.4
1	D	373	ILE	3.3
1	B	93	GLU	3.3
1	A	493	ALA	3.3
1	C	399	GLN	3.2
1	D	307	LEU	3.2
1	B	156	ALA	3.2
1	A	278	PHE	3.2
1	C	119	LEU	3.1
1	A	451	GLN	3.1
1	C	278	PHE	3.1
1	B	159	ALA	3.1
1	C	373	ILE	3.1
1	B	161	LEU	3.0
1	D	326	TRP	3.0
1	C	436	TYR	3.0
1	C	292	LYS	3.0
1	B	372	ILE	2.9
1	B	340	PRO	2.9
1	A	496	ARG	2.9
1	C	229	LEU	2.9
1	B	308	CYS	2.9
1	C	177	PRO	2.8
1	D	399	GLN	2.8
1	A	320	ALA	2.8
1	B	186	ILE	2.8
1	D	398	LEU	2.8
1	D	119	LEU	2.8
1	B	307	LEU	2.7
1	C	118	ILE	2.7
1	A	396	GLN	2.7
1	B	441	THR	2.7
1	A	187	ILE	2.7

*Continued on next page...*

*Continued from previous page...*

<b>Mol</b>	<b>Chain</b>	<b>Res</b>	<b>Type</b>	<b>RSRZ</b>
1	B	178	PHE	2.7
1	A	506	ASP	2.7
1	A	93	GLU	2.6
1	C	398	LEU	2.6
1	D	436	TYR	2.6
1	C	92	ARG	2.6
1	D	371	TYR	2.6
1	C	396	GLN	2.5
1	D	292	LYS	2.5
1	B	326	TRP	2.5
1	A	92	ARG	2.5
1	B	506	ASP	2.5
1	D	79	GLY	2.5
1	C	450	SER	2.5
1	D	440	ALA	2.5
1	D	531	ALA	2.4
1	B	229	LEU	2.4
1	C	306	TYR	2.4
1	A	186	ILE	2.4
1	C	514	PRO	2.4
1	D	230	PRO	2.4
1	A	505	ARG	2.4
1	D	514	PRO	2.4
1	C	142	ILE	2.4
1	C	185	LEU	2.4
1	D	301	ASP	2.4
1	B	451	GLN	2.4
1	D	372	ILE	2.4
1	D	361	ASN	2.3
1	C	140	ILE	2.3
1	B	436	TYR	2.3
1	A	185	LEU	2.3
1	C	476	ARG	2.3
1	A	308	CYS	2.3
1	A	399	GLN	2.3
1	C	365	GLN	2.3
1	B	119	LEU	2.3
1	C	330	MET	2.3
1	B	302	THR	2.3
1	C	327	THR	2.3
1	D	174	THR	2.3
1	D	191	ASP	2.2

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	D	451	GLN	2.2
1	C	371	TYR	2.2
1	A	118	ILE	2.2
1	B	187	ILE	2.2
1	D	235	ARG	2.2
1	D	484	GLY	2.2
1	A	135	GLN	2.1
1	C	187	ILE	2.1
1	D	421	MET	2.1
1	B	450	SER	2.1
1	B	493	ALA	2.1
1	C	301	ASP	2.1
1	A	302	THR	2.1
1	B	160	THR	2.1
1	A	372	ILE	2.1
1	C	186	ILE	2.1
1	D	476	ARG	2.1
1	D	365	GLN	2.1
1	D	112	GLU	2.0
1	A	301	ASP	2.0
1	C	435	PHE	2.0
1	A	101	ASP	2.0
1	D	229	LEU	2.0
1	D	493	ALA	2.0

## 6.2 Non-standard residues in protein, DNA, RNA chains [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(Å <sup>2</sup> )	Q<0.9
1	NEP	B	182	14/15	0.96	0.09	27,29,31,37	4
1	NEP	C	182	14/15	0.96	0.09	34,35,37,42	4
1	NEP	D	182	14/15	0.96	0.08	38,40,41,47	4
1	NEP	A	182	14/15	0.97	0.09	27,31,32,39	4

## 6.3 Carbohydrates [i](#)

There are no monosaccharides in this entry.

## 6.4 Ligands

There are no ligands in this entry.

## 6.5 Other polymers

There are no such residues in this entry.