wwPDB X-ray Structure Validation Summary Report (i)

Oct 11, 2023 – 10:31 PM EDT

PDB ID	:	7N3Y
Title	:	Crystal Structure of Saccharomyces cerevisiae Apn2 Catalytic Domain E59Q/D222N Mutant in Complex with DNA
Authors	:	Wojtaszek, J.L.; Krahn, J.; Wallace, B.D.; Williams, R.S.
Deposited on	:	2021-06-02
Resolution	:	2.73 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*
A user guide is available at
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

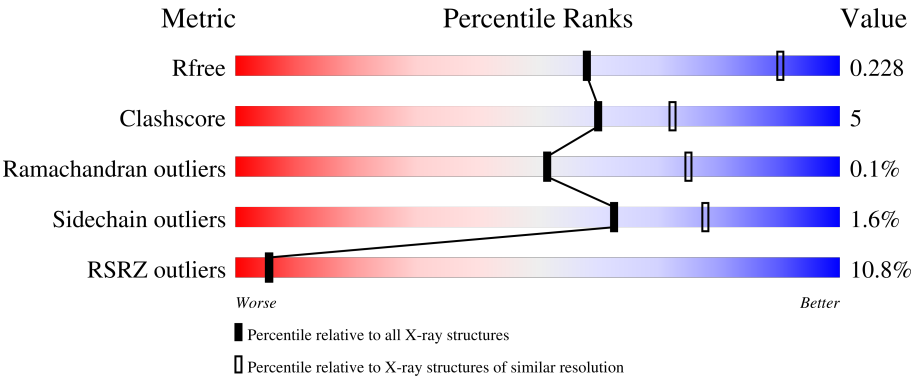
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 2.73 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	1271 (2.76-2.72)
Clashscore	141614	1322 (2.76-2.72)
Ramachandran outliers	138981	1297 (2.76-2.72)
Sidechain outliers	138945	1298 (2.76-2.72)
RSRZ outliers	127900	1243 (2.76-2.72)

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 >=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 <=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	413	% 85% 7% 8%
1	B	413	3% 81% 8% • 10%
1	C	413	17% 78% 11% 9%
1	D	413	18% 75% 12% 13%
2	E	13	8% 38% 46% 15%

Continued on next page...

Continued from previous page...

Mol	Chain	Length	Quality of chain
2	F	13	46% 38% 15%
2	G	13	8% 54% 46%
3	H	13	62% 23% 15%

2 Entry composition (i)

There are 9 unique types of molecules in this entry. The entry contains 12821 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 DNA-(apurinic or apyrimidinic site) endonuclease 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	378	Total	C	N	O	S	0	1	0
			2984	1901	523	545	15			
1	B	371	Total	C	N	O	S	0	2	0
			2919	1866	500	538	15			
1	C	374	Total	C	N	O	S	0	1	0
			2918	1856	509	537	16			
1	D	359	Total	C	N	O	S	0	1	0
			2806	1792	480	519	15			

There are 32 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	59	GLN	GLU	engineered mutation	UNP P38207
A	222	ASN	ASP	engineered mutation	UNP P38207
A	408	GLU	-	expression tag	UNP P38207
A	409	ASN	-	expression tag	UNP P38207
A	410	LEU	-	expression tag	UNP P38207
A	411	TYR	-	expression tag	UNP P38207
A	412	PHE	-	expression tag	UNP P38207
A	413	GLN	-	expression tag	UNP P38207
B	59	GLN	GLU	engineered mutation	UNP P38207
B	222	ASN	ASP	engineered mutation	UNP P38207
B	408	GLU	-	expression tag	UNP P38207
B	409	ASN	-	expression tag	UNP P38207
B	410	LEU	-	expression tag	UNP P38207
B	411	TYR	-	expression tag	UNP P38207
B	412	PHE	-	expression tag	UNP P38207
B	413	GLN	-	expression tag	UNP P38207
C	59	GLN	GLU	engineered mutation	UNP P38207
C	222	ASN	ASP	engineered mutation	UNP P38207
C	408	GLU	-	expression tag	UNP P38207
C	409	ASN	-	expression tag	UNP P38207
C	410	LEU	-	expression tag	UNP P38207

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	411	TYR	-	expression tag	UNP P38207
C	412	PHE	-	expression tag	UNP P38207
C	413	GLN	-	expression tag	UNP P38207
D	59	GLN	GLU	engineered mutation	UNP P38207
D	222	ASN	ASP	engineered mutation	UNP P38207
D	408	GLU	-	expression tag	UNP P38207
D	409	ASN	-	expression tag	UNP P38207
D	410	LEU	-	expression tag	UNP P38207
D	411	TYR	-	expression tag	UNP P38207
D	412	PHE	-	expression tag	UNP P38207
D	413	GLN	-	expression tag	UNP P38207

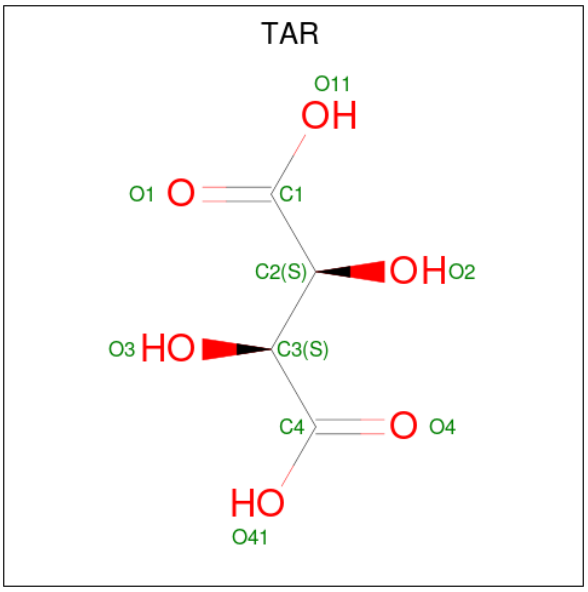
- Molecule 2 is a DNA chain called DNA (5'-D(*TP*CP*CP*GP*AP*AP*AP*TP*(PST)P*(PST)P*(SC)P*(GS)P*(GS))-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
2	E	13	Total	C	N	O	P	S	0	0	0
			263	127	47	72	12	5			
2	F	13	Total	C	N	O	P	S	0	0	0
			263	127	47	72	12	5			
2	G	13	Total	C	N	O	P	S	0	0	0
			263	127	47	72	12	5			

- Molecule 3 is a DNA chain called DNA (5'-D(*CP*CP*GP*AP*AP*AP*TP*TP*(PST)P*(SC)P*(GS)P*(GS)P*(AS))-3').

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
3	H	13	Total	C	N	O	P	S	0	0	0
			247	117	45	68	12	5			

- Molecule 4 is D(-)-TARTARIC ACID (three-letter code: TAR) (formula: $C_4H_6O_6$).

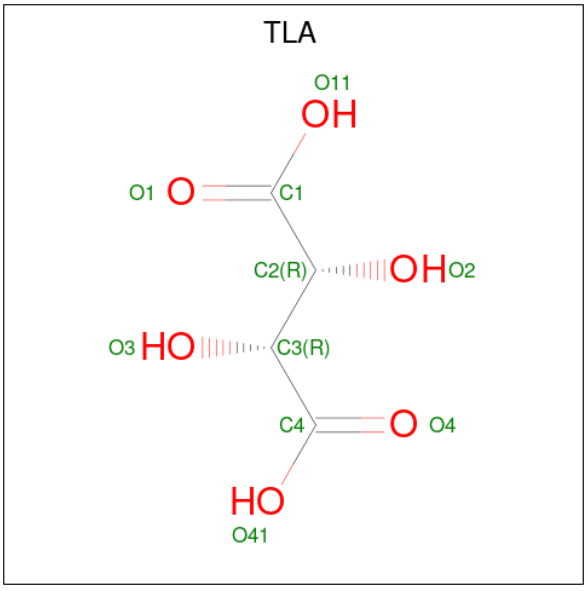


Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	1	Total C O 10 4 6	0	0
4	C	1	Total C O 10 4 6	0	0

- Molecule 5 is CALCIUM ION (three-letter code: CA) (formula: Ca).

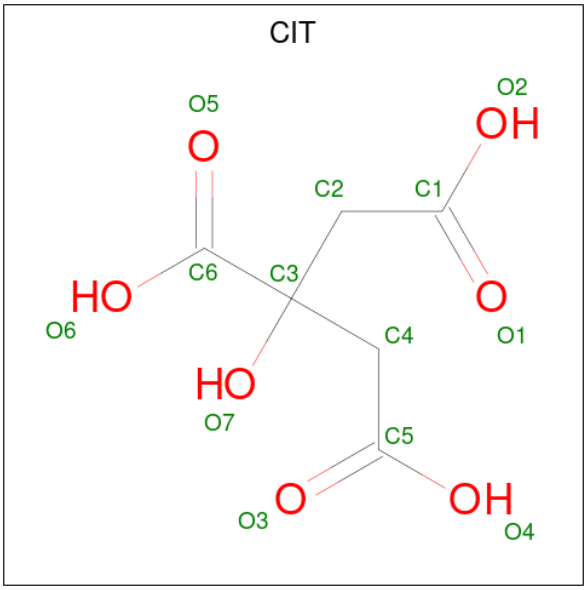
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
5	A	1	Total Ca 1 1	0	0

- Molecule 6 is L(+)-TARTARIC ACID (three-letter code: TLA) (formula: $C_4H_6O_6$).



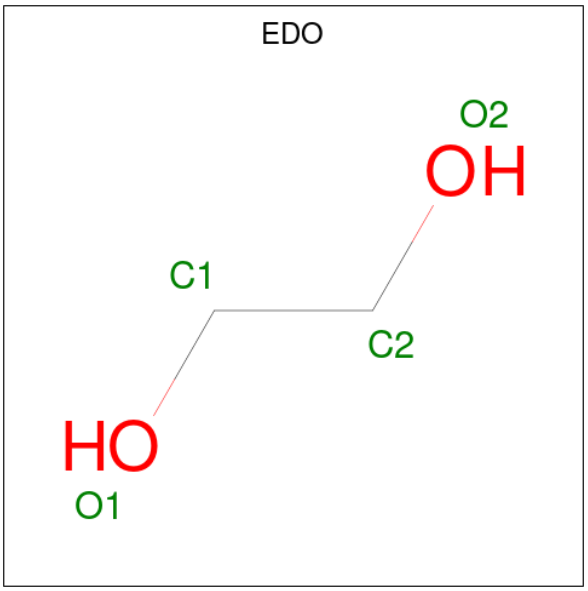
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	B	1	Total C O 10 4 6	0	0
6	D	1	Total C O 10 4 6	0	0

- Molecule 7 is CITRIC ACID (three-letter code: CIT) (formula: $C_6H_8O_7$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	1	Total C O 13 6 7	0	0

- Molecule 8 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	B	1	Total C O 4 2 2	0	0

- Molecule 9 is water.

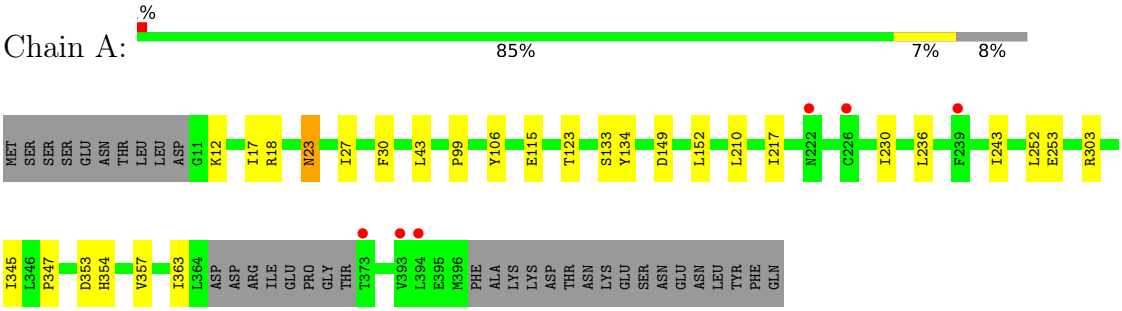
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
9	A	56	Total O 56 56	0	0
9	B	39	Total O 39 39	0	0
9	E	3	Total O 3 3	0	0
9	F	2	Total O 2 2	0	0

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.

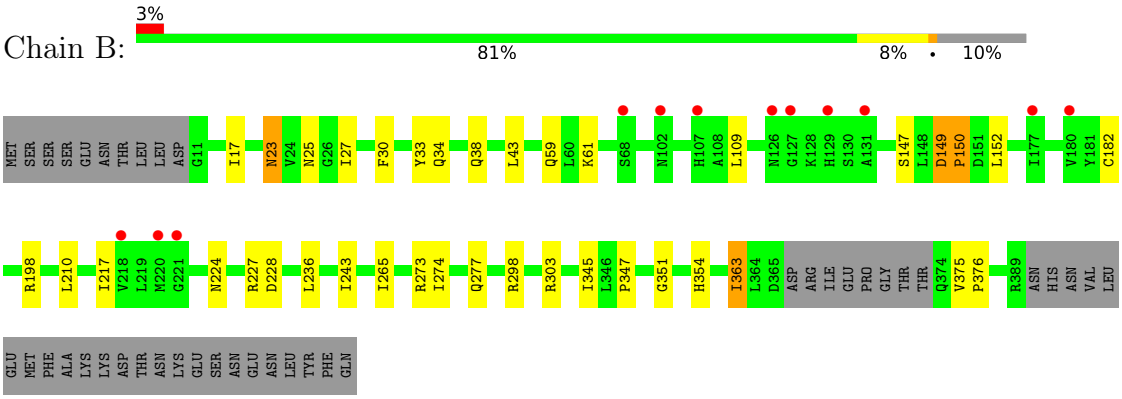
• Molecule 1: DNA-(apurinic or apyrimidinic site) endonuclease 2

Chain A:



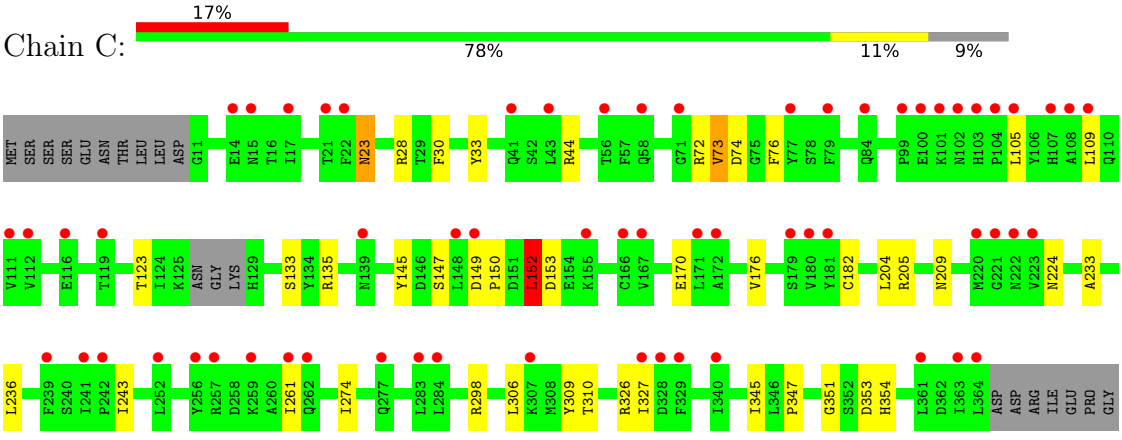
• Molecule 1: DNA-(apurinic or apyrimidinic site) endonuclease 2

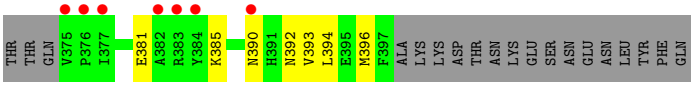
Chain B:



• Molecule 1: DNA-(apurinic or apyrimidinic site) endonuclease 2

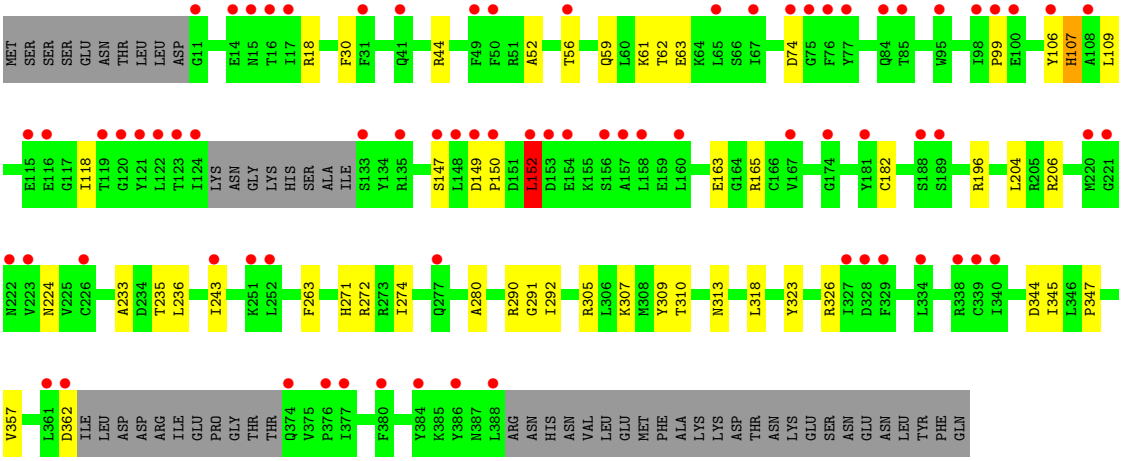
Chain C:





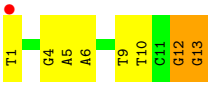
• Molecule 1: DNA-(apurinic or apyrimidinic site) endonuclease 2

Chain D:



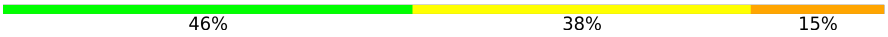
• Molecule 2: DNA (5'-D(*TP*CP*CP*GP*AP*AP*AP*TP*(PST)P*(PST)P*(SC)P*(GS)P*(GS))-3')

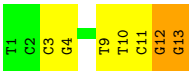
Chain E:



• Molecule 2: DNA (5'-D(*TP*CP*CP*GP*AP*AP*AP*TP*(PST)P*(PST)P*(SC)P*(GS)P*(GS))-3')

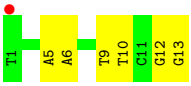
Chain F:





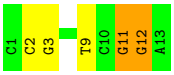
• Molecule 2: DNA (5'-D(*TP*CP*CP*GP*AP*AP*AP*TP*(PST)P*(PST)P*(SC)P*(GS)P*(GS))-3')

Chain G:



• Molecule 3: DNA (5'-D(*CP*CP*GP*AP*AP*AP*TP*TP*(PST)P*(SC)P*(GS)P*(GS)P*(AS))-3')

Chain H:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	69.41Å 101.47Å 144.97Å 90.00° 90.14° 90.00°	Depositor
Resolution (Å)	36.97 – 2.73 36.97 – 2.73	Depositor EDS
% Data completeness (in resolution range)	95.6 (36.97-2.73) 95.6 (36.97-2.73)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.06 (at 2.72Å)	Xtriage
Refinement program	PHENIX 1.19.2_4158	Depositor
R, R_{free}	0.200 , 0.228 0.199 , 0.228	Depositor DCC
R_{free} test set	1894 reflections (3.69%)	wwPDB-VP
Wilson B-factor (Å^2)	67.9	Xtriage
Anisotropy	0.279	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 53.1	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.044 for h,-k,-l	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	12821	wwPDB-VP
Average B, all atoms (Å^2)	101.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.54% of the height of the origin peak. No significant pseudotranslation is detected.*

[1] Intensities estimated from amplitudes.

[2] Theoretical values of $< |L| >$, $< L^2 >$ for acentric reflections are 0.5, 0.333 respectively for untwinned datasets, and 0.375, 0.2 for perfectly twinned datasets.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: CA, TAR, TLA, CIT, AS, SC, GS, EDO, PST

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	0.27	0/3042	0.43	0/4118
1	B	0.29	1/2979 (0.0%)	0.43	0/4035
1	C	0.29	2/2974 (0.1%)	0.42	0/4028
1	D	0.29	2/2862 (0.1%)	0.44	0/3881
2	E	0.59	0/179	0.92	0/274
2	F	0.49	0/179	0.86	0/274
2	G	0.59	0/179	0.90	0/274
3	H	0.47	0/179	0.89	0/274
All	All	0.31	5/12573 (0.0%)	0.48	0/17158

All (5) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	D	152	LEU	C-N	-5.63	1.21	1.34
1	B	147	SER	C-N	-5.46	1.21	1.34
1	D	147	SER	C-N	-5.44	1.21	1.34
1	C	147	SER	C-N	-5.18	1.22	1.34
1	C	152	LEU	C-N	5.10	1.45	1.34

There are no bond angle outliers.

There are no chirality outliers.

There are no planarity outliers.

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	2984	0	2965	15	0
1	B	2919	0	2889	25	0
1	C	2918	0	2845	34	0
1	D	2806	0	2737	29	0
2	E	263	0	151	6	0
2	F	263	0	151	3	0
2	G	263	0	151	1	0
3	H	247	0	138	2	0
4	A	10	0	4	1	0
4	C	10	0	4	0	0
5	A	1	0	0	0	0
6	B	10	0	4	0	0
6	D	10	0	4	0	0
7	B	13	0	5	1	0
8	B	4	0	6	3	0
9	A	56	0	0	0	0
9	B	39	0	0	1	0
9	E	3	0	0	0	0
9	F	2	0	0	0	0
All	All	12821	0	12054	114	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.

The worst 5 of 114 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:280:ALA:HB2	1:D:291:GLY:O	1.70	0.92
1:C:72:ARG:NH1	1:C:390:ASN:HA	1.98	0.79
1:C:152:LEU:HD23	1:C:152:LEU:O	1.82	0.78
1:B:198:ARG:HD2	8:B:503:EDO:O1	1.94	0.68
1:C:381:GLU:O	1:C:385:LYS:CB	2.42	0.68

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	375/413 (91%)	359 (96%)	16 (4%)	0	100	100
1	B	369/413 (89%)	354 (96%)	14 (4%)	1 (0%)	41	61
1	C	369/413 (89%)	352 (95%)	17 (5%)	0	100	100
1	D	354/413 (86%)	335 (95%)	19 (5%)	0	100	100
All	All	1467/1652 (89%)	1400 (95%)	66 (4%)	1 (0%)	51	75

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	150	PRO

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	324/373 (87%)	322 (99%)	2 (1%)	86	91
1	B	316/373 (85%)	312 (99%)	4 (1%)	69	82
1	C	311/373 (83%)	304 (98%)	7 (2%)	50	70
1	D	301/373 (81%)	294 (98%)	7 (2%)	50	70
All	All	1252/1492 (84%)	1232 (98%)	20 (2%)	62	78

5 of 20 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	D	56	THR
1	D	290	ARG
1	D	362	ASP
1	D	292	ILE
1	C	23	ASN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such

sidechains are listed below:

Mol	Chain	Res	Type
1	D	107	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)

20 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	$\#\|Z\| > 2$
2	GS	G	12	3,2	18,24,25	1.10	2 (11%)	19,35,38	0.71	1 (5%)
2	PST	F	9	2	18,21,22	0.32	0	26,30,33	0.76	1 (3%)
2	GS	G	13	3,2	18,24,25	1.09	2 (11%)	19,35,38	0.65	0
2	SC	F	11	2	17,20,21	0.34	0	24,28,31	0.46	0
2	GS	E	12	2	18,24,25	1.10	2 (11%)	19,35,38	0.69	1 (5%)
2	PST	E	9	2	18,21,22	0.31	0	26,30,33	0.73	1 (3%)
2	PST	F	10	2	18,21,22	0.30	0	26,30,33	0.71	1 (3%)
2	PST	G	10	3,2	18,21,22	0.30	0	26,30,33	0.66	1 (3%)
2	SC	E	11	2	17,20,21	0.36	0	24,28,31	0.51	0
3	PST	H	9	3,2	18,21,22	0.31	0	26,30,33	0.72	1 (3%)
3	SC	H	10	3,2	17,20,21	0.34	0	24,28,31	0.48	0
2	SC	G	11	3,2	17,20,21	0.32	0	24,28,31	0.43	0
3	AS	H	13	3	0,3,24	-	-	0,3,36	-	-
2	GS	F	12	2	18,24,25	1.14	3 (16%)	19,35,38	0.73	1 (5%)
2	GS	F	13	2	18,24,25	1.08	2 (11%)	19,35,38	0.64	0
3	GS	H	11	3,2	18,24,25	1.11	2 (11%)	19,35,38	0.73	1 (5%)
3	GS	H	12	3,2	18,24,25	1.08	2 (11%)	19,35,38	0.70	0
2	PST	G	9	3,2	18,21,22	0.32	0	26,30,33	0.68	1 (3%)
2	PST	E	10	2	18,21,22	0.29	0	26,30,33	0.69	1 (3%)

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
2	GS	E	13	2	18,24,25	1.11	2 (11%)	19,35,38	0.65	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
2	GS	G	12	3,2	-	0/2/21/22	0/3/3/3
2	PST	F	9	2	-	3/7/21/22	0/2/2/2
2	GS	G	13	3,2	-	0/2/21/22	0/3/3/3
2	SC	F	11	2	-	0/7/21/22	0/2/2/2
2	GS	E	12	2	-	0/2/21/22	0/3/3/3
2	PST	E	9	2	-	0/7/21/22	0/2/2/2
2	PST	F	10	2	-	0/7/21/22	0/2/2/2
2	PST	G	10	3,2	-	0/7/21/22	0/2/2/2
2	SC	E	11	2	-	0/7/21/22	0/2/2/2
3	PST	H	9	3,2	-	0/7/21/22	0/2/2/2
3	SC	H	10	3,2	-	0/7/21/22	0/2/2/2
2	SC	G	11	3,2	-	0/7/21/22	0/2/2/2
2	GS	F	12	2	-	0/2/21/22	0/3/3/3
2	GS	F	13	2	-	0/2/21/22	0/3/3/3
3	GS	H	11	3,2	-	0/2/21/22	0/3/3/3
3	GS	H	12	3,2	-	0/2/21/22	0/3/3/3
2	PST	G	9	3,2	-	0/7/21/22	0/2/2/2
2	PST	E	10	2	-	0/7/21/22	0/2/2/2
2	GS	E	13	2	-	0/2/21/22	0/3/3/3

The worst 5 of 17 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	13	GS	C5-C6	-2.74	1.41	1.47
2	G	12	GS	C5-C6	-2.73	1.41	1.47
2	G	13	GS	C5-C6	-2.73	1.41	1.47
2	F	12	GS	C8-N7	-2.73	1.30	1.35
2	F	12	GS	C5-C6	-2.70	1.41	1.47

The worst 5 of 11 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	E	9	PST	O4-C4-C5	-2.50	122.01	124.90
2	F	9	PST	O4-C4-C5	-2.37	122.16	124.90

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	H	9	PST	O4-C4-C5	-2.35	122.17	124.90
2	F	10	PST	O4-C4-C5	-2.32	122.22	124.90
2	G	10	PST	O4-C4-C5	-2.25	122.30	124.90

There are no chirality outliers.

All (3) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	F	9	PST	O4'-C4'-C5'-O5'
2	F	9	PST	C3'-C4'-C5'-O5'
2	F	9	PST	C4'-C5'-O5'-P

There are no ring outliers.

7 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	F	11	SC	1	0
2	E	12	GS	1	0
2	F	12	GS	1	0
2	F	13	GS	1	0
3	H	11	GS	1	0
3	H	12	GS	1	0
2	E	13	GS	1	0

5.5 Carbohydrates (i)

There are no monosaccharides in this entry.

5.6 Ligand geometry (i)

Of 7 ligands modelled in this entry, 1 is monoatomic - leaving 6 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	$\#\|Z\|>2$
6	TLA	D	501	-	9,9,9	1.20	0	12,12,12	1.02	0
6	TLA	B	501	-	9,9,9	1.14	0	12,12,12	1.02	0
4	TAR	C	501	-	9,9,9	1.26	0	12,12,12	0.96	0
8	EDO	B	503	-	3,3,3	0.32	0	2,2,2	0.39	0
7	CIT	B	502	-	12,12,12	1.01	0	17,17,17	1.59	1 (5%)
4	TAR	A	501	-	9,9,9	1.79	3 (33%)	12,12,12	1.25	2 (16%)

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
6	TLA	D	501	-	-	10/12/12/12	-
6	TLA	B	501	-	-	6/12/12/12	-
4	TAR	C	501	-	-	4/12/12/12	-
8	EDO	B	503	-	-	1/1/1/1	-
7	CIT	B	502	-	-	6/16/16/16	-
4	TAR	A	501	-	-	4/12/12/12	-

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	A	501	TAR	C3-C4	-3.19	1.48	1.52
4	A	501	TAR	C2-C1	-2.79	1.48	1.52
4	A	501	TAR	O11-C1	-2.39	1.22	1.30

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
7	B	502	CIT	O6-C6-C3	4.23	120.40	113.05
4	A	501	TAR	O4-C4-C3	-2.13	116.04	121.63
4	A	501	TAR	O11-C1-C2	2.08	118.89	113.27

There are no chirality outliers.

5 of 31 torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
4	C	501	TAR	O3-C3-C4-O41
6	B	501	TLA	O11-C1-C2-C3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
7	B	502	CIT	C2-C3-C6-O5
7	B	502	CIT	C2-C3-C6-O6
7	B	502	CIT	O7-C3-C6-O5

There are no ring outliers.

3 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
8	B	503	EDO	3	0
7	B	502	CIT	1	0
4	A	501	TAR	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^{th} 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^2)	Q<0.9
1	A	378/413 (91%)	-0.15	6 (1%) 72 78	39, 57, 102, 150	0
1	B	371/413 (89%)	0.01	12 (3%) 47 54	41, 59, 114, 172	0
1	C	374/413 (90%)	0.99	69 (18%) 1 1	91, 129, 173, 193	0
1	D	359/413 (86%)	1.12	75 (20%) 1 0	95, 147, 198, 218	0
2	E	8/13 (61%)	1.02	1 (12%) 3 4	102, 120, 151, 158	0
2	F	8/13 (61%)	-0.28	0 100 100	46, 55, 75, 76	0
2	G	8/13 (61%)	0.21	1 (12%) 3 4	137, 141, 157, 163	0
3	H	8/13 (61%)	-0.67	0 100 100	98, 109, 126, 127	0
All	All	1514/1704 (88%)	0.48	164 (10%) 5 5	39, 103, 175, 218	0

The worst 5 of 164 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	221	GLY	6.8
1	C	239	PHE	5.7
1	D	362	ASP	5.6
1	C	109	LEU	5.4
1	D	15	ASN	5.2

6.2 Non-standard residues in protein, DNA, RNA chains

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^{th} 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^2)	Q<0.9
2	PST	G	9	20/21	0.81	0.18	130,142,175,197	0

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	PST	F	10	20/21	0.83	0.26	80,109,136,145	0
3	SC	H	10	19/20	0.84	0.24	124,141,151,160	0
2	SC	F	11	19/20	0.85	0.20	89,117,142,148	0
3	GS	H	11	22/23	0.85	0.16	117,136,148,159	0
2	GS	F	12	22/23	0.86	0.12	107,121,140,142	0
2	PST	F	9	20/21	0.88	0.20	71,95,116,128	0
2	GS	F	13	22/23	0.88	0.17	125,141,159,163	0
3	PST	H	9	20/21	0.90	0.17	113,136,158,167	0
3	GS	H	12	22/23	0.91	0.14	122,136,150,163	0
2	PST	E	10	20/21	0.92	0.17	52,68,106,121	0
2	PST	E	9	20/21	0.93	0.20	60,83,136,139	0
2	SC	G	11	19/20	0.93	0.14	98,115,151,152	0
3	AS	H	13	4/22	0.93	0.14	183,190,193,198	0
2	PST	G	10	20/21	0.95	0.12	96,126,156,193	0
2	GS	G	12	22/23	0.95	0.12	86,101,132,135	0
2	SC	E	11	19/20	0.96	0.13	49,63,107,108	0
2	GS	E	12	22/23	0.96	0.14	42,48,85,88	0
2	GS	G	13	22/23	0.97	0.14	83,93,105,121	0
2	GS	E	13	22/23	0.99	0.12	36,42,50,59	0

6.3 Carbohydrates (i)

There are no monosaccharides 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, 95th 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^2)	Q<0.9
7	CIT	B	502	13/13	0.76	0.23	106,119,132,136	0
4	TAR	C	501	10/10	0.84	0.27	129,139,148,156	0
6	TLA	D	501	10/10	0.85	0.21	140,147,154,159	0
4	TAR	A	501	10/10	0.85	0.27	88,104,115,118	0
8	EDO	B	503	4/4	0.92	0.40	75,80,83,92	0
6	TLA	B	501	10/10	0.93	0.15	47,79,95,95	0
5	CA	A	502	1/1	0.97	0.18	108,108,108,108	0

6.5 Other polymers (i)

There are no such residues in this entry.