



wwPDB X-ray Structure Validation Summary Report i

Jun 30, 2014 – 10:45 AM EDT

PDB ID : 4PSX
Title : Crystal structure of histone acetyltransferase complex
Authors : Yang, M.; Li, Y.
Deposited on : 2014-03-08
Resolution : 2.51 Å(reported)

This is a wwPDB validation summary report for a publicly released PDB entry.

We welcome your comments at validation@mail.wwpdb.org

A user guide is available at <http://wwpdb.org/ValidationPDFNotes.html>

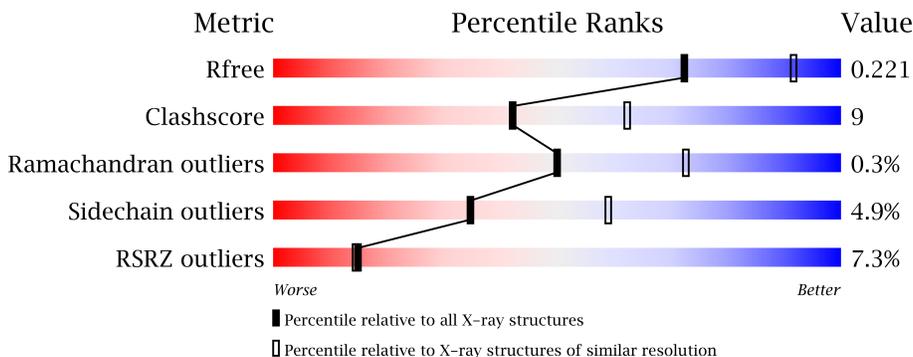
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : **FAILED**
Xtriage (Phenix) : dev-1439
EDS : stable23161
Percentile statistics : 21963
Refmac : 5.8.0049
CCP4 : 6.3.0 (Settle)
Ideal geometry (proteins) : Engh & Huber (2001)
Ideal geometry (DNA, RNA) : Parkinson et. al. (1996)
Validation Pipeline (wwPDB-VP) : stable23161

1 Overall quality at a glance

The reported resolution of this entry is 2.51 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}	66092	2784 (2.50-2.50)
Clashscore	79885	3562 (2.50-2.50)
Ramachandran outliers	78287	3480 (2.50-2.50)
Sidechain outliers	78261	3482 (2.50-2.50)
RSRZ outliers	66119	2785 (2.50-2.50)

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 ≥ 3 , 2, 1 and 0 types of geometric quality criteria. The upper red bar (where present) indicates the fraction of residues that have poor fit to the electron density.

Mol	Chain	Length	Quality of chain
1	A	320	
1	D	320	
2	B	401	
2	E	401	
3	C	48	
3	F	48	
4	P	15	
4	Y	15	

The following table lists non-polymeric compounds that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Geometry	Electron density
5	COA	A	401	-	X
5	COA	D	401	-	X
6	SO4	B	501	-	X

2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 12392 atoms, of which 0 are hydrogen and 0 are deuterium.

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 Histone acetyltransferase type B catalytic subunit.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	313	Total 2603	C 1679	N 426	O 494	S 4	0	0	0
1	D	313	Total 2603	C 1679	N 426	O 494	S 4	0	0	0

- Molecule 2 is a protein called Histone acetyltransferase type B subunit 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	361	Total 2868	C 1808	N 484	O 566	S 10	0	0	0
2	E	361	Total 2868	C 1808	N 484	O 566	S 10	0	0	0

There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
B	143	THR	VAL	ENGINEERED MUTATION	UNP P39984
E	143	THR	VAL	ENGINEERED MUTATION	UNP P39984

- Molecule 3 is a protein called Histone H4.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
3	C	39	Total 294	C 180	N 70	O 44	0	0	0
3	F	37	Total 280	C 171	N 67	O 42	0	0	0

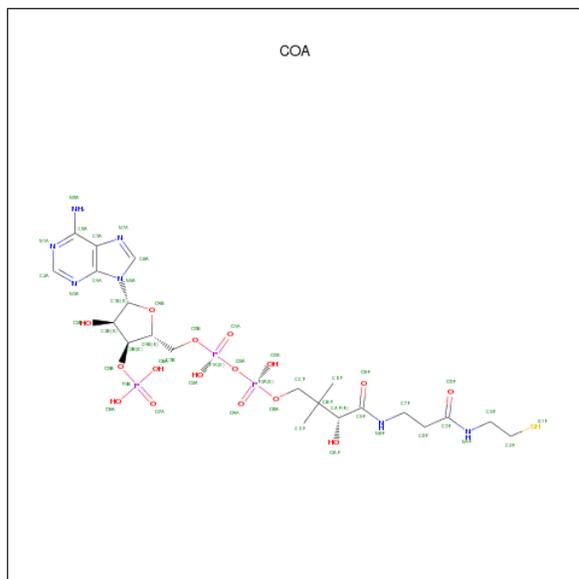
There are 2 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	21	VAL	ILE	ENGINEERED MUTATION	UNP P02309
F	21	VAL	ILE	ENGINEERED MUTATION	UNP P02309

- Molecule 4 is a protein called Histone H3.

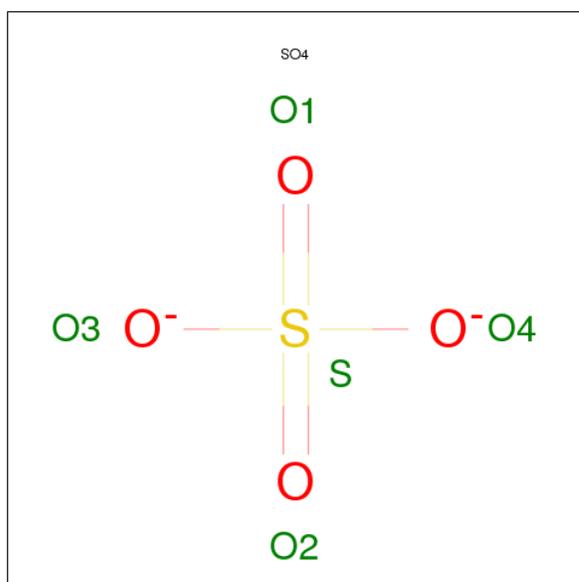
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
4	Y	12	90	52	21	17	0	0	0
4	P	10	79	46	19	14	0	0	0

- Molecule 5 is COENZYME A (three-letter code: COA) (formula: $C_{21}H_{36}N_7O_{16}P_3S$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	
			Total	C	N	O	P			S
5	A	1	48	21	7	16	3	1	0	0
5	D	1	48	21	7	16	3	1	0	0

- Molecule 6 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
6	B	1	Total	O	S	0	0
			5	4	1		

- Molecule 7 is water.

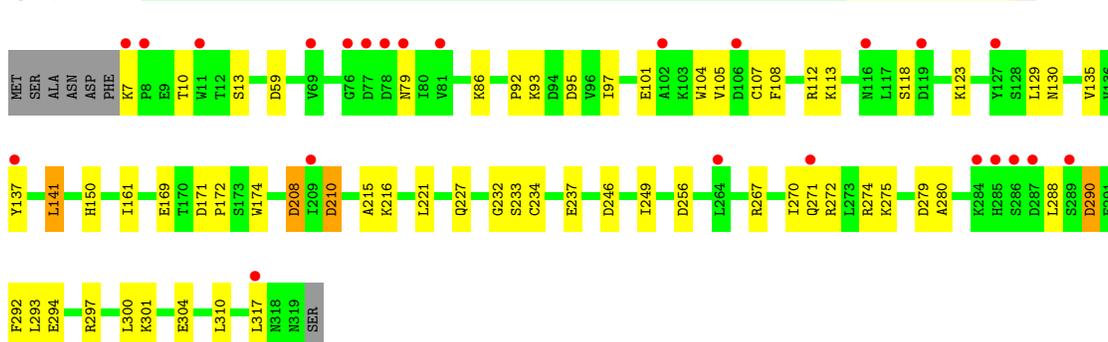
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	125	Total	O	0	0
			125	125		
7	B	153	Total	O	0	0
			153	153		
7	C	19	Total	O	0	0
			19	19		
7	D	122	Total	O	0	0
			122	122		
7	E	167	Total	O	0	0
			167	167		
7	F	18	Total	O	0	0
			18	18		
7	Y	1	Total	O	0	0
			1	1		
7	P	1	Total	O	0	0
			1	1		

3 Residue-property plots

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

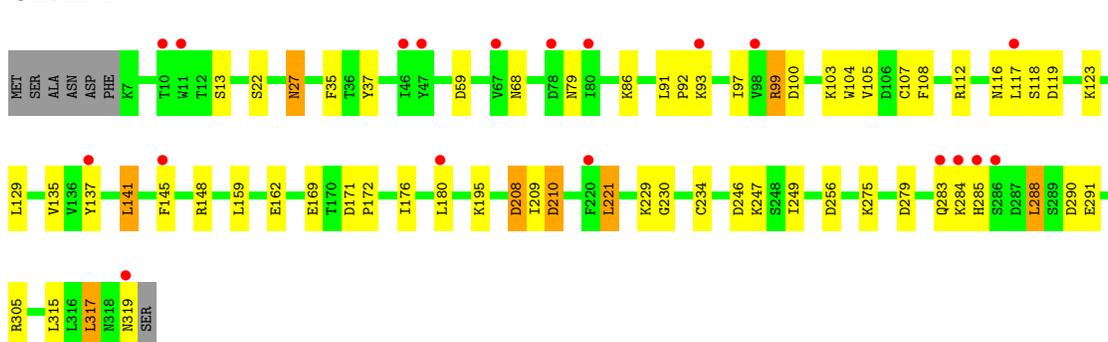
- Molecule 1: Histone acetyltransferase type B catalytic subunit

Chain A:



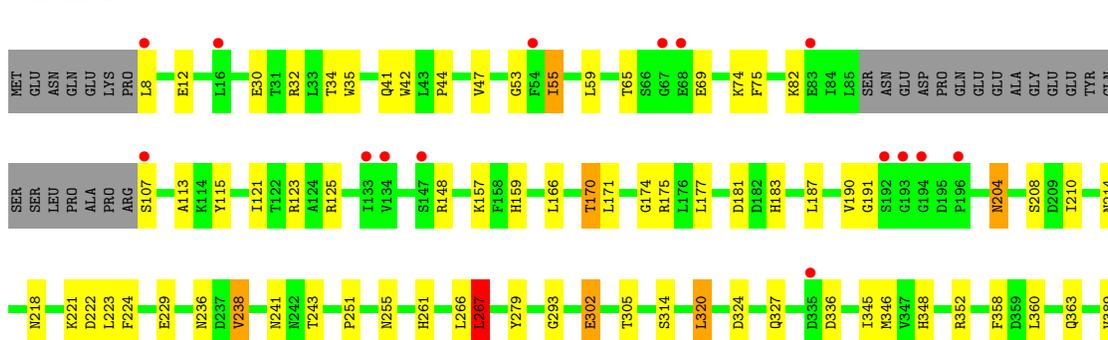
- Molecule 1: Histone acetyltransferase type B catalytic subunit

Chain D:



- Molecule 2: Histone acetyltransferase type B subunit 2

Chain B:



GLY
GLY
PRO
PRO
LYS
VAL
ASN
LYS
ASP
ILE
ILE
SER

• Molecule 2: Histone acetyltransferase type B subunit 2

Chain E:

MET	GLU	ASN	GLN	PRO	LYS	VAL	ASN	LYS	ASP	ILE	ILE	SER																													
L8	S9	V10	Y14	D26	F27	V28	S29	E30	Q41	P44	P46	V47	I65	K56	Q67	E58	F65	E69	F76	A76	E77	I78	E83	I84	L85	S86	ASN	GLU	ASP	PRO	GLN	GLU	GLU	ALA	GLY	GLU	TYR	GLN	SER	SER	LEU

PRO	ALA	PRO	ARG	SER	N108	I109	R110	I111	T112	Y115	R125	I133	R148	S149	E150	G151	L152	K157	F158	H159	S169	L301	L177	D181	D182	H183	V188	S192	G193	G194	R201	N204	H207	I210	N214	L223	F224	E229	I233	N236	D237
-----	-----	-----	-----	-----	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

V288	N241	T243	T244	L245	K249	N255	F259	L267	M272	Y279	K285	E286	H290	G293	N300	L301	E302	S314	L320	D324	Q327	A330	E331	Q332	T333	P334	D335	D336	A337	E338	E343	M346	R352	F358	D359	L360	N361	P362
------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------	------

Q363	L367	N375	S383	P387	I388	V389	GLY	GLY	PRO	PRO	LYS	VAL	ASN	LYS	ASP	ILE	ILE	SER
------	------	------	------	------	------	------	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

• Molecule 3: Histone H4

Chain C:

SER	GLY	ARG	GLY	LYS	GLY	G7	K8	K12	K16	R17	V21	L22	R23	I26	Q27	G28	I29	T30	A33	I34	R35	R36	L37	B40	K44	R45	ILE	SER	GLY
-----	-----	-----	-----	-----	-----	----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

• Molecule 3: Histone H4

Chain F:

SER	GLY	ARG	GLY	LYS	GLY	G7	K12	R17	H18	R19	K20	V21	L22	R23	D24	N25	I26	Q27	G28	I29	T30	K31	P32	A33	I34	R35	R36	R36	L37	V43	LYS	ARG	ILE	ILE	SER	GLY
-----	-----	-----	-----	-----	-----	----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

• Molecule 4: Histone H3

Chain Y:

A1	R2	Q5	R8	K9	S10	G12	GLY	LYS	ALA
----	----	----	----	----	-----	-----	-----	-----	-----

• Molecule 4: Histone H3

Chain P:

A1	R2	T3	T6	A7	R8	K9	S10	THR	GLY	GLY	LYS	ALA
----	----	----	----	----	----	----	-----	-----	-----	-----	-----	-----

4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, α , β , γ	78.55Å 85.33Å 114.51Å 75.97° 72.40° 66.06°	Depositor
Resolution (Å)	38.60 – 2.51 38.62 – 2.51	Depositor EDS
% Data completeness (in resolution range)	75.7 (38.60-2.51) 65.4 (38.62-2.51)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	3.06 (at 2.51Å)	Xtriage
Refinement program	PHENIX (phenix.refine: 1.8.1069)	Depositor
R, R_{free}	0.181 , 0.223 0.180 , 0.221	Depositor DCC
R_{free} test set	3340 reflections (5.03%)	DCC
Wilson B-factor (Å ²)	26.7	Xtriage
Anisotropy	0.277	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.33 , 87.4	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.49$, $\langle L^2 \rangle = 0.32$	Xtriage
Outliers	0 of 66429 reflections	Xtriage
F_o, F_c correlation	0.82	EDS
Total number of atoms	12392	wwPDB-VP
Average B, all atoms (Å ²)	79.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.80% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

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: COA, SO4

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	0.47	0/2665	0.59	0/3602
1	D	0.48	0/2665	0.58	0/3602
2	B	0.47	0/2938	0.64	1/4001 (0.0%)
2	E	0.46	0/2938	0.62	1/4001 (0.0%)
3	C	0.42	0/295	0.62	0/386
3	F	0.46	0/281	0.71	0/368
4	P	0.35	0/78	0.60	0/101
4	Y	0.35	0/89	0.57	0/116
All	All	0.47	0/11949	0.61	2/16177 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	E	267	LEU	CA-CB-CG	6.88	131.12	115.30
2	B	267	LEU	CA-CB-CG	6.66	130.61	115.30

There are no chirality outliers.

There are no planarity outliers.

5.2 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 hydrogens added by MolProbity. The Clashes column lists the number of clashes within the asymmetric unit, and the number in parentheses is this value normalized per 1000 atoms of the molecule in the chain. The Symm-Clashes column gives symmetry related clashes, in the same way as for the Clashes column.

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
1	A	2603	0	2550	41	0
1	D	2603	0	2550	47	0
2	B	2868	0	2744	51	0
2	E	2868	0	2744	48	0
3	C	294	0	332	9	0
3	F	280	0	317	10	0
4	P	79	0	91	3	0
4	Y	90	0	101	5	0
5	A	48	0	32	5	0
5	D	48	0	32	8	0
6	B	5	0	0	0	0
7	A	125	0	0	11	0
7	B	153	0	0	18	0
7	C	19	0	0	3	0
7	D	122	0	0	18	0
7	E	167	0	0	16	0
7	F	18	0	0	4	0
7	P	1	0	0	0	0
7	Y	1	0	0	0	0
All	All	12392	0	11493	210	0

Clashscore is defined as the number of clashes calculated for the entry per 1000 atoms (including hydrogens) of the entry. The overall clashscore for this entry is 9.

The worst 5 of 210 close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
1:A:107:CYS:SG	7:A:592:HOH:O	2.25	0.94
3:F:17:ARG:HG3	7:F:103:HOH:O	1.71	0.89
1:A:150:HIS:ND1	7:A:532:HOH:O	2.07	0.88
2:E:194:GLY:O	7:E:620:HOH:O	1.95	0.83
1:A:232:GLY:HA3	5:A:401:COA:H132	1.62	0.82

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	311/320 (97%)	300 (96%)	11 (4%)	0	100	100
1	D	311/320 (97%)	305 (98%)	5 (2%)	1 (0%)	50	73
2	B	357/401 (89%)	340 (95%)	16 (4%)	1 (0%)	50	73
2	E	357/401 (89%)	342 (96%)	14 (4%)	1 (0%)	50	73
3	C	37/48 (77%)	34 (92%)	3 (8%)	0	100	100
3	F	35/48 (73%)	33 (94%)	2 (6%)	0	100	100
4	P	8/15 (53%)	8 (100%)	0	0	100	100
4	Y	10/15 (67%)	9 (90%)	0	1 (10%)	1	1
All	All	1426/1568 (91%)	1371 (96%)	51 (4%)	4 (0%)	50	73

All (4) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	D	99	ARG
4	Y	11	THR
2	B	293	GLY
2	E	293	GLY

5.3.2 Protein sidechains ❶

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	288/294 (98%)	277 (96%)	11 (4%)	44	71
1	D	288/294 (98%)	276 (96%)	12 (4%)	40	66
2	B	323/358 (90%)	307 (95%)	16 (5%)	34	58
2	E	323/358 (90%)	310 (96%)	13 (4%)	42	68
3	C	27/33 (82%)	23 (85%)	4 (15%)	4	8
3	F	26/33 (79%)	21 (81%)	5 (19%)	2	3
4	P	8/10 (80%)	7 (88%)	1 (12%)	7	12
4	Y	9/10 (90%)	8 (89%)	1 (11%)	9	16
All	All	1292/1390 (93%)	1229 (95%)	63 (5%)	35	59

5 of 63 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
3	C	21	VAL
1	D	159	LEU
3	F	21	VAL
3	C	26	ILE
1	D	13	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (2) such sidechains are listed below:

Mol	Chain	Res	Type
2	E	108	ASN
2	E	255	ASN

5.3.3 RNA [i](#)

There are no RNA chains in this entry.

5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

Mogul failed to run properly - this section will therefore be empty.

5.5 Carbohydrates [i](#)

Mogul failed to run properly - this section will therefore be empty.

5.6 Ligand geometry [i](#)

Mogul failed to run properly - this section will therefore be empty.

5.7 Other polymers [i](#)

Mogul failed to run properly - this section will therefore be empty.

5.8 Polymer linkage issues

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, 95th 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	313/320 (97%)	1.05	24 (7%) 13 13	52, 76, 122, 158	0
1	D	313/320 (97%)	0.95	19 (6%) 21 20	49, 75, 113, 149	0
2	B	361/401 (90%)	0.96	15 (4%) 35 36	51, 72, 110, 150	0
2	E	361/401 (90%)	0.97	19 (5%) 25 26	51, 73, 108, 152	0
3	C	39/48 (81%)	1.41	6 (15%) 3 2	59, 84, 154, 190	0
3	F	37/48 (77%)	1.43	8 (21%) 1 1	55, 81, 140, 175	0
4	P	10/15 (66%)	2.33	4 (40%) 1 0	85, 116, 128, 152	0
4	Y	12/15 (80%)	2.77	8 (66%) 0 0	90, 111, 132, 163	0
All	All	1446/1568 (92%)	1.03	103 (7%) 15 15	49, 75, 119, 190	0

The worst 5 of 103 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
3	C	26	ILE	10.2
4	P	10	SER	8.0
3	F	27	GLN	7.8
4	Y	11	THR	6.5
1	A	287	ASP	5.7

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

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, 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	RSR	LLDF	B-factors(Å ²)	Q<0.9
6	SO4	B	501	5/5	0.41	14.05	111,122,126,126	0
5	COA	A	401	48/48	0.38	4.65	137,138,141,141	0
5	COA	D	401	48/48	0.36	4.31	129,132,135,135	0

6.5 Other polymers

There are no such residues in this entry.