wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 01:11 PM GMT

PDB ID	:	3T81
Title	:	Crystal Structure of diiron adenine deaminase
Authors	:	Bagaria, A.; Kumaran, D.; Burley, S.K.; Swaminathan, S.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on	:	2011-08-01
Resolution	:	2.63 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

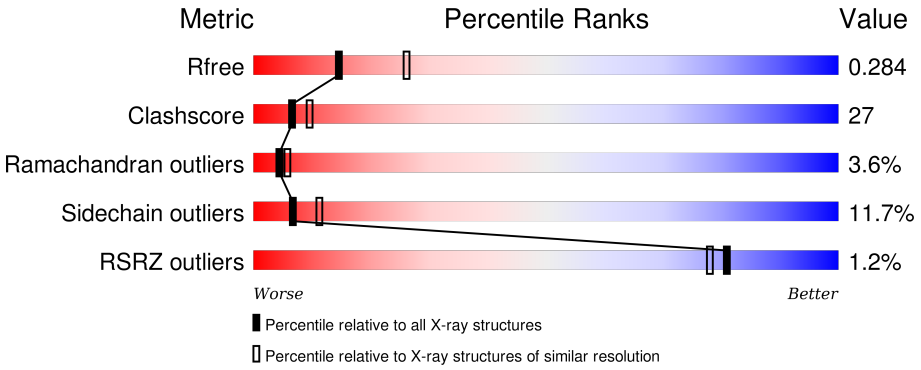
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 2.63 Å.

Percentile scores (ranging between 0-100) for global validation metrics of the entry are shown in the following graphic. The table shows the number of entries on which the scores are based.



Metric	Whole archive (#Entries)	Similar resolution (#Entries, resolution range(Å))
R_{free}	91344	3377 (2.68-2.60)
Clashscore	102246	3781 (2.68-2.60)
Ramachandran outliers	100387	3722 (2.68-2.60)
Sidechain outliers	100360	3722 (2.68-2.60)
RSRZ outliers	91569	3388 (2.68-2.60)

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. 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	608	47% 37% 12%
1	B	608	2% 40% 36% 17% 5%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 8674 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 Adenine deaminase 2.

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	A	587	Total	C	N	O	S	Se	0	0	0
			4337	2723	775	817	6	16			
1	B	579	Total	C	N	O	S	Se	0	0	0
			4277	2685	763	807	6	16			

There are 22 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-2	MSE	-	EXPRESSION TAG	UNP Q7CUX4
A	-1	SER	-	EXPRESSION TAG	UNP Q7CUX4
A	0	LEU	-	EXPRESSION TAG	UNP Q7CUX4
A	598	GLU	-	EXPRESSION TAG	UNP Q7CUX4
A	599	GLY	-	EXPRESSION TAG	UNP Q7CUX4
A	600	HIS	-	EXPRESSION TAG	UNP Q7CUX4
A	601	HIS	-	EXPRESSION TAG	UNP Q7CUX4
A	602	HIS	-	EXPRESSION TAG	UNP Q7CUX4
A	603	HIS	-	EXPRESSION TAG	UNP Q7CUX4
A	604	HIS	-	EXPRESSION TAG	UNP Q7CUX4
A	605	HIS	-	EXPRESSION TAG	UNP Q7CUX4
B	-2	MSE	-	EXPRESSION TAG	UNP Q7CUX4
B	-1	SER	-	EXPRESSION TAG	UNP Q7CUX4
B	0	LEU	-	EXPRESSION TAG	UNP Q7CUX4
B	598	GLU	-	EXPRESSION TAG	UNP Q7CUX4
B	599	GLY	-	EXPRESSION TAG	UNP Q7CUX4
B	600	HIS	-	EXPRESSION TAG	UNP Q7CUX4
B	601	HIS	-	EXPRESSION TAG	UNP Q7CUX4
B	602	HIS	-	EXPRESSION TAG	UNP Q7CUX4
B	603	HIS	-	EXPRESSION TAG	UNP Q7CUX4
B	604	HIS	-	EXPRESSION TAG	UNP Q7CUX4
B	605	HIS	-	EXPRESSION TAG	UNP Q7CUX4

- Molecule 2 is FE (III) ION (three-letter code: FE) (formula: Fe).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	3	Total Fe 3 3	0	0
2	A	3	Total Fe 3 3	0	0

- Molecule 3 is water.

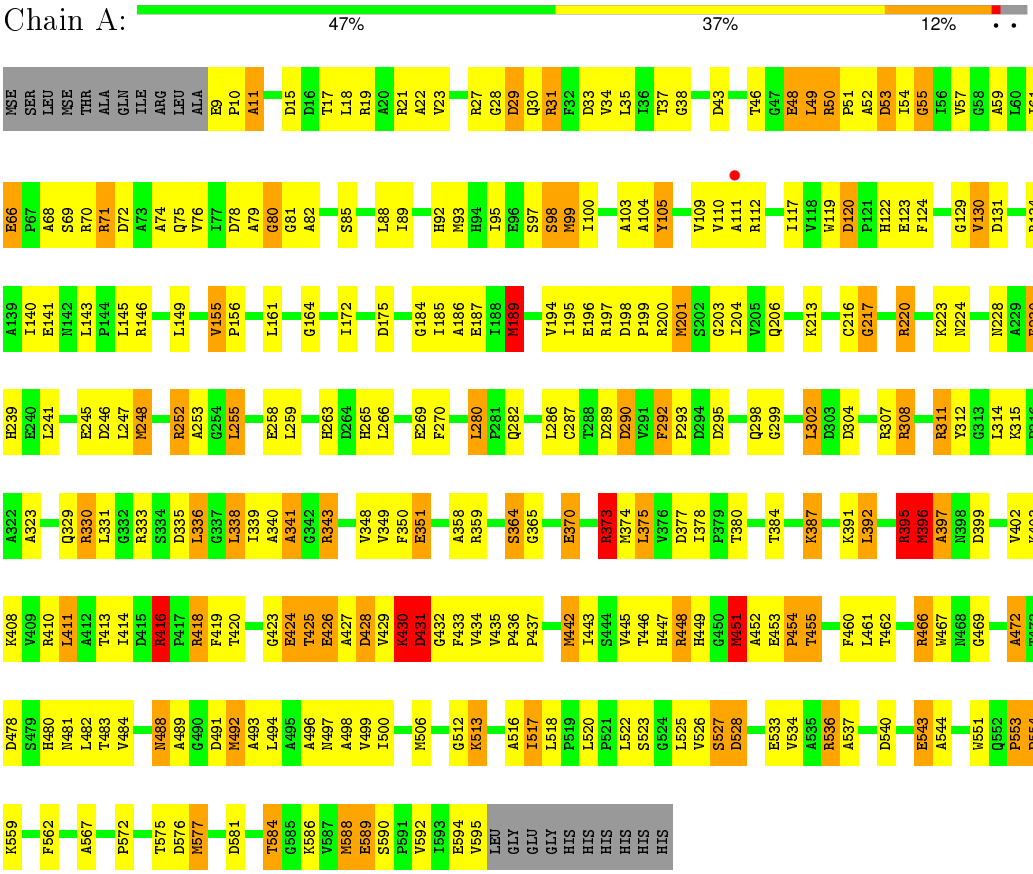
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	23	Total O 23 23	0	0
3	B	31	Total O 31 31	0	0

3 Residue-property plots (i)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of errors displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. Residues are color-coded according to the number of geometric quality criteria for which they contain at least one outlier: green = 0, yellow = 1, orange = 2 and red = 3 or more. A red dot above a residue indicates a poor fit to the electron density (RSRZ > 2). Stretches of 2 or more consecutive residues without any outlier are shown as a green connector. Residues present in the sample, but not in the model, are shown in grey.

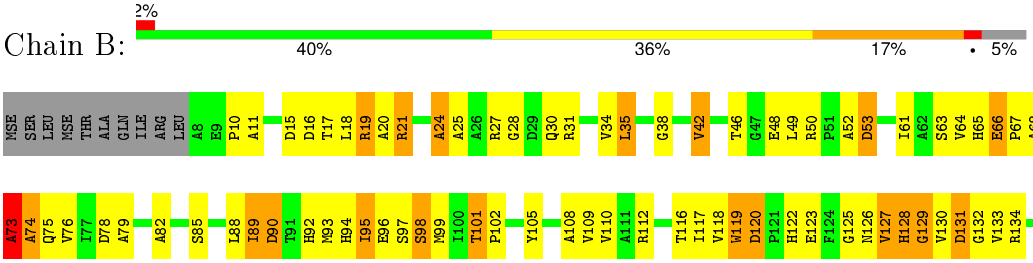
- Molecule 1: Adenine deaminase 2

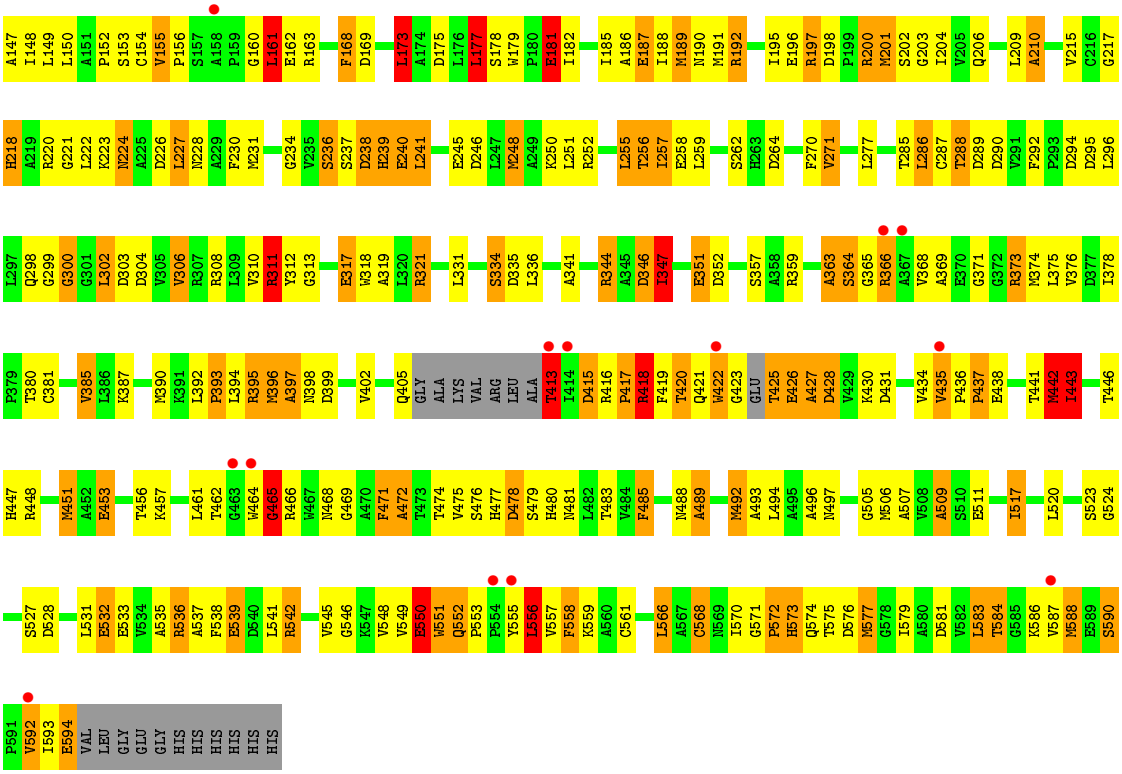
Chain A:



- Molecule 1: Adenine deaminase 2

Chain B:




A147 I148 L149 L150 A151 P152 S153 C154 V155 P156 S157 A158 P159 G160 L161 E162 R163 F168 D169 L173 A174 D175 L176 L177 S178 W179 P180 E181 I182 I185 A186 E187 I188 M189 N190 M191 R192 I195 E196 R197 D198 P199 R200 M201 S202 G203 I204 V205 Q206 L209 A210 V215 C216 G217
H218 A219 R220 G221 L222 K223 N224 A225 D226 L227 N228 A229 F230 M231 G234 V235 S236 S237 D238 H239 E240 L241 E245 D246 L247 M248 A249 K250 L251 R252 L255 T256 I257 E258 L259 S262 H263 D264 F270 V271 L277 T285 L286 C287 T288 D289 D290 V291 F292 P293 D294 D295 L296
L297 Q298 G299 G300 G301 L302 D303 D304 V305 V306 R307 R308 L309 V310 R311 Y312 G313 E317 W318 A319 L320 R321 L331 S334 D335 L336 A341 R344 A345 D346 I347 E351 D352 S357 A358 R359 A363 S364 G365 R366 A367 V368 A369 E370 G371 G372 R373 M374 L375 V376 D377 I378
P379 T380 C381 V385 L386 K387 M390 K391 L392 P393 L394 R395 M396 A397 N398 D399 V402 Q405 GLY ALA LYS VAL ARG LEU ALA T413 I414 D415 R416 P417 R418 F419 T420 Q421 W422 G423 GLU T425 E426 A427 D428 V429 K430 D431 V434 V435 P436 P437 E438 T441 M442 I443 T446
H447 R448 M451 A452 E453 T456 K457 L461 T462 G463 W464 G465 R466 W467 N468 G469 A470 F471 A472 T473 T474 V475 S476 H477 D478 S479 H480 N481 L482 T483 V484 F485 N488 A489 M492 A493 L494 A495 A496 N497 G505 M506 A507 V508 A509 S510 E511 I517 L520 S523 G524
S527 D528 L531 E532 E533 V534 A535 R536 A537 F538 E539 D540 L541 R542 V545 G546 K547 V548 V549 E550 W551 Q552 P553 P554 Y555 L556 V557 F558 K559 A560 C561 L566 A567 C568 N569 I570 G571 P572 H573 Q574 T575 D576 M577 G578 I579 A580 D581 V582 L583 T584 G585 K586 V587 M588 E589 S590
P591 V592 I593 E594 VAL LEU GLY GLU GLY HIS HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	61.66Å 131.84Å 69.63Å 90.00° 97.04° 90.00°	Depositor
Resolution (Å)	48.89 – 2.63 48.88 – 2.63	Depositor EDS
% Data completeness (in resolution range)	99.6 (48.89-2.63) 99.6 (48.88-2.63)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	2.06 (at 2.61Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.185 , 0.287 0.185 , 0.284	Depositor DCC
R_{free} test set	1663 reflections (5.35%)	DCC
Wilson B-factor (Å^2)	54.5	Xtriage
Anisotropy	0.089	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 41.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.49$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 32786 reflections	Xtriage
F_o,F_c correlation	0.95	EDS
Total number of atoms	8674	wwPDB-VP
Average B, all atoms (Å^2)	51.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.14% 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, 0.375 respectively for untwinned datasets, and 0.333, 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: FE

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	1.83	79/4402 (1.8%)	1.65	67/5964 (1.1%)
1	B	1.79	73/4340 (1.7%)	1.67	75/5878 (1.3%)
All	All	1.81	152/8742 (1.7%)	1.66	142/11842 (1.2%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	A	0	5
1	B	0	9
All	All	0	14

The worst 5 of 152 bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	A	451	MSE	SE-CE	10.07	2.54	1.95
1	A	189	MSE	SE-CE	9.45	2.51	1.95
1	B	496	ALA	CA-CB	-8.90	1.33	1.52
1	B	448	ARG	CZ-NH1	8.55	1.44	1.33
1	B	133	VAL	CB-CG2	-8.37	1.35	1.52

The worst 5 of 142 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	21	ARG	NE-CZ-NH2	-13.26	113.67	120.30
1	A	50	ARG	NE-CZ-NH2	12.07	126.34	120.30
1	B	311	ARG	NE-CZ-NH2	-11.54	114.53	120.30
1	A	241	LEU	CB-CG-CD1	-11.32	91.75	111.00

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	536	ARG	NE-CZ-NH2	-11.26	114.67	120.30

There are no chirality outliers.

5 of 14 planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	217	GLY	Peptide
1	A	396	MSE	Peptide
1	A	402	VAL	Peptide
1	A	451	MSE	Peptide
1	A	48	GLU	Peptide

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	4337	0	4344	198	0
1	B	4277	0	4273	285	0
2	A	3	0	0	0	0
2	B	3	0	0	0	0
3	A	23	0	0	4	0
3	B	31	0	0	4	0
All	All	8674	0	8617	471	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 27.

The worst 5 of 471 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:393:PRO:CB	1:B:393:PRO:CG	1.76	1.45
1:A:231:MSE:SE	1:A:231:MSE:CE	2.15	1.44
1:A:201:MSE:SE	1:A:201:MSE:CG	2.16	1.42
1:B:248:MSE:SE	1:B:248:MSE:CE	2.17	1.42
1:B:588:MSE:CE	1:B:588:MSE:SE	2.18	1.42

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	585/608 (96%)	516 (88%)	54 (9%)	15 (3%)	7 11
1	B	573/608 (94%)	474 (83%)	72 (13%)	27 (5%)	3 3
All	All	1158/1216 (95%)	990 (86%)	126 (11%)	42 (4%)	4 6

5 of 42 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	111	ALA
1	A	397	ALA
1	A	430	LYS
1	A	431	ASP
1	A	588	MSE

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	445/444 (100%)	401 (90%)	44 (10%)	10 17
1	B	439/444 (99%)	380 (87%)	59 (13%)	5 8
All	All	884/888 (100%)	781 (88%)	103 (12%)	7 11

5 of 103 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	95	ILE
1	B	181	GLU
1	B	555	TYR
1	B	97	SER
1	B	161	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 20 such sidechains are listed below:

Mol	Chain	Res	Type
1	B	65	HIS
1	B	126	ASN
1	B	447	HIS
1	A	481	ASN
1	A	488	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

Of 6 ligands modelled in this entry, 6 are monoatomic - leaving 0 for Mogul analysis.

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

No monomer is involved in short contacts.

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	571/608 (93%)	-0.55	1 (0%)	95	95	28, 46, 68, 80	0
1	B	563/608 (92%)	-0.33	13 (2%)	64	57	31, 50, 89, 100	0
All	All	1134/1216 (93%)	-0.44	14 (1%)	81	78	28, 48, 83, 100	0

The worst 5 of 14 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	435	VAL	3.5
1	B	413	THR	3.0
1	B	414	ILE	2.8
1	B	554	PRO	2.8
1	B	366	ARG	2.8

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

There are no non-standard protein/DNA/RNA residues in this entry.

6.3 Carbohydrates

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, 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	LLDF	B-factors(Å^2)	Q<0.9
2	FE	B	608	1/1	0.98	0.15	1.47	59,59,59,59	0
2	FE	A	607	1/1	0.97	0.11	0.16	45,45,45,45	0
2	FE	B	606	1/1	0.99	0.08	-2.12	52,52,52,52	0
2	FE	A	606	1/1	0.98	0.07	-2.46	49,49,49,49	0
2	FE	A	608	1/1	0.98	0.10	-	50,50,50,50	0
2	FE	B	607	1/1	0.99	0.11	-	64,64,64,64	0

6.5 Other polymers (i)

There are no such residues in this entry.