



Full wwPDB X-ray Structure Validation Report

Feb 27, 2014 – 01:33 AM GMT

PDB ID : 4HST
Title : Crystal structure of a double mutant of a class III engineered cephalosporin acylase
Authors : Vrielink, A.; Golden, E.; Patterson, R.; Tie, W.J.; Anandan, A.; Flematti, G.; Molla, G.; Rosini, E.; Pollegioni, L.
Deposited on : 2012-10-30
Resolution : 1.57 Å(reported)

This is a full wwPDB validation 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>

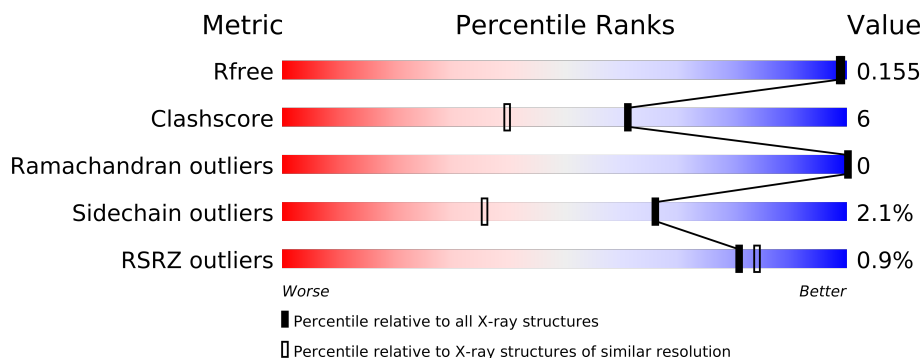
The following versions of software and data (see [references](#)) were used in the production of this report:

MolProbity : 4.02b-467
Mogul : 1.15 2013
Xtriage (Phenix) : dev-1323
EDS : stable22639
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) : stable22683

1 Overall quality at a glance



The reported resolution of this entry is 1.57 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	2778 (1.60-1.56)
Clashscore	79885	3207 (1.60-1.56)
Ramachandran outliers	78287	3107 (1.60-1.56)
Sidechain outliers	78261	3104 (1.60-1.56)
RSRZ outliers	66119	2778 (1.60-1.56)

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	229	
2	B	543	

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6863 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.

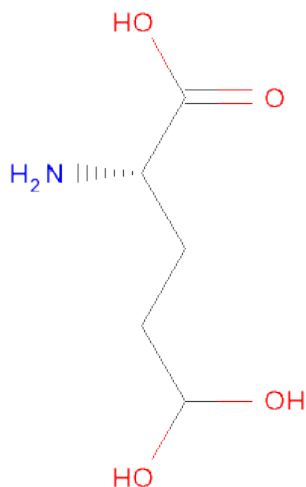
- Molecule 1 is a protein called glutaryl-7-aminocephalosporanicacid acylase alpha chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	216	1686	1082	300	293	11	0	9	0

- Molecule 2 is a protein called glutaryl-7-aminocephalosporanicacid acylase beta chain.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
2	B	535	4191	2647	765	763	16	0	17	0

- Molecule 3 is 5,5-DIHYDROXY-L-NORVALINE (three-letter code: GLJ) (formula: $C_5H_{11}NO_4$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
3	B	1	10	5	1	4	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
4	A	290	Total 290	O 290	0	0
4	B	686	Total 686	O 686	0	0

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.

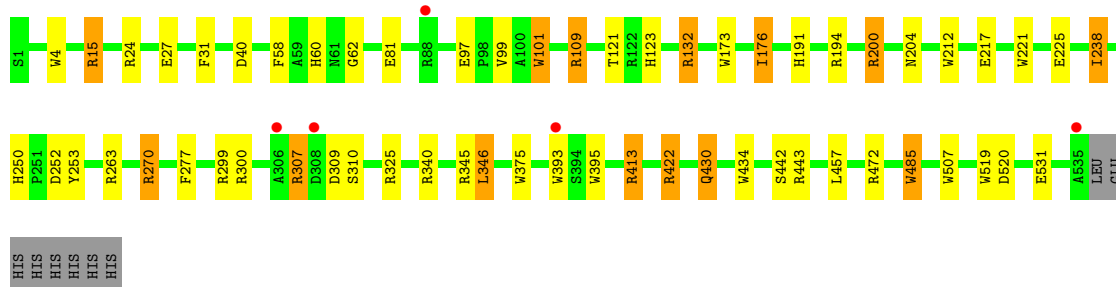
- Molecule 1: glutaryl-7-aminocephalosporanicacid acylase alpha chain

Chain A: 



- Molecule 2: glutaryl-7-aminocephalosporanicacid acylase beta chain

Chain B: 



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	68.39Å 77.84Å 191.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	96.00 – 1.57 64.42 – 1.57	Depositor EDS
% Data completeness (in resolution range)	93.4 (96.00-1.57) 93.4 (64.42-1.57)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	4.51 (at 1.57Å)	Xtriage
Refinement program	REFMAC 5.6.0117	Depositor
R, R_{free}	0.119 , 0.162 0.114 , 0.155	Depositor DCC
R_{free} test set	6717 reflections (5.02%)	DCC
Wilson B-factor (Å ²)	13.9	Xtriage
Anisotropy	0.204	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.37 , 41.8	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning	$\langle L \rangle = 0.47$, $\langle L^2 \rangle = 0.30$	Xtriage
Outliers	0 of 133826 reflections	Xtriage
F_o, F_c correlation	0.97	EDS
Total number of atoms	6863	wwPDB-VP
Average B, all atoms (Å ²)	15.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.04% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: GLJ

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	1.39	7/1753 (0.4%)	1.19	10/2375 (0.4%)
2	B	1.39	22/4366 (0.5%)	1.27	32/5968 (0.5%)
All	All	1.39	29/6119 (0.5%)	1.25	42/8343 (0.5%)

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
2	B	3	0

All (29) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	101	TRP	CD2-CE2	11.66	1.55	1.41
2	B	225	GLU	CD-OE1	8.04	1.34	1.25
2	B	443	ARG	CZ-NH1	7.82	1.43	1.33
2	B	4	TRP	CG-CD1	7.46	1.47	1.36
2	B	485	TRP	CE3-CZ3	6.62	1.49	1.38
2	B	485	TRP	CD2-CE2	6.37	1.49	1.41
2	B	212	TRP	CD2-CE2	6.25	1.48	1.41
1	A	113	GLU	CD-OE1	6.14	1.32	1.25
1	A	91	ASP	CG-OD1	6.11	1.39	1.25
1	A	191	TYR	CG-CD1	5.95	1.46	1.39
1	A	138	GLU	CD-OE2	5.92	1.32	1.25
2	B	375	TRP	CD2-CE2	5.90	1.48	1.41
2	B	253	TYR	CG-CD1	5.77	1.46	1.39
2	B	531	GLU	CD-OE2	5.71	1.31	1.25
1	A	71[A]	THR	CB-CG2	-5.70	1.33	1.52
1	A	71[B]	THR	CB-CG2	-5.70	1.33	1.52

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	430	GLN	CD-OE1	5.69	1.36	1.24
2	B	97	GLU	CD-OE1	-5.69	1.19	1.25
2	B	277	PHE	CG-CD2	5.61	1.47	1.38
2	B	221	TRP	CD2-CE2	5.55	1.48	1.41
2	B	519	TRP	CG-CD1	5.44	1.44	1.36
1	A	82	TRP	CZ3-CH2	5.29	1.48	1.40
2	B	434	TRP	CD2-CE2	5.28	1.47	1.41
2	B	194	ARG	NE-CZ	5.26	1.39	1.33
2	B	395	TRP	CD2-CE2	5.23	1.47	1.41
2	B	442[A]	SER	CB-OG	-5.19	1.35	1.42
2	B	442[B]	SER	CB-OG	-5.19	1.35	1.42
2	B	15	ARG	CD-NE	5.18	1.55	1.46
2	B	507	TRP	CG-CD1	5.04	1.43	1.36

All (42) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	109[A]	ARG	NE-CZ-NH1	13.53	127.06	120.30
2	B	109[B]	ARG	NE-CZ-NH1	13.53	127.06	120.30
2	B	443	ARG	NE-CZ-NH2	-13.07	113.77	120.30
1	A	91	ASP	CB-CG-OD2	-12.24	107.28	118.30
2	B	109[A]	ARG	NE-CZ-NH2	-11.78	114.41	120.30
2	B	109[B]	ARG	NE-CZ-NH2	-11.78	114.41	120.30
2	B	307	ARG	NE-CZ-NH1	11.65	126.12	120.30
2	B	443	ARG	NE-CZ-NH1	11.59	126.10	120.30
2	B	422	ARG	NE-CZ-NH1	11.43	126.01	120.30
1	A	91	ASP	CB-CG-OD1	9.79	127.11	118.30
2	B	307	ARG	NE-CZ-NH2	-8.66	115.97	120.30
2	B	422	ARG	NE-CZ-NH2	-8.64	115.98	120.30
2	B	238[A]	ILE	CG1-CB-CG2	8.61	130.33	111.40
2	B	238[B]	ILE	CG1-CB-CG2	8.61	130.33	111.40
1	A	159	ARG	NE-CZ-NH1	8.01	124.30	120.30
2	B	345	ARG	NE-CZ-NH1	7.47	124.04	120.30
2	B	472	ARG	NE-CZ-NH2	-7.43	116.59	120.30
1	A	119	ARG	NE-CZ-NH2	-7.21	116.70	120.30
2	B	270	ARG	NE-CZ-NH1	-7.17	116.72	120.30
2	B	252	ASP	CB-CG-OD1	-7.06	111.95	118.30
1	A	35	ARG	NE-CZ-NH2	-6.70	116.95	120.30
1	A	165	MET	CG-SD-CE	-6.62	89.61	100.20
1	A	36	ARG	NE-CZ-NH2	-6.57	117.02	120.30
2	B	300	ARG	NE-CZ-NH1	6.49	123.54	120.30
2	B	200	ARG	NE-CZ-NH2	-6.40	117.10	120.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	B	442[A]	SER	N-CA-CB	-6.39	100.92	110.50
2	B	442[B]	SER	N-CA-CB	-6.39	100.92	110.50
2	B	176[A]	ILE	CB-CA-C	-6.10	99.40	111.60
2	B	176[B]	ILE	CB-CA-C	-6.10	99.40	111.60
1	A	159	ARG	NE-CZ-NH2	-5.96	117.32	120.30
1	A	33	ARG	NE-CZ-NH2	-5.85	117.37	120.30
2	B	24[A]	ARG	NE-CZ-NH1	5.85	123.22	120.30
2	B	24[B]	ARG	NE-CZ-NH1	5.85	123.22	120.30
2	B	132	ARG	NE-CZ-NH1	5.83	123.21	120.30
2	B	340	ARG	NE-CZ-NH2	-5.53	117.53	120.30
2	B	58	PHE	CB-CG-CD2	-5.46	116.98	120.80
1	A	36	ARG	NE-CZ-NH1	5.24	122.92	120.30
2	B	413	ARG	NE-CZ-NH1	5.24	122.92	120.30
2	B	263	ARG	NE-CZ-NH2	-5.18	117.71	120.30
2	B	346	LEU	CB-CG-CD1	-5.09	102.35	111.00
2	B	309	ASP	N-CA-CB	-5.06	101.49	110.60
2	B	300	ARG	NE-CZ-NH2	-5.03	117.79	120.30

All (3) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	B	121	THR	CB
2	B	138	THR	CB
2	B	238[A]	ILE	CB

There are no planarity outliers.

5.2 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 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	1686	0	2	5	0
2	B	4191	0	16	28	0
3	B	10	0	0	2	0
4	A	290	0	0	5	1

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	686	0	0	16	1
All	All	6863	0	18	34	1

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 6.

All (34) close contacts within the same asymmetric unit are listed below.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
2:B:109[B]:ARG:NH1	2:B:430:GLN:NE2	2.10	0.99
2:B:413:ARG:NH2	4:B:1116:HOH:O	2.01	0.93
2:B:15:ARG:NH2	4:B:1055:HOH:O	2.02	0.91
2:B:109[B]:ARG:CZ	2:B:430:GLN:NE2	2.36	0.88
1:A:142:LEU:CD2	2:B:109[B]:ARG:NH1	2.37	0.86
4:A:358:HOH:O	2:B:176[A]:ILE:CD1	2.24	0.84
2:B:99[B]:VAL:CG1	2:B:101:TRP:CD1	2.61	0.84
2:B:520:ASP:CB	4:B:795:HOH:O	2.27	0.81
2:B:299:ARG:NH2	4:B:1119:HOH:O	2.13	0.81
1:A:226:LYS:NZ	4:A:521:HOH:O	2.29	0.66
2:B:132:ARG:NH1	4:B:1326:HOH:O	2.35	0.60
2:B:299:ARG:NH1	4:B:779:HOH:O	2.33	0.60
2:B:430:GLN:NE2	4:B:1195:HOH:O	2.39	0.56
2:B:109[B]:ARG:NH2	4:B:854:HOH:O	2.39	0.55
2:B:40:ASP:CB	4:B:1274:HOH:O	2.56	0.54
2:B:173:TRP:NE1	4:B:1376:HOH:O	2.34	0.53
2:B:346:LEU:CD2	2:B:393:TRP:CZ3	2.92	0.52
2:B:109[A]:ARG:NH2	4:B:917:HOH:O	2.42	0.52
2:B:99[B]:VAL:CG1	2:B:101:TRP:NE1	2.73	0.51
2:B:325:ARG:NH2	4:B:978:HOH:O	2.46	0.49
2:B:27[B]:GLU:OE2	2:B:31:PHE:CE1	2.67	0.48
2:B:123:HIS:CE1	2:B:217:GLU:OE2	2.67	0.48
2:B:250:HIS:CE1	4:B:794:HOH:O	2.68	0.46
2:B:81:GLU:OE2	2:B:121:THR:CG2	2.63	0.46
3:B:601:GLJ:CA	3:B:601:GLJ:OE1	2.63	0.46
1:A:165:MET:N	4:A:512:HOH:O	2.51	0.43
2:B:325:ARG:CZ	4:B:978:HOH:O	2.68	0.42
1:A:20[C]:SER:OG	4:A:351:HOH:O	2.22	0.42
2:B:81:GLU:OE2	2:B:123:HIS:CD2	2.72	0.42
1:A:160:ARG:NH1	4:A:590:HOH:O	2.53	0.42
2:B:60:HIS:CD2	2:B:62:GLY:N	2.88	0.41
2:B:176[A]:ILE:CD1	4:B:1378:HOH:O	2.68	0.41
2:B:270:ARG:NH2	4:B:1240:HOH:O	2.53	0.41

All (1) symmetry-related close contacts are listed below. The label for Atom-2 includes the symmetry operator and encoded unit-cell translations to be applied.

Atom-1	Atom-2	Distance(Å)	Clash(Å)
4:A:578:HOH:O	4:B:835:HOH:O[3_656]	1.14	1.06

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	224/229 (98%)	221 (99%)	3 (1%)	0	100	100
2	B	550/543 (101%)	537 (98%)	13 (2%)	0	100	100
All	All	774/772 (100%)	758 (98%)	16 (2%)	0	100	100

There are no Ramachandran outliers to report.

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	171/171 (100%)	168 (98%)	3 (2%)	71	43
2	B	430/421 (102%)	420 (98%)	10 (2%)	63	31
All	All	601/592 (102%)	588 (98%)	13 (2%)	66	33

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	108	GLU
1	A	165	MET
1	A	226	LYS
2	B	191	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
2	B	200	ARG
2	B	204	ASN
2	B	238[A]	ILE
2	B	238[B]	ILE
2	B	307	ARG
2	B	310	SER
2	B	422	ARG
2	B	457	LEU
2	B	485	TRP

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. There are no such sidechains identified.

5.3.3 RNA ⓘ

There are no RNA chains 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 ⓘ

1 ligand is 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	$\# Z > 2$
3	GLJ	B	601	-	9,9,9	3.23	3 (33%)	11,11,11	2.94	7 (63%)

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
3	GLJ	B	601	-	-	0/9/9/9	0/0/0/0

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	B	601	GLJ	CA-C	-8.39	1.22	1.53
3	B	601	GLJ	CG-CB	-3.40	1.42	1.53
3	B	601	GLJ	CB-CA	-2.06	1.48	1.54

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	B	601	GLJ	CG-CB-CA	-5.26	94.17	113.95
3	B	601	GLJ	OXT-C-O	4.35	133.90	124.07
3	B	601	GLJ	CB-CA-N	3.98	119.86	110.14
3	B	601	GLJ	C-CA-N	3.40	114.99	109.36
3	B	601	GLJ	O-C-CA	-2.82	110.23	118.36
3	B	601	GLJ	OXT-C-CA	-2.43	111.42	116.88
3	B	601	GLJ	CB-CA-C	-2.21	105.75	110.98

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

5.7 Other polymers ⓘ

There are no such residues in this entry.

5.8 Polymer linkage issues

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, 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	216/229 (94%)	-0.43	2 (0%) 81 84	4, 10, 22, 42	0
2	B	535/543 (98%)	-0.27	5 (0%) 81 84	4, 9, 22, 40	0
All	All	751/772 (97%)	-0.31	7 (0%) 81 84	4, 10, 22, 42	0

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
2	B	393	TRP	3.7
2	B	308	ASP	3.0
1	A	214	LEU	2.8
2	B	306	ALA	2.5
2	B	88	ARG	2.3
1	A	218	ARG	2.1
2	B	535	ALA	2.0

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, 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(\AA^2)	Q<0.9
3	GLJ	B	601	10/10	0.12	1.86	6,14,29,35	0

6.5 Other polymers ⓘ

There are no such residues in this entry.