Full wwPDB X-ray Structure Validation Report ⓘ

Aug 23, 2023 – 01:35 AM EDT

PDB ID	:	3BOA
Title	:	Crystal structure of yeast protein disulfide isomerase.
Authors	:	Tian, G.; Lennarz, W.J.; Schindelin, H.
Deposited on	:	2007-12-17
Resolution	:	3.70 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*
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:

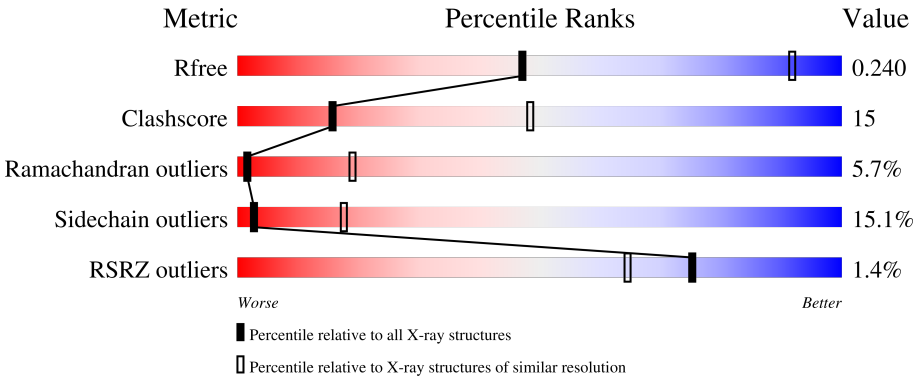
MolProbity	:	4.02b-467
Xtriage (Phenix)	:	1.13
EDS	:	2.35
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35

1 Overall quality at a glance (i)

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

The reported resolution of this entry is 3.70 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	1049 (3.88-3.52)
Clashscore	141614	1027 (3.86-3.54)
Ramachandran outliers	138981	1069 (3.88-3.52)
Sidechain outliers	138945	1065 (3.88-3.52)
RSRZ outliers	127900	1578 (3.90-3.50)

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 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	504	% 56% 34% 7% .

2 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 3864 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 Protein disulfide-isomerase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	490	3864	2464	611	775	14	0	0	0

There are 4 discrepancies between the modelled and reference sequences:

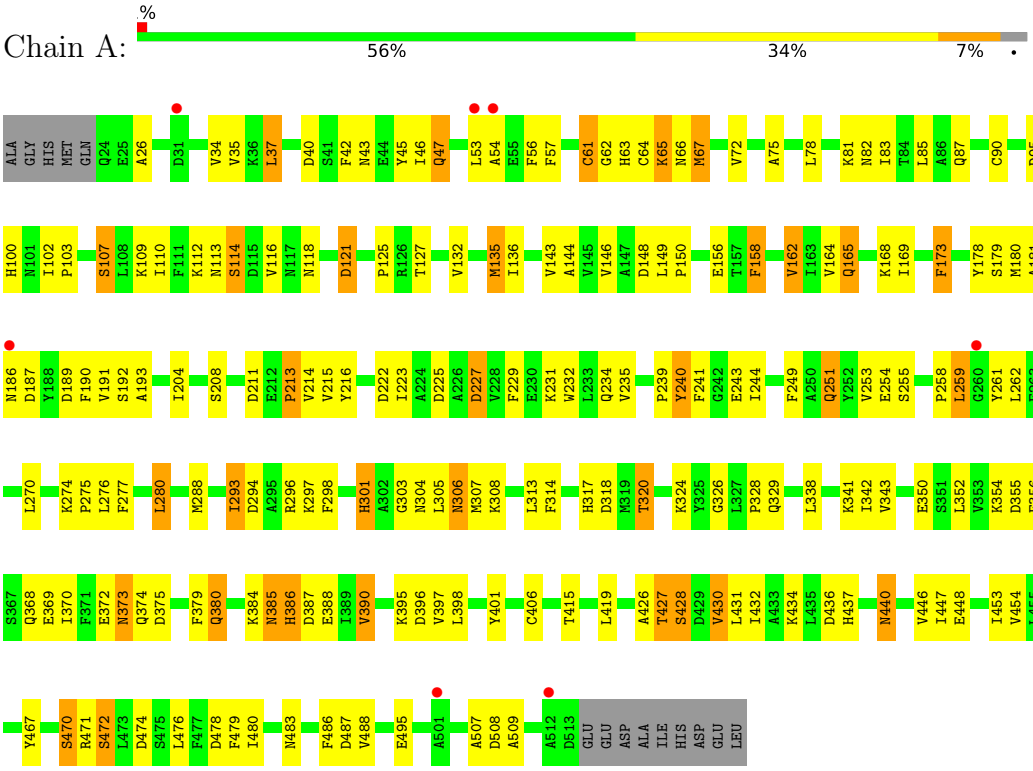
Chain	Residue	Modelled	Actual	Comment	Reference
A	19	ALA	-	expression tag	UNP P17967
A	20	GLY	-	expression tag	UNP P17967
A	21	HIS	-	expression tag	UNP P17967
A	22	MET	-	expression tag	UNP P17967

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: Protein disulfide-isomerase

Chain A:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, α, β, γ	116.92Å 123.15Å 75.72Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.81 – 3.70 19.81 – 3.70	Depositor EDS
% Data completeness (in resolution range)	100.0 (19.81-3.70) 100.0 (19.81-3.70)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.09	Depositor
$< I/\sigma(I) >$ [1]	2.08 (at 3.71Å)	Xtriage
Refinement program	REFMAC 5.2.0005	Depositor
R, R_{free}	0.239 , 0.340 0.246 , 0.240	Depositor DCC
R_{free} test set	279 reflections (4.61%)	wwPDB-VP
Wilson B-factor (Å^2)	165.5	Xtriage
Anisotropy	0.039	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.28 , 201.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	0.037 for -k,-h,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	3864	wwPDB-VP
Average B, all atoms (Å^2)	175.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.56% 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

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	0.46	0/3951	0.58	0/5363

There are no bond length outliers.

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	3864	0	3689	111	0
All	All	3864	0	3689	111	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 15.

All (111) 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:A:78:LEU:HD11	1:A:132:VAL:HG23	1.46	0.95
1:A:396:ASP:HB2	1:A:430:VAL:HG22	1.58	0.83
1:A:397:VAL:HG12	1:A:431:LEU:HB3	1.67	0.75
1:A:156:GLU:O	1:A:158:PHE:CZ	2.42	0.73
1:A:149:LEU:HD21	1:A:164:VAL:HG21	1.73	0.70

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:368:GLN:HE21	1:A:440:ASN:ND2	1.88	0.69
1:A:396:ASP:O	1:A:430:VAL:HG13	1.93	0.69
1:A:63:HIS:O	1:A:67:MET:HG3	1.93	0.69
1:A:368:GLN:HE21	1:A:440:ASN:HD22	1.42	0.67
1:A:307:MET:HG2	1:A:313:LEU:HD11	1.80	0.64
1:A:426:ALA:O	1:A:427:THR:C	2.36	0.64
1:A:436:ASP:O	1:A:440:ASN:HB2	1.98	0.62
1:A:307:MET:SD	1:A:326:GLY:HA3	2.42	0.59
1:A:318:ASP:OD2	1:A:320:THR:OG1	2.19	0.59
1:A:244:ILE:O	1:A:244:ILE:HG23	2.03	0.59
1:A:264:TYR:CD2	1:A:270:LEU:HD13	2.37	0.58
1:A:234:GLN:HE21	1:A:254:GLU:C	2.07	0.58
1:A:390:VAL:HA	1:A:397:VAL:HG21	1.85	0.58
1:A:37:LEU:HD11	1:A:45:TYR:HB3	1.83	0.58
1:A:78:LEU:HD11	1:A:132:VAL:CG2	2.26	0.58
1:A:144:ALA:O	1:A:191:VAL:HG13	2.03	0.57
1:A:259:LEU:HD11	1:A:261:TYR:CZ	2.39	0.57
1:A:293:ILE:HD11	1:A:298:PHE:HD1	1.69	0.57
1:A:149:LEU:N	1:A:150:PRO:HD2	2.20	0.56
1:A:507:ALA:O	1:A:509:ALA:N	2.38	0.56
1:A:342:ILE:HD12	1:A:342:ILE:N	2.22	0.55
1:A:274:LYS:N	1:A:275:PRO:HD2	2.22	0.55
1:A:56:PHE:CD2	1:A:102:ILE:HD11	2.43	0.54
1:A:164:VAL:HA	1:A:191:VAL:O	2.08	0.53
1:A:231:LYS:O	1:A:235:VAL:HG23	2.08	0.53
1:A:26:ALA:HB2	1:A:65:LYS:HG2	1.91	0.52
1:A:467:TYR:OH	1:A:470:SER:O	2.19	0.52
1:A:472:SER:O	1:A:476:LEU:HD12	2.09	0.52
1:A:185:PHE:C	1:A:187:ASP:H	2.13	0.51
1:A:301:HIS:O	1:A:304:ASN:N	2.43	0.51
1:A:97:CYS:HB3	1:A:102:ILE:HD12	1.92	0.51
1:A:370:ILE:HD11	1:A:384:LYS:HD3	1.91	0.51
1:A:262:LEU:HD22	1:A:277:PHE:CE1	2.46	0.51
1:A:307:MET:HG2	1:A:313:LEU:HD21	1.93	0.51
1:A:43:ASN:OD1	1:A:116:VAL:HG11	2.11	0.51
1:A:308:LYS:N	1:A:329:GLN:OE1	2.44	0.50
1:A:368:GLN:NE2	1:A:440:ASN:HD22	2.10	0.49
1:A:54:ALA:HB3	1:A:109:LYS:O	2.13	0.49
1:A:95:ASP:HA	1:A:98:MET:HG2	1.93	0.49
1:A:293:ILE:HD11	1:A:298:PHE:CD1	2.48	0.49
1:A:135:MET:SD	1:A:135:MET:N	2.85	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:162:VAL:HA	1:A:189:ASP:O	2.13	0.48
1:A:415:THR:O	1:A:419:LEU:N	2.45	0.48
1:A:204:ILE:HD11	1:A:232:TRP:CZ3	2.48	0.48
1:A:430:VAL:HG21	1:A:488:VAL:HG21	1.95	0.48
1:A:97:CYS:O	1:A:100:HIS:O	2.32	0.47
1:A:225:ASP:O	1:A:229:PHE:N	2.43	0.47
1:A:98:MET:N	1:A:98:MET:SD	2.87	0.47
1:A:305:LEU:HA	1:A:324:LYS:HG2	1.96	0.47
1:A:121:ASP:N	1:A:121:ASP:OD1	2.48	0.47
1:A:227:ASP:N	1:A:227:ASP:OD1	2.48	0.46
1:A:398:LEU:HD12	1:A:454:VAL:O	2.14	0.46
1:A:427:THR:O	1:A:428:SER:C	2.54	0.46
1:A:181:ALA:O	1:A:184:HIS:N	2.49	0.46
1:A:156:GLU:O	1:A:158:PHE:CE1	2.68	0.46
1:A:181:ALA:HB2	1:A:190:PHE:CG	2.51	0.46
1:A:307:MET:CG	1:A:313:LEU:HD11	2.46	0.46
1:A:143:VAL:HB	1:A:178:TYR:CE1	2.51	0.46
1:A:75:ALA:HA	1:A:85:LEU:HD21	1.99	0.45
1:A:355:ASP:O	1:A:356:PHE:C	2.55	0.45
1:A:57:PHE:CE2	1:A:87:GLN:HB3	2.52	0.45
1:A:78:LEU:HB3	1:A:83:ILE:HD13	1.99	0.45
1:A:62:GLY:O	1:A:66:ASN:N	2.39	0.45
1:A:158:PHE:CD1	1:A:158:PHE:N	2.85	0.45
1:A:276:LEU:HD11	1:A:350:GLU:OE2	2.17	0.45
1:A:303:GLY:O	1:A:306:ASN:N	2.44	0.44
1:A:258:PRO:O	1:A:288:MET:HB2	2.17	0.44
1:A:81:LYS:HD2	1:A:136:ILE:HD11	1.99	0.44
1:A:370:ILE:HD11	1:A:384:LYS:CD	2.48	0.44
1:A:56:PHE:HB2	1:A:107:SER:HB2	1.99	0.44
1:A:380:GLN:HE21	1:A:380:GLN:N	2.16	0.44
1:A:61:CYS:SG	1:A:64:CYS:N	2.90	0.44
1:A:112:LYS:O	1:A:114:SER:N	2.52	0.43
1:A:366:LYS:O	1:A:440:ASN:HA	2.18	0.43
1:A:165:GLN:O	1:A:193:ALA:O	2.37	0.43
1:A:386:HIS:CE1	1:A:390:VAL:HG21	2.53	0.43
1:A:42:PHE:CE2	1:A:46:ILE:HG13	2.54	0.43
1:A:143:VAL:HG13	1:A:192:SER:HB3	2.00	0.43
1:A:169:ILE:HG23	1:A:173:PHE:HD2	1.83	0.43
1:A:314:PHE:CE2	1:A:352:LEU:HD23	2.54	0.43
1:A:446:VAL:HG22	1:A:447:ILE:O	2.18	0.43
1:A:385:ASN:O	1:A:388:GLU:N	2.52	0.43

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:479:PHE:HD1	1:A:483:ASN:HD22	1.65	0.43
1:A:390:VAL:HG13	1:A:456:TYR:CG	2.54	0.42
1:A:486:PHE:HB2	1:A:488:VAL:HG13	2.01	0.42
1:A:249:PHE:O	1:A:253:VAL:HG23	2.19	0.42
1:A:454:VAL:HB	1:A:456:TYR:CE1	2.55	0.42
1:A:211:ASP:O	1:A:213:PRO:HD3	2.19	0.42
1:A:277:PHE:HA	1:A:280:LEU:HD22	2.01	0.42
1:A:401:TYR:HH	1:A:456:TYR:HH	1.62	0.42
1:A:112:LYS:C	1:A:114:SER:N	2.73	0.41
1:A:390:VAL:HG22	1:A:456:TYR:CD1	2.56	0.41
1:A:47:GLN:HA	1:A:114:SER:O	2.21	0.41
1:A:239:PRO:O	1:A:240:TYR:C	2.59	0.41
1:A:386:HIS:HE1	1:A:390:VAL:HG21	1.85	0.41
1:A:431:LEU:HD12	1:A:432:ILE:H	1.84	0.41
1:A:214:VAL:HG11	1:A:232:TRP:CZ2	2.56	0.41
1:A:239:PRO:O	1:A:241:PHE:N	2.54	0.41
1:A:397:VAL:HA	1:A:431:LEU:O	2.20	0.41
1:A:156:GLU:HB2	1:A:158:PHE:CE2	2.56	0.40
1:A:251:GLN:HA	1:A:254:GLU:HB2	2.03	0.40
1:A:72:VAL:O	1:A:75:ALA:HB3	2.21	0.40
1:A:216:TYR:OH	1:A:223:ILE:HD12	2.22	0.40
1:A:398:LEU:HD11	1:A:453:ILE:CG2	2.52	0.40
1:A:222:ASP:OD1	1:A:223:ILE:N	2.54	0.40
1:A:34:VAL:HG13	1:A:85:LEU:CB	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	488/504 (97%)	359 (74%)	101 (21%)	28 (6%)	1 19

All (28) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	82	ASN
1	A	375	ASP
1	A	103	PRO
1	A	118	ASN
1	A	184	HIS
1	A	186	ASN
1	A	113	ASN
1	A	114	SER
1	A	121	ASP
1	A	240	TYR
1	A	328	PRO
1	A	427	THR
1	A	428	SER
1	A	508	ASP
1	A	125	PRO
1	A	173	PHE
1	A	213	PRO
1	A	215	VAL
1	A	338	LEU
1	A	430	VAL
1	A	208	SER
1	A	255	SER
1	A	293	ILE
1	A	306	ASN
1	A	373	ASN
1	A	457	PRO
1	A	478	ASP
1	A	440	ASN

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	418/429 (97%)	355 (85%)	63 (15%)	3 17

All (63) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	35	VAL
1	A	37	LEU
1	A	40	ASP
1	A	47	GLN
1	A	53	LEU
1	A	61	CYS
1	A	65	LYS
1	A	67	MET
1	A	90	CYS
1	A	98	MET
1	A	107	SER
1	A	110	ILE
1	A	127	THR
1	A	135	MET
1	A	146	VAL
1	A	148	ASP
1	A	158	PHE
1	A	162	VAL
1	A	165	GLN
1	A	168	LYS
1	A	179	SER
1	A	180	MET
1	A	182	ASN
1	A	227	ASP
1	A	243	GLU
1	A	251	GLN
1	A	259	LEU
1	A	266	ASP
1	A	267	GLU
1	A	280	LEU
1	A	294	ASP
1	A	296	ARG
1	A	297	LYS
1	A	301	HIS
1	A	317	HIS
1	A	320	THR
1	A	341	LYS
1	A	343	VAL
1	A	354	LYS
1	A	364	ILE
1	A	369	GLU
1	A	372	GLU
1	A	373	ASN

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	374	GLN
1	A	379	PHE
1	A	380	GLN
1	A	385	ASN
1	A	386	HIS
1	A	387	ASP
1	A	390	VAL
1	A	395	LYS
1	A	406	CYS
1	A	434	LYS
1	A	437	HIS
1	A	448	GLU
1	A	463	GLU
1	A	470	SER
1	A	471	ARG
1	A	472	SER
1	A	474	ASP
1	A	480	ILE
1	A	487	ASP
1	A	495	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (12) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	94	GLN
1	A	165	GLN
1	A	174	ASN
1	A	182	ASN
1	A	184	HIS
1	A	265	ASN
1	A	368	GLN
1	A	373	ASN
1	A	374	GLN
1	A	380	GLN
1	A	437	HIS
1	A	483	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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

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

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^{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	490/504 (97%)	-0.21	7 (1%)	75	64	166, 176, 181, 184	1 (0%)

All (7) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	53	LEU	3.3
1	A	260	GLY	2.7
1	A	501	ALA	2.6
1	A	31	ASP	2.4
1	A	186	ASN	2.2
1	A	54	ALA	2.2
1	A	512	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 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.