Full wwPDB X-ray Structure Validation Report (i)

Feb 1, 2016 – 09:06 AM GMT

PDB ID	:	3H7A
Title	:	CRYSTAL STRUCTURE OF SHORT-CHAIN DEHYDROGENASE FROM Rhodopseudomonas palustris
Authors	:	Patskovsky, Y.; Toro, R.; Morano, C.; Freeman, J.; Miller, S.; Sauder, J.M.; Burley, S.K.; Almo, S.C.; New York SGX Research Center for Structural Genomics (NYSGXRC)
Deposited on	:	2009-04-24
Resolution	:	1.87 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
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:

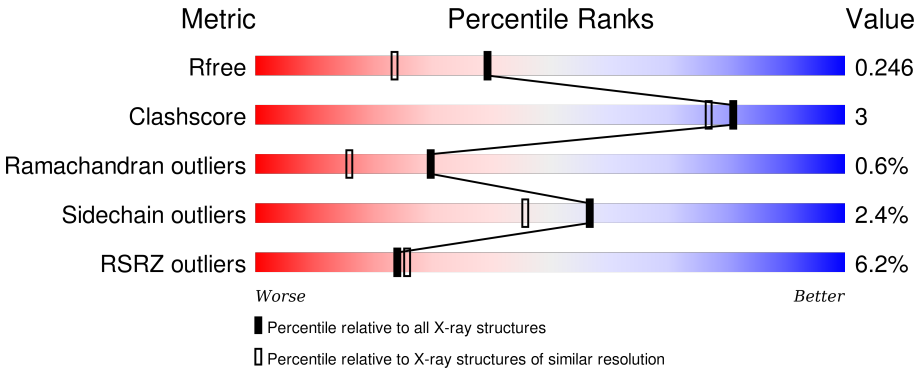
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	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 1.87 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	6965 (1.90-1.86)
Clashscore	102246	7778 (1.90-1.86)
Ramachandran outliers	100387	7691 (1.90-1.86)
Sidechain outliers	100360	7692 (1.90-1.86)
RSRZ outliers	91569	6979 (1.90-1.86)

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	252	2% 79% 5% 16%
1	B	252	3% 75% 8% 15%
1	C	252	10% 75% 9% 16%
1	D	252	7% 74% 10% 16%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 6980 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 short chain dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	212	Total	C	N	O	S	0	2	0
			1625	1041	291	286	7			
1	B	213	Total	C	N	O	S	0	4	0
			1643	1054	293	288	8			
1	C	212	Total	C	N	O	S	0	4	0
			1633	1046	291	288	8			
1	D	212	Total	C	N	O	S	0	3	0
			1633	1046	294	286	7			

There are 40 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	-1	MET	-	expression tag	UNP Q6NDT3
A	0	SER	-	expression tag	UNP Q6NDT3
A	1	LEU	-	expression tag	UNP Q6NDT3
A	244	GLY	-	expression tag	UNP Q6NDT3
A	245	HIS	-	expression tag	UNP Q6NDT3
A	246	HIS	-	expression tag	UNP Q6NDT3
A	247	HIS	-	expression tag	UNP Q6NDT3
A	248	HIS	-	expression tag	UNP Q6NDT3
A	249	HIS	-	expression tag	UNP Q6NDT3
A	250	HIS	-	expression tag	UNP Q6NDT3
B	-1	MET	-	expression tag	UNP Q6NDT3
B	0	SER	-	expression tag	UNP Q6NDT3
B	1	LEU	-	expression tag	UNP Q6NDT3
B	244	GLY	-	expression tag	UNP Q6NDT3
B	245	HIS	-	expression tag	UNP Q6NDT3
B	246	HIS	-	expression tag	UNP Q6NDT3
B	247	HIS	-	expression tag	UNP Q6NDT3
B	248	HIS	-	expression tag	UNP Q6NDT3
B	249	HIS	-	expression tag	UNP Q6NDT3
B	250	HIS	-	expression tag	UNP Q6NDT3
C	-1	MET	-	expression tag	UNP Q6NDT3

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	0	SER	-	expression tag	UNP Q6NDT3
C	1	LEU	-	expression tag	UNP Q6NDT3
C	244	GLY	-	expression tag	UNP Q6NDT3
C	245	HIS	-	expression tag	UNP Q6NDT3
C	246	HIS	-	expression tag	UNP Q6NDT3
C	247	HIS	-	expression tag	UNP Q6NDT3
C	248	HIS	-	expression tag	UNP Q6NDT3
C	249	HIS	-	expression tag	UNP Q6NDT3
C	250	HIS	-	expression tag	UNP Q6NDT3
D	-1	MET	-	expression tag	UNP Q6NDT3
D	0	SER	-	expression tag	UNP Q6NDT3
D	1	LEU	-	expression tag	UNP Q6NDT3
D	244	GLY	-	expression tag	UNP Q6NDT3
D	245	HIS	-	expression tag	UNP Q6NDT3
D	246	HIS	-	expression tag	UNP Q6NDT3
D	247	HIS	-	expression tag	UNP Q6NDT3
D	248	HIS	-	expression tag	UNP Q6NDT3
D	249	HIS	-	expression tag	UNP Q6NDT3
D	250	HIS	-	expression tag	UNP Q6NDT3

- Molecule 2 is UNKNOWN LIGAND (three-letter code: UNL) (formula:).

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
2	D	2	Total	C	O	0	0
			14	8	6		
2	C	1	Total	C	O	0	0
			7	4	3		

- Molecule 3 is water.

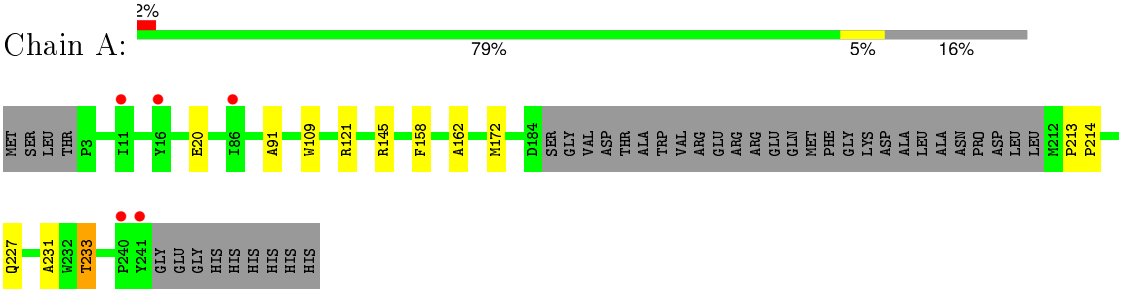
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
3	A	138	Total	O	0	1
			139	139		
3	B	122	Total	O	0	0
			122	122		
3	C	64	Total	O	0	0
			64	64		
3	D	100	Total	O	0	0
			100	100		

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.

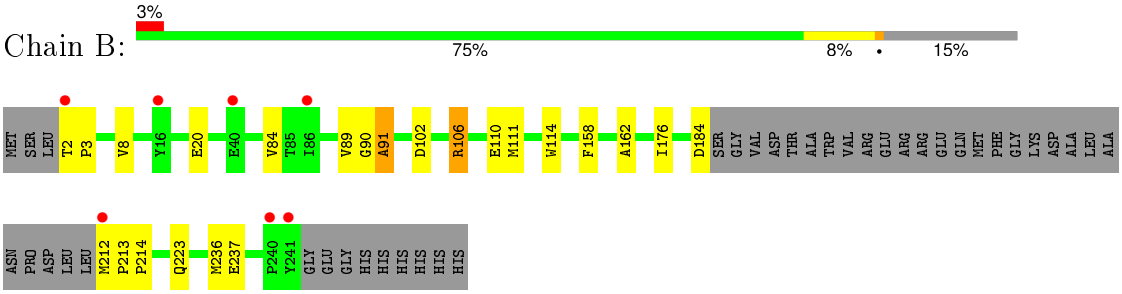
- Molecule 1: short chain dehydrogenase

Chain A:



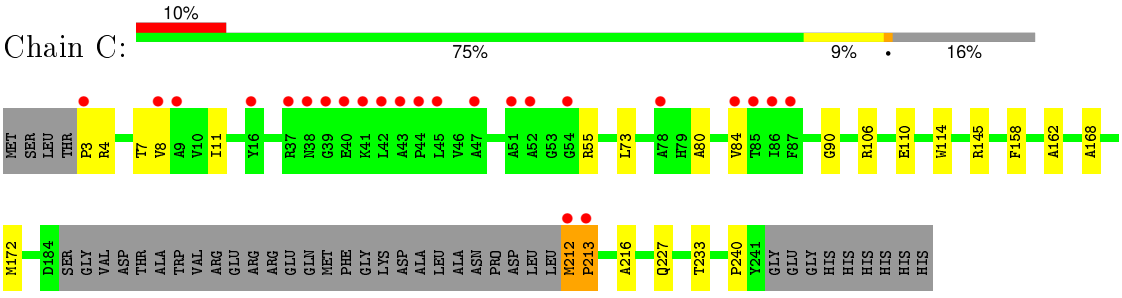
- Molecule 1: short chain dehydrogenase

Chain B:



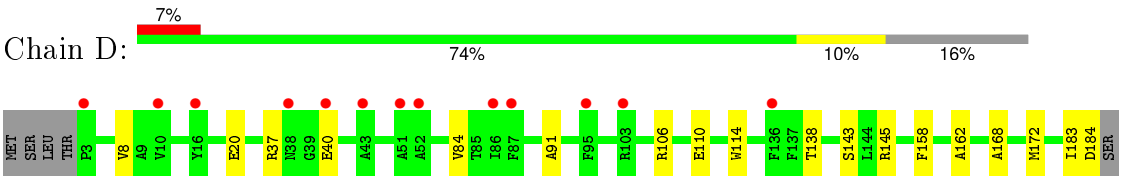
- Molecule 1: short chain dehydrogenase

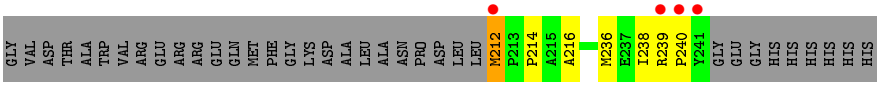
Chain C:



- Molecule 1: short chain dehydrogenase

Chain D:




GLY
VAL
ASP
THR
ALA
TRP
VAL
ARG
GLU
ARG
ARG
GLU
GLN
MET
PHE
GLY
LYS
ASP
ALA
LEU
ALA
ASN
PRO
ASP
LEU
LEU
M212
P213
P214
A215
A216
M236
E237
I238
R239
P240
Y241
GLY
GLU
GLY
HIS
HIS
HIS
HIS
HIS
HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	91.43Å 91.43Å 270.71Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 1.87 30.30 – 1.87	Depositor EDS
% Data completeness (in resolution range)	98.2 (20.00-1.87) 98.2 (30.30-1.87)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	0.13	Depositor
$< I/\sigma(I) >$ [1]	1.44 (at 1.87Å)	Xtriage
Refinement program	REFMAC 5.3.0034	Depositor
R, R_{free}	0.205 , 0.241 0.208 , 0.246	Depositor DCC
R_{free} test set	2830 reflections (3.11%)	DCC
Wilson B-factor (Å^2)	24.0	Xtriage
Anisotropy	0.417	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 63.0	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 94033 reflections	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	6980	wwPDB-VP
Average B, all atoms (Å^2)	34.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 6.22% 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: UNL

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	0.51	0/1668	0.61	0/2255
1	B	0.50	0/1692	0.60	0/2287
1	C	0.47	0/1682	0.60	0/2275
1	D	0.49	0/1679	0.58	0/2269
All	All	0.49	0/6721	0.60	0/9086

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	C	0	2

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (2) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	212	MET	Peptide
1	C	3	PRO	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	1625	0	1624	12	0
1	B	1643	0	1652	16	0
1	C	1633	0	1633	15	0
1	D	1633	0	1637	15	0
2	C	7	0	0	0	0
2	D	14	0	0	0	0
3	A	139	0	0	1	0
3	B	122	0	0	0	0
3	C	64	0	0	0	0
3	D	100	0	0	0	0
All	All	6980	0	6546	43	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 3.

All (43) 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:20:GLU:HG3	1:D:214:PRO:HB2	1.71	0.71
1:A:20:GLU:HG3	1:A:214:PRO:HB2	1.75	0.68
1:A:227:GLN:HE22	1:A:233:THR:HG23	1.59	0.68
1:A:233:THR:HG22	1:D:238:ILE:HG22	1.80	0.63
1:A:121[A]:ARG:NH1	1:B:102:ASP:OD1	2.27	0.60
1:B:223:GLN:HE22	1:C:227:GLN:HE21	1.52	0.58
1:C:212:MET:HE2	1:C:240:PRO:HD3	1.85	0.56
1:B:236:MET:HG2	1:C:233[B]:THR:HG21	1.88	0.56
1:D:8:VAL:HG22	1:D:84:VAL:HB	1.86	0.55
1:D:184:ASP:HB3	1:D:239:ARG:HG2	1.88	0.55
1:C:168:ALA:HB1	1:C:172:MET:HE2	1.89	0.54
1:A:231:ALA:HB2	1:D:216:ALA:HB1	1.89	0.53
1:D:212:MET:HB3	1:D:240:PRO:HB3	1.92	0.52
1:C:8:VAL:HG22	1:C:84:VAL:HB	1.91	0.51
1:C:11:ILE:HD11	1:C:73:LEU:HD21	1.93	0.50
1:D:110:GLU:HA	1:D:114:TRP:HB3	1.94	0.50
1:A:227:GLN:NE2	1:A:233:THR:HG23	2.25	0.49
1:D:168:ALA:HB1	1:D:172:MET:HE2	1.94	0.49
1:B:8:VAL:HG22	1:B:84:VAL:HB	1.94	0.49
1:C:158:PHE:HB3	1:D:158:PHE:HB3	1.98	0.46
1:A:158:PHE:HB3	1:B:158:PHE:HB3	1.97	0.46
1:B:213:PRO:HA	1:B:214:PRO:HD3	1.85	0.46
1:B:110:GLU:HA	1:B:114:TRP:HB3	1.97	0.45
1:C:162:ALA:HB2	1:D:158:PHE:CG	2.51	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
3:A:388[B]:HOH:O	1:B:106[B]:ARG:NH1	2.50	0.45
1:C:158:PHE:CG	1:D:162:ALA:HB2	2.53	0.44
1:B:237:GLU:O	1:C:233[B]:THR:HG22	2.18	0.44
1:A:233:THR:HG21	1:D:236:MET:HG2	1.99	0.44
1:B:20:GLU:HG3	1:B:214:PRO:HB2	2.00	0.44
1:C:110:GLU:HA	1:C:114:TRP:HB3	1.98	0.44
1:C:7:THR:HG23	1:C:80:ALA:HB3	2.00	0.43
1:A:213:PRO:HA	1:A:214:PRO:HD3	1.91	0.43
1:D:20:GLU:CG	1:D:214:PRO:HB2	2.43	0.43
1:C:212:MET:HA	1:C:213:PRO:HD2	1.43	0.43
1:A:158:PHE:CG	1:B:162:ALA:HB2	2.54	0.42
1:B:90:GLY:O	1:B:91:ALA:HB2	2.20	0.42
1:C:212:MET:HG2	1:C:216:ALA:HB3	2.02	0.42
1:D:138:THR:HG23	1:D:183:ILE:HD13	2.01	0.42
1:A:162:ALA:HB2	1:B:158:PHE:CG	2.55	0.41
1:B:2:THR:HA	1:B:3:PRO:HD2	1.95	0.41
1:A:109:TRP:HH2	1:B:106[A]:ARG:HG2	1.86	0.40
1:B:89:VAL:CG1	1:B:111[A]:MET:HG3	2.52	0.40
1:C:106:ARG:HD3	1:D:106:ARG:HH22	1.86	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	210/252 (83%)	203 (97%)	6 (3%)	1 (0%)	34 20
1	B	213/252 (84%)	206 (97%)	6 (3%)	1 (0%)	34 20
1	C	212/252 (84%)	202 (95%)	8 (4%)	2 (1%)	21 8
1	D	211/252 (84%)	202 (96%)	8 (4%)	1 (0%)	34 20
All	All	846/1008 (84%)	813 (96%)	28 (3%)	5 (1%)	30 16

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	91	ALA
1	C	90	GLY
1	C	213	PRO
1	A	91	ALA
1	D	91	ALA

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	156/187 (83%)	153 (98%)	3 (2%)	65	56
1	B	159/187 (85%)	154 (97%)	5 (3%)	47	34
1	C	158/187 (84%)	155 (98%)	3 (2%)	65	56
1	D	157/187 (84%)	152 (97%)	5 (3%)	46	33
All	All	630/748 (84%)	614 (98%)	16 (2%)	57	44

All (16) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	145	ARG
1	A	172	MET
1	A	233	THR
1	B	106[A]	ARG
1	B	106[B]	ARG
1	B	176	ILE
1	B	184	ASP
1	B	212	MET
1	C	4	ARG
1	C	55	ARG
1	C	145	ARG
1	D	37	ARG
1	D	40	GLU
1	D	143	SER
1	D	145	ARG

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	D	212	MET

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (5) such sidechains are listed below:

Mol	Chain	Res	Type
1	B	5	ASN
1	B	223	GLN
1	C	132	GLN
1	D	5	ASN
1	D	226	GLN

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 3 ligands modelled in this entry, 3 are unknown - 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

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, 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	212/252 (84%)	-0.14	5 (2%) 62 64	15, 25, 50, 65	0
1	B	213/252 (84%)	-0.06	7 (3%) 50 52	15, 27, 54, 69	0
1	C	212/252 (84%)	0.52	24 (11%) 7 7	17, 40, 69, 84	0
1	D	212/252 (84%)	0.30	17 (8%) 15 16	18, 33, 65, 76	0
All	All	849/1008 (84%)	0.15	53 (6%) 24 26	15, 30, 63, 84	0

All (53) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	D	51	ALA	5.9
1	C	40	GLU	4.8
1	C	42	LEU	4.6
1	C	43	ALA	4.5
1	D	241	TYR	4.5
1	C	38	ASN	4.3
1	C	16	TYR	4.2
1	C	51	ALA	4.2
1	D	40	GLU	4.1
1	A	241	TYR	4.1
1	C	47	ALA	4.1
1	D	212	MET	3.7
1	B	241	TYR	3.7
1	B	2	THR	3.6
1	D	240	PRO	3.6
1	D	16	TYR	3.6
1	C	86	ILE	3.4
1	C	44	PRO	3.4
1	C	37	ARG	3.4
1	C	45	LEU	3.3
1	D	86	ILE	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	C	52	ALA	3.1
1	D	52	ALA	3.1
1	B	86	ILE	3.0
1	B	40	GLU	2.9
1	D	3	PRO	2.8
1	B	16	TYR	2.8
1	B	212	MET	2.8
1	D	38	ASN	2.8
1	C	9	ALA	2.7
1	C	39	GLY	2.7
1	C	54	GLY	2.7
1	D	95	PHE	2.7
1	A	16	TYR	2.7
1	C	84	VAL	2.6
1	C	213	PRO	2.6
1	C	8	VAL	2.6
1	C	3	PRO	2.5
1	A	240	PRO	2.5
1	D	239	ARG	2.4
1	B	240	PRO	2.4
1	D	10	VAL	2.4
1	D	87	PHE	2.4
1	C	212	MET	2.3
1	C	41	LYS	2.3
1	C	85	THR	2.2
1	C	87	PHE	2.2
1	D	136	PHE	2.2
1	C	78	ALA	2.1
1	A	11	ILE	2.1
1	D	43	ALA	2.1
1	A	86	ILE	2.0
1	D	103	ARG	2.0

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, 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	UNL	D	251	7/-	0.85	0.14	0.16	48,48,55,55	0
2	UNL	D	252	7/-	0.87	0.15	0.09	30,42,47,48	0
2	UNL	C	251	7/-	0.89	0.12	0.03	40,44,50,52	0

6.5 Other polymers

There are no such residues in this entry.