wwPDB X-ray Structure Validation Summary Report (i)

Jun 27, 2023 – 07:20 pm BST

PDB ID	:	8A4R
Title	:	Proline Racemase (ProR) from the Gram-positive bacterium Acetoanaerobium sticklandii from isotropic orthorhombic data at 3.59 A
Authors	:	Najmudin, S.; Pan, X.-S.; McAuley, K.E.; Fisher, L.M.; Sanderson, M.R.
Deposited on	:	2022-06-13
Resolution	:	3.59 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
https://www.wwpdb.org/validation/2017/XrayValidationReportHelp
with specific help available everywhere you see the (i) 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 (i)) were used in the production of this report:

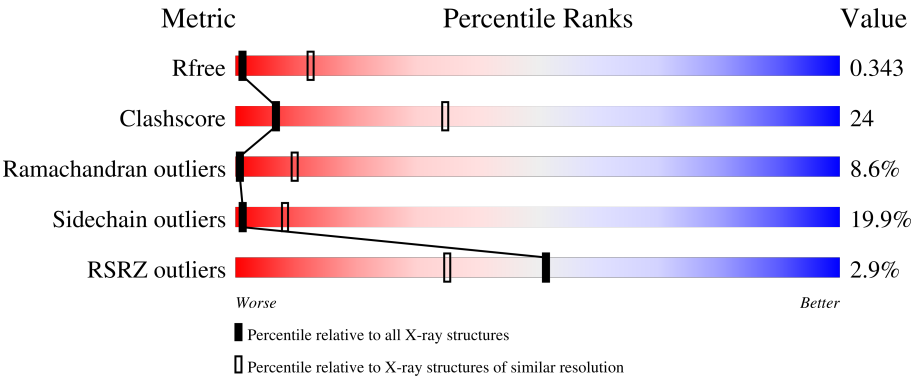
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.33
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.33

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 3.59 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	1257 (3.70-3.50)
Clashscore	141614	1353 (3.70-3.50)
Ramachandran outliers	138981	1307 (3.70-3.50)
Sidechain outliers	138945	1307 (3.70-3.50)
RSRZ outliers	127900	1161 (3.70-3.50)

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	AAA	343	47% 38% 13%
1	BBB	343	48% 36% 13%
1	CCC	343	4% 45% 38% 14%
1	DDD	343	6% 48% 36% 12%

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard

residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
2	PYC	AAA	401	-	-	X	-

2 Entry composition (i)

There are 2 unique types of molecules in this entry. The entry contains 20776 atoms, of which 10420 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 Proline racemase A (AsProR).

Mol	Chain	Residues	Atoms						ZeroOcc	AltConf	Trace
1	AAA	335	Total	C	H	N	O	S	111	0	0
			5188	1649	2603	421	500	15			
1	BBB	335	Total	C	H	N	O	S	111	0	0
			5188	1649	2603	421	500	15			
1	CCC	335	Total	C	H	N	O	S	111	0	0
			5188	1649	2603	421	500	15			
1	DDD	335	Total	C	H	N	O	S	111	0	0
			5188	1649	2603	421	500	15			

There are 32 discrepancies between the modelled and reference sequences:

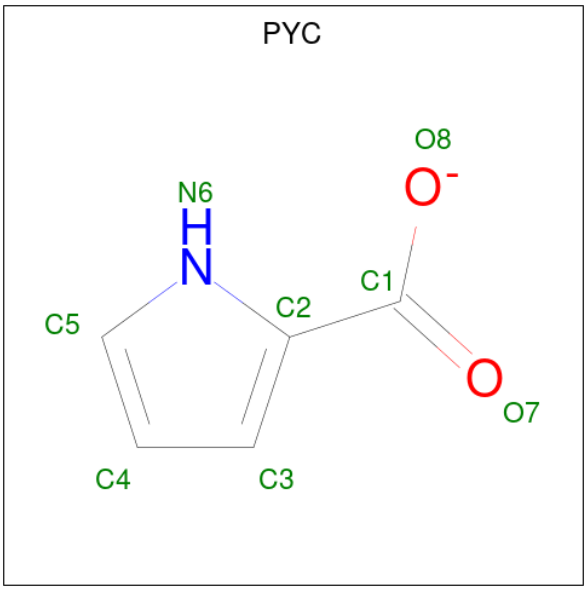
Chain	Residue	Modelled	Actual	Comment	Reference
AAA	336	LEU	-	expression tag	UNP E3PTZ4
AAA	337	GLU	-	expression tag	UNP E3PTZ4
AAA	338	HIS	-	expression tag	UNP E3PTZ4
AAA	339	HIS	-	expression tag	UNP E3PTZ4
AAA	340	HIS	-	expression tag	UNP E3PTZ4
AAA	341	HIS	-	expression tag	UNP E3PTZ4
AAA	342	HIS	-	expression tag	UNP E3PTZ4
AAA	343	HIS	-	expression tag	UNP E3PTZ4
BBB	336	LEU	-	expression tag	UNP E3PTZ4
BBB	337	GLU	-	expression tag	UNP E3PTZ4
BBB	338	HIS	-	expression tag	UNP E3PTZ4
BBB	339	HIS	-	expression tag	UNP E3PTZ4
BBB	340	HIS	-	expression tag	UNP E3PTZ4
BBB	341	HIS	-	expression tag	UNP E3PTZ4
BBB	342	HIS	-	expression tag	UNP E3PTZ4
BBB	343	HIS	-	expression tag	UNP E3PTZ4
CCC	336	LEU	-	expression tag	UNP E3PTZ4
CCC	337	GLU	-	expression tag	UNP E3PTZ4
CCC	338	HIS	-	expression tag	UNP E3PTZ4
CCC	339	HIS	-	expression tag	UNP E3PTZ4
CCC	340	HIS	-	expression tag	UNP E3PTZ4

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
CCC	341	HIS	-	expression tag	UNP E3PTZ4
CCC	342	HIS	-	expression tag	UNP E3PTZ4
CCC	343	HIS	-	expression tag	UNP E3PTZ4
DDD	336	LEU	-	expression tag	UNP E3PTZ4
DDD	337	GLU	-	expression tag	UNP E3PTZ4
DDD	338	HIS	-	expression tag	UNP E3PTZ4
DDD	339	HIS	-	expression tag	UNP E3PTZ4
DDD	340	HIS	-	expression tag	UNP E3PTZ4
DDD	341	HIS	-	expression tag	UNP E3PTZ4
DDD	342	HIS	-	expression tag	UNP E3PTZ4
DDD	343	HIS	-	expression tag	UNP E3PTZ4

- Molecule 2 is PYRROLE-2-CARBOXYLATE (three-letter code: PYC) (formula: $C_5H_4NO_2$).



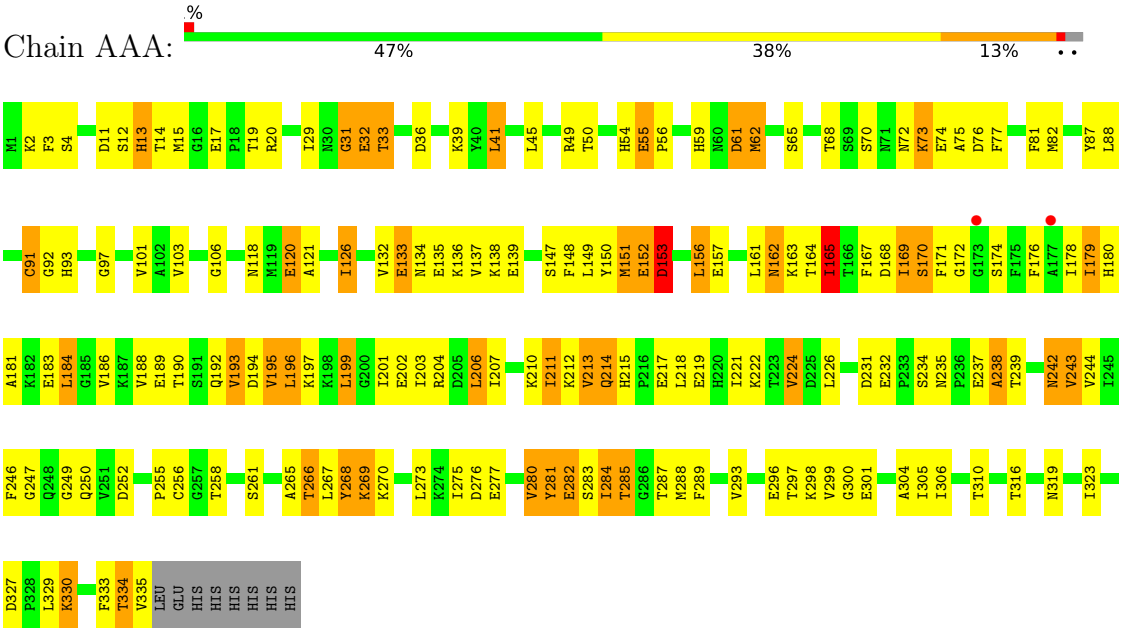
Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
2	AAA	1	Total	C	H	N	O	0	0
			12	5	4	1	2		
2	BBB	1	Total	C	H	N	O	0	0
			12	5	4	1	2		

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.

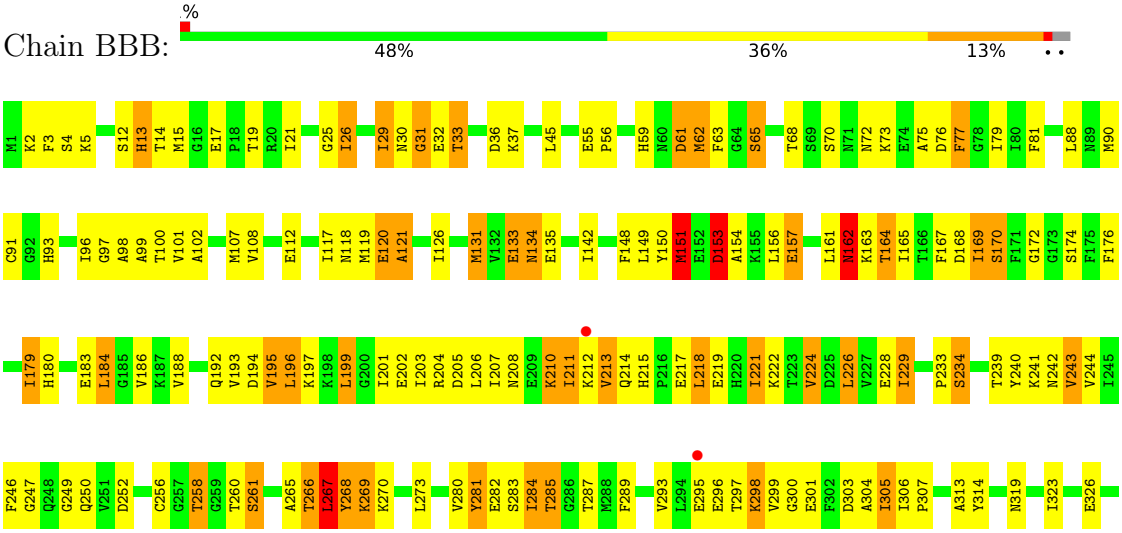
• Molecule 1: Proline racemase A (AsProR)

Chain AAA:



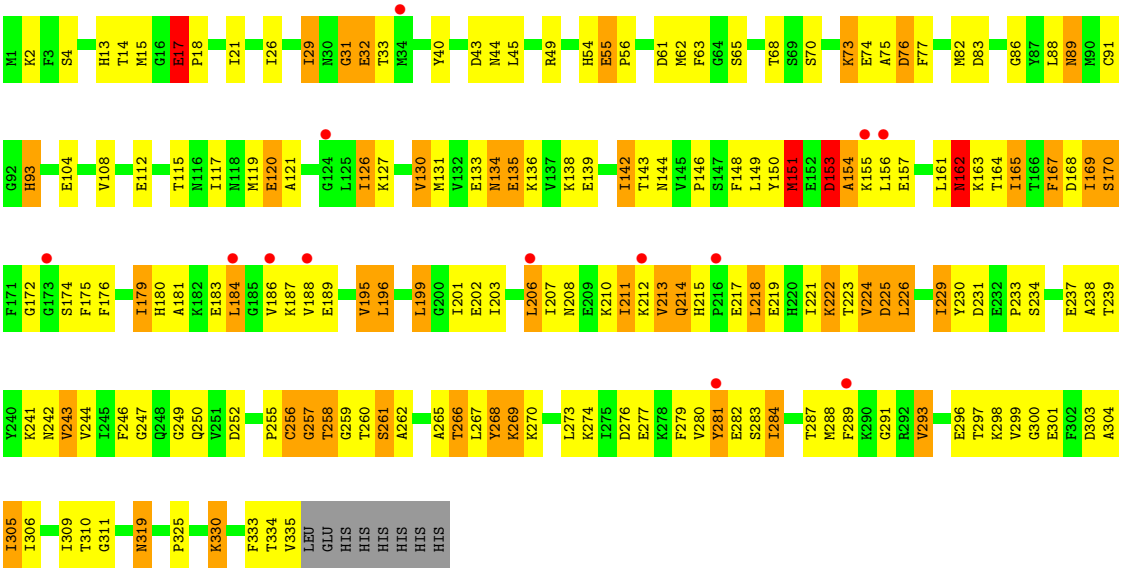
• Molecule 1: Proline racemase A (AsProR)

Chain BBB:



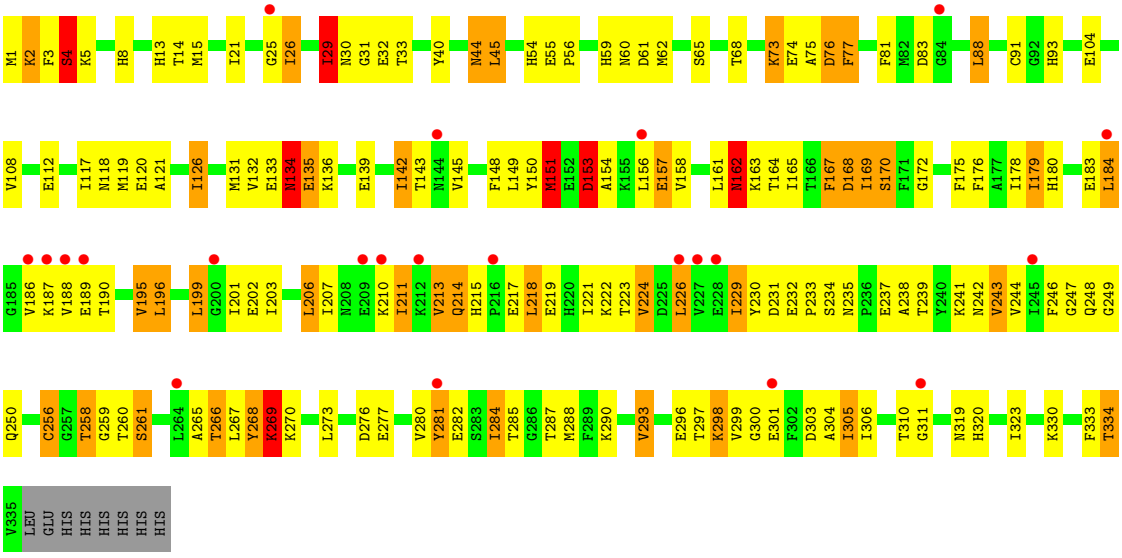
• Molecule 1: Proline racemase A (AsProR)

Chain CCC:



• Molecule 1: Proline racemase A (AsProR)

Chain DDD:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	104.45Å 107.48Å 109.05Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	76.55 – 3.59 76.55 – 3.59	Depositor EDS
% Data completeness (in resolution range)	100.0 (76.55-3.59) 96.0 (76.55-3.59)	Depositor EDS
R_{merge}	0.20	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.99 (at 3.58Å)	Xtriage
Refinement program	REFMAC 5.8.0258	Depositor
R, R_{free}	0.260 , 0.360 0.260 , 0.343	Depositor DCC
R_{free} test set	747 reflections (5.02%)	wwPDB-VP
Wilson B-factor (Å^2)	119.1	Xtriage
Anisotropy	0.117	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.39 , 132.0	EDS
L-test for twinning[2]	$< \|L\| > = 0.43$, $< L^2 > = 0.26$	Xtriage
Estimated twinning fraction	0.199 for -h,l,k 0.046 for -l,-k,-h 0.059 for k,h,-l 0.043 for k,l,h 0.043 for l,h,k	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	20776	wwPDB-VP
Average B, all atoms (Å^2)	153.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.71% 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

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

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	AAA	0.75	0/2634	0.95	0/3556
1	BBB	0.71	0/2634	0.92	0/3556
1	CCC	0.73	3/2634 (0.1%)	0.87	0/3556
1	DDD	0.69	0/2634	0.87	0/3556
All	All	0.72	3/10536 (0.0%)	0.90	0/14224

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	AAA	0	1
1	BBB	0	1
1	CCC	0	2
1	DDD	0	1
All	All	0	5

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	CCC	55	GLU	CD-OE1	7.07	1.33	1.25
1	CCC	55	GLU	CD-OE2	6.33	1.32	1.25
1	CCC	17	GLU	CD-OE1	5.31	1.31	1.25

There are no bond angle outliers.

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	AAA	266	THR	Peptide
1	BBB	266	THR	Peptide
1	CCC	266	THR	Peptide
1	CCC	319	ASN	Peptide
1	DDD	266	THR	Peptide

5.2 Too-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 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	AAA	2585	2603	2593	117	1
1	BBB	2585	2603	2593	139	1
1	CCC	2585	2603	2593	124	1
1	DDD	2585	2603	2593	113	1
2	AAA	8	4	4	4	0
2	BBB	8	4	4	2	0
All	All	10356	10420	10380	488	2

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 24.

The worst 5 of 488 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:CCC:91:CYS:SG	1:CCC:93:HIS:HD2	1.58	1.24
1:CCC:91:CYS:SG	1:CCC:93:HIS:CD2	2.51	1.04
1:BBB:70:SER:O	1:BBB:120:GLU:OE1	1.92	0.88
1:AAA:153:ASP:HA	1:AAA:167:PHE:O	1.76	0.86
1:BBB:170:SER:OG	1:BBB:266:THR:HG21	1.76	0.86

All (2) 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	Interatomic distance (Å)	Clash overlap (Å)
1:AAA:39:LYS:HZ2	1:BBB:153:ASP:OD2[2_555]	1.59	0.01
1:CCC:319:ASN:OD1	1:DDD:319:ASN:OD1[2_454]	2.19	0.01

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	AAA	333/343 (97%)	247 (74%)	57 (17%)	29 (9%)	1 9
1	BBB	333/343 (97%)	252 (76%)	54 (16%)	27 (8%)	1 11
1	CCC	333/343 (97%)	251 (75%)	53 (16%)	29 (9%)	1 9
1	DDD	333/343 (97%)	247 (74%)	57 (17%)	29 (9%)	1 9
All	All	1332/1372 (97%)	997 (75%)	221 (17%)	114 (9%)	1 10

5 of 114 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	AAA	32	GLU
1	AAA	134	ASN
1	AAA	153	ASP
1	AAA	163	LYS
1	AAA	218	LEU

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	AAA	285/293 (97%)	227 (80%)	58 (20%)	1 8
1	BBB	285/293 (97%)	233 (82%)	52 (18%)	1 10
1	CCC	285/293 (97%)	225 (79%)	60 (21%)	1 7
1	DDD	285/293 (97%)	228 (80%)	57 (20%)	1 8
All	All	1140/1172 (97%)	913 (80%)	227 (20%)	1 8

5 of 227 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	CCC	33	THR
1	DDD	298	LYS
1	CCC	202	GLU
1	DDD	290	LYS
1	DDD	202	GLU

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

5.3.3 RNA (i)

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)

2 ligands are 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	$\#\|Z\| > 2$
2	PYC	BBB	401	-	8,8,8	0.91	0	8,10,10	1.32	1 (12%)
2	PYC	AAA	401	-	8,8,8	1.00	1 (12%)	8,10,10	1.07	0

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
2	PYC	BBB	401	-	-	2/2/4/4	0/1/1/1
2	PYC	AAA	401	-	-	0/2/4/4	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	AAA	401	PYC	O8-C1	-2.00	1.24	1.30

All (1) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	BBB	401	PYC	O8-C1-C2	2.22	119.74	114.69

There are no chirality outliers.

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	BBB	401	PYC	O7-C1-C2-N6
2	BBB	401	PYC	O7-C1-C2-C3

There are no ring outliers.

2 monomers are involved in 6 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	BBB	401	PYC	2	0
2	AAA	401	PYC	4	0

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	AAA	335/343 (97%)	0.02	2 (0%)	89	81	77, 117, 170, 189	0
1	BBB	335/343 (97%)	0.06	2 (0%)	89	81	87, 128, 185, 216	0
1	CCC	335/343 (97%)	0.33	13 (3%)	39	25	105, 174, 246, 270	0
1	DDD	335/343 (97%)	0.38	22 (6%)	18	10	102, 181, 242, 264	0
All	All	1340/1372 (97%)	0.20	39 (2%)	51	35	77, 144, 231, 270	0

The worst 5 of 39 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	DDD	212	LYS	5.0
1	CCC	173	GLY	4.4
1	DDD	200	GLY	4.1
1	DDD	188	VAL	3.9
1	CCC	212	LYS	3.8

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 monosaccharides 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. 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	B-factors(Å^2)	Q<0.9
2	PYC	BBB	401	8/8	0.96	0.38	116,155,163,163	0
2	PYC	AAA	401	8/8	0.98	0.36	95,120,129,129	0

6.5 Other polymers (i)

There are no such residues in this entry.