Full wwPDB X-ray Structure Validation Report ⓘ

Sep 13, 2020 – 09:41 PM BST

PDB ID	:	5O1O
Title	:	Crystal structure of human aminoadipate semialdehyde synthase, saccharopine dehydrogenase domain with proline bound.
Authors	:	Kopec, J.; Rembeza, E.; Pena, I.A.; Mathea, S.; Velupillai, S.; Strain-Damerell, C.; Goubin, S.; Kupinska, K.; Talon, R.; Collins, P.; Krojer, T.; Burgess-Brown, N.; Arrowsmith, C.; Edwards, A.; Bountra, C.; von Delft, F.; Arruda, P.; Yue, W.W.
Deposited on	:	2017-05-18
Resolution	:	2.48 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 following versions of software and data (see references ①) were used in the production of this report:

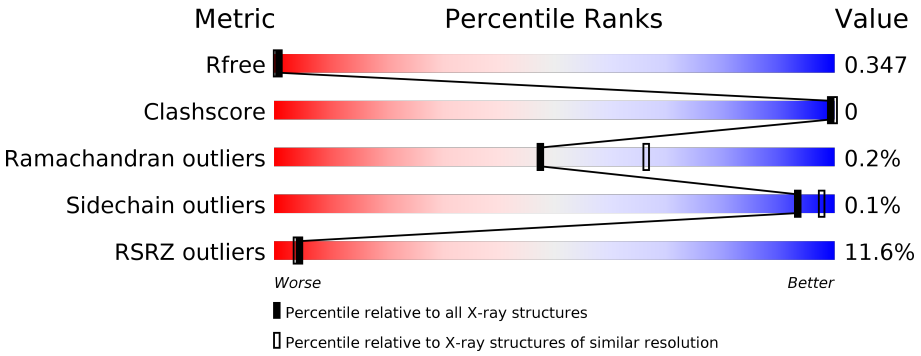
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.14.4.dev1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.14.4.dev1

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 2.48 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	5857 (2.50-2.46)
Clashscore	141614	6594 (2.50-2.46)
Ramachandran outliers	138981	6469 (2.50-2.46)
Sidechain outliers	138945	6471 (2.50-2.46)
RSRZ outliers	127900	5738 (2.50-2.46)

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 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	497	9% 88% 12%
1	B	497	11% 87% 12%

2 Entry composition

There are 4 unique types of molecules in this entry. The entry contains 6529 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 Alpha-aminoadipic semialdehyde synthase, mitochondrial.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	439	Total	C	N	O	S	0	0	0
			3262	2087	530	626	19			
1	B	435	Total	C	N	O	S	0	1	0
			3231	2068	523	620	20			

There are 52 discrepancies between the modelled and reference sequences:

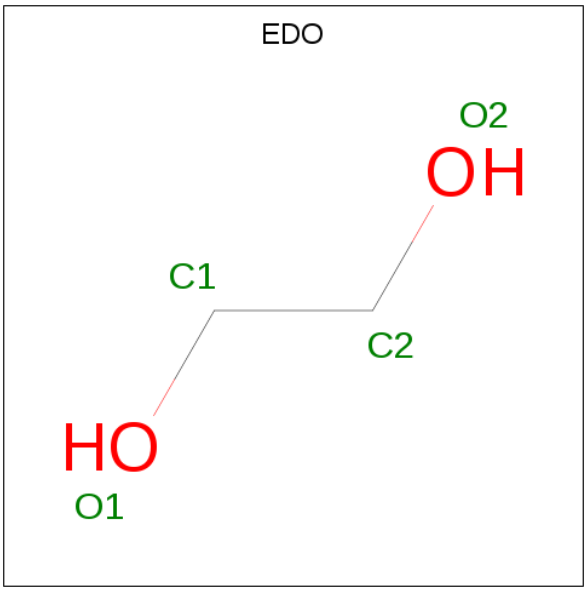
Chain	Residue	Modelled	Actual	Comment	Reference
A	430	MET	-	initiating methionine	UNP Q9UDR5
A	431	GLY	-	expression tag	UNP Q9UDR5
A	432	HIS	-	expression tag	UNP Q9UDR5
A	433	HIS	-	expression tag	UNP Q9UDR5
A	434	HIS	-	expression tag	UNP Q9UDR5
A	435	HIS	-	expression tag	UNP Q9UDR5
A	436	HIS	-	expression tag	UNP Q9UDR5
A	437	HIS	-	expression tag	UNP Q9UDR5
A	438	SER	-	expression tag	UNP Q9UDR5
A	439	SER	-	expression tag	UNP Q9UDR5
A	440	GLY	-	expression tag	UNP Q9UDR5
A	441	VAL	-	expression tag	UNP Q9UDR5
A	442	ASP	-	expression tag	UNP Q9UDR5
A	443	LEU	-	expression tag	UNP Q9UDR5
A	444	GLY	-	expression tag	UNP Q9UDR5
A	445	THR	-	expression tag	UNP Q9UDR5
A	446	GLU	-	expression tag	UNP Q9UDR5
A	447	ASN	-	expression tag	UNP Q9UDR5
A	448	LEU	-	expression tag	UNP Q9UDR5
A	449	TYR	-	expression tag	UNP Q9UDR5
A	450	PHE	-	expression tag	UNP Q9UDR5
A	451	GLN	-	expression tag	UNP Q9UDR5
A	452	SER	-	expression tag	UNP Q9UDR5
A	453	MET	-	expression tag	UNP Q9UDR5
A	454	ALA	-	expression tag	UNP Q9UDR5

Continued on next page...

Continued from previous page...

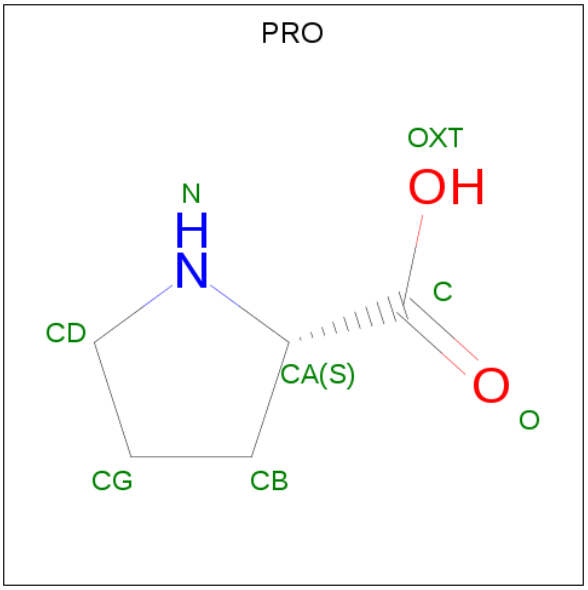
Chain	Residue	Modelled	Actual	Comment	Reference
A	615	SER	THR	cloning artifact	UNP Q9UDR5
B	430	MET	-	initiating methionine	UNP Q9UDR5
B	431	GLY	-	expression tag	UNP Q9UDR5
B	432	HIS	-	expression tag	UNP Q9UDR5
B	433	HIS	-	expression tag	UNP Q9UDR5
B	434	HIS	-	expression tag	UNP Q9UDR5
B	435	HIS	-	expression tag	UNP Q9UDR5
B	436	HIS	-	expression tag	UNP Q9UDR5
B	437	HIS	-	expression tag	UNP Q9UDR5
B	438	SER	-	expression tag	UNP Q9UDR5
B	439	SER	-	expression tag	UNP Q9UDR5
B	440	GLY	-	expression tag	UNP Q9UDR5
B	441	VAL	-	expression tag	UNP Q9UDR5
B	442	ASP	-	expression tag	UNP Q9UDR5
B	443	LEU	-	expression tag	UNP Q9UDR5
B	444	GLY	-	expression tag	UNP Q9UDR5
B	445	THR	-	expression tag	UNP Q9UDR5
B	446	GLU	-	expression tag	UNP Q9UDR5
B	447	ASN	-	expression tag	UNP Q9UDR5
B	448	LEU	-	expression tag	UNP Q9UDR5
B	449	TYR	-	expression tag	UNP Q9UDR5
B	450	PHE	-	expression tag	UNP Q9UDR5
B	451	GLN	-	expression tag	UNP Q9UDR5
B	452	SER	-	expression tag	UNP Q9UDR5
B	453	MET	-	expression tag	UNP Q9UDR5
B	454	ALA	-	expression tag	UNP Q9UDR5
B	615	SER	THR	cloning artifact	UNP Q9UDR5

- Molecule 2 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: $C_2H_6O_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	A	1	Total C O 4 2 2	0	0
2	B	1	Total C O 4 2 2	0	0

- Molecule 3 is PROLINE (three-letter code: PRO) (formula: $C_5H_9NO_2$).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total C N O 8 5 1 2	0	0
3	B	1	Total C N O 8 5 1 2	0	0

- Molecule 4 is water.

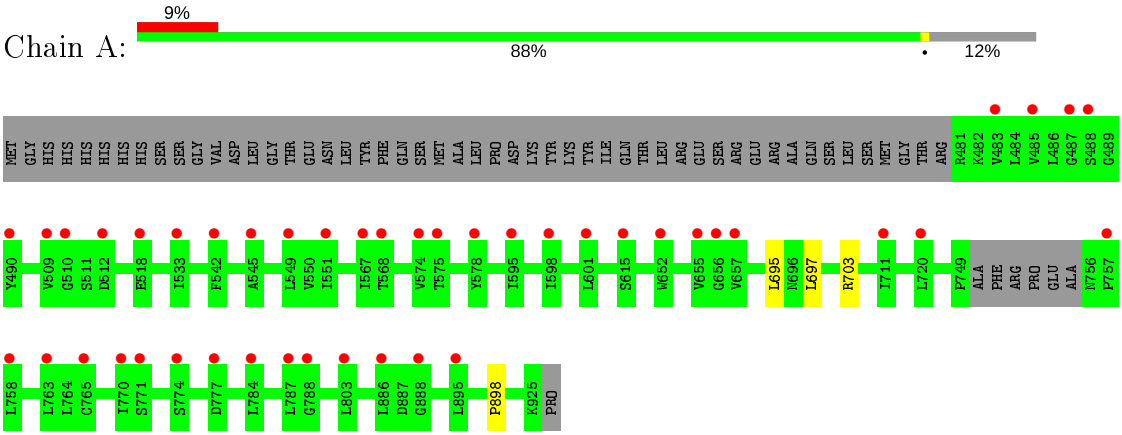
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	6	Total O 6 6	0	0
4	B	6	Total O 6 6	0	0

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.

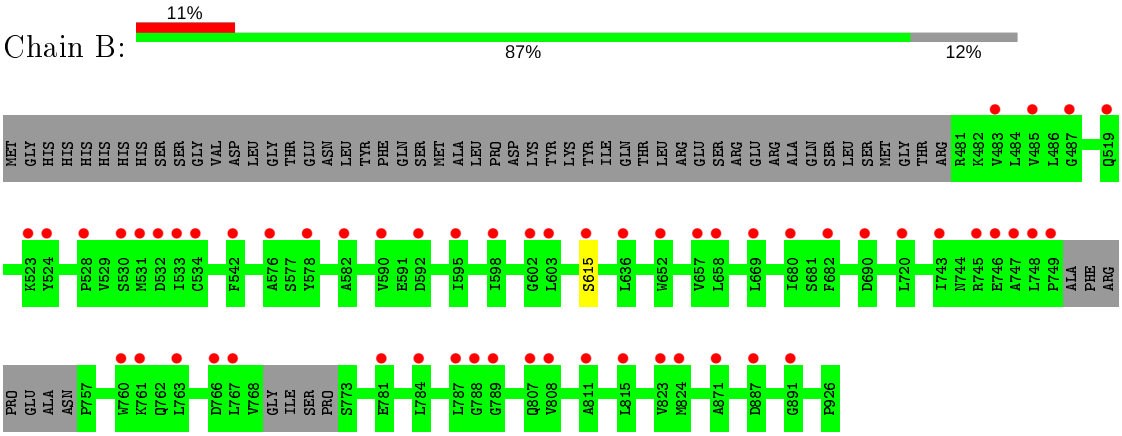
- Molecule 1: Alpha-aminoadipic semialdehyde synthase, mitochondrial

Chain A:



- Molecule 1: Alpha-aminoadipic semialdehyde synthase, mitochondrial

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	92.37Å 102.41Å 142.07Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	83.08 – 2.48 83.08 – 2.48	Depositor EDS
% Data completeness (in resolution range)	92.8 (83.08-2.48) 92.8 (83.08-2.48)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.56 (at 2.48Å)	Xtriage
Refinement program	REFMAC 5.8.0158	Depositor
R, R_{free}	0.322 , 0.348 0.322 , 0.347	Depositor DCC
R_{free} test set	2169 reflections (4.80%)	wwPDB-VP
Wilson B-factor (Å^2)	49.0	Xtriage
Anisotropy	0.361	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.33 , 34.5	EDS
L-test for twinning[2]	$< \|L\| > = 0.53$, $< L^2 > = 0.37$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.90	EDS
Total number of atoms	6529	wwPDB-VP
Average B, all atoms (Å^2)	63.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 3.75% 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: EDO

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	0.34	0/3322	0.47	0/4514
1	B	0.34	0/3290	0.47	0/4468
All	All	0.34	0/6612	0.47	0/8982

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	3262	0	3227	1	0
1	B	3231	0	3184	0	0
2	A	4	0	6	0	0
2	B	4	0	6	0	0
3	A	8	0	7	0	0
3	B	8	0	7	0	0
4	A	6	0	0	0	0
4	B	6	0	0	0	0
All	All	6529	0	6437	1	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 0.

All (1) 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:695:LEU:HD23	1:A:697:LEU:HD11	1.93	0.50

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	435/497 (88%)	412 (95%)	21 (5%)	2 (0%)	29 46
1	B	430/497 (86%)	407 (95%)	23 (5%)	0	100 100
All	All	865/994 (87%)	819 (95%)	44 (5%)	2 (0%)	47 66

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	703	ARG
1	A	898	PRO

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	343/418 (82%)	343 (100%)	0	100 100

Continued on next page...

Continued from previous page...

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
1	B	338/418 (81%)	337 (100%)	1 (0%)	92 97
All	All	681/836 (82%)	680 (100%)	1 (0%)	93 97

All (1) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	615	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	702	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 monosaccharides in this entry.

5.6 Ligand geometry

4 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	$\#\|Z\|>2$
3	PRO	B	1002	-	5,8,8	0.58	0	6,10,10	0.98	0
2	EDO	A	1001	-	3,3,3	0.46	0	2,2,2	0.27	0
3	PRO	A	1002	-	5,8,8	0.58	0	6,10,10	1.01	0
2	EDO	B	1001	-	3,3,3	0.46	0	2,2,2	0.26	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
3	PRO	B	1002	-	-	0/0/11/11	0/1/1/1
2	EDO	A	1001	-	-	1/1/1/1	-
3	PRO	A	1002	-	-	0/0/11/11	0/1/1/1
2	EDO	B	1001	-	-	0/1/1/1	-

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (1) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	1001	EDO	O1-C1-C2-O2

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	439/497 (88%)	0.98	44 (10%) 7 6	36, 61, 88, 95	0
1	B	435/497 (87%)	1.00	57 (13%) 3 2	39, 64, 97, 105	0
All	All	874/994 (87%)	0.99	101 (11%) 4 4	36, 62, 92, 105	0

All (101) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	787	LEU	9.7
1	A	788	GLY	5.2
1	B	788	GLY	4.9
1	B	595	ILE	4.2
1	B	749	PRO	3.8
1	A	655	VAL	3.6
1	B	528	PRO	3.5
1	B	891	GLY	3.5
1	B	542	PHE	3.4
1	B	763	LEU	3.4
1	A	784	LEU	3.4
1	B	602	GLY	3.4
1	B	657	VAL	3.4
1	A	777	ASP	3.4
1	A	533	ILE	3.3
1	A	652	TRP	3.3
1	A	771	SER	3.3
1	B	767	LEU	3.2
1	B	487	GLY	3.1
1	B	824	MET	3.1
1	A	549	LEU	3.1
1	A	485	VAL	3.1
1	A	657	VAL	3.0
1	A	765	CYS	3.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	575	THR	2.9
1	B	807	GLN	2.9
1	A	656	GLY	2.9
1	B	761	LYS	2.9
1	B	781	GLU	2.9
1	B	811	ALA	2.9
1	B	789	GLY	2.8
1	B	680	ILE	2.8
1	A	568	THR	2.8
1	B	532	ASP	2.8
1	A	615	SER	2.8
1	B	743	ILE	2.8
1	B	766	ASP	2.8
1	B	784	LEU	2.8
1	B	531	MET	2.8
1	A	770	ILE	2.8
1	A	488	SER	2.7
1	A	601	LEU	2.7
1	B	533	ILE	2.7
1	A	567	ILE	2.7
1	B	746	GLU	2.7
1	B	582	ALA	2.7
1	A	774	SER	2.6
1	B	823	VAL	2.6
1	B	530	SER	2.6
1	B	808	VAL	2.6
1	A	598	ILE	2.6
1	A	757	PRO	2.5
1	A	803	LEU	2.5
1	B	603	LEU	2.5
1	A	487	GLY	2.5
1	B	578	TYR	2.5
1	A	509	VAL	2.5
1	B	690	ASP	2.5
1	A	551	ILE	2.5
1	A	578	TYR	2.5
1	A	886	LEU	2.5
1	B	592	ASP	2.5
1	B	658	LEU	2.4
1	A	574	VAL	2.4
1	A	595	ILE	2.3
1	A	720	LEU	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	888	GLY	2.3
1	B	682	PHE	2.3
1	A	510	GLY	2.3
1	A	542	PHE	2.2
1	B	669	LEU	2.2
1	B	590	VAL	2.2
1	A	711	ILE	2.2
1	B	598	ILE	2.2
1	B	524	TYR	2.2
1	B	534	CYS	2.2
1	A	512	ASP	2.2
1	B	523	LYS	2.2
1	A	545	ALA	2.2
1	B	652	TRP	2.2
1	B	519	GLN	2.2
1	B	871	ALA	2.2
1	B	887	ASP	2.2
1	A	490	TYR	2.2
1	B	485	VAL	2.2
1	A	763	LEU	2.2
1	B	815	LEU	2.1
1	B	636	LEU	2.1
1	B	615	SER	2.1
1	A	483	VAL	2.1
1	A	758	LEU	2.1
1	B	720	LEU	2.1
1	A	518	GLU	2.1
1	B	760	TRP	2.1
1	B	576	ALA	2.0
1	B	748	LEU	2.0
1	B	787	LEU	2.0
1	B	747	ALA	2.0
1	A	895	LEU	2.0
1	B	483	VAL	2.0
1	B	745	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 monosaccharides in this entry.

6.4 Ligands (i)

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	EDO	B	1001	4/4	0.75	0.25	60,60,60,60	0
2	EDO	A	1001	4/4	0.86	0.22	61,62,62,62	0
3	PRO	A	1002	8/8	0.93	0.23	56,58,58,58	0
3	PRO	B	1002	8/8	0.95	0.16	46,47,47,47	0

6.5 Other polymers (i)

There are no such residues in this entry.