



Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 02:51 PM GMT

PDB ID : 4APS
Title : Crystal structure of a POT family peptide transporter in an inward open conformation.
Authors : Solcan, N.; Kwok, J.; Fowler, P.W.; Cameron, A.D.; Drew, D.; Iwata, S.; Newstead, S.
Deposited on : 2012-04-05
Resolution : 3.30 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 ⓘ symbol.

The following versions of software and data (see [references ⓘ](#)) 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

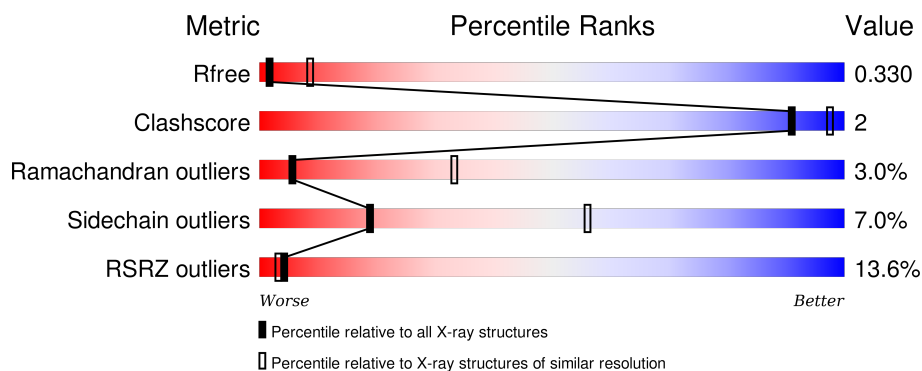
1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

X-RAY DIFFRACTION

The reported resolution of this entry is 3.30 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	2060 (3.40-3.20)
Clashscore	102246	1058 (3.38-3.22)
Ramachandran outliers	100387	1038 (3.38-3.22)
Sidechain outliers	100360	1037 (3.38-3.22)
RSRZ outliers	91569	2070 (3.40-3.20)

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 $\leq 5\%$. 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	491	 12% 79% 11% • 9%
1	B	491	 13% 80% 10% • 9%

2 Entry composition

There are 2 unique types of molecules in this entry. The entry contains 6891 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 DI-OR TRIPEPTIDE H⁺ SYMPORTER.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	446	Total	C	N	O	S	0	0	0
			3445	2326	528	574	17			
1	B	446	Total	C	N	O	S	0	0	0
			3445	2326	528	574	17			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	484	GLY	-	EXPRESSION TAG	UNP Q5M4H8
A	485	SER	-	EXPRESSION TAG	UNP Q5M4H8
A	486	GLU	-	EXPRESSION TAG	UNP Q5M4H8
A	487	ASN	-	EXPRESSION TAG	UNP Q5M4H8
A	488	LEU	-	EXPRESSION TAG	UNP Q5M4H8
A	489	TYR	-	EXPRESSION TAG	UNP Q5M4H8
A	490	PHE	-	EXPRESSION TAG	UNP Q5M4H8
A	491	GLN	-	EXPRESSION TAG	UNP Q5M4H8
B	484	GLY	-	EXPRESSION TAG	UNP Q5M4H8
B	485	SER	-	EXPRESSION TAG	UNP Q5M4H8
B	486	GLU	-	EXPRESSION TAG	UNP Q5M4H8
B	487	ASN	-	EXPRESSION TAG	UNP Q5M4H8
B	488	LEU	-	EXPRESSION TAG	UNP Q5M4H8
B	489	TYR	-	EXPRESSION TAG	UNP Q5M4H8
B	490	PHE	-	EXPRESSION TAG	UNP Q5M4H8
B	491	GLN	-	EXPRESSION TAG	UNP Q5M4H8

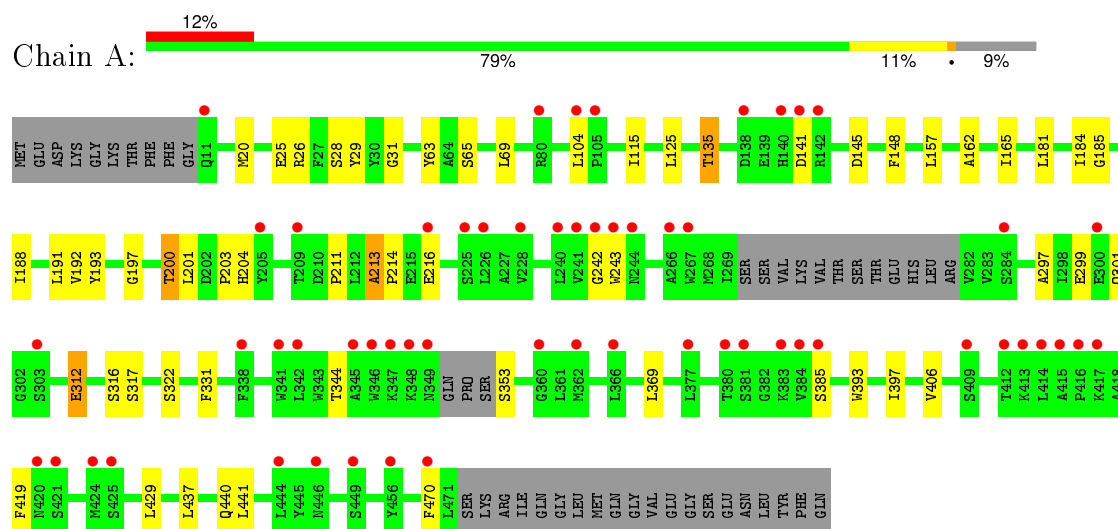
- Molecule 2 is CADMIUM ION (three-letter code: CD) (formula: Cd).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
2	B	1	Total	Cd	0	0
			1	1		

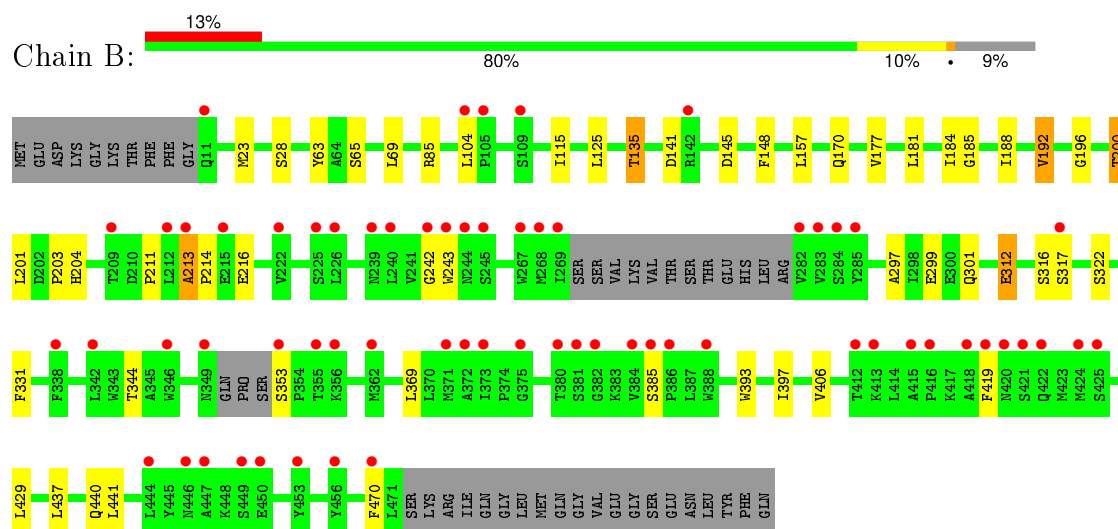
3 Residue-property plots

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: DI-OR TRIPEPTIDE H⁺ SYMPORTER



• Molecule 1: DI-OR TRIPEPTIDE H⁺ SYMPORTER



4 Data and refinement statistics

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α , β , γ	89.39Å 112.99Å 215.49Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	22.62 – 3.30 22.48 – 3.30	Depositor EDS
% Data completeness (in resolution range)	83.5 (22.62-3.30) 83.8 (22.48-3.30)	Depositor EDS
R_{merge}	0.10	Depositor
R_{sym}	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ ¹	5.94 (at 3.30Å)	Xtriage
Refinement program	BUSTER 2.10.0	Depositor
R, R_{free}	0.272 , 0.289 0.315 , 0.330	Depositor DCC
R_{free} test set	1420 reflections (5.09%)	DCC
Wilson B-factor (Å ²)	66.9	Xtriage
Anisotropy	0.195	Xtriage
Bulk solvent k_{sol} (e/Å ³), B_{sol} (Å ²)	0.26 , 137.9	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning ²	$\langle L \rangle = 0.39$, $\langle L^2 \rangle = 0.22$	Xtriage
Outliers	11 of 27943 reflections (0.039%)	Xtriage
F_o, F_c correlation	0.80	EDS
Total number of atoms	6891	wwPDB-VP
Average B, all atoms (Å ²)	140.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 7.13% of the height of the origin peak. No significant pseudotranslation is detected.*

¹Intensities estimated from amplitudes.

²Theoretical values of $\langle |L| \rangle$, $\langle L^2 \rangle$ 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: CD

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	0.52	0/3549	0.61	0/4844
1	B	0.51	0/3549	0.60	0/4844
All	All	0.52	0/7098	0.61	0/9688

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	3445	0	3534	19	0
1	B	3445	0	3534	13	0
2	B	1	0	0	0	0
All	All	6891	0	7068	30	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 2.

All (30) 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:165:ILE:HG22	1:A:181:LEU:HD21	1.73	0.70
1:A:331:PHE:HB2	1:A:397:ILE:HD13	1.76	0.67
1:B:331:PHE:HB2	1:B:397:ILE:HD13	1.76	0.66
1:A:135:THR:HG21	1:A:204:HIS:CD2	2.32	0.65
1:A:135:THR:HG21	1:A:204:HIS:HD2	1.62	0.64
1:A:185:GLY:HA2	1:A:188:ILE:HD12	1.81	0.62
1:A:393:TRP:O	1:A:397:ILE:HD12	2.07	0.55
1:B:393:TRP:O	1:B:397:ILE:HD12	2.07	0.54
1:A:20:MET:HB2	1:A:192:VAL:HG11	1.89	0.54
1:A:135:THR:OG1	1:A:204:HIS:NE2	2.37	0.53
1:A:184:ILE:HG23	1:B:184:ILE:HG21	1.93	0.50
1:B:63:TYR:CD1	1:B:115:ILE:HG23	2.48	0.48
1:A:63:TYR:CD1	1:A:115:ILE:HG23	2.49	0.48
1:B:185:GLY:HA2	1:B:188:ILE:HD12	1.96	0.48
1:B:297:ALA:O	1:B:301:GLN:HB2	2.14	0.47
1:A:297:ALA:O	1:A:301:GLN:HB2	2.14	0.47
1:A:20:MET:CB	1:A:192:VAL:HG11	2.46	0.46
1:A:28:SER:OG	1:A:29:TYR:N	2.50	0.45
1:A:213:ALA:HB3	1:A:214:PRO:HD3	1.99	0.44
1:A:25:GLU:OE1	1:A:26:ARG:NE	2.51	0.44
1:B:213:ALA:HB3	1:B:214:PRO:HD3	1.99	0.44
1:B:213:ALA:HB3	1:B:214:PRO:CD	2.49	0.42
1:A:213:ALA:HB3	1:A:214:PRO:CD	2.49	0.42
1:B:135:THR:HG21	1:B:204:HIS:HD2	1.84	0.42
1:A:193:TYR:O	1:A:197:GLY:N	2.50	0.42
1:B:135:THR:HG21	1:B:204:HIS:CD2	2.54	0.42
1:B:192:VAL:O	1:B:196:GLY:N	2.51	0.41
1:B:135:THR:HG1	1:B:204:HIS:CD2	2.39	0.40
1:A:191:LEU:HG	1:B:177:VAL:HG13	2.04	0.40
1:A:31:GLY:CA	1:A:162:ALA:HB1	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	440/491 (90%)	406 (92%)	21 (5%)	13 (3%)	5	33
1	B	440/491 (90%)	403 (92%)	24 (6%)	13 (3%)	5	33
All	All	880/982 (90%)	809 (92%)	45 (5%)	26 (3%)	5	33

All (26) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	203	PRO
1	A	317	SER
1	B	200	THR
1	B	203	PRO
1	B	317	SER
1	A	141	ASP
1	A	200	THR
1	A	385	SER
1	B	141	ASP
1	B	385	SER
1	A	213	ALA
1	A	344	THR
1	A	419	PHE
1	B	213	ALA
1	B	344	THR
1	B	419	PHE
1	A	104	LEU
1	A	312	GLU
1	B	104	LEU
1	B	312	GLU
1	A	125	LEU
1	B	125	LEU
1	A	211	PRO
1	B	211	PRO
1	A	242	GLY
1	B	242	GLY

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	358/398 (90%)	336 (94%)	22 (6%)	23	62
1	B	358/398 (90%)	330 (92%)	28 (8%)	16	50
All	All	716/796 (90%)	666 (93%)	50 (7%)	19	56

All (50) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	65	SER
1	A	69	LEU
1	A	135	THR
1	A	145	ASP
1	A	148	PHE
1	A	157	LEU
1	A	200	THR
1	A	201	LEU
1	A	216	GLU
1	A	243	TRP
1	A	299	GLU
1	A	312	GLU
1	A	316	SER
1	A	322	SER
1	A	353	SER
1	A	369	LEU
1	A	406	VAL
1	A	429	LEU
1	A	437	LEU
1	A	440	GLN
1	A	441	LEU
1	A	470	PHE
1	B	23	MET
1	B	28	SER
1	B	65	SER
1	B	69	LEU
1	B	85	ARG
1	B	135	THR
1	B	145	ASP
1	B	148	PHE
1	B	157	LEU
1	B	170	GLN
1	B	181	LEU
1	B	192	VAL
1	B	200	THR

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	201	LEU
1	B	216	GLU
1	B	243	TRP
1	B	299	GLU
1	B	312	GLU
1	B	316	SER
1	B	322	SER
1	B	353	SER
1	B	369	LEU
1	B	406	VAL
1	B	429	LEU
1	B	437	LEU
1	B	440	GLN
1	B	441	LEU
1	B	470	PHE

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 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 1 ligands modelled in this entry, 1 is monoatomic - 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, 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	446/491 (90%)	0.59	57 (12%) 5 4	11, 134, 219, 253	0
1	B	446/491 (90%)	0.69	64 (14%) 4 3	28, 141, 235, 273	0
All	All	892/982 (90%)	0.64	121 (13%) 4 3	11, 138, 230, 273	0

All (121) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	420	ASN	9.7
1	B	446	ASN	9.0
1	B	419	PHE	8.3
1	B	11	GLN	8.0
1	A	416	PRO	7.8
1	A	11	GLN	7.5
1	A	348	LYS	7.5
1	A	424	MET	7.5
1	B	449	SER	7.5
1	B	421	SER	7.4
1	A	412	THR	7.3
1	B	282	VAL	6.9
1	A	417	LYS	6.9
1	A	449	SER	6.9
1	A	384	VAL	6.7
1	B	384	VAL	6.7
1	B	225	SER	6.7
1	A	243	TRP	6.7
1	A	381	SER	6.5
1	A	446	ASN	6.3
1	B	240	LEU	6.1
1	B	385	SER	6.0
1	B	338	PHE	5.5
1	B	470	PHE	5.4

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	105	PRO	5.3
1	B	416	PRO	5.2
1	B	353	SER	5.2
1	B	418	ALA	5.1
1	B	424	MET	5.1
1	A	346	TRP	5.1
1	A	241	VAL	5.0
1	A	242	GLY	4.9
1	A	413	LYS	4.5
1	B	412	THR	4.5
1	A	415	ALA	4.4
1	B	243	TRP	4.2
1	B	349	ASN	4.2
1	A	347	LYS	4.0
1	A	345	ALA	3.9
1	A	470	PHE	3.8
1	B	381	SER	3.8
1	B	346	TRP	3.8
1	A	425	SER	3.8
1	A	209	THR	3.7
1	B	356	LYS	3.7
1	B	209	THR	3.7
1	A	456	TYR	3.6
1	A	385	SER	3.6
1	A	226	LEU	3.6
1	A	267	TRP	3.6
1	B	226	LEU	3.6
1	A	414	LEU	3.6
1	A	105	PRO	3.6
1	A	380	THR	3.5
1	B	447	ALA	3.5
1	B	415	ALA	3.4
1	B	375	GLY	3.4
1	B	283	VAL	3.3
1	A	205	TYR	3.2
1	A	349	ASN	3.2
1	B	450	GLU	3.2
1	B	284	SER	3.2
1	B	142	ARG	3.2
1	B	239	ASN	3.1
1	B	268	MET	3.1
1	A	362	MET	3.1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	213	ALA	3.1
1	A	244	ASN	3.1
1	B	267	TRP	3.1
1	B	317	SER	3.1
1	A	80	ARG	3.0
1	B	382	GLY	3.0
1	B	342	LEU	3.0
1	B	413	LYS	3.0
1	A	444	LEU	3.0
1	B	422	GLN	2.9
1	B	245	SER	2.9
1	B	355	THR	2.9
1	A	104	LEU	2.8
1	B	380	THR	2.8
1	B	362	MET	2.8
1	B	425	SER	2.8
1	A	300	GLU	2.7
1	A	240	LEU	2.7
1	B	388	TRP	2.7
1	A	421	SER	2.6
1	B	212	LEU	2.6
1	A	216	GLU	2.5
1	A	142	ARG	2.5
1	A	377	LEU	2.5
1	A	360	GLY	2.5
1	B	373	ILE	2.4
1	A	338	PHE	2.4
1	B	285	TYR	2.4
1	A	409	SER	2.4
1	A	228	VAL	2.4
1	A	420	ASN	2.4
1	B	371	MET	2.4
1	B	456	TYR	2.4
1	A	366	LEU	2.3
1	B	109	SER	2.3
1	A	225	SER	2.3
1	B	215	GLU	2.3
1	B	386	PRO	2.3
1	B	242	GLY	2.3
1	A	140	HIS	2.2
1	B	453	TYR	2.2
1	A	303	SER	2.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	383	LYS	2.2
1	A	342	LEU	2.2
1	A	266	ALA	2.2
1	A	138	ASP	2.2
1	A	141	ASP	2.1
1	B	222	VAL	2.1
1	B	444	LEU	2.1
1	B	104	LEU	2.1
1	B	372	ALA	2.1
1	A	341	TRP	2.1
1	B	244	ASN	2.1
1	B	269	ILE	2.0
1	A	284	SER	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 [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. 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	RSCC	RSR	LLDF	B-factors(\AA^2)	Q<0.9
2	CD	B	1472	1/1	0.96	0.09	-	99,99,99,99	0

6.5 Other polymers [i](#)

There are no such residues in this entry.