Full wwPDB X-ray Structure Validation Report (i)

Mar 10, 2026 – 12:44 AM UTC

PDB ID	:	5U02 / pdb_00005u02
Title	:	Crystal structure of S. aureus TarS 217-571
Authors	:	Worrall, L.J.; Sobhanifar, S.; King, D.T.; Strynadka, N.C.
Deposited on	:	2016-11-22
Resolution	:	2.30 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
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 (1)) were used in the production of this report:

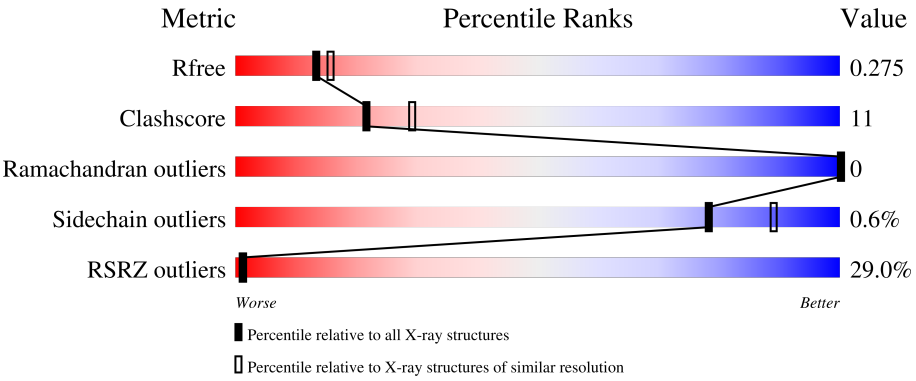
MolProbity	:	4-5-2 with Phenix2.0
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	2.0
EDS	:	3.0
Percentile statistics	:	20250101.v01 (using entries in the PDB archive January 1st 2025)
CCP4	:	9.0.010 (Gargrove)
Density-Fitness	:	1.0.12
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.49

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 2.30 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}	180053	6319 (2.30-2.30)
Clashscore	190562	6919 (2.30-2.30)
Ramachandran outliers	187476	6854 (2.30-2.30)
Sidechain outliers	187428	6854 (2.30-2.30)
RSRZ outliers	180081	6325 (2.30-2.30)

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	356	29%, 80%, 20%

2 Entry composition

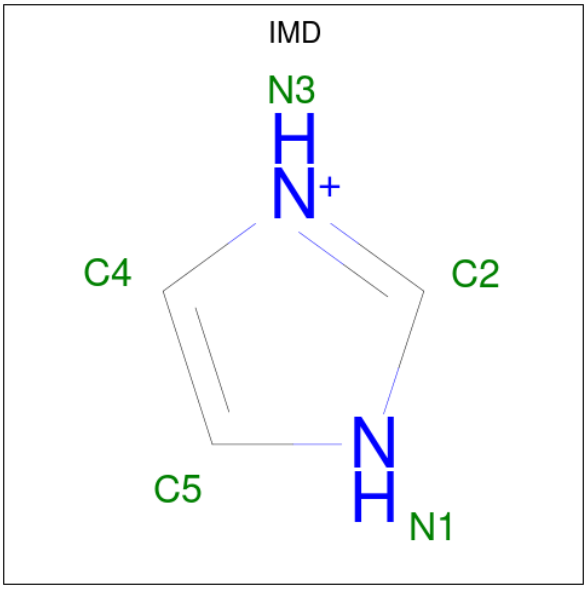
There are 3 unique types of molecules in this entry. The entry contains 3100 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Glycosyl transferase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	355	2916	1875	489	540	12	0	0	0

- Molecule 2 is IMIDAZOLE (CCD ID: IMD) (formula: $C_3H_5N_2$).



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

- Molecule 3 is water.

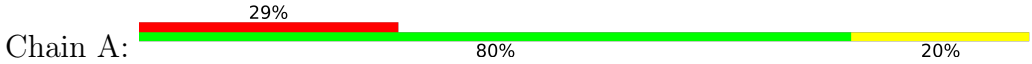
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
3	A	179	179	179	0	0

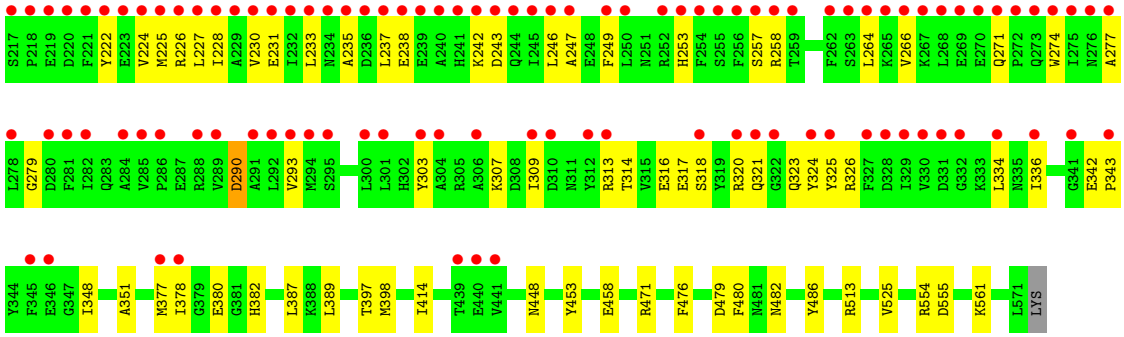
3 Residue-property plots (i)

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: Glycosyl transferase

Chain A:





4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, α, β, γ	105.69Å 105.69Å 80.50Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	52.83 – 2.30 52.83 – 2.30	Depositor EDS
% Data completeness (in resolution range)	98.1 (52.83-2.30) 98.3 (52.83-2.30)	Depositor EDS
R_{merge}	0.14	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.61 (at 2.29Å)	Xtriage
Refinement program	PHENIX (1.10_2155)	Depositor
R, R_{free}	0.224 , 0.268 0.227 , 0.275	Depositor DCC
R_{free} test set	1229 reflections (5.34%)	wwPDB-VP
Wilson B-factor (Å^2)	33.5	Xtriage
Anisotropy	0.019	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 49.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.34$	Xtriage
Estimated twinning fraction	0.029 for -h,-k,l	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	3100	wwPDB-VP
Average B, all atoms (Å^2)	53.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.35% 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: IMD

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	0.20	0/2976	0.43	0/4013

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	2916	0	2911	62	0
2	A	5	0	5	3	0
3	A	179	0	0	8	3
All	All	3100	0	2916	63	3

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 11.

All (63) 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:476:PHE:O	3:A:701:HOH:O	1.86	0.93

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:471:ARG:HH12	2:A:601:IMD:H2	1.36	0.91
1:A:380:GLU:OE1	3:A:702:HOH:O	1.97	0.81
1:A:321:GLN:HE21	1:A:323:GLN:HB3	1.46	0.79
1:A:448:ASN:HD22	1:A:482:ASN:HD21	1.33	0.75
1:A:321:GLN:NE2	1:A:323:GLN:HB3	2.03	0.73
1:A:266:VAL:O	1:A:271:GLN:NE2	2.21	0.72
1:A:321:GLN:HE21	1:A:323:GLN:H	1.37	0.71
1:A:513:ARG:NH2	3:A:708:HOH:O	2.22	0.71
1:A:317:GLU:HB3	1:A:320:ARG:HH21	1.56	0.71
1:A:453:TYR:O	3:A:703:HOH:O	2.10	0.69
1:A:317:GLU:O	1:A:320:ARG:HG2	1.95	0.67
1:A:233:LEU:O	1:A:242:LYS:NZ	2.30	0.64
1:A:316:GLU:OE2	3:A:705:HOH:O	2.14	0.64
1:A:321:GLN:HE21	1:A:323:GLN:N	1.95	0.64
1:A:342:GLU:HG3	1:A:343:PRO:HD2	1.81	0.63
1:A:235:ALA:O	1:A:242:LYS:NZ	2.22	0.61
1:A:321:GLN:NE2	1:A:323:GLN:H	1.99	0.60
1:A:293:VAL:O	3:A:706:HOH:O	2.16	0.60
1:A:525:VAL:HB	1:A:561:LYS:HB3	1.85	0.59
1:A:448:ASN:HD22	1:A:482:ASN:ND2	2.01	0.58
1:A:318:SER:HA	1:A:321:GLN:HE22	1.68	0.58
1:A:264:LEU:HD11	1:A:316:GLU:HB3	1.85	0.58
1:A:318:SER:HB3	1:A:324:TYR:CD1	2.39	0.57
1:A:321:GLN:HE21	1:A:323:GLN:CB	2.17	0.57
1:A:237:LEU:HG	1:A:238:GLU:H	1.69	0.57
1:A:227:LEU:O	1:A:231:GLU:HG3	2.06	0.56
1:A:243:ASP:O	1:A:247:ALA:N	2.34	0.55
1:A:313:ARG:HG2	1:A:317:GLU:HG3	1.88	0.55
1:A:336:ILE:HD12	1:A:348:ILE:HD12	1.89	0.54
1:A:303:TYR:OH	1:A:342:GLU:HG2	2.08	0.53
1:A:471:ARG:NH1	2:A:601:IMD:H2	2.16	0.52
2:A:601:IMD:N1	3:A:704:HOH:O	2.11	0.52
1:A:314:THR:O	1:A:318:SER:N	2.40	0.52
1:A:389:LEU:HB2	1:A:398:MET:HB2	1.93	0.51
1:A:351:ALA:H	1:A:377:MET:CE	2.23	0.51
1:A:290:ASP:N	1:A:290:ASP:OD1	2.44	0.50
1:A:387:LEU:HD23	1:A:414:ILE:HD11	1.93	0.49
1:A:226:ARG:O	1:A:230:VAL:HG23	2.12	0.49
1:A:486:TYR:OH	3:A:707:HOH:O	2.19	0.48
1:A:271:GLN:HB3	1:A:309:ILE:CD1	2.44	0.46
1:A:554:ARG:HG3	1:A:555:ASP:OD2	2.16	0.45

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:279:GLY:HA3	1:A:307:LYS:HE2	1.99	0.45
1:A:253:HIS:O	1:A:257:SER:OG	2.33	0.45
1:A:318:SER:HB3	1:A:324:TYR:CE1	2.52	0.45
1:A:378:ILE:HD11	1:A:382:HIS:HB2	1.99	0.44
1:A:237:LEU:CG	1:A:238:GLU:H	2.31	0.44
1:A:334:LEU:HD23	1:A:336:ILE:HD11	2.00	0.44
1:A:321:GLN:HE21	1:A:323:GLN:CA	2.30	0.43
1:A:561:LYS:HE3	1:A:561:LYS:HB2	1.85	0.43
1:A:224:VAL:O	1:A:228:ILE:HG13	2.19	0.43
1:A:325:TYR:O	1:A:326:ARG:HD2	2.18	0.43
1:A:222:TYR:HB2	1:A:277:ALA:HB1	2.01	0.42
1:A:225:MET:HE3	1:A:249:PHE:HZ	1.84	0.42
1:A:258:ARG:HD3	1:A:274:TRP:HH2	1.83	0.42
1:A:313:ARG:O	1:A:317:GLU:HG3	2.20	0.42
1:A:317:GLU:HA	1:A:320:ARG:HE	1.83	0.42
1:A:243:ASP:HA	1:A:246:LEU:HB3	2.01	0.42
1:A:317:GLU:HB3	1:A:320:ARG:NH2	2.28	0.42
1:A:458:GLU:HB3	1:A:513:ARG:HH11	1.85	0.41
1:A:479:ASP:OD1	1:A:480:PHE:N	2.53	0.41
1:A:351:ALA:H	1:A:377:MET:HE2	1.86	0.41
1:A:324:TYR:CE1	1:A:336:ILE:HD13	2.57	0.40

All (3) 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)
3:A:772:HOH:O	3:A:856:HOH:O[4_556]	1.76	0.44
3:A:709:HOH:O	3:A:824:HOH:O[3_565]	2.01	0.19
3:A:707:HOH:O	3:A:707:HOH:O[4_556]	2.14	0.06

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	353/356 (99%)	343 (97%)	10 (3%)	0	100 100

There are no Ramachandran outliers to report.

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	319/322 (99%)	317 (99%)	2 (1%)	78 89

All (2) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	290	ASP
1	A	397	THR

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (6) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	311	ASN
1	A	321	GLN
1	A	339	ASN
1	A	365	HIS
1	A	396	HIS
1	A	448	ASN

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 oligosaccharides in this entry.

5.6 Ligand geometry (i)

1 ligand is 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	IMD	A	601	-	5,5,5	0.60	0	5,5,5	0.45	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	IMD	A	601	-	-	-	0/1/1/1

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.

1 monomer is involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	601	IMD	3	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	A	355/356 (99%)	1.18	103 (29%) 1 1	21, 42, 105, 134	0

All (103) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	A	237	LEU	6.6
1	A	250	LEU	5.2
1	A	227	LEU	5.2
1	A	221	PHE	5.0
1	A	240	ALA	4.9
1	A	329	ILE	4.8
1	A	321	GLN	4.5
1	A	241	HIS	4.5
1	A	243	ASP	4.4
1	A	441	VAL	4.3
1	A	274	TRP	4.2
1	A	258	ARG	4.2
1	A	217	SER	4.1
1	A	233	LEU	4.1
1	A	341	GLY	4.0
1	A	292	LEU	4.0
1	A	273	GLN	3.9
1	A	264	LEU	3.9
1	A	225	MET	3.8
1	A	330	VAL	3.8
1	A	218	PRO	3.7
1	A	252	ARG	3.7
1	A	293	VAL	3.7
1	A	310	ASP	3.7
1	A	229	ALA	3.7
1	A	228	ILE	3.6
1	A	230	VAL	3.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	245	ILE	3.6
1	A	309	ILE	3.6
1	A	327	PHE	3.5
1	A	268	LEU	3.4
1	A	254	PHE	3.4
1	A	219	GLU	3.4
1	A	236	ASP	3.4
1	A	224	VAL	3.4
1	A	378	ILE	3.3
1	A	304	ALA	3.3
1	A	275	ILE	3.2
1	A	222	TYR	3.2
1	A	320	ARG	3.2
1	A	331	ASP	3.2
1	A	232	ILE	3.2
1	A	220	ASP	3.1
1	A	256	PHE	3.1
1	A	312	TYR	3.1
1	A	306	ALA	3.0
1	A	300	LEU	3.0
1	A	238	GLU	3.0
1	A	234	ASN	3.0
1	A	259	THR	3.0
1	A	272	PRO	3.0
1	A	289	VAL	3.0
1	A	282	ILE	2.9
1	A	235	ALA	2.9
1	A	239	GLU	2.9
1	A	265	LYS	2.9
1	A	267	LYS	2.8
1	A	246	LEU	2.8
1	A	223	GLU	2.8
1	A	244	GLN	2.8
1	A	255	SER	2.8
1	A	345	PHE	2.8
1	A	377	MET	2.7
1	A	303	TYR	2.7
1	A	285	VAL	2.6
1	A	294	MET	2.6
1	A	313	ARG	2.6
1	A	281	PHE	2.6
1	A	253	HIS	2.6

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	288	ARG	2.6
1	A	324	TYR	2.6
1	A	343	PRO	2.5
1	A	269	GLU	2.5
1	A	276	ASN	2.5
1	A	247	ALA	2.5
1	A	277	ALA	2.4
1	A	270	GLU	2.4
1	A	266	VAL	2.4
1	A	334	LEU	2.4
1	A	263	SER	2.4
1	A	322	GLY	2.4
1	A	440	GLU	2.4
1	A	291	ALA	2.3
1	A	226	ARG	2.3
1	A	295	SER	2.3
1	A	346	GLU	2.3
1	A	332	GLY	2.3
1	A	325	TYR	2.3
1	A	249	PHE	2.2
1	A	231	GLU	2.2
1	A	286	PRO	2.2
1	A	318	SER	2.2
1	A	301	LEU	2.2
1	A	242	LYS	2.2
1	A	328	ASP	2.1
1	A	278	LEU	2.1
1	A	280	ASP	2.1
1	A	262	PHE	2.1
1	A	271	GLN	2.1
1	A	257	SER	2.0
1	A	439	THR	2.0
1	A	284	ALA	2.0
1	A	336	ILE	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 oligosaccharides 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	IMD	A	601	5/5	0.85	0.20	49,50,54,55	0

6.5 Other polymers (i)

There are no such residues in this entry.