Full wwPDB X-ray Structure Validation Report (i)

Sep 12, 2023 – 05:48 PM EDT

PDB ID	:	4NEX
Title	:	Structure of the N-acetyltransferase domain of X. fastidiosa NAGS/K
Authors	:	Zhao, G.; Jin, Z.; Allewell, N.M.; Tuchman, M.; Shi, D.
Deposited on	:	2013-10-30
Resolution	:	1.70 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 (i)) were used in the production of this report:

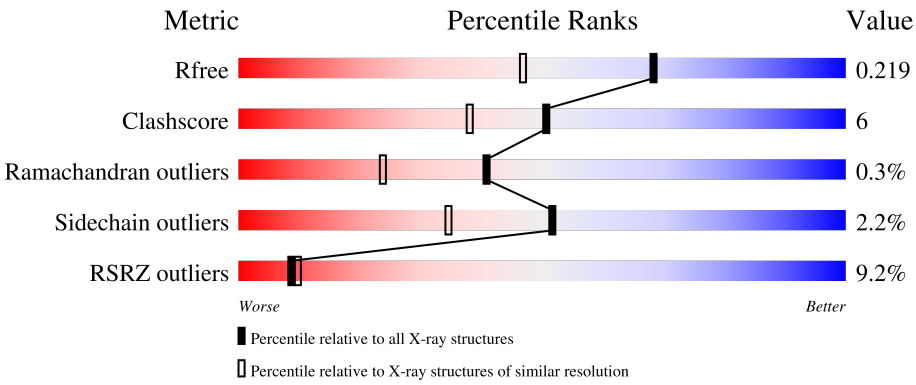
MolProbity	:	4.02b-467
Mogul	:	1.8.5 (274361), CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.35.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.35.1

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 1.70 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	4298 (1.70-1.70)
Clashscore	141614	4695 (1.70-1.70)
Ramachandran outliers	138981	4610 (1.70-1.70)
Sidechain outliers	138945	4610 (1.70-1.70)
RSRZ outliers	127900	4222 (1.70-1.70)

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	167	2%, 84%, 8%, • 6%
1	B	167	14%, 78%, 9%, • 13%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 2817 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 Acetylglutamate kinase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	157	1320	845	240	231	4	0	3	0
			Total	C	N	O	S			
1	B	146	1226	790	219	213	4	0	1	0

There are 32 discrepancies between the modelled and reference sequences:

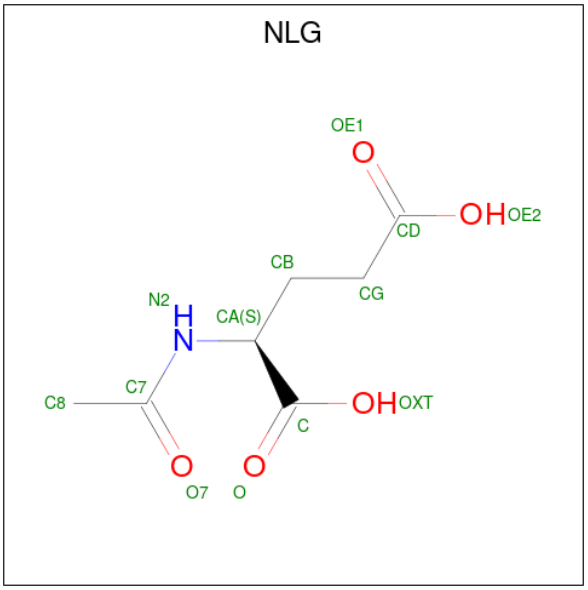
Chain	Residue	Modelled	Actual	Comment	Reference
A	272	MET	-	expression tag	UNP Q87EL2
A	273	GLY	-	expression tag	UNP Q87EL2
A	274	SER	-	expression tag	UNP Q87EL2
A	275	SER	-	expression tag	UNP Q87EL2
A	276	HIS	-	expression tag	UNP Q87EL2
A	277	HIS	-	expression tag	UNP Q87EL2
A	278	HIS	-	expression tag	UNP Q87EL2
A	279	HIS	-	expression tag	UNP Q87EL2
A	281	HIS	LYS	conflict	UNP Q87EL2
A	282	SER	GLY	conflict	UNP Q87EL2
A	?	-	THR	deletion	UNP Q87EL2
A	286	VAL	ILE	conflict	UNP Q87EL2
A	287	PRO	ARG	conflict	UNP Q87EL2
A	290	SER	-	expression tag	UNP Q87EL2
A	291	HIS	GLU	conflict	UNP Q87EL2
A	292	MET	ARG	conflict	UNP Q87EL2
B	272	MET	-	expression tag	UNP Q87EL2
B	273	GLY	-	expression tag	UNP Q87EL2
B	274	SER	-	expression tag	UNP Q87EL2
B	275	SER	-	expression tag	UNP Q87EL2
B	276	HIS	-	expression tag	UNP Q87EL2
B	277	HIS	-	expression tag	UNP Q87EL2
B	278	HIS	-	expression tag	UNP Q87EL2
B	279	HIS	-	expression tag	UNP Q87EL2
B	281	HIS	LYS	conflict	UNP Q87EL2

Continued on next page...

Continued from previous page...

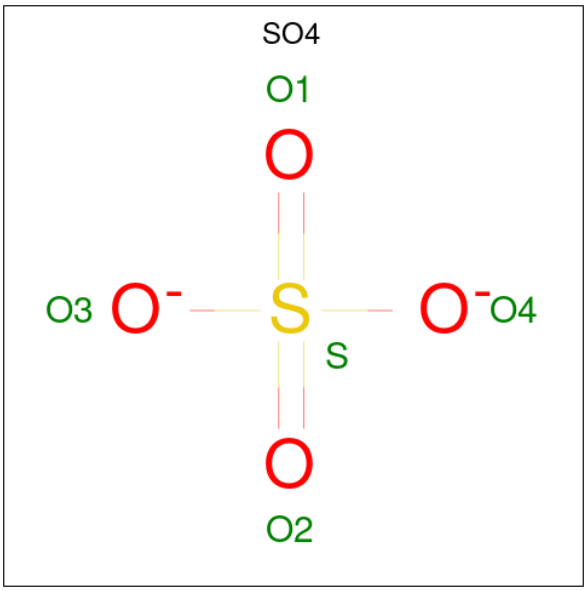
Chain	Residue	Modelled	Actual	Comment	Reference
B	282	SER	GLY	conflict	UNP Q87EL2
B	?	-	THR	deletion	UNP Q87EL2
B	286	VAL	ILE	conflict	UNP Q87EL2
B	287	PRO	ARG	conflict	UNP Q87EL2
B	290	SER	-	expression tag	UNP Q87EL2
B	291	HIS	GLU	conflict	UNP Q87EL2
B	292	MET	ARG	conflict	UNP Q87EL2

- Molecule 2 is N-ACETYL-L-GLUTAMATE (three-letter code: NLG) (formula: $C_7H_{11}NO_5$).



Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	13	7	1	5	0	0
			Total	C	N	O		
2	B	1	13	7	1	5	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	1	Total O S 5 4 1	0	0
3	A	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	155	Total O 155 155	0	0
4	B	75	Total O 75 75	0	0

3 Residue-property plots (i)

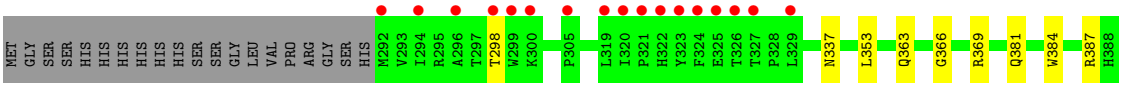
These plots are drawn for all protein, RNA, DNA and oligosaccharide chains in the entry. The first graphic for a chain summarises the proportions of the various outlier classes displayed in the second graphic. The second graphic shows the sequence view annotated by issues in geometry and electron density. 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: Acetylglutamate kinase

Chain A:

- Molecule 1: Acetylglutamate kinase

Chain B:



4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	51.72Å 51.72Å 242.31Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	39.33 – 1.70 39.33 – 1.70	Depositor EDS
% Data completeness (in resolution range)	98.5 (39.33-1.70) 95.9 (39.33-1.70)	Depositor EDS
R_{merge}	0.08	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	3.28 (at 1.70Å)	Xtriage
Refinement program	PHENIX 1.8.2_1309	Depositor
R, R_{free}	0.189 , 0.219 0.189 , 0.219	Depositor DCC
R_{free} test set	1866 reflections (5.00%)	wwPDB-VP
Wilson B-factor (Å^2)	26.1	Xtriage
Anisotropy	0.156	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.36 , 49.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.30$	Xtriage
Estimated twinning fraction	No twinning to report.	Xtriage
F_o,F_c correlation	0.96	EDS
Total number of atoms	2817	wwPDB-VP
Average B, all atoms (Å^2)	39.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 5.58% 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 (i)

5.1 Standard geometry (i)

Bond lengths and bond angles in the following residue types are not validated in this section: SO4, NLG

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	0.38	0/1370	0.56	0/1862
1	B	0.33	0/1268	0.49	0/1726
All	All	0.35	0/2638	0.53	0/3588

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 (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	A	1320	0	1281	17	0
1	B	1226	0	1187	12	1
2	A	13	0	9	4	0
2	B	13	0	9	0	0
3	A	10	0	0	0	0
3	B	5	0	0	1	0
4	A	155	0	0	6	1
4	B	75	0	0	7	2
All	All	2817	0	2486	29	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 6.

All (29) 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:369:ARG:NE	4:A:798:HOH:O	1.91	0.94
1:A:339:ARG:NE	4:A:829:HOH:O	2.02	0.76
1:A:282:SER:HA	4:A:820:HOH:O	1.90	0.72
1:B:381:GLN:NE2	4:B:768:HOH:O	2.22	0.71
1:B:366:GLY:HA2	1:B:369:ARG:HH11	1.56	0.70
1:A:387:ARG:NH1	2:A:600:NLG:OE2	2.26	0.69
1:B:363:GLN:NE2	4:B:771:HOH:O	2.26	0.69
1:B:420:GLN:NE2	4:B:723:HOH:O	2.27	0.67
1:A:288:ARG:HB2	1:A:374:ILE:HD11	1.75	0.67
1:A:288:ARG:NH1	1:A:377:GLU:OE2	2.29	0.66
1:A:282:SER:N	4:A:813:HOH:O	2.30	0.64
1:B:369:ARG:NH1	4:B:703:HOH:O	2.37	0.57
1:A:418:HIS:CE1	1:A:420:GLN:HB3	2.41	0.56
1:A:339:ARG:NH2	1:A:361:ASP:OD2	2.38	0.56
1:A:396:TYR:OH	2:A:600:NLG:H8C3	2.06	0.55
1:B:366:GLY:HA2	1:B:369:ARG:NH1	2.22	0.55
1:A:405:LYS:NZ	4:A:845:HOH:O	2.39	0.54
1:A:283:SER:HA	1:A:295:ARG:O	2.09	0.53
1:B:391:GLN:NE2	4:B:724:HOH:O	2.43	0.51
1:B:389:ASN:OD1	1:B:389:ASN:N	2.41	0.51
1:A:402:GLY:HA3	1:A:413:TRP:CE2	2.49	0.47
1:B:337:ASN:ND2	3:B:601:SO4:O4	2.30	0.45
1:B:298:THR:HA	4:B:769:HOH:O	2.17	0.45
1:A:387:ARG:HH11	2:A:600:NLG:CD	2.29	0.45
1:B:431:GLN:HG3	4:B:745:HOH:O	2.19	0.43
1:A:369:ARG:CZ	4:A:798:HOH:O	2.50	0.42
1:A:353:LEU:CD2	2:A:600:NLG:H8C2	2.50	0.42
1:B:353:LEU:HB3	1:B:384:TRP:HB3	2.02	0.41
1:A:339:ARG:HH22	1:A:361:ASP:CG	2.25	0.41

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)
4:B:714:HOH:O	4:B:729:HOH:O[7_555]	1.93	0.27
1:B:406:GLN:OE1	4:A:793:HOH:O[6_544]	2.01	0.19
4:B:749:HOH:O	4:B:750:HOH:O[7_555]	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	158/167 (95%)	155 (98%)	2 (1%)	1 (1%)	25 11
1	B	145/167 (87%)	139 (96%)	6 (4%)	0	100 100
All	All	303/334 (91%)	294 (97%)	8 (3%)	1 (0%)	41 24

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	283	SER

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	141/147 (96%)	139 (99%)	2 (1%)	67 53
1	B	130/147 (88%)	126 (97%)	4 (3%)	40 21
All	All	271/294 (92%)	265 (98%)	6 (2%)	52 34

All (6) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	282	SER
1	A	361	ASP
1	B	387	ARG
1	B	389	ASN
1	B	408	HIS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	418	HIS

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)

5 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	NLG	A	600	-	12,12,12	1.39	1 (8%)	15,15,15	2.08	4 (26%)
3	SO4	A	601	-	4,4,4	0.22	0	6,6,6	0.26	0
2	NLG	B	600	-	12,12,12	1.44	1 (8%)	15,15,15	1.49	1 (6%)
3	SO4	A	602	-	4,4,4	0.15	0	6,6,6	0.13	0
3	SO4	B	601	-	4,4,4	0.13	0	6,6,6	0.14	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	NLG	A	600	-	-	6/13/13/13	-
2	NLG	B	600	-	-	2/13/13/13	-

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	B	600	NLG	C7-N2	3.63	1.46	1.34
2	A	600	NLG	C7-N2	3.04	1.44	1.34

All (5) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	A	600	NLG	C-CA-N2	-5.35	97.88	110.55
2	A	600	NLG	CG-CB-CA	3.38	119.47	113.16
2	B	600	NLG	CG-CB-CA	3.32	119.37	113.16
2	A	600	NLG	C8-C7-N2	-2.23	112.32	116.10
2	A	600	NLG	OE2-CD-CG	2.06	120.66	114.03

There are no chirality outliers.

All (8) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	600	NLG	N2-CA-CB-CG
2	A	600	NLG	OXT-C-CA-N2
2	A	600	NLG	O-C-CA-N2
2	B	600	NLG	OE2-CD-CG-CB
2	B	600	NLG	OE1-CD-CG-CB
2	A	600	NLG	OE2-CD-CG-CB
2	A	600	NLG	OE1-CD-CG-CB
2	A	600	NLG	C-CA-N2-C7

There are no ring outliers.

2 monomers are involved in 5 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	600	NLG	4	0
3	B	601	SO4	1	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 (i)

6.1 Protein, DNA and RNA chains (i)

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	157/167 (94%)	-0.01	4 (2%)	57	61	18, 28, 59, 76	0
1	B	146/167 (87%)	0.67	24 (16%)	1	1	20, 43, 74, 87	0
All	All	303/334 (90%)	0.32	28 (9%)	9	10	18, 34, 66, 87	0

All (28) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	324	PHE	7.5
1	B	320	ILE	6.2
1	B	299	TRP	4.9
1	A	282	SER	4.0
1	B	323	TYR	4.0
1	B	300	LYS	4.0
1	B	321	PRO	4.0
1	B	326	THR	3.5
1	A	283	SER	3.5
1	A	289	GLY	3.3
1	B	327	THR	3.0
1	B	298	THR	3.0
1	B	296	ALA	3.0
1	B	322	HIS	2.9
1	B	431	GLN	2.9
1	B	319	LEU	2.8
1	B	389	ASN	2.8
1	B	325	GLU	2.6
1	B	434	PRO	2.4
1	B	294	ILE	2.3
1	B	437	ILE	2.3
1	B	305	PRO	2.3
1	B	430	THR	2.3
1	B	408	HIS	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	407	ASP	2.2
1	B	292	MET	2.2
1	A	284	GLY	2.2
1	B	329	LEU	2.1

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, 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	B-factors(Å^2)	Q<0.9
2	NLG	A	600	13/13	0.89	0.21	27,48,74,83	0
2	NLG	B	600	13/13	0.90	0.16	28,44,55,60	0
3	SO4	A	601	5/5	0.92	0.16	37,43,48,48	0
3	SO4	A	602	5/5	0.92	0.10	71,72,77,79	0
3	SO4	B	601	5/5	0.98	0.10	54,60,75,77	0

6.5 Other polymers (i)

There are no such residues in this entry.