wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 11:04 AM GMT

PDB ID	:	3NZQ
Title	:	Crystal Structure of Biosynthetic arginine decarboxylase ADC (SpeA) from Escherichia coli, Northeast Structural Genomics Consortium Target ER600
Authors	:	Forouhar, F.; Lew, S.; Seetharaman, J.; Sahdev, S.; Xiao, R.; Ciccosanti, C.; Belote, R.L.; Everett, J.K.; Nair, R.; Acton, T.B.; Rost, B.; Montelione, G.T.; Hunt, J.F.; Tong, L.; Northeast Structural Genomics Consortium (NESG)
Deposited on	:	2010-07-16
Resolution	:	3.10 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
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (1)) were used in the production of this report:

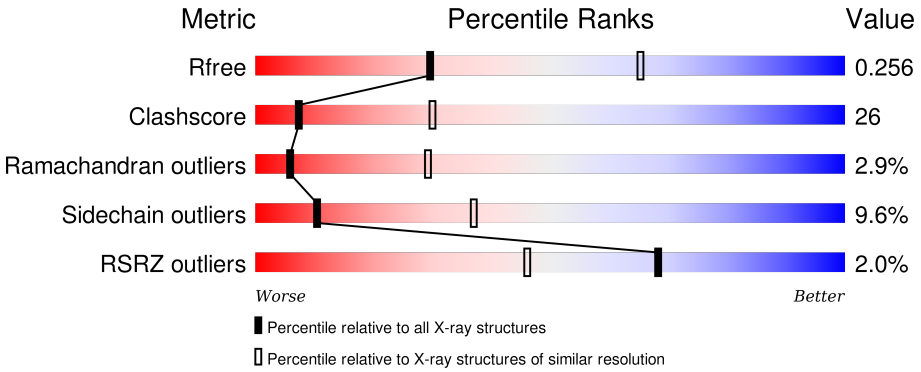
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	1.9-1692
EDS	:	rb-20026688
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Refmac	:	5.8.0135
CCP4	:	6.5.0
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	trunk26865

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
X-RAY DIFFRACTION

The reported resolution of this entry is 3.10 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	1114 (3.14-3.06)
Clashscore	102246	1222 (3.14-3.06)
Ramachandran outliers	100387	1174 (3.14-3.06)
Sidechain outliers	100360	1174 (3.14-3.06)
RSRZ outliers	91569	1119 (3.14-3.06)

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 <=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	666	2% 52% 36% 6% 6%
1	B	666	2% 51% 37% 6% 6%

2 Entry composition (i)

There are 3 unique types of molecules in this entry. The entry contains 10064 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.

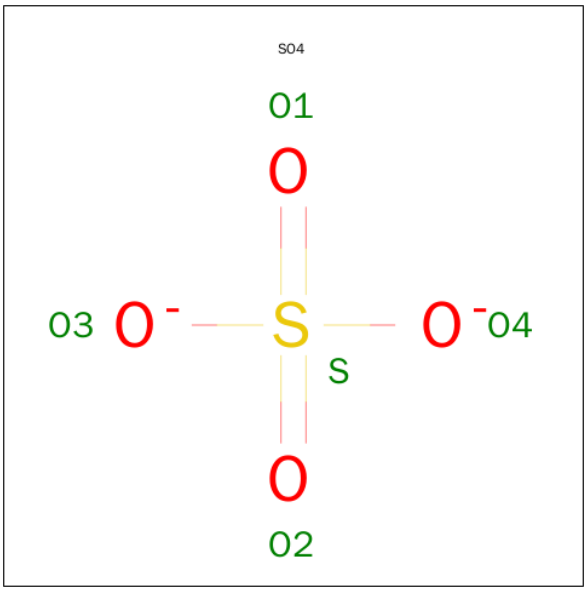
- Molecule 1 is a protein called Biosynthetic arginine decarboxylase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	628	Total	C	N	O	S	0	0	0
			4984	3142	862	953	27			
1	B	628	Total	C	N	O	S	0	0	0
			4984	3142	862	953	27			

There are 16 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	659	LEU	-	expression tag	UNP P21170
A	660	GLU	-	expression tag	UNP P21170
A	661	HIS	-	expression tag	UNP P21170
A	662	HIS	-	expression tag	UNP P21170
A	663	HIS	-	expression tag	UNP P21170
A	664	HIS	-	expression tag	UNP P21170
A	665	HIS	-	expression tag	UNP P21170
A	666	HIS	-	expression tag	UNP P21170
B	659	LEU	-	expression tag	UNP P21170
B	660	GLU	-	expression tag	UNP P21170
B	661	HIS	-	expression tag	UNP P21170
B	662	HIS	-	expression tag	UNP P21170
B	663	HIS	-	expression tag	UNP P21170
B	664	HIS	-	expression tag	UNP P21170
B	665	HIS	-	expression tag	UNP P21170
B	666	HIS	-	expression tag	UNP P21170

- Molecule 2 is SULFATE ION (three-letter code: SO4) (formula: O_4S).



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

- Molecule 3 is water.

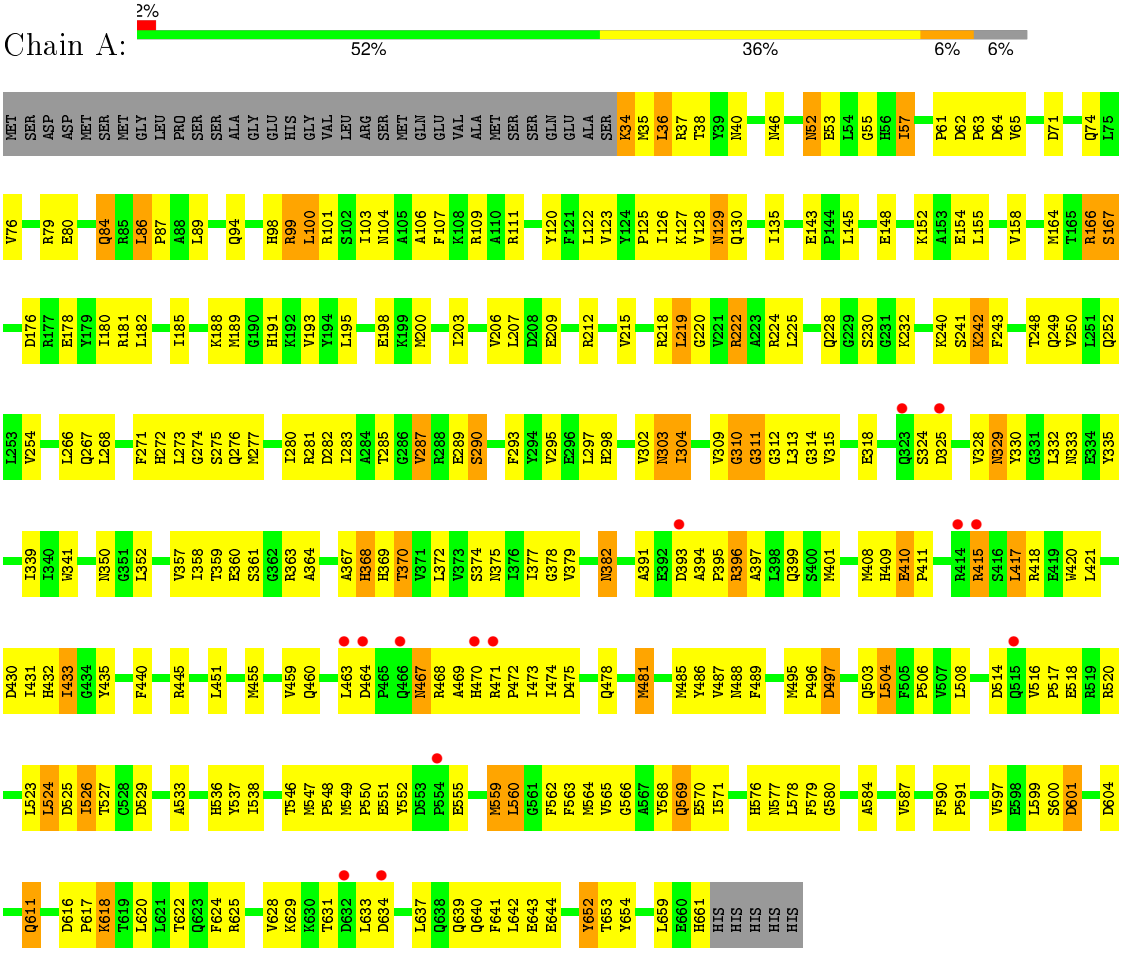
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
3	A	36	Total O 36 36	0	0
3	B	40	Total O 40 40	0	0

3 Residue-property plots (i)

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.

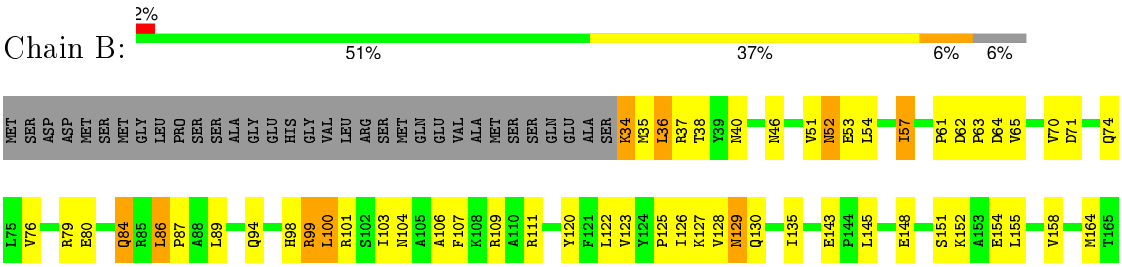
- Molecule 1: Biosynthetic arginine decarboxylase

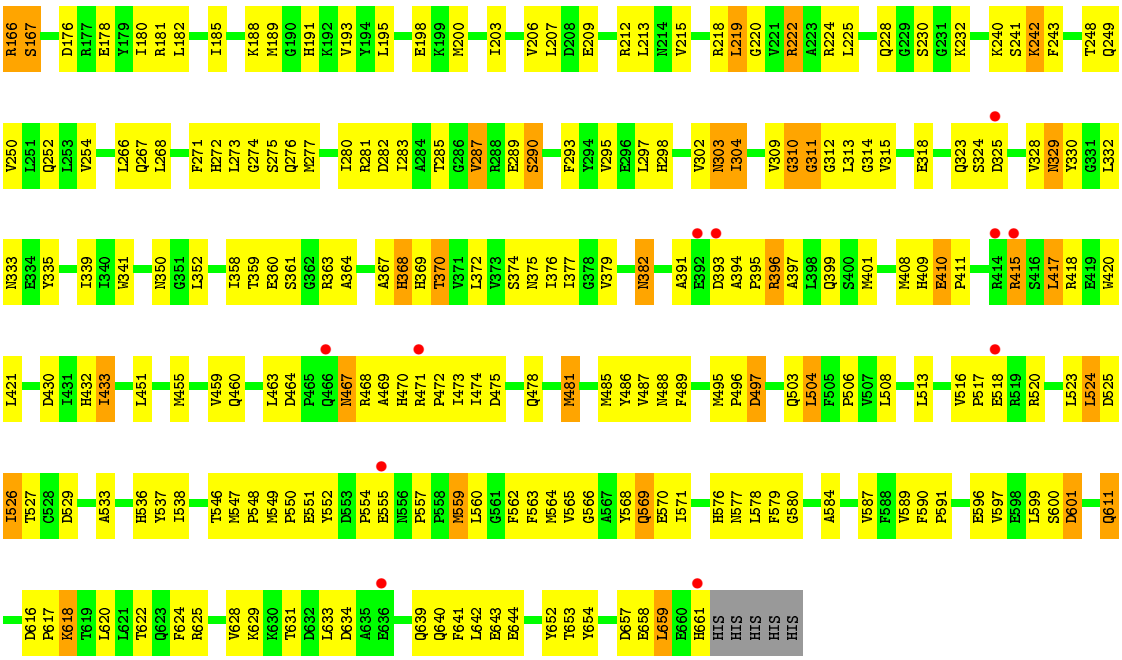
Chain A:



- Molecule 1: Biosynthetic arginine decarboxylase

Chain B:




R166 S167 D176 R177 E178 Y179 I180 R181 L182 I185 K188 M189 G190 H191 K192 V193 Y194 L195 E198 K199 M200 I203 V206 L207 D208 E209 R212 L213 W214 V215 R218 L219 G220 V221 R222 A223 R224 L225 Q228 G229 S230 G231 K232 K240 S241 K242 F243 T248 Q249
V250 L251 Q252 L253 V254 L266 Q267 L268 F271 H272 L273 G274 S275 Q276 M277 I280 R281 D282 I283 A284 T285 G286 V287 R288 E289 S290 P293 Y294 V295 E296 L297 H298 V302 N303 I304 V309 G310 G311 G312 L313 G314 V315 E318 Q323 S324 D325 V328 N329 Y330 G331 L332
N333 E334 Y335 I339 I340 W341 N350 G351 L352 I358 T359 E360 S361 G362 R363 A364 A367 H368 H369 T370 V371 L372 V373 S374 N375 I376 I377 G378 V379 N382 A391 E392 D393 A394 P395 R396 A397 L398 Q399 S400 M401 M408 H409 E410 P411 R414 R415 S416 L417 R418 E419 W420
L421 D430 I431 H432 I433 L451 M455 V459 Q460 L463 D464 P465 Q466 N467 R468 A469 H470 R471 P472 I473 I474 D475 Q478 M481 M485 Y486 V487 N488 F489 M495 P496 D497 Q503 L504 P505 P506 V507 L508 L513 V516 P517 E518 R519 R520 L523 L524 D525
I526 T527 C528 D529 A533 H536 Y537 I538 T546 M547 P548 M549 P550 E551 Y552 D553 P554 E555 N556 P557 P558 M559 L560 G561 F562 F563 M564 V565 G566 A567 Y568 Q569 E570 I571 H576 N577 L578 F579 G580 A584 V587 F588 V589 F590 P591 E596 V597 E598 L599 S600 D601 Q611
D616 P617 K618 T619 L620 L621 T622 Q623 F624 R625 V628 K629 K630 T631 D632 L633 D634 A635 E636 Q639 Q640 F641 L642 E643 E644 Y652 T653 Y654 D657 E658 L659 E660 H661 HIS HIS HIS HIS HIS

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 41 21 2	Depositor
Cell constants a, b, c, α, β, γ	192.52Å 192.52Å 119.89Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	19.94 – 3.10 29.97 – 3.09	Depositor EDS
% Data completeness (in resolution range)	82.2 (19.94-3.10) 92.0 (29.97-3.09)	Depositor EDS
R_{merge}	0.17	Depositor
R_{sym}	0.13	Depositor
$< I/\sigma(I) >$ [1]	3.58 (at 3.11Å)	Xtriage
Refinement program	CNS 1.2	Depositor
R, R_{free}	0.200 , 0.239 0.219 , 0.256	Depositor DCC
R_{free} test set	4133 reflections (12.08%)	DCC
Wilson B-factor (Å^2)	65.3	Xtriage
Anisotropy	0.566	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.32 , 63.0	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning[2]	$< \|L\| > = 0.48$, $< L^2 > = 0.32$	Xtriage
Outliers	0 of 41549 reflections	Xtriage
F_o,F_c correlation	0.92	EDS
Total number of atoms	10064	wwPDB-VP
Average B, all atoms (Å^2)	74.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 2.65% 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, 0.375 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: SO4

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	0.47	0/5094	0.61	1/6910 (0.0%)
1	B	0.44	0/5094	0.61	1/6910 (0.0%)
All	All	0.45	0/10188	0.61	2/13820 (0.0%)

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	290	SER	CB-CA-C	-6.64	97.48	110.10
1	B	290	SER	CB-CA-C	-6.64	97.48	110.10

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	4984	0	4830	257	0
1	B	4984	0	4830	249	0
2	A	10	0	0	0	0
2	B	10	0	0	0	0
3	A	36	0	0	3	0
3	B	40	0	0	1	0
All	All	10064	0	9660	506	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 26.

The worst 5 of 506 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:B:212:ARG:HB3	1:B:212:ARG:NH2	1.86	0.90
1:B:126:ILE:HG22	1:B:130:GLN:HG3	1.53	0.90
1:A:212:ARG:NH2	1:A:212:ARG:HB3	1.86	0.90
1:B:297:LEU:O	1:B:302:VAL:HG12	1.73	0.89
1:A:631:THR:HG23	1:A:633:LEU:H	1.37	0.88

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	626/666 (94%)	552 (88%)	56 (9%)	18 (3%)	6	29
1	B	626/666 (94%)	549 (88%)	59 (9%)	18 (3%)	6	29
All	All	1252/1332 (94%)	1101 (88%)	115 (9%)	36 (3%)	6	29

5 of 36 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	86	LEU
1	A	230	SER
1	A	576	HIS
1	B	86	LEU
1	B	230	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	534/566 (94%)	483 (90%)	51 (10%)	10 37
1	B	534/566 (94%)	483 (90%)	51 (10%)	10 37
All	All	1068/1132 (94%)	966 (90%)	102 (10%)	10 37

5 of 102 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	601	ASP
1	B	100	LEU
1	B	559	MET
1	A	611	GLN
1	B	34	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 50 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	569	GLN
1	B	52	ASN
1	B	569	GLN
1	A	611	GLN
1	B	40	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 carbohydrates in this entry.

5.6 Ligand geometry (i)

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	$\#\|Z\| > 2$
2	SO4	A	701	-	4,4,4	0.24	0	6,6,6	0.10	0
2	SO4	A	702	-	4,4,4	0.10	0	6,6,6	0.07	0
2	SO4	B	701	-	4,4,4	0.34	0	6,6,6	0.10	0
2	SO4	B	702	-	4,4,4	0.11	0	6,6,6	0.08	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	SO4	A	701	-	-	0/0/0/0	0/0/0/0
2	SO4	A	702	-	-	0/0/0/0	0/0/0/0
2	SO4	B	701	-	-	0/0/0/0	0/0/0/0
2	SO4	B	702	-	-	0/0/0/0	0/0/0/0

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, 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	628/666 (94%)	-0.18	14 (2%)	65	42	45, 70, 111, 144	0
1	B	628/666 (94%)	-0.21	11 (1%)	71	50	40, 71, 109, 139	0
All	All	1256/1332 (94%)	-0.20	25 (1%)	68	46	40, 71, 110, 144	0

The worst 5 of 25 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	414	ARG	4.4
1	A	414	ARG	4.2
1	A	415	ARG	4.0
1	A	634	ASP	3.8
1	A	393	ASP	3.5

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 carbohydrates 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. 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, 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	LLDF	B-factors(Å^2)	Q<0.9
2	SO4	B	702	5/5	0.91	0.24	0.93	55,56,59,64	5
2	SO4	A	702	5/5	0.92	0.23	0.67	60,62,66,72	5
2	SO4	A	701	5/5	0.97	0.20	0.30	60,62,65,70	0
2	SO4	B	701	5/5	0.96	0.18	0.11	65,68,76,76	0

6.5 Other polymers (i)

There are no such residues in this entry.