wwPDB X-ray Structure Validation Summary Report (i)

Jun 22, 2024 – 05:38 PM EDT

PDB ID	:	6ELQ
Title	:	Carbon Monoxide Dehydrogenase IV from Carboxydothermus hydrogenoformans
Authors	:	Domnik, L.; Goetzl, S.; Jeoung, J.H.; Dobbek, H.
Deposited on	:	2017-09-29
Resolution	:	2.52 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
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:

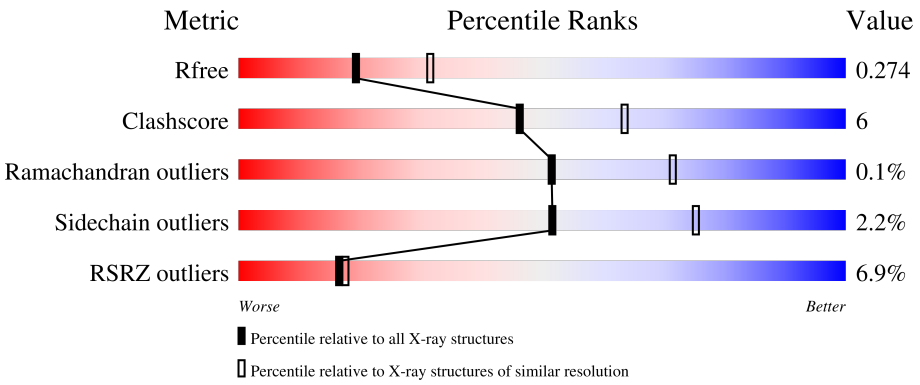
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.37.1

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 2.52 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	5743 (2.54-2.50)
Clashscore	141614	6463 (2.54-2.50)
Ramachandran outliers	138981	6335 (2.54-2.50)
Sidechain outliers	138945	6337 (2.54-2.50)
RSRZ outliers	127900	5630 (2.54-2.50)

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	633	4% 85% 14%
1	B	633	15% 76% 22% •
1	X	633	% 88% 11% •

The following table lists non-polymeric compounds, carbohydrate monomers and non-standard residues in protein, DNA, RNA chains that are outliers for geometric or electron-density-fit criteria:

Mol	Type	Chain	Res	Chirality	Geometry	Clashes	Electron density
3	BF8	B	5002	-	-	X	-

2 Entry composition (i)

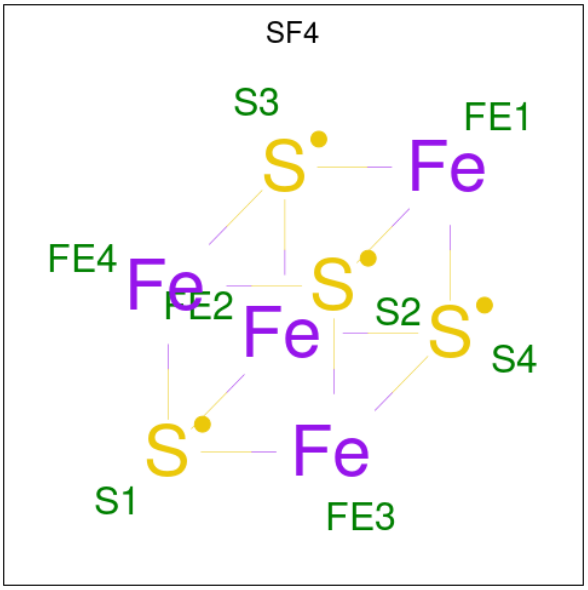
There are 4 unique types of molecules in this entry. The entry contains 14594 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Carbon Monoxide Dehydrogenase.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	X	630	4762	3020	821	894	27	0	0	0
1	A	630	4762	3020	821	894	27	0	0	0
1	B	630	4762	3020	821	894	27	0	0	0

- Molecule 2 is IRON/SULFUR CLUSTER (three-letter code: SF4) (formula: Fe_4S_4).



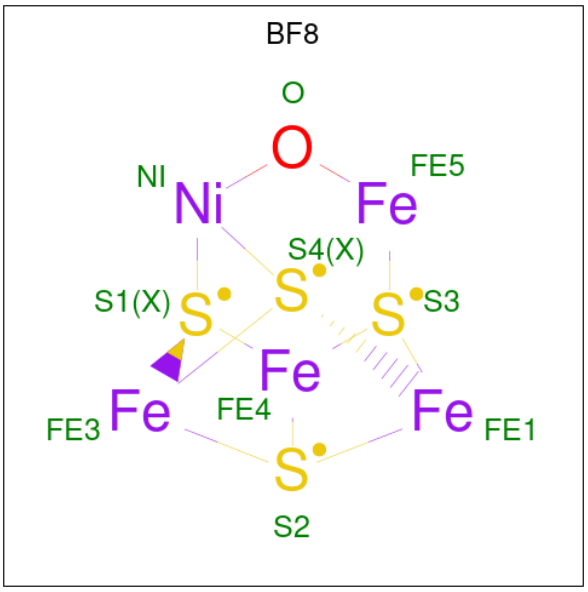
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	Fe	S		
2	X	1	8	4	4	0	0
2	A	1	8	4	4	0	0
2	A	1	8	4	4	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
			Total	Fe	S		
2	B	1	8	4	4	0	0
			Total	Fe	S		
2	B	1	8	4	4	0	0

- Molecule 3 is FE(4)-NI(1)-S(5) CLUSTER with Oxygen (three-letter code: BF8) (formula: Fe_4NiOS_4).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
			Total	Fe	Ni	O	S		
3	X	1	10	4	1	1	4	0	0
			Total	Fe	Ni	O	S		
3	A	1	10	4	1	1	4	0	0
			Total	Fe	Ni	O	S		
3	B	1	10	4	1	1	4	0	0

- Molecule 4 is water.

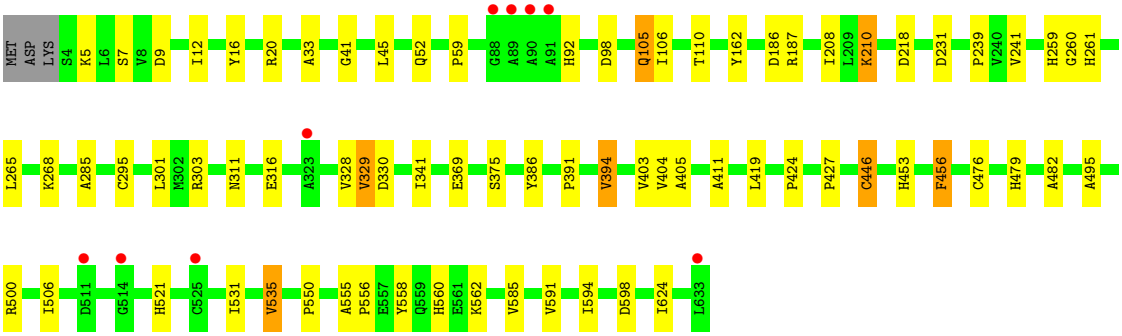
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	X	123	123	123	0	0
			Total	O		
4	A	77	77	77	0	0
			Total	O		
4	B	38	38	38	0	0

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.

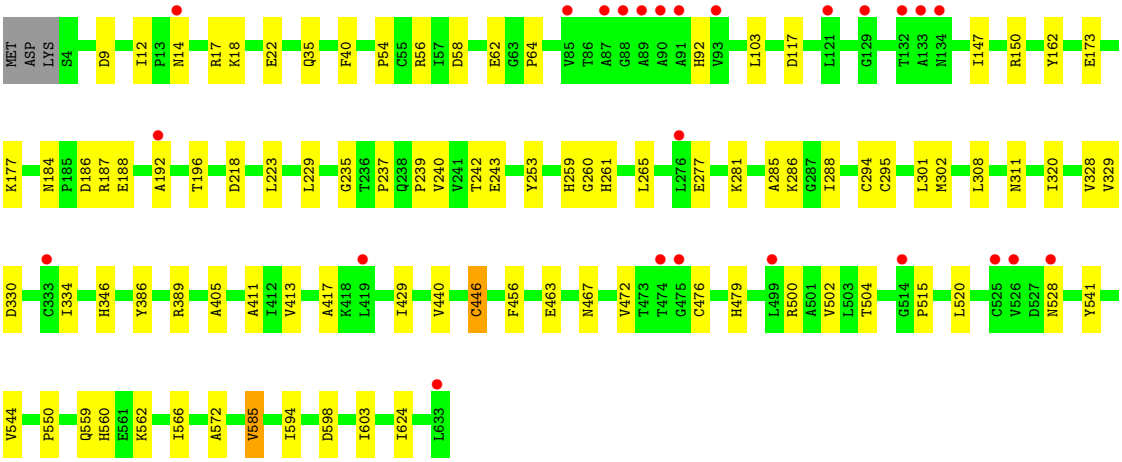
• Molecule 1: Carbon Monoxide Dehydrogenase

Chain X:



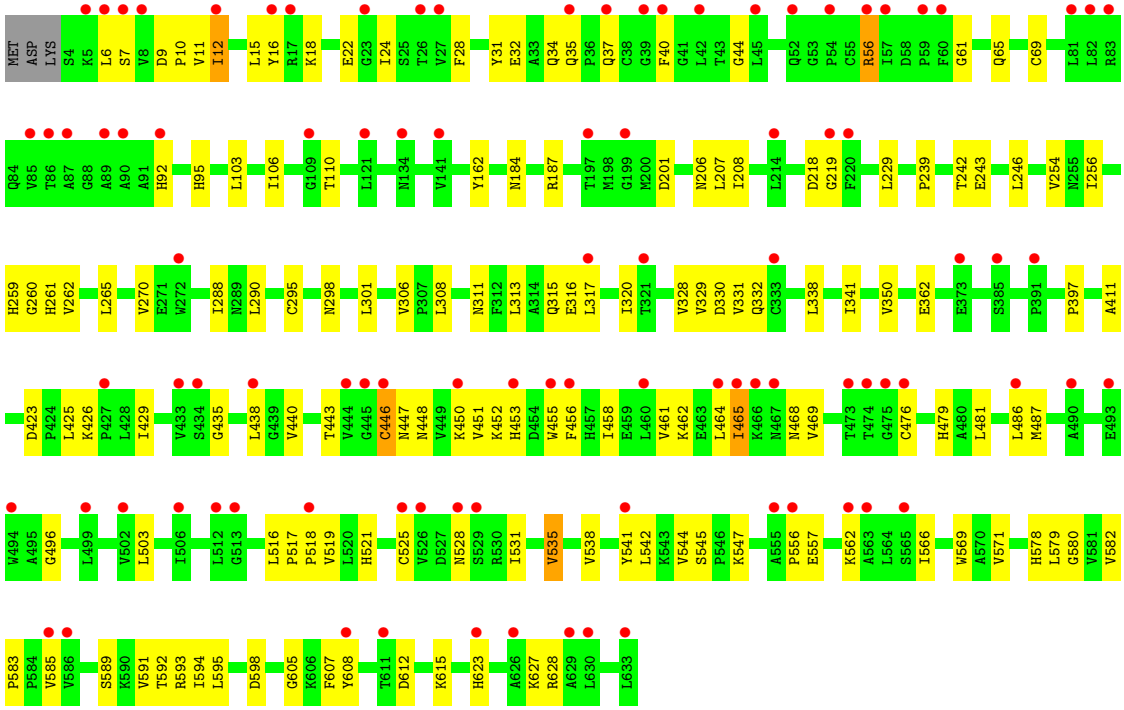
• Molecule 1: Carbon Monoxide Dehydrogenase

Chain A:



• Molecule 1: Carbon Monoxide Dehydrogenase

Chain B:


MET ASP LYS S4 K5 L6 S7 V8 D9 P10 V11 I12 L15 Y16 R17 K18 E22 G23 I24 S25 T26 V27 F28 Y31 E32 A33 Q34 Q35 P36 Q37 C38 G39 F40 G41 L42 T43 G44 L45 Q52 G53 P54 C55 R56 I57 D58 P59 F60 G61 Q65 C69 L81 L82 R83
Q84 V85 T86 A87 G88 A89 A90 A91 H92 H95 L103 I106 G109 T110 L121 N134 V141 Y162 N184 R187 T197 M198 G199 M200 D201 N206 L207 I208 L214 D218 G219 F220 L229 P239 T242 E243 L246 V254 N255 I256
H259 G260 H261 V262 L265 V270 E271 W272 I288 N289 L290 C295 N298 L301 V306 P307 L308 N311 F312 L313 A314 Q315 E316 L317 I320 T321 V328 V329 D330 V331 Q332 G333 L338 I341 V350 E362 E373 S385 P391 P397 A411
D423 P424 L425 K426 P427 L428 I429 V433 S434 G435 L438 G439 V440 T443 V444 G445 C446 N447 N448 V449 K450 V451 K452 H453 D454 W455 F456 H457 I458 E459 L460 V461 K462 E463 L464 I465 K466 N467 N468 V469 T473 T474 G475 C476 H479 A480 L481 L486 M487 A490 E493
W494 A495 G496 L499 V502 L503 I506 L512 G513 L516 P517 P518 V519 L520 H521 C525 V526 D527 N528 S529 R530 I531 V535 V538 Y541 L542 K543 V544 S545 P546 K547 A555 P556 E557 K562 A563 L564 S565 I566 W569 A570 V571 H578 L579 G580 V581 V582
P583 P584 V585 V586 S589 K590 V591 T592 R593 I594 L595 D598 G605 K606 F607 Y608 T611 D612 K615 H623 A626 K627 R628 A629 L630 L633

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 3 2 1	Depositor
Cell constants a, b, c, α, β, γ	209.15Å 209.15Å 93.41Å 90.00° 90.00° 120.00°	Depositor
Resolution (Å)	45.63 – 2.52 45.63 – 2.52	Depositor EDS
% Data completeness (in resolution range)	99.9 (45.63-2.52) 99.9 (45.63-2.52)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.53 (at 2.51Å)	Xtriage
Refinement program	PHENIX (1.11.1-2575_1168: ???)	Depositor
R, R_{free}	0.235 , 0.275 0.236 , 0.274	Depositor DCC
R_{free} test set	2113 reflections (2.67%)	wwPDB-VP
Wilson B-factor (Å^2)	49.0	Xtriage
Anisotropy	0.578	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 40.4	EDS
L-test for twinning[2]	$< \|L\| > = 0.47$, $< L^2 > = 0.29$	Xtriage
Estimated twinning fraction	0.092 for -h,-k,l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	14594	wwPDB-VP
Average B, all atoms (Å^2)	57.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.62% 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: BF8, SF4

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	0.24	0/4846	0.42	0/6591
1	B	0.25	0/4846	0.42	0/6591
1	X	0.25	0/4846	0.42	0/6591
All	All	0.25	0/14538	0.42	0/19773

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	4762	0	4867	49	0
1	B	4762	0	4866	97	1
1	X	4762	0	4866	40	0
2	A	16	0	0	0	0
2	B	16	0	0	2	0
2	X	8	0	0	0	0
3	A	10	0	0	0	0
3	B	10	0	0	4	0
3	X	10	0	0	1	0
4	A	77	0	0	1	0

Continued on next page...

Continued from previous page...

Mol	Chain	Non-H	H(model)	H(added)	Clashes	Symm-Clashes
4	B	38	0	0	0	0
4	X	123	0	0	0	0
All	All	14594	0	14599	182	1

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.

The worst 5 of 182 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:261:HIS:CE1	3:B:5002:BF8:S3	2.73	0.80
1:B:435:GLY:HA2	1:B:628:ARG:HH22	1.47	0.79
1:B:446:CYS:HA	1:B:476:CYS:HB2	1.69	0.74
1:B:594:ILE:HA	1:B:598:ASP:HB2	1.70	0.73
1:B:612:ASP:HB3	1:B:615:LYS:HB2	1.72	0.72

All (1) 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)
1:B:34:GLN:NE2	1:B:56:ARG:O[5_675]	2.19	0.01

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	628/633 (99%)	605 (96%)	22 (4%)	1 (0%)	47 67
1	B	628/633 (99%)	597 (95%)	30 (5%)	1 (0%)	47 67
1	X	628/633 (99%)	604 (96%)	24 (4%)	0	100 100
All	All	1884/1899 (99%)	1806 (96%)	76 (4%)	2 (0%)	51 71

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	585	VAL
1	B	332	GLN

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	512/515 (99%)	503 (98%)	9 (2%)	59	80
1	B	512/515 (99%)	500 (98%)	12 (2%)	50	74
1	X	512/515 (99%)	499 (98%)	13 (2%)	47	72
All	All	1536/1545 (99%)	1502 (98%)	34 (2%)	52	75

5 of 34 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	B	465	ILE
1	B	535	VAL
1	B	589	SER
1	X	585	VAL
1	X	535	VAL

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)

8 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	SF4	X	701	1	0,12,12	-	-	-		
2	SF4	B	5001	1	0,12,12	-	-	-		
3	BF8	X	702	1	0,14,14	-	-	-		
3	BF8	A	5003	1	0,14,14	-	-	-		
3	BF8	B	5002	1	0,14,14	-	-	-		
2	SF4	A	5001	1	0,12,12	-	-	-		
2	SF4	A	5002	1	0,12,12	-	-	-		
2	SF4	B	5000	1	0,12,12	-	-	-		

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	SF4	X	701	1	-	-	0/6/5/5
2	SF4	B	5001	1	-	-	0/6/5/5
3	BF8	X	702	1	-	-	0/4/5/5
3	BF8	A	5003	1	-	-	0/4/5/5
3	BF8	B	5002	1	-	-	0/4/5/5
2	SF4	A	5001	1	-	-	0/6/5/5
2	SF4	A	5002	1	-	-	0/6/5/5
2	SF4	B	5000	1	-	-	0/6/5/5

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.

4 monomers are involved in 7 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	B	5001	SF4	1	0
3	X	702	BF8	1	0
3	B	5002	BF8	4	0
2	B	5000	SF4	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	630/633 (99%)	0.19	25 (3%) 38 42	38, 54, 73, 88	0
1	B	630/633 (99%)	0.96	96 (15%) 2 2	43, 69, 95, 106	0
1	X	630/633 (99%)	-0.06	9 (1%) 75 78	36, 46, 60, 75	0
All	All	1890/1899 (99%)	0.37	130 (6%) 16 17	36, 54, 86, 106	0

The worst 5 of 130 RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	5	LYS	6.6
1	B	633	LEU	6.2
1	B	40	PHE	6.0
1	B	317	LEU	5.8
1	A	525	CYS	5.2

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	SF4	B	5001	8/8	0.93	0.14	77,83,104,104	8
3	BF8	B	5002	10/10	0.95	0.22	60,66,72,73	6
2	SF4	A	5001	8/8	0.97	0.10	42,44,51,55	0
3	BF8	X	702	10/10	0.97	0.18	45,47,55,59	8
3	BF8	A	5003	10/10	0.97	0.20	46,50,59,63	7
2	SF4	B	5000	8/8	0.97	0.12	64,67,71,76	0
2	SF4	X	701	8/8	0.98	0.14	34,45,50,50	0
2	SF4	A	5002	8/8	0.99	0.12	34,41,45,48	0

6.5 Other polymers (i)

There are no such residues in this entry.