



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 08:19 AM GMT

PDB ID : 3E8L  
Title : The Crystal Structure of the Double-headed Arrowhead Protease Inhibitor A  
in Complex with Two Trypsins  
Authors : Bao, R.; Jiang, C.-H.; Chi, C.W.; Lin, S.X.; Chen, Y.X.  
Deposited on : 2008-08-20  
Resolution : 2.48 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](mailto: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

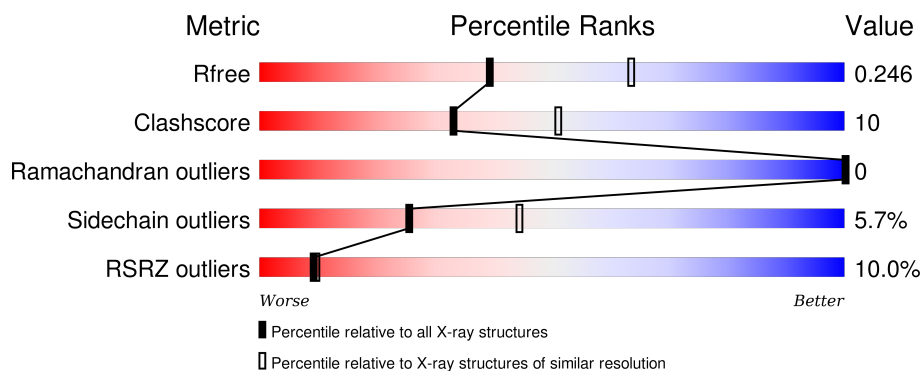
# 1 Overall quality at a glance ⓘ

The following experimental techniques were used to determine the structure:

## *X-RAY DIFFRACTION*

The reported resolution of this entry is 2.48 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	4309 (2.50-2.46)
Clashscore	102246	5050 (2.50-2.46)
Ramachandran outliers	100387	4961 (2.50-2.46)
Sidechain outliers	100360	4963 (2.50-2.46)
RSRZ outliers	91569	4319 (2.50-2.46)

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  $\geq 3$ , 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	C	185	<div> <div>9%</div> <div>66%</div> <div>24%</div> <div>5%</div> <div>5%</div> </div>
2	A	223	<div> <div>7%</div> <div>83%</div> <div>17%</div> </div>
2	B	223	<div> <div>13%</div> <div>81%</div> <div>18%</div> <div>.</div> </div>

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	GOL	A	239	-	-	-	X
3	GOL	C	179	-	-	X	X
3	GOL	C	180	-	-	-	X
4	ACT	A	6	-	-	X	X
4	ACT	B	2	-	-	X	X
4	ACT	C	181	-	-	-	X
4	ACT	C	182	-	-	-	X
6	NA	A	7	-	-	X	-
7	CA	A	5	-	-	-	X
8	EDO	B	3	-	-	-	X
8	EDO	B	7[A]	-	-	-	X
8	EDO	B	7[B]	-	-	-	X
8	EDO	B	8	-	-	X	-

## 2 Entry composition

There are 10 unique types of molecules in this entry. The entry contains 4908 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 Serine proteinase inhibitor A.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	C	176	Total	C	N	O	S	36	2	0
			1359	867	222	262	8			

There are 11 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
C	-7	MET	-	EXPRESSION TAG	UNP Q7M1P4
C	-6	GLY	-	EXPRESSION TAG	UNP Q7M1P4
C	-5	HIS	-	EXPRESSION TAG	UNP Q7M1P4
C	-4	HIS	-	EXPRESSION TAG	UNP Q7M1P4
C	-3	HIS	-	EXPRESSION TAG	UNP Q7M1P4
C	-2	HIS	-	EXPRESSION TAG	UNP Q7M1P4
C	-1	HIS	-	EXPRESSION TAG	UNP Q7M1P4
C	0	HIS	-	EXPRESSION TAG	UNP Q7M1P4
C	1	MET	-	EXPRESSION TAG	UNP Q7M1P4
C	39	ARG	HIS	ENGINEERED	UNP Q7M1P4
C	172	GLN	ARG	ENGINEERED	UNP Q7M1P4

- Molecule 2 is a protein called Cationic trypsin.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	A	223	Total	C	N	O	S	0	6	0
			1654	1027	281	332	14			
2	B	223	Total	C	N	O	S	0	1	0
			1630	1013	279	324	14			

- Molecule 3 is GLYCEROL (three-letter code: GOL) (formula: C<sub>3</sub>H<sub>8</sub>O<sub>3</sub>).



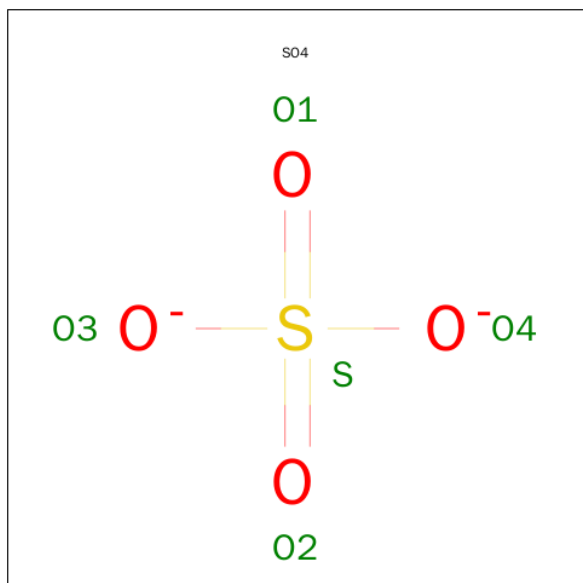
Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
3	C	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		
3	C	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	A	1	Total	C	O	0	0
			6	3	3		
3	B	1	Total	C	O	0	0
			6	3	3		

- Molecule 4 is ACETATE ION (three-letter code: ACT) (formula: C<sub>2</sub>H<sub>3</sub>O<sub>2</sub>).



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

- Molecule 5 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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

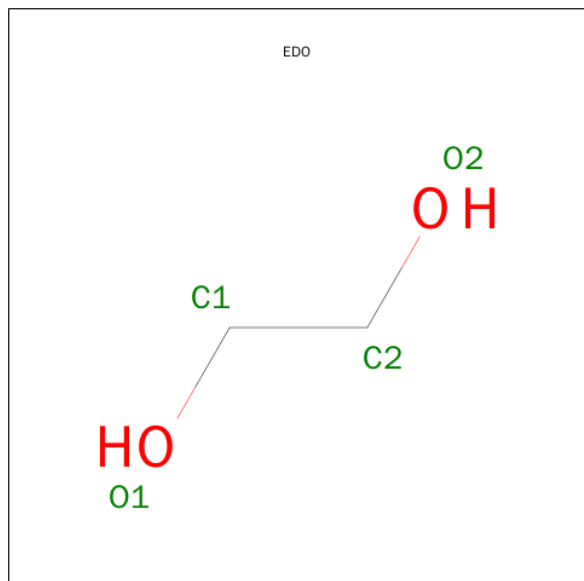
- Molecule 6 is SODIUM ION (three-letter code: NA) (formula: Na).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
6	A	3	Total Na 3 3	0	0

- Molecule 7 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
7	B	2	Total Ca 2 2	0	0
7	A	5	Total Ca 5 5	0	0

- Molecule 8 is 1,2-ETHANEDIOL (three-letter code: EDO) (formula: C<sub>2</sub>H<sub>6</sub>O<sub>2</sub>).



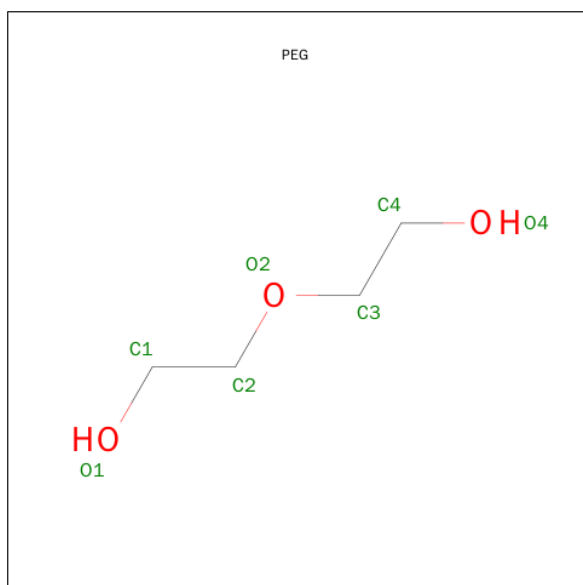
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
8	B	1	Total C O 4 2 2	0	0
8	B	1	Total C O 4 2 2	0	0

*Continued on next page...*

Continued from previous page...

Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
8	B	1	Total	C	O	0	1
			8	4	4		
8	B	1	Total	C	O	0	0
			4	2	2		

- Molecule 9 is DI(HYDROXYETHYL)ETHER (three-letter code: PEG) (formula: C<sub>4</sub>H<sub>10</sub>O<sub>3</sub>).



Mol	Chain	Residues	Atoms			ZeroOcc	AltConf
9	B	1	Total	C	O	0	0
			7	4	3		

- Molecule 10 is water.

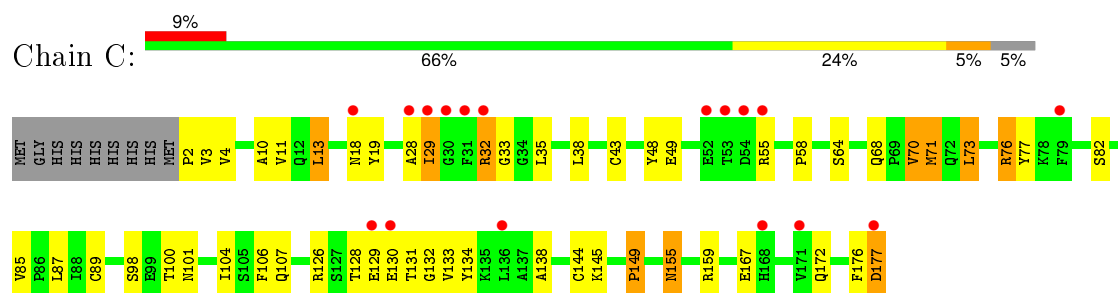
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
10	C	42	Total	O	0	0
			42	42		
10	A	77	Total	O	0	0
			77	77		
10	B	41	Total	O	0	0
			41	41		



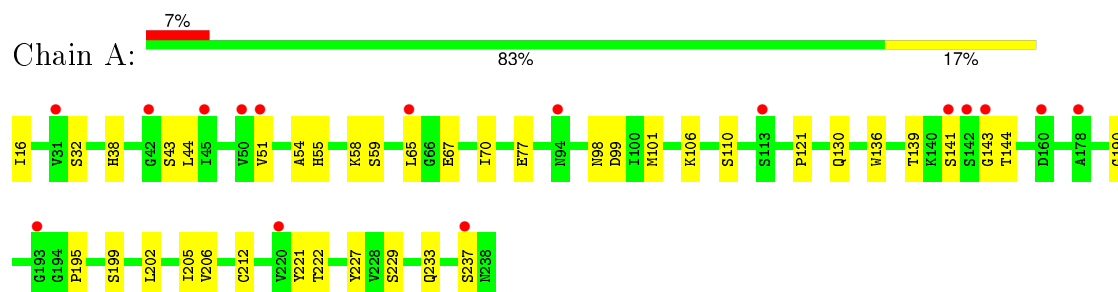
### 3 Residue-property plots [i](#)

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.

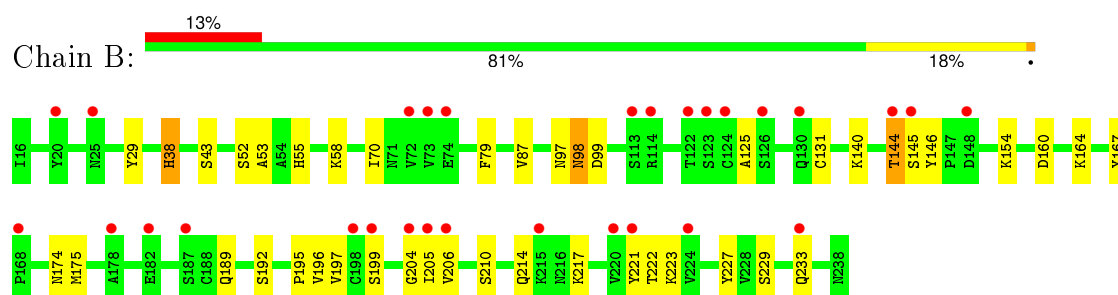
- Molecule 1: Serine proteinase inhibitor A



- Molecule 2: Cationic trypsin



- Molecule 2: Cationic trypsin



## 4 Data and refinement statistics

Property	Value	Source
Space group	C 2 2 21	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	76.63Å 110.86Å 152.99Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	30.00 – 2.48 21.10 – 2.48	Depositor EDS
% Data completeness (in resolution range)	99.7 (30.00-2.48) 99.8 (21.10-2.48)	Depositor EDS
$R_{merge}$	0.12	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.82 (at 2.47Å)	Xtriage
Refinement program	REFMAC 5.4.0066	Depositor
R, $R_{free}$	0.194 , 0.245 0.207 , 0.246	Depositor DCC
$R_{free}$ test set	1395 reflections (6.34%)	DCC
Wilson B-factor (Å <sup>2</sup> )	34.1	Xtriage
Anisotropy	0.186	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 44.5	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.52$ , $\langle L^2 \rangle = 0.35$	Xtriage
Outliers	1 of 23426 reflections (0.004%)	Xtriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	4908	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	36.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.66% of the height of the origin peak. No significant pseudotranslation is detected.*

<sup>1</sup>Intensities estimated from amplitudes.

<sup>2</sup>Theoretical 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: GOL, NA, CA, EDO, SO4, ACT, PEG

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	C	0.69	3/1393 (0.2%)	0.77	3/1887 (0.2%)
2	A	0.56	0/1701	0.64	0/2305
2	B	0.52	0/1665	0.63	0/2258
All	All	0.59	3/4759 (0.1%)	0.68	3/6450 (0.0%)

Chiral center outliers are detected by calculating the chiral volume of a chiral center and verifying if the center is modelled as a planar moiety or with the opposite hand. A planarity outlier is detected by checking planarity of atoms in a peptide group, atoms in a mainchain group or atoms of a sidechain that are expected to be planar.

Mol	Chain	#Chirality outliers	#Planarity outliers
1	C	0	1

All (3) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
1	C	29	ILE	C-O	9.02	1.40	1.23
1	C	28	ALA	C-N	-5.95	1.20	1.34
1	C	29	ILE	C-N	-5.22	1.23	1.33

All (3) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	C	33	GLY	N-CA-C	-7.68	93.89	113.10
1	C	28	ALA	C-N-CA	6.08	136.91	121.70
1	C	29	ILE	CA-C-N	5.07	126.34	116.20

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	29	ILE	Mainchain

## 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	C	1359	0	1297	35	0
2	A	1654	0	1607	26	0
2	B	1630	0	1589	35	0
3	A	18	0	24	1	0
3	B	6	0	8	0	0
3	C	18	0	24	4	0
4	A	4	0	3	2	0
4	B	4	0	3	2	0
4	C	8	0	6	1	0
5	A	5	0	0	1	0
5	C	5	0	0	0	0
6	A	3	0	0	2	0
7	A	5	0	0	0	0
7	B	2	0	0	0	0
8	B	20	0	30	11	0
9	B	7	0	10	1	0
10	A	77	0	0	3	0
10	B	41	0	0	2	0
10	C	42	0	0	1	0
All	All	4908	0	4601	96	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 10.

All (96) 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)
2:B:160:ASP:OD2	8:B:8:EDO:H11	1.29	1.26
2:B:160:ASP:OD2	8:B:8:EDO:C1	1.88	1.21
6:A:7:NA:NA	4:A:6:ACT:CH3	1.63	1.02
1:C:32:ARG:O	1:C:49:GLU:OE2	1.83	0.96

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:11:VAL:HG11	1:C:71:MET:HE3	1.54	0.87
1:C:2:PRO:O	3:C:179:GOL:O2	1.92	0.87
2:A:77:GLU:OE2	10:A:314:HOH:O	1.91	0.86
1:C:76:ARG:NH2	1:C:144:CYS:O	2.14	0.80
1:C:176:PHE:O	1:C:177:ASP:HB3	1.82	0.79
6:A:7:NA:NA	4:A:6:ACT:H1	1.38	0.77
2:B:87:VAL:H	8:B:239:EDO:H12	1.48	0.77
2:B:192:SER:HB2	10:B:242:HOH:O	1.84	0.76
1:C:11:VAL:HG11	1:C:71:MET:CE	2.17	0.75
2:A:55:HIS:HD2	2:A:99:ASP:OD2	1.70	0.73
1:C:2:PRO:N	10:C:225:HOH:O	2.25	0.70
2:B:175:MET:O	8:B:8:EDO:H22	1.92	0.69
2:A:106:LYS:HE3	5:A:240:SO4:O2	1.96	0.66
1:C:155:ASN:HD22	1:C:155:ASN:C	1.99	0.66
1:C:11:VAL:CG1	1:C:71:MET:HE3	2.24	0.66
1:C:64:SER:HB3	4:C:181:ACT:H2	1.76	0.66
2:B:160:ASP:OD2	8:B:8:EDO:H12	1.91	0.65
8:B:7[A]:EDO:C1	8:B:8:EDO:H12	2.26	0.65
1:C:87:LEU:HB2	2:B:192:SER:OG	1.96	0.65
2:B:55:HIS:HB2	4:B:2:ACT:H1	1.78	0.65
1:C:3:VAL:HG22	1:C:73:LEU:HD13	1.79	0.65
1:C:132:GLY:O	1:C:172:GLN:HG2	1.97	0.65
2:B:38:HIS:CE1	2:B:70:ILE:HD12	2.33	0.64
1:C:126:ARG:HG2	2:A:144:THR:HG21	1.80	0.63
2:B:58:LYS:HE3	10:B:272:HOH:O	2.00	0.61
1:C:134:TYR:OH	3:C:179:GOL:H31	2.01	0.61
2:B:223:LYS:HE3	8:B:8:EDO:H21	1.82	0.61
2:B:196:VAL:HB	2:B:204:GLY:HA3	1.86	0.57
2:B:38:HIS:CD2	2:B:38:HIS:H	2.23	0.57
1:C:176:PHE:O	1:C:177:ASP:CB	2.53	0.56
2:A:55:HIS:CD2	2:A:99:ASP:OD2	2.56	0.56
2:A:38:HIS:CE1	2:A:70:ILE:HB	2.41	0.55
2:B:229:SER:O	2:B:233:GLN:HG2	2.05	0.55
2:A:205:ILE:HB	2:A:222:THR:HB	1.88	0.55
2:B:55:HIS:HB2	4:B:2:ACT:CH3	2.37	0.55
1:C:58:PRO:HB2	1:C:82:SER:HB3	1.90	0.54
1:C:18:ASN:HA	2:B:214:GLN:OE1	2.07	0.54
1:C:10:ALA:HB1	1:C:70:VAL:CG2	2.38	0.54
8:B:7[A]:EDO:H12	8:B:8:EDO:H12	1.89	0.53
1:C:100:THR:HA	1:C:104:ILE:O	2.08	0.53
2:A:130:GLN:HG3	10:A:285:HOH:O	2.08	0.53

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:C:43:CYS:HB3	1:C:89:CYS:HB3	1.90	0.52
2:A:98:ASN:ND2	2:A:227:TYR:OH	2.34	0.51
2:A:143:GLY:O	2:A:144:THR:CG2	2.58	0.51
8:B:7[A]:EDO:H11	8:B:8:EDO:H12	1.92	0.50
2:B:206:VAL:HA	2:B:221:TYR:CD1	2.46	0.50
1:C:38:LEU:HD11	1:C:159:ARG:HD2	1.93	0.49
2:B:210:SER:HB3	2:B:217:LYS:HD3	1.95	0.49
2:B:55:HIS:HD2	2:B:99:ASP:OD2	1.96	0.49
1:C:10:ALA:HB1	1:C:70:VAL:HG21	1.95	0.48
2:B:98:ASN:HA	2:B:227:TYR:OH	2.13	0.48
2:A:143:GLY:C	2:A:144:THR:HG23	2.33	0.48
2:A:121:PRO:HD3	2:A:202:LEU:O	2.13	0.48
2:A:143:GLY:O	2:A:144:THR:HG23	2.14	0.48
1:C:138:ALA:O	1:C:149:PRO:HD2	2.13	0.48
2:B:97:ASN:O	2:B:98:ASN:HB2	2.14	0.48
1:C:71:MET:HG3	1:C:77:TYR:CE2	2.49	0.47
2:B:223:LYS:CE	8:B:8:EDO:H21	2.45	0.47
2:B:205:ILE:HB	2:B:222:THR:HB	1.96	0.46
2:A:43:SER:OG	2:A:195:PRO:HB3	2.16	0.46
2:B:144:THR:HG23	2:B:146:TYR:CE2	2.51	0.46
2:A:54:ALA:HA	2:A:101:MET:HB2	1.97	0.46
2:A:229:SER:O	2:A:233:GLN:HG2	2.16	0.46
2:B:196:VAL:HB	2:B:204:GLY:CA	2.46	0.46
1:C:98:SER:HB3	1:C:107:GLN:HG2	1.98	0.46
1:C:4:VAL:HG22	3:C:179:GOL:H12	1.99	0.45
2:A:32:SER:HB2	2:A:38:HIS:HD2	1.81	0.45
2:B:160:ASP:O	2:B:164:LYS:HG3	2.17	0.45
2:B:125:ALA:H	8:B:7[B]:EDO:H21	1.81	0.45
1:C:13:LEU:HD22	1:C:70:VAL:C	2.37	0.44
2:B:43:SER:OG	2:B:195:PRO:HB3	2.16	0.44
1:C:128:THR:HB	1:C:133:VAL:O	2.17	0.44
2:A:67:GLU:HG3	10:A:263:HOH:O	2.16	0.44
1:C:128:THR:HG22	1:C:130:GLU:H	1.82	0.44
1:C:145:LYS:O	2:A:190:GLY:N	2.48	0.44
1:C:87:LEU:O	2:B:189:GLN:HA	2.17	0.43
2:B:174:ASN:HD21	9:B:240:PEG:H32	1.83	0.43
1:C:3:VAL:HG22	1:C:73:LEU:CD1	2.47	0.43
2:B:140:LYS:HE2	2:B:145:SER:HB2	2.01	0.42
2:B:131:CYS:O	2:B:154:LYS:HA	2.19	0.42
2:A:206:VAL:HA	2:A:221:TYR:CD1	2.55	0.42
2:A:199:SER:N	3:A:12:GOL:H32	2.35	0.42

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:A:206:VAL:HG22	2:A:221:TYR:HE1	1.85	0.41
2:B:29:TYR:CZ	2:B:197:VAL:HG21	2.55	0.41
2:A:141:SER:HB2	2:A:212:CYS:O	2.21	0.41
1:C:128:THR:HG22	1:C:129:GLU:N	2.36	0.41
2:B:52:SER:OG	2:B:53:ALA:N	2.54	0.41
2:B:167:TYR:HE1	2:B:210:SER:HB2	1.85	0.40
1:C:134:TYR:HH	3:C:179:GOL:H31	1.86	0.40
2:A:70:ILE:HG23	2:A:136:TRP:NE1	2.36	0.40
2:A:44:LEU:HD13	2:A:65:LEU:HD21	2.03	0.40
2:A:16:ILE:O	2:A:139:THR:HA	2.21	0.40

There are no symmetry-related clashes.

## 5.3 Torsion angles ⓘ

### 5.3.1 Protein backbone ⓘ

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	C	176/185 (95%)	169 (96%)	7 (4%)	0	100	100
2	A	227/223 (102%)	220 (97%)	7 (3%)	0	100	100
2	B	222/223 (100%)	215 (97%)	7 (3%)	0	100	100
All	All	625/631 (99%)	604 (97%)	21 (3%)	0	100	100

There are no Ramachandran outliers to report.

### 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	C	146/152 (96%)	127 (87%)	19 (13%)	5	8
2	A	190/184 (103%)	184 (97%)	6 (3%)	46	72
2	B	185/184 (100%)	180 (97%)	5 (3%)	52	78
All	All	521/520 (100%)	491 (94%)	30 (6%)	25	43

All (30) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	C	13	LEU
1	C	19	TYR
1	C	32	ARG
1	C	35	LEU
1	C	48	TYR
1	C	55	ARG
1	C	68	GLN
1	C	70	VAL
1	C	71	MET
1	C	73	LEU
1	C	76	ARG
1	C	85	VAL
1	C	101	ASN
1	C	106	PHE
1	C	131	THR
1	C	149	PRO
1	C	155	ASN
1	C	167	GLU
1	C	177	ASP
2	A	51	VAL
2	A	58	LYS
2	A	59	SER
2	A	110[A]	SER
2	A	110[B]	SER
2	A	237	SER
2	B	38	HIS
2	B	79	PHE
2	B	98	ASN
2	B	144	THR
2	B	199	SER

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (11) such sidechains are listed below:



Mol	Chain	Res	Type
1	C	68	GLN
1	C	101	ASN
1	C	155	ASN
2	A	30	GLN
2	A	38	HIS
2	A	55	HIS
2	A	98	ASN
2	B	30	GLN
2	B	38	HIS
2	B	55	HIS
2	B	233	GLN

### 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](#)

Of 29 ligands modelled in this entry, 10 are monoatomic - leaving 19 for Mogul analysis.

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	$\# Z  > 2$
3	GOL	A	12	-	5,5,5	0.33	0	5,5,5	0.36	0
3	GOL	A	2	-	5,5,5	0.33	0	5,5,5	0.37	0
3	GOL	A	239	-	5,5,5	0.58	0	5,5,5	0.97	0
5	SO4	A	240	-	4,4,4	0.37	0	6,6,6	0.14	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	# Z  > 2	Counts	RMSZ	# Z  > 2
4	ACT	A	6	-	1,3,3	1.45	0	0,3,3	0.00	-
3	GOL	B	1	-	5,5,5	0.61	0	5,5,5	0.63	0
4	ACT	B	2	-	1,3,3	0.65	0	0,3,3	0.00	-
8	EDO	B	239	-	3,3,3	0.32	0	2,2,2	0.54	0
9	PEG	B	240	-	6,6,6	0.47	0	5,5,5	0.29	0
8	EDO	B	3	-	3,3,3	0.46	0	2,2,2	0.53	0
8	EDO	B	7[A]	-	3,3,3	0.49	0	2,2,2	0.41	0
8	EDO	B	7[B]	-	3,3,3	0.48	0	2,2,2	0.46	0
8	EDO	B	8	-	3,3,3	0.34	0	2,2,2	0.30	0
3	GOL	C	178	-	5,5,5	0.55	0	5,5,5	0.19	0
3	GOL	C	179	-	5,5,5	0.69	0	5,5,5	0.36	0
3	GOL	C	180	-	5,5,5	0.48	0	5,5,5	0.35	0
4	ACT	C	181	-	1,3,3	0.70	0	0,3,3	0.00	-
4	ACT	C	182	-	1,3,3	1.60	0	0,3,3	0.00	-
5	SO4	C	183	-	4,4,4	0.18	0	6,6,6	0.16	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
3	GOL	A	12	-	-	0/4/4/4	0/0/0/0
3	GOL	A	2	-	-	0/4/4/4	0/0/0/0
3	GOL	A	239	-	-	0/4/4/4	0/0/0/0
5	SO4	A	240	-	-	0/0/0/0	0/0/0/0
4	ACT	A	6	-	-	0/0/0/0	0/0/0/0
3	GOL	B	1	-	-	0/4/4/4	0/0/0/0
4	ACT	B	2	-	-	0/0/0/0	0/0/0/0
8	EDO	B	239	-	-	0/1/1/1	0/0/0/0
9	PEG	B	240	-	-	0/4/4/4	0/0/0/0
8	EDO	B	3	-	-	0/1/1/1	0/0/0/0
8	EDO	B	7[A]	-	-	0/1/1/1	0/0/0/0
8	EDO	B	7[B]	-	-	0/1/1/1	0/0/0/0
8	EDO	B	8	-	-	0/1/1/1	0/0/0/0
3	GOL	C	178	-	-	0/4/4/4	0/0/0/0
3	GOL	C	179	-	-	0/4/4/4	0/0/0/0
3	GOL	C	180	-	-	0/4/4/4	0/0/0/0
4	ACT	C	181	-	-	0/0/0/0	0/0/0/0
4	ACT	C	182	-	-	0/0/0/0	0/0/0/0
5	SO4	C	183	-	-	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.

11 monomers are involved in 23 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	12	GOL	1	0
5	A	240	SO4	1	0
4	A	6	ACT	2	0
4	B	2	ACT	2	0
8	B	239	EDO	1	0
9	B	240	PEG	1	0
8	B	7[A]	EDO	3	0
8	B	7[B]	EDO	1	0
8	B	8	EDO	9	0
3	C	179	GOL	4	0
4	C	181	ACT	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 ⓘ

### 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<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
1	C	176/185 (95%)	0.65	17 (9%)	10 10	18, 28, 40, 46	7 (3%)
2	A	223/223 (100%)	0.52	16 (7%)	18 20	34, 39, 46, 55	0
2	B	223/223 (100%)	0.85	29 (13%)	5 4	33, 39, 45, 56	0
All	All	622/631 (98%)	0.68	62 (9%)	9 10	18, 38, 45, 56	7 (1%)

All (62) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	53	THR	9.3
1	C	177	ASP	6.1
1	C	55	ARG	6.0
2	B	73	VAL	5.9
1	C	54	ASP	5.3
1	C	30	GLY	5.2
1	C	29	ILE	5.2
2	A	142[A]	SER	5.1
1	C	52	GLU	4.5
2	B	206	VAL	4.1
2	B	144	THR	4.1
1	C	129	GLU	3.6
2	A	51	VAL	3.5
2	B	199	SER	3.5
2	B	25	ASN	3.5
2	A	160	ASP	3.4
2	B	145	SER	3.3
2	A	42	GLY	3.3
2	A	94	ASN	3.2
2	A	220	VAL	3.1
2	B	130	GLN	3.0
2	A	141	SER	3.0
1	C	31	PHE	3.0

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
2	B	198	CYS	2.9
2	A	31	VAL	2.8
2	B	178	ALA	2.8
2	A	143	GLY	2.7
2	A	178	ALA	2.6
2	B	233	GLN	2.6
2	B	187	SER	2.6
2	B	72	VAL	2.6
2	B	114	ARG	2.6
2	A	113	SER	2.6
2	B	20	TYR	2.6
2	B	148	ASP	2.5
1	C	18	ASN	2.5
2	B	126	SER	2.5
1	C	32	ARG	2.5
2	A	50	VAL	2.5
2	B	220	VAL	2.5
2	B	74	GLU	2.4
2	B	113	SER	2.4
2	A	65	LEU	2.4
2	A	45[A]	ILE	2.4
2	B	182	GLU	2.4
1	C	79	PHE	2.3
2	B	205	ILE	2.3
2	A	193	GLY	2.3
2	B	124	CYS	2.2
2	B	221	TYR	2.2
1	C	28	ALA	2.2
1	C	130	GLU	2.2
2	B	168	PRO	2.2
2	B	224	VAL	2.2
1	C	168	HIS	2.2
1	C	171	VAL	2.1
2	B	123	SER	2.1
1	C	136	LEU	2.1
2	B	204	GLY	2.0
2	A	237	SER	2.0
2	B	122	THR	2.0
2	B	215	LYS	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, 95<sup>th</sup> 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 <sup>2</sup> )	Q<0.9
3	GOL	C	180	6/6	0.75	0.54	15.38	35,35,35,35	6
4	ACT	C	182	4/4	0.66	0.44	5.11	35,35,35,35	4
7	CA	A	5	1/1	0.84	0.32	5.10	35,35,35,35	1
4	ACT	B	2	4/4	0.82	0.26	3.18	35,35,35,35	4
3	GOL	C	179	6/6	0.81	0.39	3.16	35,35,35,35	6
8	EDO	B	3	4/4	0.88	0.28	3.04	35,35,35,35	4
3	GOL	A	239	6/6	0.83	0.24	3.01	35,35,35,35	6
4	ACT	C	181	4/4	0.73	0.26	2.67	35,35,35,35	4
4	ACT	A	6	4/4	0.83	0.20	2.56	35,35,35,35	4
8	EDO	B	7[A]	4/4	0.80	0.32	2.17	36,38,40,41	4
8	EDO	B	7[B]	4/4	0.80	0.32	2.02	27,29,32,33	4
8	EDO	B	8	4/4	0.82	0.23	1.78	35,35,35,35	4
9	PEG	B	240	7/7	0.84	0.27	1.56	35,35,35,35	7
7	CA	A	8	1/1	0.73	0.24	1.51	35,35,35,35	1
3	GOL	A	12	6/6	0.96	0.18	1.41	40,43,44,45	0
6	NA	A	3	1/1	0.79	0.19	1.15	35,35,35,35	1
5	SO4	C	183	5/5	0.86	0.24	0.86	35,35,35,35	5
6	NA	A	7	1/1	0.90	0.17	0.31	45,45,45,45	0
3	GOL	C	178	6/6	0.81	0.26	0.19	35,35,35,35	6
3	GOL	B	1	6/6	0.87	0.18	0.01	35,35,35,35	6
3	GOL	A	2	6/6	0.88	0.19	-0.16	35,35,35,35	6
7	CA	A	9	1/1	0.89	0.16	-0.78	35,35,35,35	1
7	CA	A	4	1/1	0.90	0.10	-	35,35,35,35	1
8	EDO	B	239	4/4	0.80	0.25	-	35,35,35,35	4
5	SO4	A	240	5/5	0.97	0.21	-	35,35,35,35	5

*Continued on next page...*

*Continued from previous page...*

Mol	Type	Chain	Res	Atoms	RSCC	RSR	LLDF	B-factors( $\text{\AA}^2$ )	Q<0.9
6	NA	A	1	1/1	0.89	0.15	-	43,43,43,43	0
7	CA	B	6	1/1	0.85	0.12	-	35,35,35,35	1
7	CA	B	11	1/1	0.85	0.15	-	35,35,35,35	1
7	CA	A	10	1/1	0.86	0.15	-	35,35,35,35	1

## 6.5 Other polymers [i](#)

There are no such residues in this entry.