



# Full wwPDB X-ray Structure Validation Report ⓘ

May 16, 2020 – 11:32 pm BST

PDB ID : 2FO5  
Title : Crystal structure of recombinant barley cysteine endoprotease B isoform 2 (EP-B2) in complex with leupeptin  
Authors : Bethune, M.T.; Strop, P.; Brunger, A.T.; Khosla, C.  
Deposited on : 2006-01-12  
Resolution : 2.20 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

<https://www.wwpdb.org/validation/2017/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 : 1.8.5 (274361), CSD as541be (2020)  
Xtriage (Phenix) : 1.13  
EDS : 2.11  
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) : 2.11

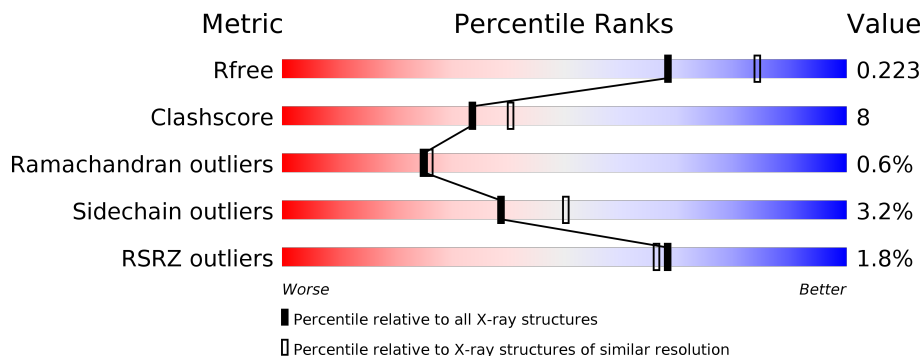
# 1 Overall quality at a glance

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 2.20 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	4898 (2.20-2.20)
Clashscore	141614	5594 (2.20-2.20)
Ramachandran outliers	138981	5503 (2.20-2.20)
Sidechain outliers	138945	5504 (2.20-2.20)
RSRZ outliers	127900	4800 (2.20-2.20)

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 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  $\leq 5\%$ . 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	A	262	<p>2% 72% 13% 15%</p>
1	B	262	<p>2% 71% 12% 15%</p>
1	C	262	<p>2% 73% 12% 15%</p>
1	D	262	<p>2% 71% 14% 15%</p>
2	E	4	<p>25% 50% 25%</p>
2	F	4	<p>25% 75%</p>

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	G	4	 75% 25%
2	H	4	 50% 25% 25%

## 2 Entry composition i

There are 4 unique types of molecules in this entry. The entry contains 7294 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 Cysteine proteinase EP-B 2.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	224	1655	1026	283	336	10	0	0	0
1	B	224	1655	1026	283	336	10	27	0	0
1	C	224	1655	1026	283	336	10	0	0	0
1	D	224	1655	1026	283	336	10	30	0	0

There are 76 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	244	ASN	-	EXPRESSION TAG	UNP P25250
A	245	SER	-	EXPRESSION TAG	UNP P25250
A	246	SER	-	EXPRESSION TAG	UNP P25250
A	247	SER	-	EXPRESSION TAG	UNP P25250
A	248	VAL	-	EXPRESSION TAG	UNP P25250
A	249	ASP	-	EXPRESSION TAG	UNP P25250
A	250	LYS	-	EXPRESSION TAG	UNP P25250
A	251	LEU	-	EXPRESSION TAG	UNP P25250
A	252	ALA	-	EXPRESSION TAG	UNP P25250
A	253	ALA	-	EXPRESSION TAG	UNP P25250
A	254	ALA	-	EXPRESSION TAG	UNP P25250
A	255	LEU	-	EXPRESSION TAG	UNP P25250
A	256	GLU	-	EXPRESSION TAG	UNP P25250
A	257	HIS	-	EXPRESSION TAG	UNP P25250
A	258	HIS	-	EXPRESSION TAG	UNP P25250
A	259	HIS	-	EXPRESSION TAG	UNP P25250
A	260	HIS	-	EXPRESSION TAG	UNP P25250
A	261	HIS	-	EXPRESSION TAG	UNP P25250
A	262	HIS	-	EXPRESSION TAG	UNP P25250
B	244	ASN	-	EXPRESSION TAG	UNP P25250
B	245	SER	-	EXPRESSION TAG	UNP P25250

*Continued on next page...*

*Continued from previous page...*

Chain	Residue	Modelled	Actual	Comment	Reference
B	246	SER	-	EXPRESSION TAG	UNP P25250
B	247	SER	-	EXPRESSION TAG	UNP P25250
B	248	VAL	-	EXPRESSION TAG	UNP P25250
B	249	ASP	-	EXPRESSION TAG	UNP P25250
B	250	LYS	-	EXPRESSION TAG	UNP P25250
B	251	LEU	-	EXPRESSION TAG	UNP P25250
B	252	ALA	-	EXPRESSION TAG	UNP P25250
B	253	ALA	-	EXPRESSION TAG	UNP P25250
B	254	ALA	-	EXPRESSION TAG	UNP P25250
B	255	LEU	-	EXPRESSION TAG	UNP P25250
B	256	GLU	-	EXPRESSION TAG	UNP P25250
B	257	HIS	-	EXPRESSION TAG	UNP P25250
B	258	HIS	-	EXPRESSION TAG	UNP P25250
B	259	HIS	-	EXPRESSION TAG	UNP P25250
B	260	HIS	-	EXPRESSION TAG	UNP P25250
B	261	HIS	-	EXPRESSION TAG	UNP P25250
B	262	HIS	-	EXPRESSION TAG	UNP P25250
C	244	ASN	-	EXPRESSION TAG	UNP P25250
C	245	SER	-	EXPRESSION TAG	UNP P25250
C	246	SER	-	EXPRESSION TAG	UNP P25250
C	247	SER	-	EXPRESSION TAG	UNP P25250
C	248	VAL	-	EXPRESSION TAG	UNP P25250
C	249	ASP	-	EXPRESSION TAG	UNP P25250
C	250	LYS	-	EXPRESSION TAG	UNP P25250
C	251	LEU	-	EXPRESSION TAG	UNP P25250
C	252	ALA	-	EXPRESSION TAG	UNP P25250
C	253	ALA	-	EXPRESSION TAG	UNP P25250
C	254	ALA	-	EXPRESSION TAG	UNP P25250
C	255	LEU	-	EXPRESSION TAG	UNP P25250
C	256	GLU	-	EXPRESSION TAG	UNP P25250
C	257	HIS	-	EXPRESSION TAG	UNP P25250
C	258	HIS	-	EXPRESSION TAG	UNP P25250
C	259	HIS	-	EXPRESSION TAG	UNP P25250
C	260	HIS	-	EXPRESSION TAG	UNP P25250
C	261	HIS	-	EXPRESSION TAG	UNP P25250
C	262	HIS	-	EXPRESSION TAG	UNP P25250
D	244	ASN	-	EXPRESSION TAG	UNP P25250
D	245	SER	-	EXPRESSION TAG	UNP P25250
D	246	SER	-	EXPRESSION TAG	UNP P25250
D	247	SER	-	EXPRESSION TAG	UNP P25250
D	248	VAL	-	EXPRESSION TAG	UNP P25250
D	249	ASP	-	EXPRESSION TAG	UNP P25250

*Continued on next page...*

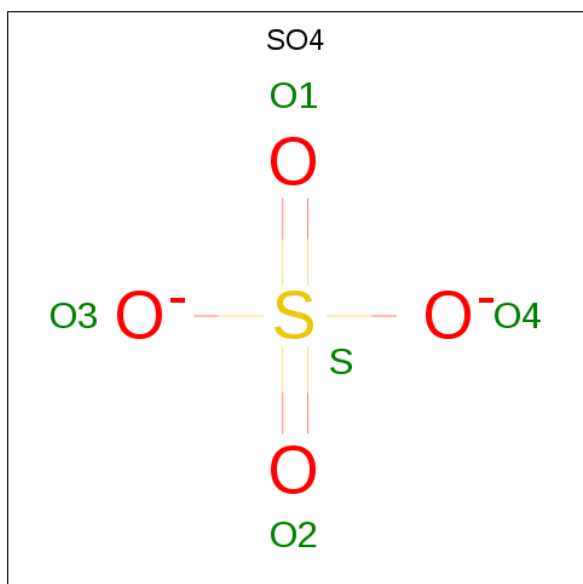
Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
D	250	LYS	-	EXPRESSION TAG	UNP P25250
D	251	LEU	-	EXPRESSION TAG	UNP P25250
D	252	ALA	-	EXPRESSION TAG	UNP P25250
D	253	ALA	-	EXPRESSION TAG	UNP P25250
D	254	ALA	-	EXPRESSION TAG	UNP P25250
D	255	LEU	-	EXPRESSION TAG	UNP P25250
D	256	GLU	-	EXPRESSION TAG	UNP P25250
D	257	HIS	-	EXPRESSION TAG	UNP P25250
D	258	HIS	-	EXPRESSION TAG	UNP P25250
D	259	HIS	-	EXPRESSION TAG	UNP P25250
D	260	HIS	-	EXPRESSION TAG	UNP P25250
D	261	HIS	-	EXPRESSION TAG	UNP P25250
D	262	HIS	-	EXPRESSION TAG	UNP P25250

- Molecule 2 is a protein called ACE-LEU-LEU-argininal (leupeptin).

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf	Trace
			Total	C	N	O			
2	E	4	Total 30	C 20	N 6	O 4	0	0	0
2	F	4	Total 30	C 20	N 6	O 4	0	0	0
2	G	4	Total 30	C 20	N 6	O 4	0	0	0
2	H	4	Total 30	C 20	N 6	O 4	0	0	0

- Molecule 3 is SULFATE ION (three-letter code: SO4) (formula: O<sub>4</sub>S).



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	A	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0
3	B	1	Total O S 5 4 1	0	0
3	C	1	Total O S 5 4 1	0	0
3	C	1	Total O S 5 4 1	0	0
3	D	1	Total O S 5 4 1	0	0
3	D	1	Total O S 5 4 1	0	0

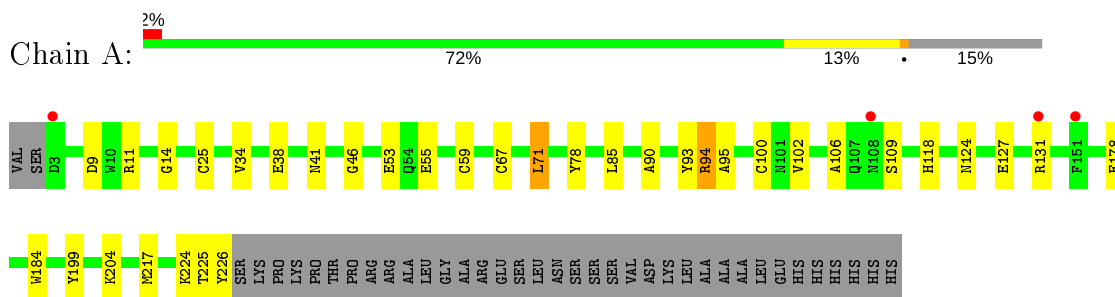
- Molecule 4 is water.

Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
4	A	138	Total O 138 138	0	0
4	B	125	Total O 125 125	0	0
4	C	131	Total O 131 131	0	0
4	D	113	Total O 113 113	0	0
4	E	1	Total O 1 1	0	0
4	H	1	Total O 1 1	0	0

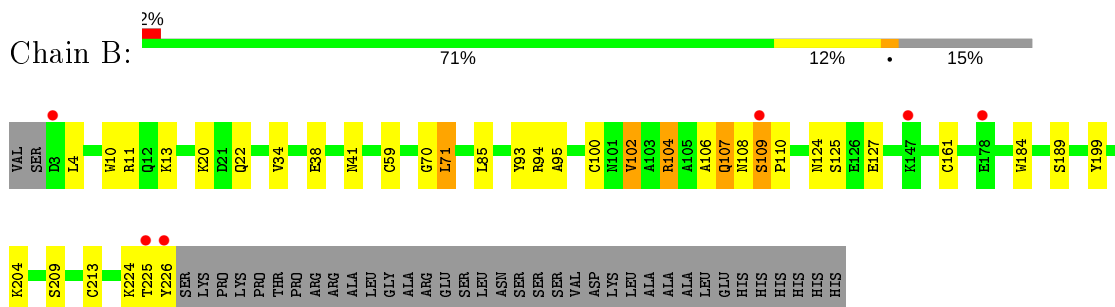
### 3 Residue-property plots [i](#)

These plots are drawn for all protein, RNA and DNA chains in the entry. The first graphic for a chain summarises the proportions of 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.

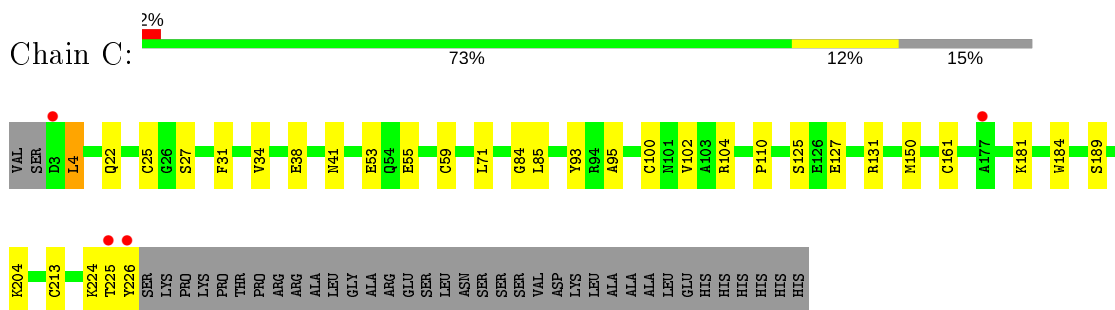
- Molecule 1: Cysteine proteinase EP-B 2



- Molecule 1: Cysteine proteinase EP-B 2



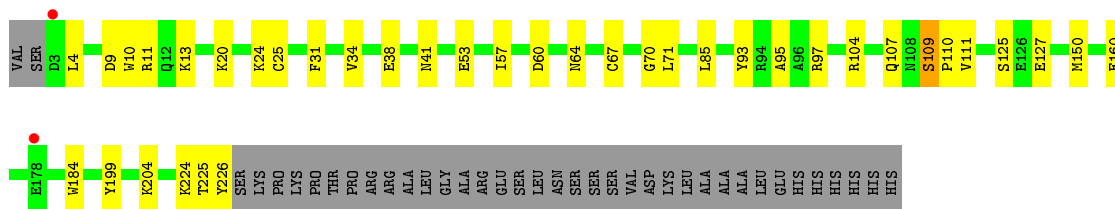
- Molecule 1: Cysteine proteinase EP-B 2



- Molecule 1: Cysteine proteinase EP-B 2







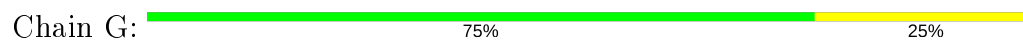
- Molecule 2: ACE-LEU-LEU-argininal (leupeptin)



- Molecule 2: ACE-LEU-LEU-argininal (leupeptin)



- Molecule 2: ACE-LEU-LEU-argininal (leupeptin)



- Molecule 2: ACE-LEU-LEU-argininal (leupeptin)



## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	44.37Å 51.73Å 115.66Å 86.32° 86.46° 79.63°	Depositor
Resolution (Å)	15.00 – 2.20 29.88 – 2.00	Depositor EDS
% Data completeness (in resolution range)	92.0 (15.00-2.20) 79.6 (29.88-2.00)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	2.28 (at 2.01Å)	Xtrriage
Refinement program	CNS	Depositor
R, $R_{free}$	0.201 , 0.233 0.190 , 0.223	Depositor DCC
$R_{free}$ test set	3842 reflections (7.13%)	wwPDB-VP
Wilson B-factor (Å <sup>2</sup> )	22.3	Xtrriage
Anisotropy	0.676	Xtrriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 46.6	EDS
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtrriage
Estimated twinning fraction	No twinning to report.	Xtrriage
$F_o, F_c$ correlation	0.94	EDS
Total number of atoms	7294	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	26.0	wwPDB-VP

Xtrriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 29.37 % of the origin peak, indicating pseudo-translational symmetry. The chance of finding a peak of this or larger height randomly in a structure without pseudo-translational symmetry is equal to 1.5679e-03. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

<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, 0.333 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, ACE, AR7

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	0.34	0/1689	0.60	0/2296
1	B	0.37	0/1689	0.68	2/2296 (0.1%)
1	C	0.34	0/1689	0.61	0/2296
1	D	0.36	0/1689	0.66	1/2296 (0.0%)
2	E	0.58	0/16	2.52	3/21 (14.3%)
2	F	0.60	0/16	1.24	0/21
2	G	0.61	0/16	1.82	0/21
2	H	0.59	0/16	2.48	1/21 (4.8%)
All	All	0.36	0/6820	0.67	7/9268 (0.1%)

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
2	E	0	2
2	F	0	1
2	G	0	1
2	H	0	1
All	All	0	5

There are no bond length outliers.

All (7) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	D	109	SER	N-CA-C	7.31	130.73	111.00
2	E	300	LEU	CA-C-N	-6.79	102.26	117.20
2	H	300	LEU	CA-C-N	-6.44	103.03	117.20
2	E	300	LEU	C-N-CA	-5.76	107.30	121.70

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
1	B	107	GLN	C-N-CA	5.54	135.54	121.70
1	B	110	PRO	N-CA-C	5.37	126.06	112.10
2	E	300	LEU	O-C-N	5.18	130.99	122.70

There are no chirality outliers.

All (5) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
2	E	300	LEU	Mainchain
2	E	301	LEU	Mainchain
2	F	301	LEU	Mainchain
2	G	301	LEU	Mainchain
2	H	300	LEU	Mainchain

## 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	1655	0	1558	29	1
1	B	1655	0	1558	23	1
1	C	1655	0	1558	28	0
1	D	1655	0	1558	22	0
2	E	30	0	37	3	0
2	F	30	0	37	3	0
2	G	30	0	37	0	0
2	H	30	0	37	3	0
3	A	15	0	0	1	0
3	B	10	0	0	0	0
3	C	10	0	0	0	0
3	D	10	0	0	1	0
4	A	138	0	0	4	0
4	B	125	0	0	5	0
4	C	131	0	0	2	0
4	D	113	0	0	2	0
4	E	1	0	0	0	0
4	H	1	0	0	0	0
All	All	7294	0	6380	100	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 8.

All (100) 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:71:LEU:HD11	2:E:299:ACE:H1	1.41	1.01
1:C:55:GLU:HG3	1:C:102:VAL:HG11	1.50	0.93
1:A:55:GLU:HG3	1:A:102:VAL:HG11	1.54	0.90
1:C:110:PRO:HB3	2:H:302:AR7:NH1	1.91	0.86
1:A:78:TYR:HA	3:A:608:SO4:O2	1.80	0.82
1:C:225:THR:O	1:C:226:TYR:HB2	1.79	0.81
1:A:225:THR:O	1:A:226:TYR:HB2	1.83	0.79
1:A:71:LEU:HD13	2:E:300:LEU:HA	1.67	0.77
1:C:59:CYS:HG	1:C:100:CYS:HG	0.76	0.76
1:B:225:THR:O	1:B:226:TYR:HB2	1.86	0.75
1:C:161:CYS:HG	1:C:213:CYS:HG	0.73	0.72
1:C:84:GLY:O	1:D:24:LYS:NZ	2.22	0.72
1:A:59:CYS:HG	1:A:100:CYS:HG	0.76	0.70
1:B:124:ASN:O	4:B:714:HOH:O	2.09	0.69
1:D:225:THR:O	1:D:226:TYR:HB2	1.94	0.68
1:B:209:SER:O	4:B:714:HOH:O	2.13	0.67
1:A:34:VAL:HG21	1:A:53:GLU:HG3	1.79	0.65
1:D:10:TRP:CE3	1:D:13:LYS:HE3	2.36	0.60
1:A:55:GLU:CG	1:A:102:VAL:HG11	2.30	0.60
1:D:41:ASN:HD22	1:D:224:LYS:HE2	1.68	0.59
1:A:106:ALA:HA	1:A:109:SER:O	2.04	0.58
1:A:46:GLY:HA3	4:B:623:HOH:O	2.03	0.58
1:B:161:CYS:HG	1:B:213:CYS:HG	0.77	0.58
1:D:25:CYS:HG	1:D:67:CYS:CB	2.17	0.57
1:B:225:THR:O	1:B:226:TYR:CB	2.54	0.56
1:C:4:LEU:HD21	1:C:131:ARG:HG2	1.88	0.56
1:B:125:SER:OG	1:B:127:GLU:HG2	2.07	0.55
1:B:204:LYS:HE3	4:B:649:HOH:O	2.06	0.55
1:C:204:LYS:HE3	4:C:719:HOH:O	2.07	0.55
1:A:131:ARG:HD2	4:A:692:HOH:O	2.07	0.54
1:A:204:LYS:HE3	4:A:675:HOH:O	2.07	0.54
1:B:41:ASN:HD22	1:B:224:LYS:HE2	1.73	0.54
1:C:41:ASN:HD22	1:C:224:LYS:HE2	1.71	0.53
1:C:4:LEU:HD21	1:C:131:ARG:CG	2.39	0.53
1:D:125:SER:OG	1:D:127:GLU:HG2	2.08	0.52
1:C:55:GLU:CG	1:C:102:VAL:HG11	2.31	0.52
1:B:71:LEU:HD11	2:F:299:ACE:H1	1.91	0.51

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:A:34:VAL:CG2	1:A:53:GLU:HG3	2.39	0.51
1:C:125:SER:OG	1:C:127:GLU:HG2	2.11	0.51
1:A:184:TRP:CH2	1:A:204:LYS:HD2	2.46	0.51
1:C:225:THR:O	1:C:226:TYR:CB	2.56	0.50
1:D:31:PHE:CZ	1:D:53:GLU:HG3	2.46	0.50
1:B:184:TRP:CH2	1:B:204:LYS:HD2	2.47	0.50
1:D:60:ASP:OD2	1:D:64:ASN:ND2	2.42	0.50
1:A:41:ASN:HD22	1:A:224:LYS:HE2	1.77	0.50
1:D:34:VAL:O	1:D:38:GLU:HG3	2.10	0.50
1:C:55:GLU:HG3	1:C:102:VAL:CG1	2.33	0.49
1:B:70:GLY:N	2:F:300:LEU:HB3	2.27	0.49
1:C:31:PHE:CZ	1:C:53:GLU:HG3	2.48	0.49
1:B:184:TRP:CZ2	1:B:204:LYS:HD2	2.48	0.49
1:D:13:LYS:HE2	4:D:717:HOH:O	2.13	0.48
1:B:34:VAL:O	1:B:38:GLU:HG3	2.11	0.48
1:C:34:VAL:O	1:C:38:GLU:HG3	2.14	0.48
1:C:184:TRP:CH2	1:C:204:LYS:HD2	2.49	0.48
1:C:184:TRP:CZ2	1:C:204:LYS:HD2	2.49	0.48
1:B:11:ARG:HD3	1:B:199:TYR:CZ	2.48	0.48
1:A:90:ALA:HB2	1:A:109:SER:HB3	1.95	0.47
1:A:11:ARG:HD3	1:A:199:TYR:CZ	2.50	0.47
1:C:161:CYS:CB	1:C:213:CYS:HG	2.26	0.47
1:D:184:TRP:CH2	1:D:204:LYS:HD2	2.50	0.47
1:A:93:TYR:CZ	1:A:95:ALA:HA	2.50	0.46
1:B:104:ARG:HG2	1:B:104:ARG:H	1.21	0.46
1:A:225:THR:O	1:A:226:TYR:CB	2.58	0.46
1:A:55:GLU:HG3	1:A:102:VAL:CG1	2.35	0.46
1:D:11:ARG:HD3	1:D:199:TYR:CZ	2.50	0.46
1:A:94:ARG:HH11	1:A:94:ARG:HG3	1.81	0.45
1:A:124:ASN:OD1	1:A:217:MET:HG2	2.16	0.45
1:C:181:LYS:HE3	1:C:181:LYS:HB2	1.78	0.45
1:D:57:ILE:HG22	1:D:97:ARG:HG3	1.99	0.45
1:A:118:HIS:HD2	4:A:651:HOH:O	1.98	0.45
1:B:59:CYS:CB	1:B:100:CYS:HG	2.25	0.45
1:D:70:GLY:N	2:H:300:LEU:HB3	2.31	0.45
1:D:24:LYS:HB2	1:D:24:LYS:HE3	1.80	0.44
1:D:184:TRP:CZ2	1:D:204:LYS:HD2	2.52	0.44
1:B:20:LYS:HE2	4:B:663:HOH:O	2.18	0.44
1:A:14:GLY:HA2	4:A:654:HOH:O	2.17	0.44
1:A:184:TRP:CZ2	1:A:204:LYS:HD2	2.53	0.44
1:C:131:ARG:NE	4:C:641:HOH:O	2.51	0.44

*Continued on next page...*

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:93:TYR:CZ	1:B:95:ALA:HA	2.53	0.43
1:D:204:LYS:HE3	4:D:694:HOH:O	2.18	0.43
1:C:104:ARG:HD2	1:D:150:MET:HG2	2.01	0.43
1:C:110:PRO:HB3	2:H:302:AR7:HN12	1.76	0.43
1:A:71:LEU:CD1	2:E:300:LEU:HA	2.43	0.43
1:C:22:GLN:NE2	1:C:189:SER:OG	2.52	0.43
1:B:71:LEU:HD13	2:F:300:LEU:HA	2.01	0.42
1:A:34:VAL:O	1:A:38:GLU:HG3	2.19	0.42
1:B:59:CYS:SG	1:B:102:VAL:HG22	2.59	0.42
1:B:10:TRP:CE3	1:B:13:LYS:HE3	2.55	0.42
1:C:25:CYS:SG	1:C:27:SER:HB3	2.58	0.42
1:D:93:TYR:CZ	1:D:95:ALA:HA	2.55	0.42
1:C:4:LEU:HD11	1:C:131:ARG:HG2	2.02	0.42
1:A:25:CYS:HG	1:A:67:CYS:HG	0.45	0.41
1:D:9:ASP:OD1	1:D:11:ARG:HB2	2.19	0.41
1:C:127:GLU:H	1:C:127:GLU:CD	2.24	0.41
1:D:127:GLU:CD	1:D:127:GLU:H	2.23	0.41
1:B:127:GLU:H	1:B:127:GLU:CD	2.23	0.41
1:A:9:ASP:OD1	1:A:11:ARG:HB2	2.21	0.41
1:B:22:GLN:NE2	1:B:189:SER:OG	2.48	0.41
1:C:93:TYR:CZ	1:C:95:ALA:HA	2.56	0.40
1:D:20:LYS:NZ	3:D:607:SO4:O4	2.43	0.40

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:A:178:GLU:OE1	1:B:109:SER:OG[1_545]	2.16	0.04

## 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	222/262 (85%)	214 (96%)	8 (4%)	0	100	100
1	B	222/262 (85%)	208 (94%)	11 (5%)	3 (1%)	11	8
1	C	222/262 (85%)	214 (96%)	8 (4%)	0	100	100
1	D	222/262 (85%)	209 (94%)	11 (5%)	2 (1%)	17	16
2	E	2/4 (50%)	2 (100%)	0	0	100	100
2	F	2/4 (50%)	2 (100%)	0	0	100	100
2	G	2/4 (50%)	2 (100%)	0	0	100	100
2	H	2/4 (50%)	2 (100%)	0	0	100	100
All	All	896/1064 (84%)	853 (95%)	38 (4%)	5 (1%)	25	26

All (5) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	B	108	ASN
1	D	109	SER
1	D	111	VAL
1	B	106	ALA
1	B	109	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	171/203 (84%)	167 (98%)	4 (2%)	50	63
1	B	171/203 (84%)	164 (96%)	7 (4%)	30	39
1	C	171/203 (84%)	167 (98%)	4 (2%)	50	63
1	D	171/203 (84%)	164 (96%)	7 (4%)	30	39
2	E	2/2 (100%)	2 (100%)	0	100	100
2	F	2/2 (100%)	2 (100%)	0	100	100
2	G	2/2 (100%)	2 (100%)	0	100	100
2	H	2/2 (100%)	2 (100%)	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles
All	All	692/820 (84%)	670 (97%)	22 (3%)	39 50

All (22) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	71	LEU
1	A	85	LEU
1	A	94	ARG
1	A	127	GLU
1	B	4	LEU
1	B	71	LEU
1	B	85	LEU
1	B	94	ARG
1	B	102	VAL
1	B	104	ARG
1	B	107	GLN
1	C	4	LEU
1	C	71	LEU
1	C	85	LEU
1	C	150	MET
1	D	4	LEU
1	D	71	LEU
1	D	85	LEU
1	D	104	ARG
1	D	107	GLN
1	D	110	PRO
1	D	160	GLU

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

Mol	Chain	Res	Type
1	A	22	GLN
1	A	41	ASN
1	A	118	HIS
1	B	22	GLN
1	B	41	ASN
1	B	101	ASN
1	C	22	GLN
1	C	41	ASN
1	C	82	ASN
1	D	41	ASN

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
1	D	118	HIS

### 5.3.3 RNA [i](#)

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains [i](#)

4 non-standard protein/DNA/RNA residues 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	AR7	E	302	-	10,10,11	1.46	1 (10%)	9,11,13	1.03	1 (11%)
2	AR7	G	302	-	10,10,11	1.46	1 (10%)	9,11,13	1.01	1 (11%)
2	AR7	F	302	-	10,10,11	1.43	1 (10%)	9,11,13	0.98	1 (11%)
2	AR7	H	302	-	10,10,11	1.45	1 (10%)	9,11,13	0.96	1 (11%)

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	AR7	E	302	-	-	1/9/9/11	-
2	AR7	G	302	-	-	1/9/9/11	-
2	AR7	F	302	-	-	2/9/9/11	-
2	AR7	H	302	-	-	1/9/9/11	-

All (4) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	G	302	AR7	O-C	-4.52	1.23	1.42

*Continued on next page...*

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	E	302	AR7	O-C	-4.52	1.23	1.42
2	H	302	AR7	O-C	-4.51	1.23	1.42
2	F	302	AR7	O-C	-4.43	1.23	1.42

All (4) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	F	302	AR7	O-C-CA	2.34	120.78	111.52
2	H	302	AR7	O-C-CA	2.31	120.66	111.52
2	G	302	AR7	O-C-CA	2.29	120.58	111.52
2	E	302	AR7	O-C-CA	2.26	120.44	111.52

There are no chirality outliers.

All (5) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	F	302	AR7	CG-CD-NE-CZ
2	E	302	AR7	O-C-CA-N
2	G	302	AR7	O-C-CA-N
2	F	302	AR7	O-C-CA-N
2	H	302	AR7	O-C-CA-N

There are no ring outliers.

1 monomer is involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	H	302	AR7	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

9 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	$\# Z  > 2$
3	SO4	C	605	-	4,4,4	0.23	0	6,6,6	0.12	0
3	SO4	B	603	-	4,4,4	0.22	0	6,6,6	0.11	0
3	SO4	D	607	-	4,4,4	0.21	0	6,6,6	0.07	0
3	SO4	D	606	-	4,4,4	0.26	0	6,6,6	0.05	0
3	SO4	C	604	-	4,4,4	0.25	0	6,6,6	0.08	0
3	SO4	A	600	-	4,4,4	0.26	0	6,6,6	0.06	0
3	SO4	B	602	-	4,4,4	0.27	0	6,6,6	0.07	0
3	SO4	A	608	-	4,4,4	0.22	0	6,6,6	0.08	0
3	SO4	A	601	-	4,4,4	0.25	0	6,6,6	0.15	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.

2 monomers are involved in 2 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	D	607	SO4	1	0
3	A	608	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, 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	A	224/262 (85%)	-0.20	4 (1%) 68 66	14, 23, 38, 55	0
1	B	220/262 (83%)	-0.27	6 (2%) 54 52	13, 24, 41, 69	0
1	C	224/262 (85%)	-0.23	4 (1%) 68 66	14, 25, 38, 63	0
1	D	220/262 (83%)	-0.11	2 (0%) 84 83	15, 24, 40, 56	0
2	E	2/4 (50%)	-0.16	0 100 100	22, 22, 22, 32	0
2	F	2/4 (50%)	-0.50	0 100 100	22, 22, 22, 33	0
2	G	2/4 (50%)	-0.25	0 100 100	26, 26, 26, 38	0
2	H	2/4 (50%)	0.04	0 100 100	20, 20, 20, 36	0
All	All	896/1064 (84%)	-0.20	16 (1%) 68 66	13, 24, 40, 69	0

All (16) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	226	TYR	7.9
1	B	226	TYR	7.6
1	D	3	ASP	5.1
1	B	178	GLU	4.5
1	C	225	THR	3.8
1	A	3	ASP	3.6
1	A	108	ASN	3.4
1	A	131	ARG	3.1
1	B	225	THR	2.9
1	D	178	GLU	2.8
1	B	147	LYS	2.8
1	C	3	ASP	2.6
1	B	3	ASP	2.6
1	A	151	PHE	2.4
1	B	109	SER	2.3
1	C	177	ALA	2.2

## 6.2 Non-standard residues in protein, DNA, RNA chains [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<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	B-factors(Å <sup>2</sup> )	Q<0.9
2	AR7	H	302	11/12	0.87	0.19	22,28,48,49	0
2	AR7	G	302	11/12	0.90	0.17	26,28,45,46	0
2	AR7	F	302	11/12	0.92	0.15	18,19,38,42	0
2	AR7	E	302	11/12	0.93	0.14	24,27,39,40	0

## 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. 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	B-factors(Å <sup>2</sup> )	Q<0.9
3	SO4	D	606	5/5	0.83	0.20	77,79,80,82	0
3	SO4	B	602	5/5	0.86	0.29	72,74,75,76	0
3	SO4	C	604	5/5	0.89	0.23	55,58,61,63	0
3	SO4	A	608	5/5	0.89	0.25	88,88,89,90	0
3	SO4	B	603	5/5	0.92	0.12	53,56,57,58	0
3	SO4	A	601	5/5	0.92	0.17	55,56,59,61	0
3	SO4	C	605	5/5	0.93	0.20	57,59,61,62	0
3	SO4	A	600	5/5	0.94	0.21	62,62,65,65	0
3	SO4	D	607	5/5	0.96	0.17	58,60,62,62	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.