



# Full wwPDB X-ray Structure Validation Report ⓘ

Feb 1, 2016 – 03:38 PM GMT

PDB ID : 4CH8  
Title : High-salt crystal structure of a thrombin-GpIbalpha peptide complex  
Authors : Lechtenberg, B.C.; Freund, S.M.V.; Huntington, J.A.  
Deposited on : 2013-11-29  
Resolution : 1.75 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

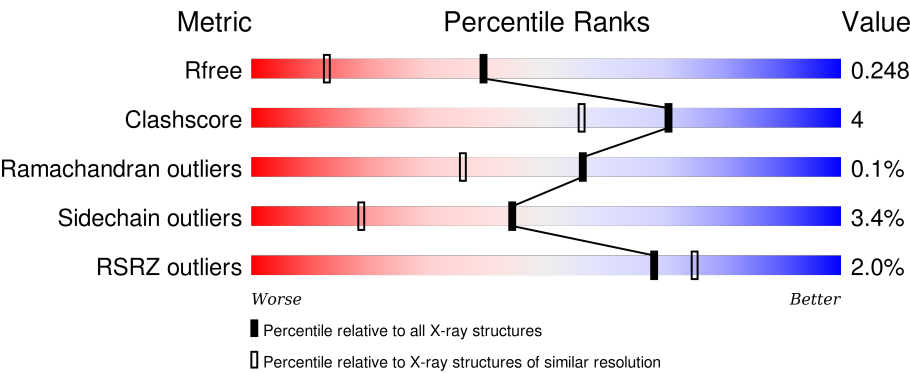
# 1 Overall quality at a glance i

The following experimental techniques were used to determine the structure:

*X-RAY DIFFRACTION*

The reported resolution of this entry is 1.75 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 <sub>free</sub>	91344	1609 (1.76-1.76)
Clashscore	102246	1730 (1.76-1.76)
Ramachandran outliers	100387	1711 (1.76-1.76)
Sidechain outliers	100360	1711 (1.76-1.76)
RSRZ outliers	91569	1610 (1.76-1.76)

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	36	<div><div>14%</div><div>67%6%28%</div></div>
1	C	36	<div><div>11%</div><div>78%11%11%</div></div>
1	E	36	<div><div>6%</div><div>72%28%</div></div>
1	G	36	<div><div></div><div>67%6%25%</div></div>
2	B	259	<div><div>2%</div><div>82%13%.</div></div>

Continued on next page...

*Continued from previous page...*

Mol	Chain	Length	Quality of chain
2	D	259	
2	F	259	
2	H	259	
3	P	14	
3	Q	14	
3	R	14	
3	S	14	

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
4	NA	B	1578	-	-	-	X
4	NA	F	1578	-	-	-	X
6	GOL	B	1580	-	-	-	X

## 2 Entry composition

There are 7 unique types of molecules in this entry. The entry contains 9557 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 THROMBIN, LIGHT CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
1	A	26	Total	C	N	O	S	0	0	0
			184	115	29	39	1			
1	C	32	Total	C	N	O	S	0	0	0
			243	151	40	51	1			
1	E	26	Total	C	N	O	S	0	0	0
			202	128	31	42	1			
1	G	27	Total	C	N	O	S	0	1	0
			222	140	34	47	1			

- Molecule 2 is a protein called THROMBIN, HEAVY CHAIN.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
2	B	249	Total	C	N	O	S	1	4	0
			1995	1274	350	356	15			
2	D	252	Total	C	N	O	S	0	2	0
			2014	1289	348	363	14			
2	F	253	Total	C	N	O	S	3	2	0
			1985	1278	339	354	14			
2	H	259	Total	C	N	O	S	0	0	0
			2016	1289	356	357	14			

- Molecule 3 is a protein called PLATELET GLYCOPROTEIN IB ALPHA CHAIN, RESIDUES 287-300.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
3	P	3	Total	C	N	O	P	0	0	0
			32	19	3	9	1			
3	Q	4	Total	C	N	O	P	0	0	0
			40	23	4	12	1			
3	R	3	Total	C	N	O	P	0	0	0
			32	19	3	9	1			
3	S	3	Total	C	N	O	P	0	0	0
			32	19	3	9	1			

- | Mol | Chain | Residues | Atoms           | ZeroOcc | AltConf |
|-----|-------|----------|-----------------|---------|---------|
| 4   | H     | 1        | Total Na<br>1 1 | 0       | 0       |
| 4   | B     | 1        | Total Na<br>1 1 | 0       | 0       |
| 4   | D     | 1        | Total Na<br>1 1 | 0       | 0       |
| 4   | F     | 1        | Total Na<br>1 1 | 0       | 0       |

- 
- Chemical structure of OG6, showing a complex molecule with a benzene ring, a pyrrolidine ring, and a guanidinium group. The structure is labeled with various atoms and bonds, including a central nitrogen atom (N1) and a carbonyl group (C=O). The molecule is shown in a 3D representation with stereochemistry indicated by wedges and dashes.

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
5	B	1	Total 30	C 21	N 6	O 3	0	0
5	D	1	Total 30	C 21	N 6	O 3	0	0
5	F	1	Total 30	C 21	N 6	O 3	0	0
5	H	1	Total 30	C 21	N 6	O 3	0	0

- 



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

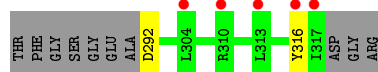
- Molecule 7 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
7	A	6	Total	O	0	0
			6	6		
7	B	94	Total	O	0	0
			94	94		
7	C	14	Total	O	0	0
			14	14		
7	D	108	Total	O	0	0
			108	108		
7	E	7	Total	O	0	0
			7	7		
7	F	94	Total	O	0	0
			94	94		
7	G	11	Total	O	0	0
			11	11		
7	H	90	Total	O	0	0
			90	90		

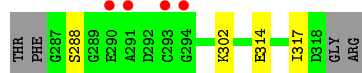
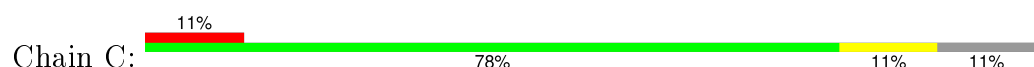
### 3 Residue-property plots [i](#)

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.

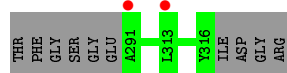
#### • Molecule 1: THROMBIN, LIGHT CHAIN



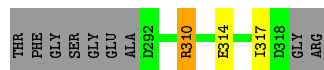
#### • Molecule 1: THROMBIN, LIGHT CHAIN



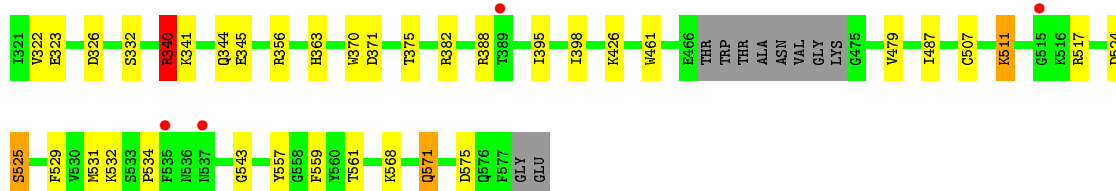
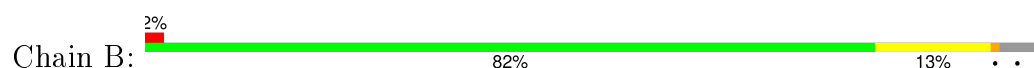
#### • Molecule 1: THROMBIN, LIGHT CHAIN



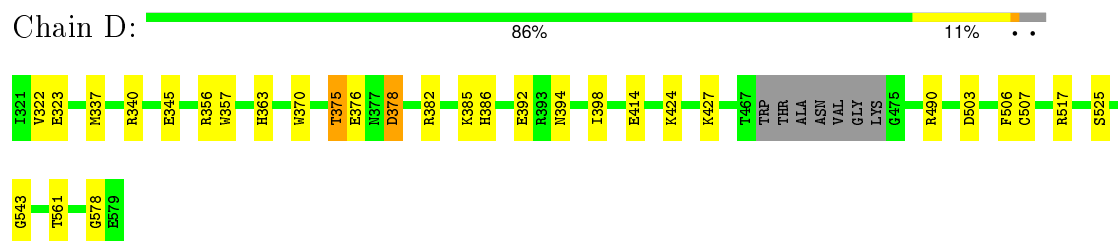
#### • Molecule 1: THROMBIN, LIGHT CHAIN



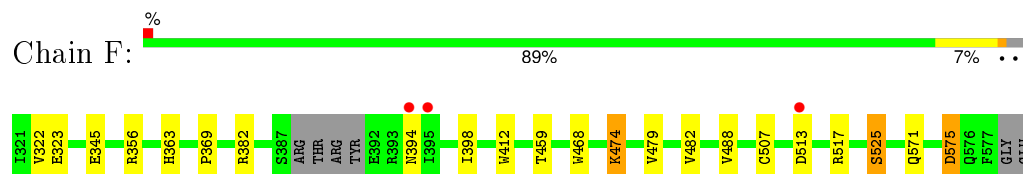
#### • Molecule 2: THROMBIN, HEAVY CHAIN



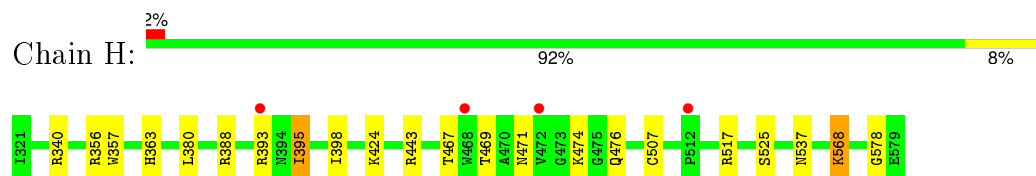
- Molecule 2: THROMBIN, HEAVY CHAIN



- Molecule 2: THROMBIN, HEAVY CHAIN



- Molecule 2: THROMBIN, HEAVY CHAIN



- Molecule 3: PLATELET GLYCOPROTEIN IB ALPHA CHAIN, RESIDUES 287-300



- Molecule 3: PLATELET GLYCOPROTEIN IB ALPHA CHAIN, RESIDUES 287-300



- Molecule 3: PLATELET GLYCOPROTEIN IB ALPHA CHAIN, RESIDUES 287-300



- Molecule 3: PLATELET GLYCOPROTEIN IB ALPHA CHAIN, RESIDUES 287-300





GLY	ASP	THR	ASP	L275	Y276	D277	PTR	PRQ	GLU	GLU	ASP	THR
-----	-----	-----	-----	------	------	------	-----	-----	-----	-----	-----	-----

## 4 Data and refinement statistics

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, $\alpha$ , $\beta$ , $\gamma$	74.22Å 51.33Å 150.08Å 90.00° 96.13° 90.00°	Depositor
Resolution (Å)	42.14 – 1.75 42.14 – 1.75	Depositor EDS
% Data completeness (in resolution range)	92.9 (42.14-1.75) 92.9 (42.14-1.75)	Depositor EDS
$R_{merge}$	0.06	Depositor
$R_{sym}$	(Not available)	Depositor
$\langle I/\sigma(I) \rangle$ <sup>1</sup>	1.81 (at 1.75Å)	Xtriage
Refinement program	REFMAC 5.7.0029	Depositor
R, $R_{free}$	0.210 , 0.242 0.215 , 0.248	Depositor DCC
$R_{free}$ test set	5305 reflections (5.28%)	DCC
Wilson B-factor (Å <sup>2</sup> )	23.3	Xtriage
Anisotropy	0.499	Xtriage
Bulk solvent $k_{sol}$ (e/Å <sup>3</sup> ), $B_{sol}$ (Å <sup>2</sup> )	0.37 , 48.6	EDS
Estimated twinning fraction	No twinning to report.	Xtriage
L-test for twinning <sup>2</sup>	$\langle  L  \rangle = 0.48$ , $\langle L^2 \rangle = 0.31$	Xtriage
Outliers	0 of 105837 reflections	Xtriage
$F_o, F_c$ correlation	0.95	EDS
Total number of atoms	9557	wwPDB-VP
Average B, all atoms (Å <sup>2</sup> )	35.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The analyses of the Patterson function reveals a significant off-origin peak that is 25.79 % 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 2.9676e-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, 0.375 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: 0G6, GOL, PTR, NA

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	0.70	0/186	0.71	0/253
1	C	0.75	0/245	0.85	0/327
1	E	0.67	0/204	0.67	0/273
1	G	0.69	0/227	0.83	0/304
2	B	1.03	2/2046 (0.1%)	1.07	11/2772 (0.4%)
2	D	1.06	2/2069 (0.1%)	1.03	5/2802 (0.2%)
2	F	0.91	1/2043 (0.0%)	0.84	1/2774 (0.0%)
2	H	0.81	1/2068 (0.0%)	0.82	2/2803 (0.1%)
3	P	1.30	0/14	1.03	0/16
3	Q	0.90	0/22	1.10	0/27
3	R	0.87	0/14	1.07	0/16
3	S	1.40	0/14	1.52	0/16
All	All	0.94	6/9152 (0.1%)	0.94	19/12383 (0.2%)

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
2	D	0	1
2	H	0	1
All	All	0	3

All (6) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	F	525	SER	CB-OG	22.00	1.70	1.42
2	B	525	SER	CB-OG	17.40	1.64	1.42
2	H	525	SER	CB-OG	10.58	1.56	1.42

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
2	D	525	SER	CB-OG	10.14	1.55	1.42
2	D	506	PHE	CD2-CE2	-5.30	1.28	1.39
2	B	557	TYR	CG-CD1	-5.16	1.32	1.39

All (19) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
2	D	356	ARG	NE-CZ-NH2	-11.53	114.54	120.30
2	B	340[A]	ARG	NE-CZ-NH1	10.12	125.36	120.30
2	B	340[B]	ARG	NE-CZ-NH1	10.12	125.36	120.30
2	D	356	ARG	NE-CZ-NH1	9.31	124.96	120.30
2	B	571	GLN	CG-CD-OE1	-9.15	103.30	121.60
2	B	575	ASP	CB-CG-OD1	8.48	125.93	118.30
2	B	356	ARG	NE-CZ-NH2	-8.24	116.18	120.30
2	B	356	ARG	NE-CZ-NH1	7.17	123.88	120.30
2	B	340[A]	ARG	NE-CZ-NH2	-6.86	116.87	120.30
2	B	340[B]	ARG	NE-CZ-NH2	-6.86	116.87	120.30
2	B	524	ASP	CB-CG-OD1	6.50	124.15	118.30
2	F	356	ARG	NE-CZ-NH2	-5.76	117.42	120.30
2	D	506	PHE	CG-CD2-CE2	5.67	127.04	120.80
2	D	503	ASP	CB-CG-OD2	-5.63	113.23	118.30
2	B	575	ASP	CB-CG-OD2	-5.53	113.32	118.30
2	D	490	ARG	NE-CZ-NH2	-5.20	117.70	120.30
2	B	559	PHE	CB-CG-CD1	-5.12	117.22	120.80
2	H	340	ARG	NE-CZ-NH2	-5.11	117.75	120.30
2	H	356	ARG	NE-CZ-NH2	-5.04	117.78	120.30

There are no chirality outliers.

All (3) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	C	288	SER	Peptide
2	D	578	GLY	Peptide
2	H	578	GLY	Peptide

## 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	184	0	144	2	0
1	C	243	0	228	1	0
1	E	202	0	189	0	0
1	G	222	0	207	2	0
2	B	1995	0	1920	31	0
2	D	2014	0	1944	18	0
2	F	1985	0	1916	16	0
2	H	2016	0	1939	11	0
3	P	32	0	21	0	0
3	Q	40	0	25	0	0
3	R	32	0	21	1	0
3	S	32	0	21	3	0
4	B	1	0	0	0	0
4	D	1	0	0	0	0
4	F	1	0	0	0	0
4	H	1	0	0	0	0
5	B	30	0	31	4	0
5	D	30	0	31	2	0
5	F	30	0	31	4	0
5	H	30	0	31	3	0
6	B	6	0	8	2	0
6	F	6	0	8	2	0
7	A	6	0	0	0	0
7	B	94	0	0	2	0
7	C	14	0	0	0	0
7	D	108	0	0	3	0
7	E	7	0	0	0	0
7	F	94	0	0	1	0
7	G	11	0	0	1	0
7	H	90	0	0	1	0
All	All	9557	0	8715	79	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 4.

All (79) 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:525:SER:OG	2:B:525:SER:CB	1.64	1.43
2:F:363:HIS:NE2	5:F:1579:OG6:C3	1.86	1.38
2:F:525:SER:OG	2:F:525:SER:CB	1.70	1.36
2:H:363:HIS:NE2	5:H:1581:OG6:C3	1.98	1.25

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:B:363:HIS:NE2	5:B:1579:0G6:C3	2.00	1.25
2:B:487:ILE:HD11	2:B:531[A]:MET:CE	1.67	1.22
2:D:363:HIS:NE2	5:D:1581:0G6:C3	2.02	1.22
2:B:487:ILE:HD11	2:B:531[A]:MET:HE1	1.29	1.08
2:B:487:ILE:CD1	2:B:531[A]:MET:HE1	1.93	0.99
2:B:487:ILE:HD11	2:B:531[A]:MET:HE3	1.50	0.92
2:F:525:SER:CB	5:F:1579:0G6:C2	2.59	0.80
2:F:459:THR:HG22	2:F:482[B]:VAL:HG22	1.64	0.80
2:D:340[B]:ARG:NH1	2:D:345:GLU:OE1	2.15	0.79
2:B:345:GLU:OE1	7:B:2010:HOH:O	2.05	0.75
2:B:487:ILE:CD1	2:B:531[A]:MET:CE	2.55	0.74
1:G:310:ARG:NH1	7:G:2006:HOH:O	2.20	0.74
2:B:525:SER:CB	5:B:1579:0G6:C2	2.66	0.71
2:F:363:HIS:CD2	5:F:1579:0G6:C3	2.73	0.71
2:D:414:GLU:OE2	7:D:2045:HOH:O	2.09	0.70
2:B:340[A]:ARG:NH2	7:B:2009:HOH:O	2.25	0.70
2:B:326:ASP:CG	2:B:479:VAL:HG21	2.13	0.69
2:F:382:ARG:HH12	2:F:398:ILE:HD11	1.58	0.68
2:F:345:GLU:OE1	7:F:2017:HOH:O	2.12	0.67
1:A:316:TYR:OH	2:B:532:LYS:O	2.08	0.65
2:B:382:ARG:HH12	2:B:398:ILE:HD11	1.62	0.64
2:D:382:ARG:HH12	2:D:398:ILE:HD11	1.62	0.63
2:B:511:LYS:H	2:B:511:LYS:HD2	1.64	0.62
2:D:375:THR:OG1	2:D:376:GLU:OE1	2.17	0.61
2:B:371:ASP:O	2:D:340[B]:ARG:NH2	2.34	0.60
2:D:427:LYS:O	7:D:2034:HOH:O	2.16	0.60
2:B:363:HIS:CD2	5:B:1579:0G6:C3	2.84	0.59
2:H:363:HIS:CD2	5:H:1581:0G6:C3	2.82	0.59
2:B:525:SER:OG	2:B:525:SER:CA	2.50	0.56
2:H:388:ARG:NH1	2:H:476:GLN:OE1	2.37	0.56
1:G:314[A]:GLU:OE2	1:G:317:ILE:HD11	2.06	0.55
2:F:488:VAL:O	6:F:1580:GOL:C3	2.56	0.54
2:B:487:ILE:HD12	2:B:531[A]:MET:HE1	1.88	0.53
2:H:568:LYS:HE3	3:S:276:PTR:CE1	2.37	0.53
2:H:443:ARG:NH1	3:S:276:PTR:O2P	2.42	0.52
2:D:386:HIS:HB3	2:D:392:GLU:OE2	2.11	0.51
2:B:388:ARG:HG3	2:B:461[B]:TRP:CZ3	2.46	0.51
2:H:393:ARG:O	2:H:395:ILE:HD13	2.11	0.51
2:D:376:GLU:CD	2:D:376:GLU:H	2.15	0.50
2:B:326:ASP:OD2	2:B:479:VAL:HG21	2.13	0.49
2:F:525:SER:OG	2:F:525:SER:CA	2.51	0.48

*Continued on next page...*

*Continued from previous page...*

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
2:F:382:ARG:NH1	2:F:398:ILE:HD11	2.27	0.48
2:F:488:VAL:O	6:F:1580:GOL:H32	2.13	0.48
2:B:363:HIS:NE2	5:B:1579:0G6:C2	2.73	0.47
2:H:380:LEU:C	2:H:380:LEU:HD12	2.34	0.47
2:D:363:HIS:CD2	5:D:1581:0G6:C3	2.92	0.47
2:D:394:ASN:O	7:D:2031:HOH:O	2.20	0.47
2:B:332:SER:C	6:B:1580:GOL:H11	2.36	0.46
2:H:357:TRP:CZ2	2:H:424:LYS:HD2	2.50	0.46
2:H:537:ASN:HB3	7:H:2079:HOH:O	2.15	0.46
2:F:363:HIS:NE2	5:F:1579:0G6:C2	2.75	0.46
2:H:471:ASN:ND2	2:H:474:LYS:O	2.49	0.45
2:D:375:THR:N	2:D:378:ASP:OD2	2.39	0.44
2:F:571:GLN:O	2:F:575:ASP:HB2	2.18	0.44
2:B:382:ARG:NH1	2:B:398:ILE:HD11	2.30	0.44
2:D:357:TRP:CZ2	2:D:424:LYS:HD2	2.53	0.44
2:B:370:TRP:HB3	2:D:370:TRP:HB3	2.00	0.44
1:C:314:GLU:O	1:C:317:ILE:HG13	2.17	0.43
2:H:363:HIS:NE2	5:H:1581:0G6:C2	2.76	0.43
2:D:543:GLY:HA2	2:D:561:THR:O	2.17	0.43
2:F:468:TRP:HA	2:F:474:LYS:HB2	2.00	0.43
2:B:322:VAL:HG12	2:B:323:GLU:HG2	1.99	0.43
2:B:543:GLY:HA2	2:B:561:THR:O	2.19	0.43
1:A:316:TYR:CE2	2:B:534:PRO:HD3	2.55	0.42
2:B:332:SER:H	6:B:1580:GOL:C1	2.33	0.42
2:D:322:VAL:HG12	2:D:323:GLU:HG2	2.02	0.41
2:D:382:ARG:NH1	2:D:398:ILE:HD11	2.31	0.41
3:R:276:PTR:O	3:R:277:ASP:C	2.57	0.41
2:B:529:PHE:HE1	2:B:531[A]:MET:CE	2.33	0.41
3:S:275:LEU:HA	3:S:275:LEU:HD12	1.92	0.41
2:F:369:PRO:HD3	2:F:412:TRP:CE3	2.56	0.41
2:B:568:LYS:HA	2:B:571:GLN:HE21	1.86	0.41
2:D:337:MET:SD	2:D:385:LYS:HD3	2.61	0.41
2:B:511:LYS:CD	2:B:511:LYS:H	2.34	0.41
2:F:322:VAL:HG12	2:F:323:GLU:HG2	2.02	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	A	24/36 (67%)	23 (96%)	1 (4%)	0	100	100
1	C	30/36 (83%)	29 (97%)	1 (3%)	0	100	100
1	E	24/36 (67%)	23 (96%)	1 (4%)	0	100	100
1	G	26/36 (72%)	25 (96%)	1 (4%)	0	100	100
2	B	249/259 (96%)	245 (98%)	4 (2%)	0	100	100
2	D	250/259 (96%)	245 (98%)	5 (2%)	0	100	100
2	F	251/259 (97%)	244 (97%)	6 (2%)	1 (0%)	39	19
2	H	257/259 (99%)	252 (98%)	5 (2%)	0	100	100
3	Q	1/14 (7%)	1 (100%)	0	0	100	100
All	All	1112/1194 (93%)	1087 (98%)	24 (2%)	1 (0%)	56	36

All (1) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
2	F	474	LYS

### 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	A	16/31 (52%)	15 (94%)	1 (6%)	22	5
1	C	25/31 (81%)	24 (96%)	1 (4%)	38	14
1	E	21/31 (68%)	21 (100%)	0	100	100

*Continued on next page...*



*Continued from previous page...*

Mol	Chain	Analysed	Rotameric	Outliers	Percentiles	
1	G	24/31 (77%)	23 (96%)	1 (4%)	36	12
2	B	207/225 (92%)	197 (95%)	10 (5%)	31	9
2	D	211/225 (94%)	207 (98%)	4 (2%)	65	43
2	F	205/225 (91%)	199 (97%)	6 (3%)	50	24
2	H	205/225 (91%)	198 (97%)	7 (3%)	44	18
3	P	2/10 (20%)	2 (100%)	0	100	100
3	Q	3/10 (30%)	3 (100%)	0	100	100
3	R	2/10 (20%)	1 (50%)	1 (50%)	0	0
3	S	2/10 (20%)	1 (50%)	1 (50%)	0	0
All	All	923/1064 (87%)	891 (96%)	32 (4%)	44	17

All (32) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	292	ASP
2	B	340[A]	ARG
2	B	340[B]	ARG
2	B	341	LYS
2	B	344	GLN
2	B	375	THR
2	B	395	ILE
2	B	426	LYS
2	B	507	CYS
2	B	511	LYS
2	B	517	ARG
1	C	302	LYS
2	D	375	THR
2	D	378	ASP
2	D	507	CYS
2	D	517	ARG
2	F	394	ASN
2	F	479	VAL
2	F	507	CYS
2	F	513	ASP
2	F	517	ARG
2	F	575	ASP
1	G	310	ARG
2	H	395	ILE
2	H	398	ILE

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type
2	H	467	THR
2	H	469	THR
2	H	507	CYS
2	H	517	ARG
2	H	568	LYS
3	R	275	LEU
3	S	275	LEU

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (3) such sidechains are listed below:

Mol	Chain	Res	Type
2	B	537	ASN
2	B	571	GLN
2	F	394	ASN

### 5.3.3 RNA ⓘ

There are no RNA molecules in this entry.

## 5.4 Non-standard residues in protein, DNA, RNA chains ⓘ

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
3	PTR	P	276	3	14,16,17	0.93	0	18,22,24	1.13	2 (11%)
3	PTR	Q	276	3	14,16,17	1.07	0	18,22,24	1.31	2 (11%)
3	PTR	R	276	3	14,16,17	0.77	0	18,22,24	1.15	1 (5%)
3	PTR	S	276	3	14,16,17	1.09	1 (7%)	18,22,24	1.37	3 (16%)

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	PTR	P	276	3	-	0/9/11/13	0/1/1/1
3	PTR	Q	276	3	-	0/9/11/13	0/1/1/1
3	PTR	R	276	3	-	0/9/11/13	0/1/1/1
3	PTR	S	276	3	-	0/9/11/13	0/1/1/1

All (1) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	S	276	PTR	CD2-CG	2.12	1.43	1.38

All (8) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
3	S	276	PTR	O3P-P-OH	-2.21	97.28	105.22
3	S	276	PTR	O-C-CA	-2.11	119.99	125.49
3	Q	276	PTR	OH-CZ-CE1	2.05	125.53	119.22
3	P	276	PTR	O2P-P-OH	2.12	112.83	105.22
3	P	276	PTR	O3P-P-O1P	2.32	118.06	110.58
3	S	276	PTR	O3P-P-O2P	2.70	117.66	107.38
3	R	276	PTR	O3P-P-O2P	3.05	118.98	107.38
3	Q	276	PTR	O3P-P-O2P	3.47	120.59	107.38

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

2 monomers are involved in 3 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	R	276	PTR	1	0
3	S	276	PTR	2	0

## 5.5 Carbohydrates [i](#)

There are no carbohydrates in this entry.

## 5.6 Ligand geometry [i](#)

Of 10 ligands modelled in this entry, 4 are monoatomic - leaving 6 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	$\# Z  > 2$
5	0G6	B	1579	2	27,31,32	1.45	2 (7%)	35,41,42	1.66	10 (28%)
6	GOL	B	1580	-	5,5,5	0.69	0	5,5,5	0.98	0
5	0G6	D	1581	2	27,31,32	1.95	2 (7%)	35,41,42	1.53	6 (17%)
5	0G6	F	1579	2	27,31,32	1.41	2 (7%)	35,41,42	1.46	6 (17%)
6	GOL	F	1580	-	5,5,5	0.50	0	5,5,5	0.84	0
5	0G6	H	1581	2	27,31,32	2.85	2 (7%)	35,41,42	1.85	6 (17%)

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
5	0G6	B	1579	2	-	0/29/41/43	0/2/2/2
6	GOL	B	1580	-	-	0/4/4/4	0/0/0/0
5	0G6	D	1581	2	-	0/29/41/43	0/2/2/2
5	0G6	F	1579	2	-	0/29/41/43	0/2/2/2
6	GOL	F	1580	-	-	0/4/4/4	0/0/0/0
5	0G6	H	1581	2	-	0/29/41/43	0/2/2/2

All (8) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
5	B	1579	0G6	CB2-CA2	-2.20	1.47	1.52
5	F	1579	0G6	CD2-CG	2.16	1.43	1.38
5	H	1581	0G6	O-C	2.53	1.27	1.22
5	D	1581	0G6	CA2-N2	3.46	1.52	1.46
5	B	1579	0G6	CD3-NE	4.73	1.56	1.47
5	F	1579	0G6	C1-N2	5.57	1.46	1.34
5	D	1581	0G6	C1-N2	8.71	1.54	1.34
5	H	1581	0G6	C1-N2	14.04	1.66	1.34

All (28) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(°)	Ideal(°)
5	H	1581	0G6	CA2-N2-C1	-6.40	111.23	123.13
5	H	1581	0G6	CB2-CA2-N2	-5.04	103.25	110.41
5	B	1579	0G6	CB2-CA2-N2	-4.32	104.26	110.41
5	F	1579	0G6	CB2-CA2-N2	-3.87	104.91	110.41
5	H	1581	0G6	O2-C2-C3	-3.71	98.88	109.61
5	F	1579	0G6	CG2-CB2-CA2	-3.66	106.38	114.10
5	D	1581	0G6	O2-C2-C3	-3.36	99.89	109.61
5	B	1579	0G6	CB2-CG2-CD3	-2.98	102.94	112.13
5	B	1579	0G6	O2-C2-C3	-2.93	101.14	109.61
5	B	1579	0G6	CB1-CA1-C1	-2.80	106.11	111.26
5	B	1579	0G6	CG2-CD3-NE	-2.69	103.03	112.35
5	D	1581	0G6	CA2-N2-C1	-2.62	118.26	123.13
5	F	1579	0G6	O2-C2-C3	-2.40	102.66	109.61
5	F	1579	0G6	CB2-CG2-CD3	-2.34	104.91	112.13
5	H	1581	0G6	CA1-C1-N2	-2.29	111.49	116.64
5	B	1579	0G6	O1-C1-N2	-2.23	118.56	122.93
5	F	1579	0G6	CE1-CD1-CG	-2.17	117.20	120.65
5	B	1579	0G6	CA1-N1-C	-2.09	114.27	121.64
5	D	1581	0G6	CG2-CB2-CA2	-2.08	109.71	114.10
5	D	1581	0G6	CD-N1-C	2.17	133.37	125.58
5	B	1579	0G6	CD-N1-C	2.25	133.65	125.58
5	H	1581	0G6	CD2-CG-CD1	2.38	121.94	118.13
5	B	1579	0G6	CA1-C1-N2	2.49	122.25	116.64
5	D	1581	0G6	O1-C1-N2	2.56	127.94	122.93
5	F	1579	0G6	CD2-CG-CD1	2.68	122.42	118.13
5	H	1581	0G6	O2-C2-CA2	2.70	114.65	108.58
5	B	1579	0G6	C2-CA2-N2	3.11	115.94	110.11
5	D	1581	0G6	C2-CA2-N2	4.13	117.84	110.11

There are no chirality outliers.

There are no torsion outliers.

There are no ring outliers.

6 monomers are involved in 17 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
5	B	1579	0G6	4	0
6	B	1580	GOL	2	0
5	D	1581	0G6	2	0
5	F	1579	0G6	4	0
6	F	1580	GOL	2	0
5	H	1581	0G6	3	0

## 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 [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	26/36 (72%)	1.06	5 (19%) 2 2	41, 60, 79, 102	0
1	C	32/36 (88%)	0.44	4 (12%) 5 7	23, 42, 71, 72	0
1	E	26/36 (72%)	0.39	2 (7%) 16 20	36, 46, 57, 66	0
1	G	27/36 (75%)	0.05	0 100 100	29, 39, 55, 60	0
2	B	249/259 (96%)	0.07	4 (1%) 74 81	14, 33, 54, 65	1 (0%)
2	D	252/259 (97%)	-0.10	0 100 100	14, 29, 51, 59	0
2	F	253/259 (97%)	-0.14	3 (1%) 81 86	19, 35, 53, 66	2 (0%)
2	H	259/259 (100%)	-0.02	4 (1%) 76 82	17, 35, 63, 70	0
3	P	2/14 (14%)	0.76	0 100 100	36, 36, 36, 54	0
3	Q	3/14 (21%)	-0.07	0 100 100	32, 32, 39, 42	0
3	R	2/14 (14%)	0.85	1 (50%) 0 0	54, 54, 54, 68	0
3	S	2/14 (14%)	0.63	0 100 100	35, 35, 35, 52	0
All	All	1133/1236 (91%)	0.01	23 (2%) 68 75	14, 34, 59, 102	3 (0%)

All (23) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	290	GLU	4.4
1	A	316	TYR	4.3
2	H	472	VAL	4.3
1	E	291	ALA	3.6
2	B	515	GLY	3.4
1	A	317	ILE	3.4
1	A	313	LEU	3.0
1	C	291	ALA	3.0
2	B	537	ASN	2.9
2	F	394	ASN	2.8
2	H	512	PRO	2.7

*Continued on next page...*

*Continued from previous page...*

Mol	Chain	Res	Type	RSRZ
1	C	293	CYS	2.6
2	H	468	TRP	2.6
2	F	513	ASP	2.6
1	A	310	ARG	2.5
1	E	313	LEU	2.5
2	B	535	PHE	2.4
1	A	304	LEU	2.3
1	C	294	GLY	2.3
2	H	393	ARG	2.2
2	F	395	ILE	2.2
2	B	389	THR	2.1
3	R	275	LEU	2.0

## 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. 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	PTR	Q	276	16/17	0.97	0.07	-	26,30,41,47	0
3	PTR	S	276	16/17	0.92	0.09	-	33,40,48,56	0
3	PTR	P	276	16/17	0.92	0.10	-	35,40,50,52	0
3	PTR	R	276	16/17	0.89	0.15	-	45,66,73,76	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. 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( $\text{\AA}^2$ )	Q<0.9
4	NA	B	1578	1/1	0.95	0.20	2.82	35,35,35,35	0
6	GOL	B	1580	6/6	0.84	0.17	2.15	33,40,44,47	0
4	NA	F	1578	1/1	0.95	0.11	2.13	38,38,38,38	0
6	GOL	F	1580	6/6	0.87	0.14	1.28	33,46,51,56	0
5	0G6	B	1579	30/31	0.95	0.09	0.58	16,21,28,28	0
5	0G6	D	1581	30/31	0.95	0.09	0.57	18,22,25,29	0
5	0G6	F	1579	30/31	0.96	0.07	0.17	20,23,26,26	0
4	NA	D	1580	1/1	0.94	0.09	-0.28	45,45,45,45	0
5	0G6	H	1581	30/31	0.94	0.08	-0.37	24,29,33,34	0
4	NA	H	1580	1/1	0.95	0.06	-2.39	45,45,45,45	0

## 6.5 Other polymers [i](#)

There are no such residues in this entry.