Full wwPDB X-ray Structure Validation Report (i)

Jun 17, 2024 – 08:00 PM EDT

PDB ID	:	3H0C
Title	:	Crystal Structure of Human Dipeptidyl Peptidase IV (CD26) in Complex with a Reversed Amide Inhibitor
Authors	:	Nordhoff, S.; Cerezo-Galvez, S.; Deppe, H.; Hill, O.; Lopez-Canet, M.; Rummey, C.; Thiemann, M.; Matassa, V.G.; Edwards, P.J.; Feurer, A.
Deposited on	:	2009-04-09
Resolution	:	2.66 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*
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:

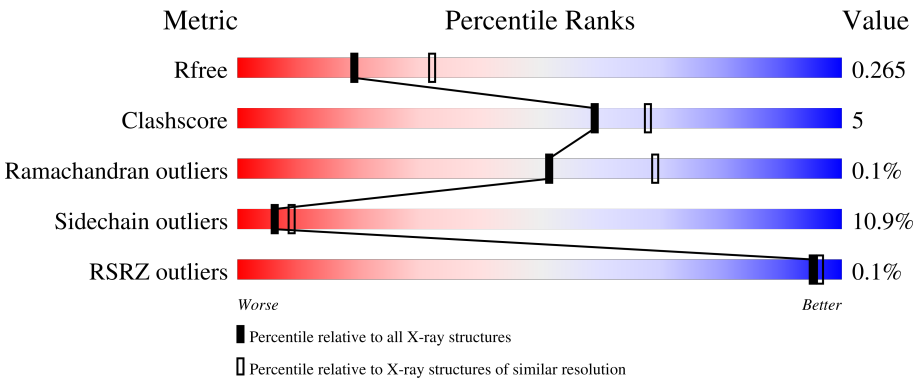
MolProbity	:	4.02b-467
Mogul	:	2022.3.0, CSD as543be (2022)
Xtriage (Phenix)	:	1.20.1
EDS	:	2.37.1
buster-report	:	1.1.7 (2018)
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (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 2.66 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	1332 (2.68-2.64)
Clashscore	141614	1374 (2.68-2.64)
Ramachandran outliers	138981	1349 (2.68-2.64)
Sidechain outliers	138945	1349 (2.68-2.64)
RSRZ outliers	127900	1318 (2.68-2.64)

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	728	78% 19% •
1	B	728	80% 17% •

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
2	NAG	A	795	X	-	-	-
2	NAG	B	797	X	-	-	-

2 Entry composition (i)

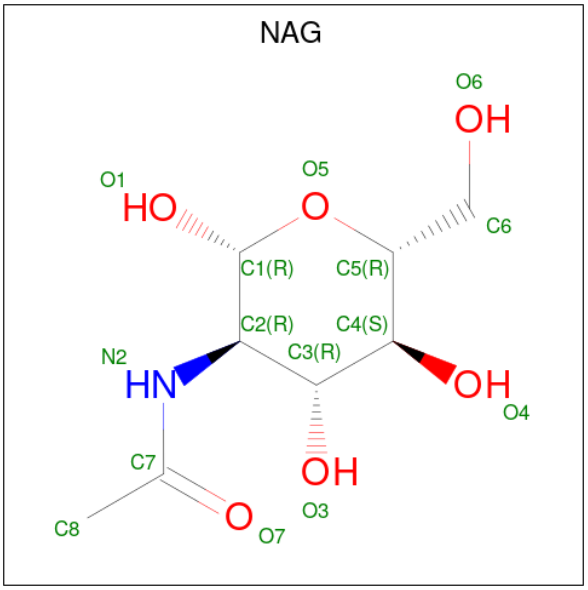
There are 4 unique types of molecules in this entry. The entry contains 12460 atoms, of which 0 are hydrogens and 0 are deuteriums.

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 Dipeptidyl peptidase 4.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	728	5963	3827	982	1128	26	0	0	0
1	B	728	5963	3827	982	1128	26	0	0	0

- Molecule 2 is 2-acetamido-2-deoxy-beta-D-glucopyranose (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



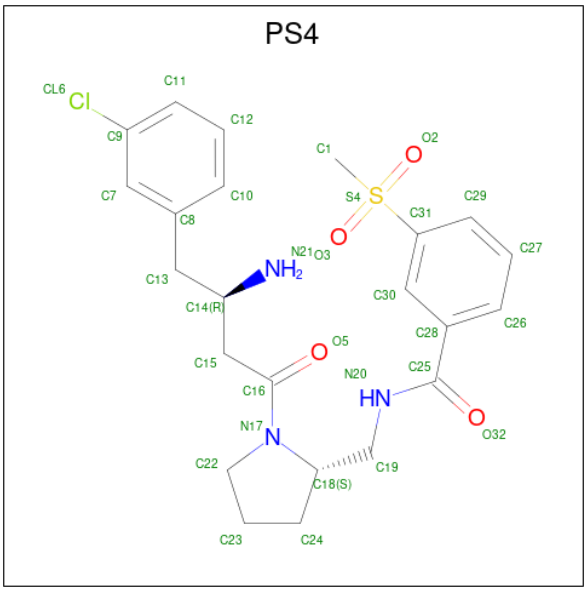
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	A	1	14	8	1	5	0	0
2	A	1	14	8	1	5	0	0
2	A	1	14	8	1	5	0	0
2	A	1	14	8	1	5	0	0

Continued on next page...

Continued from previous page...

Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
			Total	C	N	O		
2	B	1	14	8	1	5	0	0
			Total	C	N	O		
2	B	1	14	8	1	5	0	0
			Total	C	N	O		
2	B	1	14	8	1	5	0	0
			Total	C	N	O		
2	B	1	14	8	1	5	0	0

- Molecule 3 is N-({(2S)-1-[(3R)-3-amino-4-(3-chlorophenyl)butanoyl]pyrrolidin-2-yl}methyl)-3-(methylsulfonyl)benzamide (three-letter code: PS4) (formula: $C_{23}H_{28}ClN_3O_4S$).



Mol	Chain	Residues	Atoms						ZeroOcc	AltConf
			Total	C	Cl	N	O	S		
3	A	1	32	23	1	3	4	1	0	0
			Total	C	Cl	N	O	S		
3	B	1	32	23	1	3	4	1	0	0

- Molecule 4 is water.

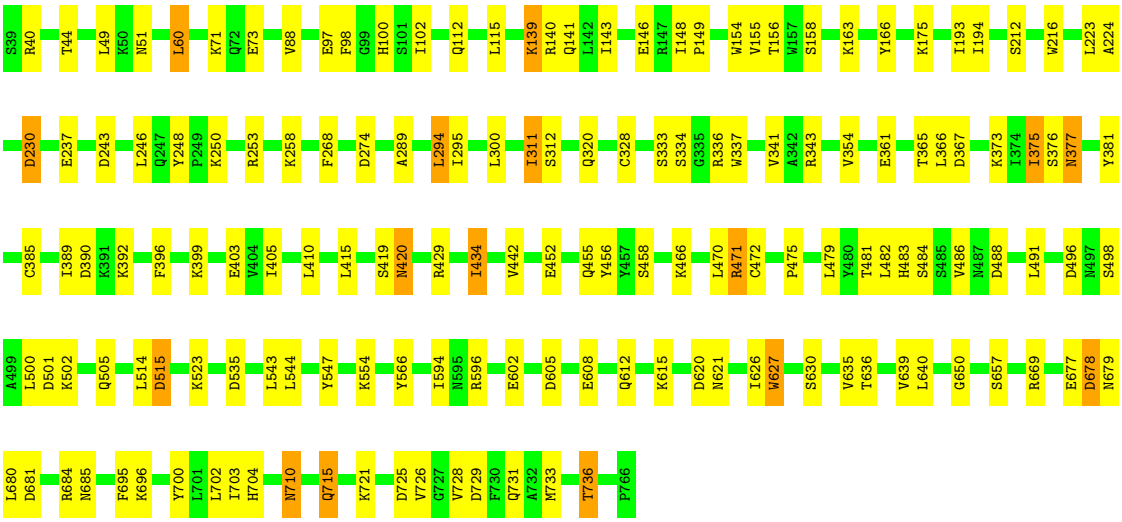
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	203	203	203	0	0
			Total	O		
4	B	155	155	155	0	0

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.

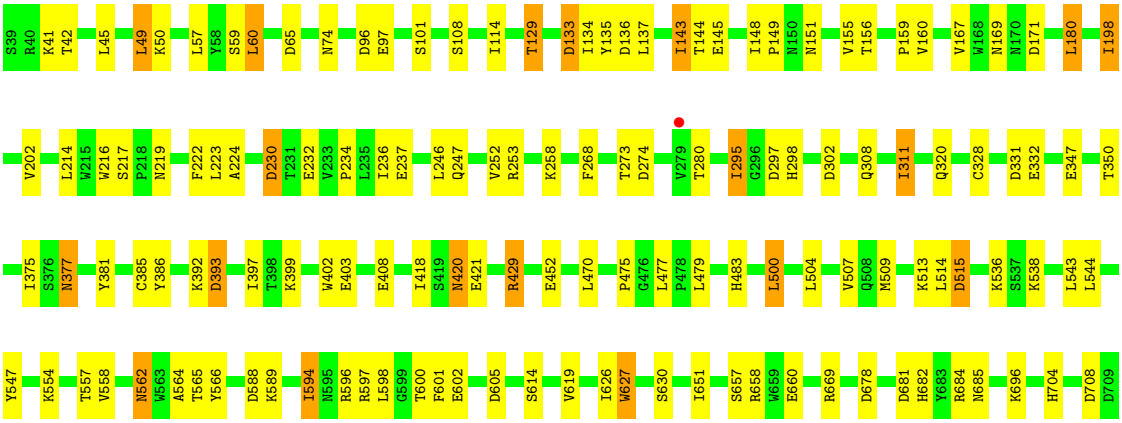
• Molecule 1: Dipeptidyl peptidase 4

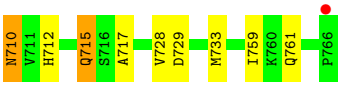
Chain A:



• Molecule 1: Dipeptidyl peptidase 4

Chain B:




N710
V711
H712
Q715
S716
A717
V728
D729
M733
I759
K760
Q761
P766

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	65.69Å 67.07Å 425.13Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.66 19.99 – 2.60	Depositor EDS
% Data completeness (in resolution range)	92.7 (20.00-2.66) 92.9 (19.99-2.60)	Depositor EDS
R_{merge}	(Not available)	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	7.97 (at 2.59Å)	Xtriage
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.204 , 0.271 0.204 , 0.265	Depositor DCC
R_{free} test set	1697 reflections (3.09%)	wwPDB-VP
Wilson B-factor (Å^2)	29.1	Xtriage
Anisotropy	0.401	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.30 , 11.2	EDS
L-test for twinning[2]	$< \|L\| > = 0.41$, $< L^2 > = 0.24$	Xtriage
Estimated twinning fraction	0.065 for k,h,-l	Xtriage
F_o,F_c correlation	0.91	EDS
Total number of atoms	12460	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

Xtriage's analysis on translational NCS is as follows: *The largest off-origin peak in the Patterson function is 4.06% 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: PS4, NAG

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	0.43	0/6135	0.72	16/8344 (0.2%)
1	B	0.41	0/6135	0.71	18/8344 (0.2%)
All	All	0.42	0/12270	0.72	34/16688 (0.2%)

There are no bond length outliers.

All (34) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	60	LEU	CA-CB-CG	7.28	132.04	115.30
1	A	729	ASP	CB-CG-OD2	6.47	124.12	118.30
1	A	488	ASP	CB-CG-OD2	6.43	124.09	118.30
1	A	678	ASP	CB-CG-OD2	6.21	123.88	118.30
1	B	681	ASP	CB-CG-OD2	6.14	123.83	118.30
1	B	729	ASP	CB-CG-OD2	6.09	123.78	118.30
1	B	302	ASP	CB-CG-OD2	5.99	123.69	118.30
1	B	133	ASP	CB-CG-OD2	5.96	123.66	118.30
1	A	60	LEU	CA-CB-CG	5.94	128.97	115.30
1	A	230	ASP	CB-CG-OD2	5.62	123.36	118.30
1	A	390	ASP	CB-CG-OD2	5.56	123.31	118.30
1	A	501	ASP	CB-CG-OD2	5.55	123.29	118.30
1	A	620	ASP	CB-CG-OD2	5.42	123.17	118.30
1	A	681	ASP	CB-CG-OD2	5.41	123.17	118.30
1	A	274	ASP	CB-CG-OD2	5.39	123.15	118.30
1	A	243	ASP	CB-CG-OD2	5.34	123.11	118.30
1	B	274	ASP	CB-CG-OD2	5.33	123.10	118.30
1	B	331	ASP	CB-CG-OD2	5.31	123.08	118.30
1	A	605	ASP	CB-CG-OD2	5.30	123.07	118.30
1	B	588	ASP	CB-CG-OD2	5.25	123.03	118.30
1	A	496	ASP	CB-CG-OD2	5.24	123.01	118.30
1	B	515	ASP	CB-CG-OD2	5.24	123.01	118.30

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	605	ASP	CB-CG-OD2	5.24	123.01	118.30
1	B	65	ASP	CB-CG-OD2	5.23	123.00	118.30
1	A	515	ASP	CB-CG-OD2	5.20	122.98	118.30
1	B	96	ASP	CB-CG-OD2	5.18	122.97	118.30
1	B	297	ASP	CB-CG-OD2	5.16	122.94	118.30
1	B	393	ASP	CB-CG-OD2	5.16	122.94	118.30
1	B	136	ASP	CB-CG-OD2	5.14	122.93	118.30
1	B	171	ASP	CB-CG-OD2	5.09	122.88	118.30
1	A	535	ASP	CB-CG-OD2	5.06	122.85	118.30
1	A	725	ASP	CB-CG-OD2	5.05	122.85	118.30
1	B	708	ASP	CB-CG-OD2	5.02	122.81	118.30
1	B	230	ASP	CB-CG-OD2	5.01	122.81	118.30

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	5963	0	5680	57	0
1	B	5963	0	5680	57	0
2	A	56	0	52	0	0
2	B	56	0	52	0	0
3	A	32	0	28	2	0
3	B	32	0	28	2	0
4	A	203	0	0	3	0
4	B	155	0	0	1	0
All	All	12460	0	11520	111	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 5.

All (111) 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:543:LEU:HD21	1:A:627:TRP:CD1	2.09	0.87
1:B:600:THR:HG23	1:B:601:PHE:H	1.42	0.84
1:A:543:LEU:HD21	1:A:627:TRP:HD1	1.47	0.80
1:B:386:TYR:HB2	1:B:397:ILE:HD11	1.65	0.78
1:B:600:THR:HG23	1:B:601:PHE:N	2.01	0.75
1:B:311:ILE:HD12	1:B:328:CYS:HB2	1.69	0.74
1:A:657:SER:H	1:A:715:GLN:NE2	1.94	0.66
1:A:434:ILE:HG23	1:A:442:VAL:HG22	1.77	0.66
1:B:429:ARG:HG2	1:B:429:ARG:HH11	1.61	0.66
1:A:343:ARG:HD2	1:A:389:ILE:HG23	1.79	0.65
1:A:403:GLU:H	1:A:420:ASN:HD21	1.45	0.65
1:A:237:GLU:HG2	1:A:253:ARG:HG2	1.77	0.64
1:B:600:THR:CG2	1:B:601:PHE:H	2.09	0.64
1:B:598:LEU:O	1:B:682:HIS:HE1	1.81	0.64
1:B:597:ARG:O	1:B:600:THR:HG22	1.98	0.63
1:B:403:GLU:H	1:B:420:ASN:HD21	1.46	0.62
1:B:224:ALA:HB1	1:B:268:PHE:CZ	2.38	0.59
1:A:415:LEU:HB3	1:A:434:ILE:HD12	1.83	0.59
1:A:736:THR:HG21	1:B:717:ALA:O	2.02	0.59
1:A:596:ARG:NH2	1:A:678:ASP:OD1	2.38	0.57
1:A:343:ARG:HD2	1:A:389:ILE:CG2	2.35	0.57
1:A:458:SER:OG	1:A:471:ARG:NE	2.33	0.57
1:B:596:ARG:NH2	1:B:678:ASP:OD1	2.38	0.56
1:B:377:ASN:HD21	1:B:381:TYR:H	1.54	0.56
1:A:311:ILE:HD12	1:A:328:CYS:HB2	1.88	0.55
1:A:258:LYS:HD2	1:B:247:GLN:HG2	1.87	0.54
1:A:377:ASN:HD22	1:A:377:ASN:C	2.10	0.54
1:B:547:TYR:CG	3:B:1:PS4:H1	2.44	0.53
1:B:129:THR:HG23	1:B:151:ASN:HA	1.89	0.53
1:A:486:VAL:HG23	4:A:910:HOH:O	2.07	0.53
1:B:219:ASN:N	1:B:308:GLN:OE1	2.41	0.53
1:B:470:LEU:HD12	1:B:483:HIS:CE1	2.44	0.53
1:A:726:VAL:HG12	1:A:728:VAL:HG23	1.91	0.52
1:A:554:LYS:HG2	3:A:1:PS4:H1B	1.92	0.52
1:A:594:ILE:HD11	1:A:602:GLU:OE1	2.11	0.51
1:B:377:ASN:ND2	1:B:381:TYR:H	2.08	0.51
1:B:402:TRP:CD1	1:B:421:GLU:HG3	2.46	0.50
1:B:429:ARG:HG2	1:B:429:ARG:NH1	2.27	0.49
1:B:167:VAL:HG11	1:B:198:ILE:HG12	1.93	0.49
1:B:477:LEU:HD22	1:B:500:LEU:HD13	1.95	0.49
1:A:470:LEU:HD12	1:A:483:HIS:CE1	2.48	0.49
1:A:547:TYR:CD2	3:A:1:PS4:H1	2.47	0.48

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:544:LEU:HD23	1:B:626:ILE:HD12	1.95	0.48
1:A:700:TYR:OH	1:A:702:LEU:HD13	2.13	0.48
1:A:703:ILE:HG12	1:A:733:MET:HB3	1.94	0.48
1:B:320:GLN:OE1	1:B:669:ARG:HD3	2.14	0.48
1:A:139:LYS:O	1:A:141:GLN:N	2.46	0.48
1:A:158:SER:HB3	1:A:163:LYS:HB2	1.95	0.48
1:A:193:ILE:HG22	1:A:194:ILE:HG12	1.95	0.47
1:A:481:THR:OG1	1:A:483:HIS:HE1	1.98	0.47
1:B:408:GLU:HG3	1:B:418:ILE:HD12	1.94	0.47
1:B:598:LEU:O	1:B:682:HIS:CE1	2.64	0.47
1:B:614:SER:HA	1:B:619:VAL:HG22	1.97	0.47
1:A:248:TYR:CZ	1:B:234:PRO:HB2	2.50	0.47
1:A:455:GLN:HB2	1:A:475:PRO:HD3	1.98	0.46
1:B:377:ASN:C	1:B:377:ASN:HD22	2.19	0.46
1:A:710:ASN:C	1:A:710:ASN:HD22	2.19	0.46
1:A:377:ASN:ND2	1:A:381:TYR:H	2.14	0.46
1:B:562:ASN:HD21	1:B:564:ALA:HB3	1.81	0.45
1:A:289:ALA:HB3	1:A:294:LEU:HD13	1.97	0.45
1:A:320:GLN:OE1	1:A:669:ARG:HG3	2.17	0.45
1:B:547:TYR:CD2	3:B:1:PS4:H1	2.52	0.45
1:A:626:ILE:HG12	1:A:636:THR:HG23	1.99	0.45
1:B:600:THR:CG2	1:B:601:PHE:N	2.69	0.45
1:B:236:ILE:HG21	1:B:712:HIS:CE1	2.52	0.45
1:B:562:ASN:C	1:B:562:ASN:HD22	2.20	0.45
1:B:134:ILE:HB	1:B:143:ILE:HG12	1.99	0.45
1:A:155:VAL:HG13	1:A:166:TYR:HB3	2.00	0.44
1:A:375:ILE:HG22	1:A:396:PHE:HZ	1.82	0.44
1:B:237:GLU:HA	1:B:252:VAL:O	2.18	0.44
1:A:429:ARG:HD2	1:A:456:TYR:CZ	2.52	0.44
1:B:295:ILE:O	1:B:295:ILE:HD13	2.18	0.44
1:A:156:THR:HG23	1:A:216:TRP:HE1	1.83	0.43
1:A:373:LYS:HD3	1:A:375:ILE:HD11	2.01	0.43
1:B:45:LEU:HG	1:B:49:LEU:HD22	2.00	0.43
1:B:696:LYS:HG3	1:B:728:VAL:HG22	2.00	0.43
1:A:311:ILE:HD11	1:A:337:TRP:CE3	2.54	0.43
1:B:403:GLU:H	1:B:420:ASN:ND2	2.15	0.43
1:A:626:ILE:O	1:A:650:GLY:HA2	2.19	0.43
1:A:224:ALA:HB1	1:A:268:PHE:CZ	2.53	0.43
1:A:696:LYS:HG3	1:A:728:VAL:HG22	2.01	0.42
1:A:98:PHE:CE2	1:A:100:HIS:HB2	2.53	0.42
1:A:680:LEU:HD21	1:A:684:ARG:NH2	2.34	0.42

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:148:ILE:HA	1:B:149:PRO:HD3	1.89	0.42
1:B:682:HIS:HA	1:B:685:ASN:HB2	2.00	0.42
1:B:180:LEU:HD12	1:B:180:LEU:HA	1.91	0.42
1:B:298:HIS:HE1	4:B:876:HOH:O	2.01	0.42
1:B:114:ILE:HG23	1:B:135:TYR:HB3	2.02	0.42
1:A:680:LEU:HD21	1:A:684:ARG:CZ	2.49	0.42
1:A:148:ILE:HA	1:A:149:PRO:HD3	1.93	0.42
1:A:365:THR:HG22	4:A:884:HOH:O	2.20	0.42
1:B:657:SER:H	1:B:715:GLN:NE2	2.18	0.42
1:B:507:VAL:HG13	1:B:509:MET:HG2	2.01	0.41
1:B:594:ILE:HD11	1:B:602:GLU:H	1.85	0.41
1:A:154:TRP:CE2	1:A:212:SER:HB2	2.55	0.41
1:B:217:SER:HB2	1:B:222:PHE:HB2	2.02	0.41
1:B:562:ASN:HD22	1:B:565:THR:H	1.68	0.41
1:B:475:PRO:HA	1:B:557:THR:O	2.20	0.41
1:A:146:GLU:O	1:A:175:LYS:NZ	2.51	0.41
1:A:403:GLU:H	1:A:420:ASN:ND2	2.15	0.41
1:B:627:TRP:HB2	1:B:651:ILE:HB	2.02	0.41
1:B:710:ASN:C	1:B:710:ASN:HD22	2.23	0.41
1:A:375:ILE:HG22	1:A:396:PHE:CZ	2.55	0.41
1:A:640:LEU:HD11	1:A:650:GLY:HA3	2.01	0.41
1:B:159:PRO:HD3	1:B:216:TRP:HB3	2.02	0.41
1:A:612:GLN:HA	1:A:615:LYS:HE2	2.03	0.41
1:A:695:PHE:HB3	1:A:728:VAL:HG11	2.03	0.41
1:B:504:LEU:HA	1:B:507:VAL:HG12	2.02	0.40
1:A:635:VAL:O	1:A:639:VAL:HG23	2.21	0.40
1:A:343:ARG:NH1	4:A:859:HOH:O	2.54	0.40
1:B:237:GLU:HG2	1:B:253:ARG:HG2	2.03	0.40

There are no symmetry-related clashes.

5.3 Torsion angles (i)

5.3.1 Protein backbone (i)

In the following table, the Percentiles column shows the percent Ramachandran outliers of the chain as a percentile score with respect to all X-ray entries followed by that with respect to entries of similar resolution.

The Analysed column shows the number of residues for which the backbone conformation was analysed, and the total number of residues.

Mol	Chain	Analysed	Favoured	Allowed	Outliers	Percentiles	
1	A	726/728 (100%)	685 (94%)	39 (5%)	2 (0%)	41	56
1	B	726/728 (100%)	691 (95%)	35 (5%)	0	100	100
All	All	1452/1456 (100%)	1376 (95%)	74 (5%)	2 (0%)	51	69

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	140	ARG
1	A	73	GLU

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	653/653 (100%)	582 (89%)	71 (11%)	6	9
1	B	653/653 (100%)	582 (89%)	71 (11%)	6	9
All	All	1306/1306 (100%)	1164 (89%)	142 (11%)	6	9

All (142) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	40	ARG
1	A	44	THR
1	A	49	LEU
1	A	51	ASN
1	A	60	LEU
1	A	71	LYS
1	A	88	VAL
1	A	97	GLU
1	A	102	ILE
1	A	112	GLN
1	A	115	LEU
1	A	139	LYS
1	A	143	ILE
1	A	223	LEU

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	230	ASP
1	A	246	LEU
1	A	250	LYS
1	A	294	LEU
1	A	295	ILE
1	A	300	LEU
1	A	311	ILE
1	A	312	SER
1	A	333	SER
1	A	334	SER
1	A	336	ARG
1	A	341	VAL
1	A	354	VAL
1	A	361	GLU
1	A	366	LEU
1	A	367	ASP
1	A	375	ILE
1	A	376	SER
1	A	377	ASN
1	A	385	CYS
1	A	392	LYS
1	A	399	LYS
1	A	405	ILE
1	A	410	LEU
1	A	419	SER
1	A	420	ASN
1	A	434	ILE
1	A	452	GLU
1	A	466	LYS
1	A	471	ARG
1	A	472	CYS
1	A	479	LEU
1	A	482	LEU
1	A	484	SER
1	A	491	LEU
1	A	498	SER
1	A	500	LEU
1	A	502	LYS
1	A	505	GLN
1	A	514	LEU
1	A	515	ASP
1	A	523	LYS

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	A	544	LEU
1	A	566	TYR
1	A	608	GLU
1	A	621	ASN
1	A	627	TRP
1	A	630	SER
1	A	677	GLU
1	A	679	ASN
1	A	685	ASN
1	A	704	HIS
1	A	710	ASN
1	A	715	GLN
1	A	721	LYS
1	A	731	GLN
1	A	736	THR
1	B	41	LYS
1	B	42	THR
1	B	49	LEU
1	B	50	LYS
1	B	57	LEU
1	B	59	SER
1	B	60	LEU
1	B	74	ASN
1	B	97	GLU
1	B	101	SER
1	B	108	SER
1	B	129	THR
1	B	133	ASP
1	B	137	LEU
1	B	143	ILE
1	B	144	THR
1	B	145	GLU
1	B	155	VAL
1	B	156	THR
1	B	160	VAL
1	B	169	ASN
1	B	180	LEU
1	B	198	ILE
1	B	202	VAL
1	B	214	LEU
1	B	223	LEU
1	B	230	ASP

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	232	GLU
1	B	246	LEU
1	B	258	LYS
1	B	273	THR
1	B	280	THR
1	B	295	ILE
1	B	311	ILE
1	B	332	GLU
1	B	347	GLU
1	B	350	THR
1	B	375	ILE
1	B	377	ASN
1	B	385	CYS
1	B	392	LYS
1	B	393	ASP
1	B	399	LYS
1	B	420	ASN
1	B	429	ARG
1	B	452	GLU
1	B	479	LEU
1	B	500	LEU
1	B	513	LYS
1	B	514	LEU
1	B	515	ASP
1	B	536	LYS
1	B	538	LYS
1	B	543	LEU
1	B	554	LYS
1	B	558	VAL
1	B	562	ASN
1	B	566	TYR
1	B	589	LYS
1	B	594	ILE
1	B	627	TRP
1	B	630	SER
1	B	658	ARG
1	B	660	GLU
1	B	684	ARG
1	B	704	HIS
1	B	710	ASN
1	B	715	GLN
1	B	733	MET

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type
1	B	759	ILE
1	B	761	GLN

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (27) such sidechains are listed below:

Mol	Chain	Res	Type
1	A	141	GLN
1	A	169	ASN
1	A	298	HIS
1	A	377	ASN
1	A	420	ASN
1	A	435	GLN
1	A	483	HIS
1	A	506	ASN
1	A	592	HIS
1	A	679	ASN
1	A	710	ASN
1	A	715	GLN
1	A	750	HIS
1	B	169	ASN
1	B	219	ASN
1	B	263	ASN
1	B	298	HIS
1	B	363	HIS
1	B	377	ASN
1	B	420	ASN
1	B	435	GLN
1	B	483	HIS
1	B	562	ASN
1	B	679	ASN
1	B	682	HIS
1	B	710	ASN
1	B	715	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 monosaccharides in this entry.

5.6 Ligand geometry (i)

10 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	PS4	A	1	-	34,34,34	0.77	1 (2%)	44,48,48	1.64	5 (11%)
2	NAG	B	793	1	14,14,15	0.61	0	17,19,21	1.51	2 (11%)
2	NAG	A	795	1	14,14,15	0.52	0	17,19,21	1.23	2 (11%)
3	PS4	B	1	-	34,34,34	0.74	1 (2%)	44,48,48	1.77	3 (6%)
2	NAG	B	796	1	14,14,15	0.49	0	17,19,21	0.99	1 (5%)
2	NAG	A	796	1	14,14,15	0.47	0	17,19,21	1.12	1 (5%)
2	NAG	B	797	1	14,14,15	0.57	0	17,19,21	0.81	1 (5%)
2	NAG	A	794	1	14,14,15	0.61	0	17,19,21	1.08	1 (5%)
2	NAG	A	793	1	14,14,15	0.51	0	17,19,21	1.11	1 (5%)
2	NAG	B	794	1	14,14,15	0.51	0	17,19,21	1.24	3 (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
3	PS4	A	1	-	-	6/27/37/37	0/3/3/3
2	NAG	B	793	1	-	6/6/23/26	0/1/1/1

Continued on next page...

Continued from previous page...

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	A	795	1	1/1/5/7	3/6/23/26	0/1/1/1
3	PS4	B	1	-	-	4/27/37/37	0/3/3/3
2	NAG	B	797	1	1/1/5/7	3/6/23/26	0/1/1/1
2	NAG	A	796	1	-	0/6/23/26	0/1/1/1
2	NAG	B	796	1	-	0/6/23/26	0/1/1/1
2	NAG	A	794	1	-	0/6/23/26	0/1/1/1
2	NAG	A	793	1	-	3/6/23/26	0/1/1/1
2	NAG	B	794	1	-	0/6/23/26	0/1/1/1

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
3	A	1	PS4	C25-N20	2.91	1.40	1.33
3	B	1	PS4	C25-N20	2.69	1.39	1.33

All (20) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
3	B	1	PS4	C1-S4-C31	8.84	114.29	104.55
3	A	1	PS4	C1-S4-C31	7.50	112.81	104.55
3	B	1	PS4	O3-S4-O2	-5.02	109.48	117.99
2	B	793	NAG	C2-N2-C7	4.47	128.90	122.90
3	A	1	PS4	O3-S4-O2	-4.14	110.97	117.99
2	A	796	NAG	C1-O5-C5	3.79	117.27	112.19
2	A	795	NAG	O5-C1-C2	-3.31	106.16	111.29
2	B	794	NAG	C2-N2-C7	-2.89	119.03	122.90
2	A	794	NAG	C1-O5-C5	2.80	115.94	112.19
2	B	796	NAG	C1-O5-C5	2.68	115.78	112.19
3	A	1	PS4	O3-S4-C31	2.68	110.31	108.23
2	B	794	NAG	C1-O5-C5	2.38	115.37	112.19
2	B	797	NAG	O5-C5-C6	2.36	112.26	107.66
2	B	793	NAG	C1-C2-N2	2.30	114.06	110.43
3	A	1	PS4	C28-C25-N20	2.28	121.86	117.12
2	A	793	NAG	C4-C3-C2	2.27	114.34	111.02
3	A	1	PS4	O2-S4-C1	-2.22	105.20	108.47
2	A	795	NAG	C1-O5-C5	2.05	114.93	112.19
3	B	1	PS4	O2-S4-C1	-2.05	105.46	108.47
2	B	794	NAG	O5-C1-C2	-2.05	108.13	111.29

All (2) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
2	A	795	NAG	C1
2	B	797	NAG	C1

All (25) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	A	795	NAG	C8-C7-N2-C2
2	A	795	NAG	O7-C7-N2-C2
2	B	793	NAG	C1-C2-N2-C7
2	B	793	NAG	C8-C7-N2-C2
2	B	793	NAG	O7-C7-N2-C2
2	B	793	NAG	O5-C5-C6-O6
2	B	793	NAG	C4-C5-C6-O6
2	A	793	NAG	C8-C7-N2-C2
3	B	1	PS4	C30-C31-S4-O2
3	A	1	PS4	C29-C31-S4-O2
2	A	793	NAG	O7-C7-N2-C2
3	A	1	PS4	C30-C31-S4-O2
3	B	1	PS4	C29-C31-S4-O2
3	B	1	PS4	C30-C31-S4-C1
3	B	1	PS4	C29-C31-S4-C1
3	A	1	PS4	C30-C31-S4-C1
3	A	1	PS4	C29-C31-S4-C1
2	B	797	NAG	C8-C7-N2-C2
3	A	1	PS4	N21-C14-C15-C16
2	B	797	NAG	O7-C7-N2-C2
2	A	795	NAG	O5-C5-C6-O6
2	B	797	NAG	C4-C5-C6-O6
2	B	793	NAG	C3-C2-N2-C7
3	A	1	PS4	C8-C13-C14-N21
2	A	793	NAG	C4-C5-C6-O6

There are no ring outliers.

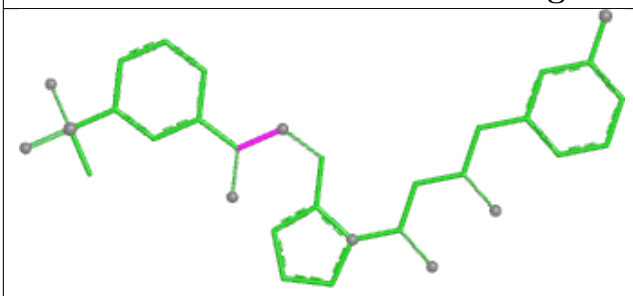
2 monomers are involved in 4 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
3	A	1	PS4	2	0
3	B	1	PS4	2	0

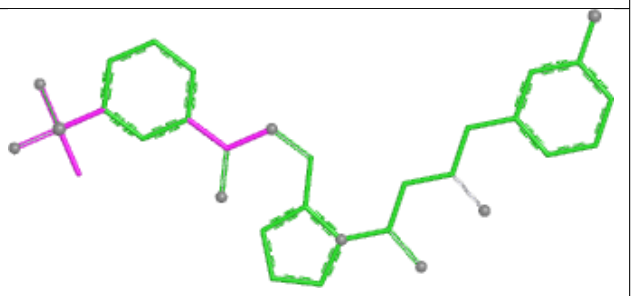
The following is a two-dimensional graphical depiction of Mogul quality analysis of bond lengths, bond angles, torsion angles, and ring geometry for all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the validation Tables will also be included. For torsion angles, if less then 5% of the Mogul distribution of torsion angles is

within 10 degrees of the torsion angle in question, then that torsion angle is considered an outlier. Any bond that is central to one or more torsion angles identified as an outlier by Mogul will be highlighted in the graph. For rings, the root-mean-square deviation (RMSD) between the ring in question and similar rings identified by Mogul is calculated over all ring torsion angles. If the average RMSD is greater than 60 degrees and the minimal RMSD between the ring in question and any Mogul-identified rings is also greater than 60 degrees, then that ring is considered an outlier. The outliers are highlighted in purple. The color gray indicates Mogul did not find sufficient equivalents in the CSD to analyse the geometry.

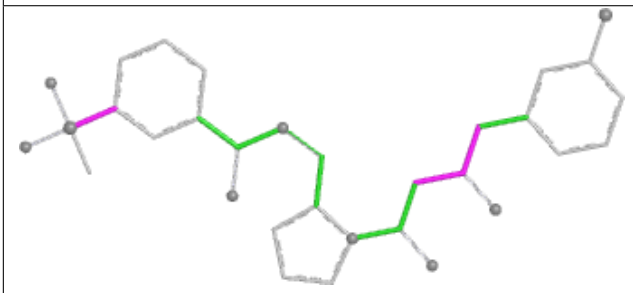
Ligand PS4 A 1



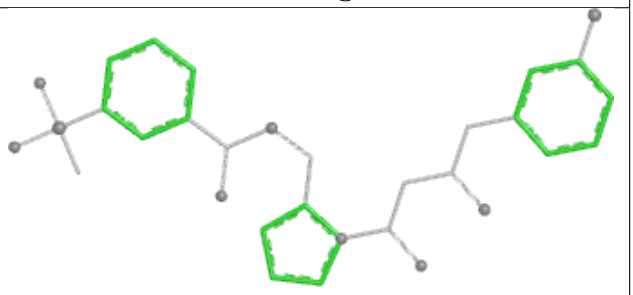
Bond lengths



Bond angles

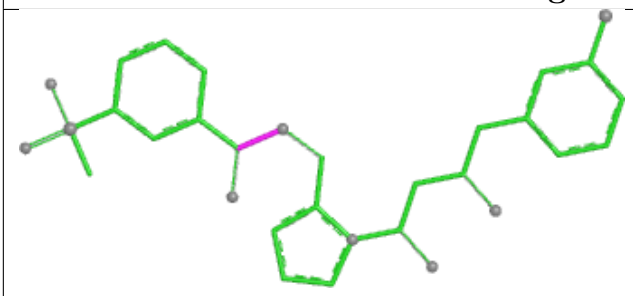


Torsions

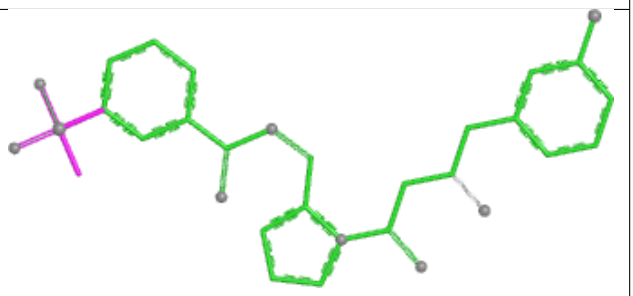


Rings

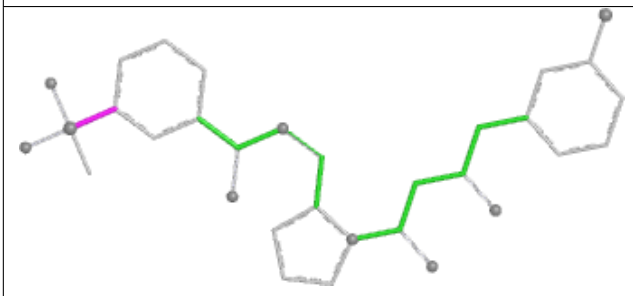
Ligand PS4 B 1



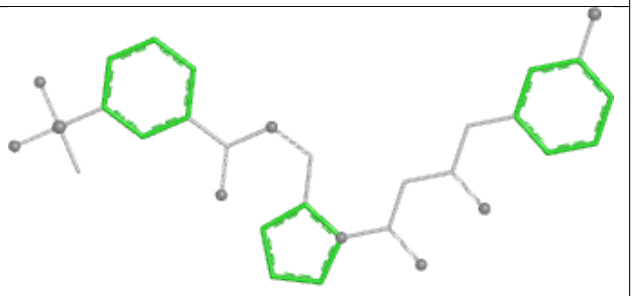
Bond lengths



Bond angles



Torsions



Rings

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	728/728 (100%)	-0.81	0 100 100	24, 35, 50, 60	0
1	B	728/728 (100%)	-0.71	2 (0%) 94 95	28, 40, 54, 61	0
All	All	1456/1456 (100%)	-0.76	2 (0%) 95 96	24, 37, 53, 61	0

All (2) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	B	279	VAL	2.6
1	B	766	PRO	2.2

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	NAG	B	797	14/15	0.85	0.22	63,64,65,65	0
2	NAG	A	795	14/15	0.88	0.19	61,62,63,63	0
2	NAG	B	793	14/15	0.89	0.17	55,58,60,60	0

Continued on next page...

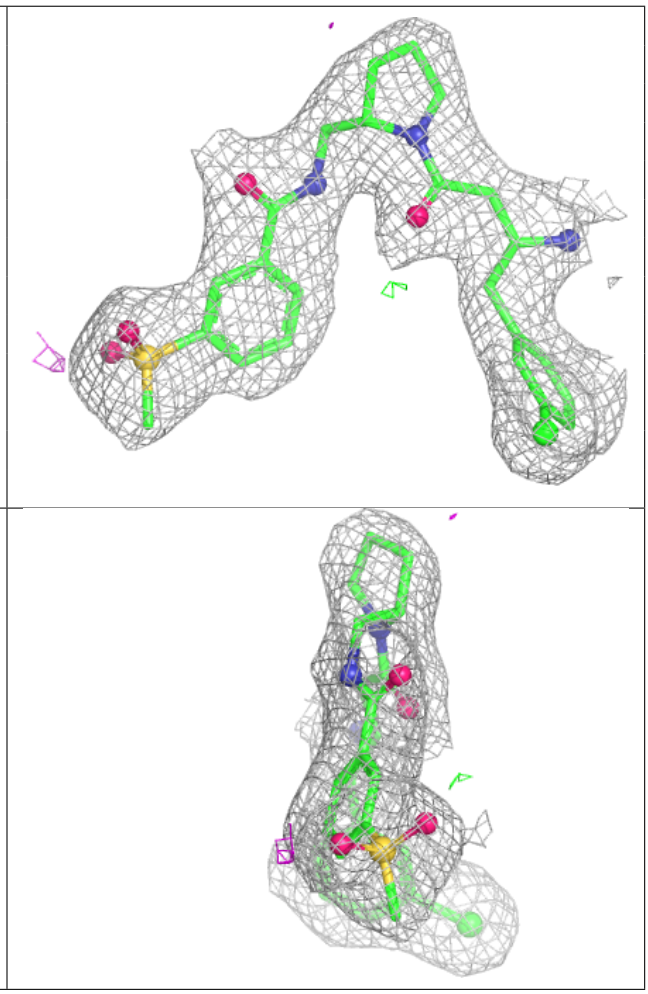
Continued from previous page...

Mol	Type	Chain	Res	Atoms	RSCC	RSR	B-factors(Å^2)	Q<0.9
2	NAG	A	793	14/15	0.92	0.19	52,53,53,53	0
2	NAG	A	796	14/15	0.92	0.14	39,42,43,44	0
2	NAG	B	796	14/15	0.93	0.12	44,45,46,47	0
2	NAG	B	794	14/15	0.95	0.13	42,44,45,45	0
2	NAG	A	794	14/15	0.97	0.12	30,32,32,32	0
3	PS4	B	1	32/32	0.97	0.13	32,35,39,40	0
3	PS4	A	1	32/32	0.98	0.11	33,35,36,38	0

The following is a graphical depiction of the model fit to experimental electron density of all instances of the Ligand of Interest. In addition, ligands with molecular weight > 250 and outliers as shown on the geometry validation Tables will also be included. Each fit is shown from different orientation to approximate a three-dimensional view.

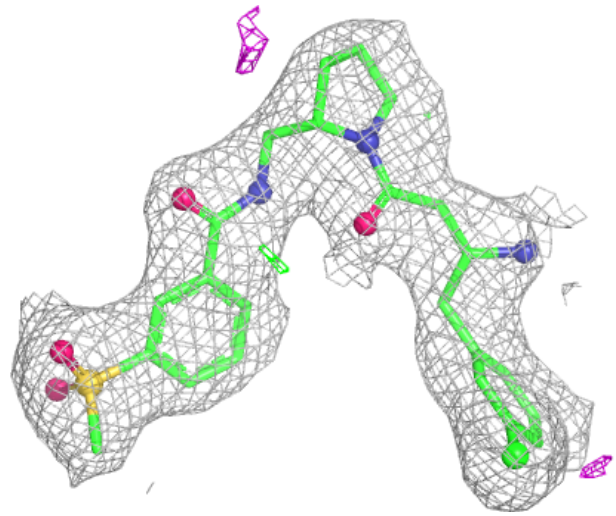
Electron density around PS4 B 1:

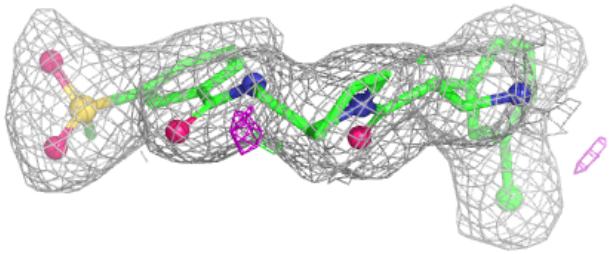
$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)

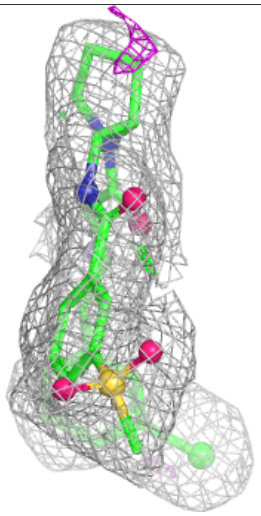


Electron density around PS4 A 1:

$2mF_o$-DF_c (at 0.7 rmsd) in gray
mF_o-DF_c (at 3 rmsd) in purple (negative)
and green (positive)







6.5 Other polymers (i)

There are no such residues in this entry.