wwPDB X-ray Structure Validation Summary Report (i)

Feb 1, 2016 – 12:50 AM GMT

PDB ID	:	2BUB
Title	:	CRYSTAL STRUCTURE OF HUMAN DIPEPTIDYL PEPTIDASE IV (CD26) IN COMPLEX WITH A REVERSED AMIDE INHIBITOR
Authors	:	Nordhoff, S.; Cerezo-Galvez, S.; Feurer, A.; Hill, O.; Matassa, V.G.; Metz, G.; Rummey, C.; Thiemann, M.; Edwards, P.J.
Deposited on	:	2005-06-09
Resolution	:	2.66 Å(reported)

This is a wwPDB X-ray Structure Validation Summary Report for a publicly released PDB entry.
We welcome your comments at validation@mail.wwpdb.org
A user guide is available at
http://wwpdb.org/validation/2016/XrayValidationReportHelp
with specific help available everywhere you see the (i) symbol.

The following versions of software and data (see references (i)) were used in the production of this report:

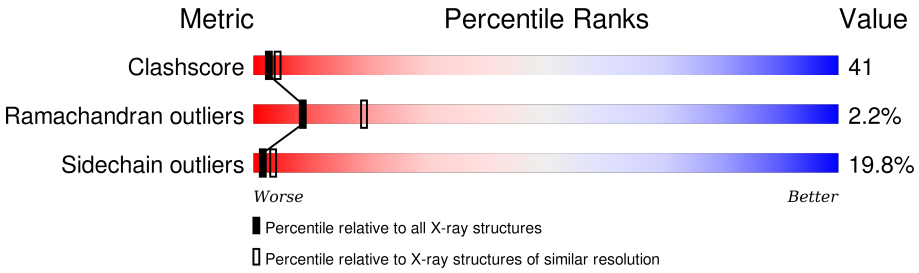
MolProbity	:	4.02b-467
Mogul	:	1.7 (RC4), CSD as536be (2015)
Xtriage (Phenix)	:	NOT EXECUTED
EDS	:	NOT EXECUTED
Percentile statistics	:	20151230.v01 (using entries in the PDB archive December 30th 2015)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
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 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))
Clashscore	102246	3524 (2.70-2.62)
Ramachandran outliers	100387	3469 (2.70-2.62)
Sidechain outliers	100360	3469 (2.70-2.62)

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.

Note EDS was not executed.

Mol	Chain	Length	Quality of chain
1	A	728	42% 46% 12%
1	B	728	36% 49% 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
3	NDG	A	1769	X	-	-	-

2 Entry composition

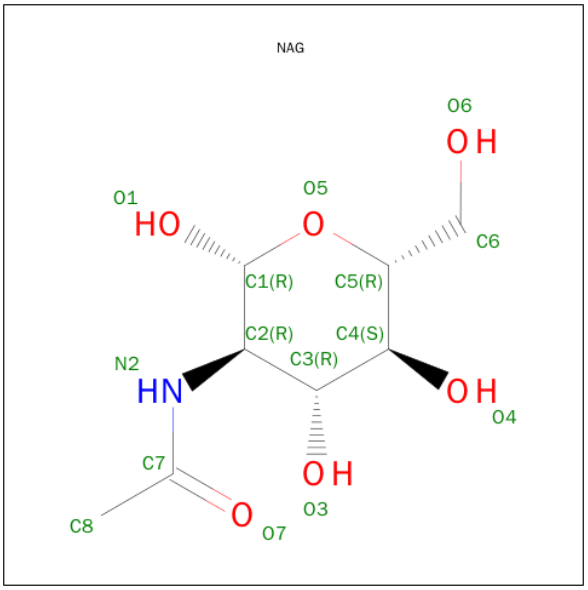
There are 5 unique types of molecules in this entry. The entry contains 12414 atoms, of which 0 are hydrogens and 0 are deuteriums.

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
1	A	728	Total	C	N	O	S	0	0	0
			5963	3827	982	1128	26			
1	B	728	Total	C	N	O	S	0	0	0
			5963	3827	982	1128	26			

- Molecule 2 is SUGAR (N-ACETYL-D-GLUCOSAMINE) (three-letter code: NAG) (formula: $C_8H_{15}NO_6$).



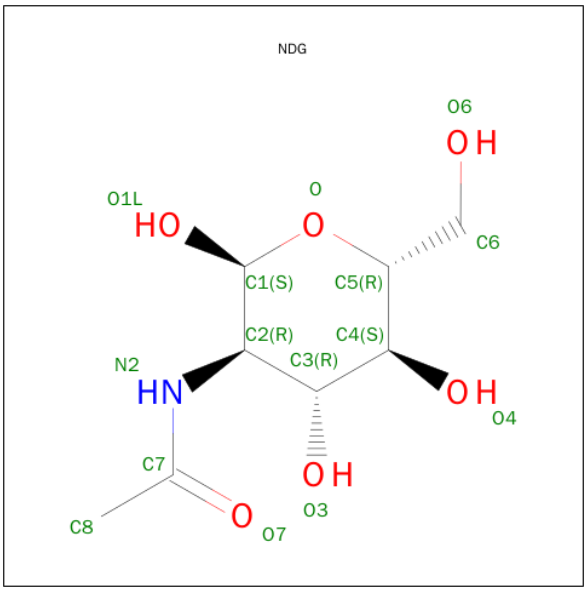
Mol	Chain	Residues	Atoms				ZeroOcc	AltConf
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	A	1	Total	C	N	O	0	0
			14	8	1	5		
2	B	1	Total	C	N	O	0	0
			14	8	1	5		

Continued on next page...

Continued from previous page...

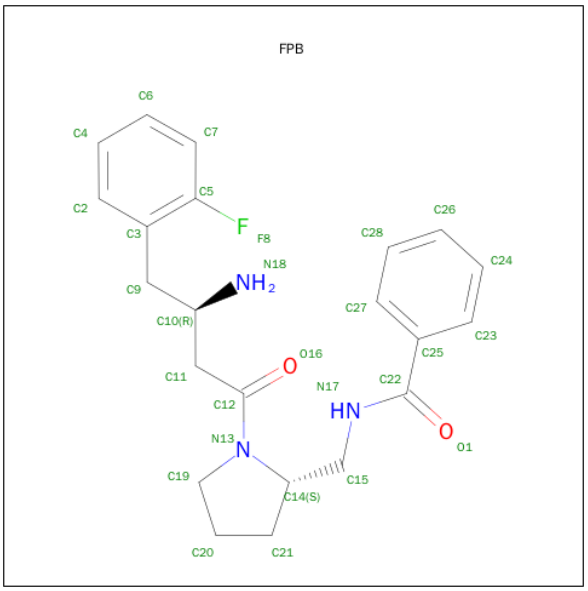
Mol	Chain	Residues	Atoms	ZeroOcc	AltConf
2	B	1	Total C N O 14 8 1 5	0	0
2	B	1	Total C N O 14 8 1 5	0	0

- Molecule 3 is SUGAR (2-(ACETYLAMINO)-2-DEOXY-A-D-GLUCOPYRANOSE) (three-letter code: NDG) (formula: $C_8H_{15}NO_6$).



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

- Molecule 4 is N-({(2S)-1-[(3R)-3-AMINO-4-(2-FLUOROPHENYL)BUTANOYL]PYRROLIDIN-2-YL}METHYL)BENZAMIDE (three-letter code: FPB) (formula: $C_{22}H_{26}FN_3O_2$).



Mol	Chain	Residues	Atoms					ZeroOcc	AltConf
4	A	1	Total	C	F	N	O	0	0
			28	22	1	3	2		
4	B	1	Total	C	F	N	O	0	0
			28	22	1	3	2		

- Molecule 5 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
5	A	161	Total	O	0	0
			161	161		
5	B	159	Total	O	0	0
			159	159		

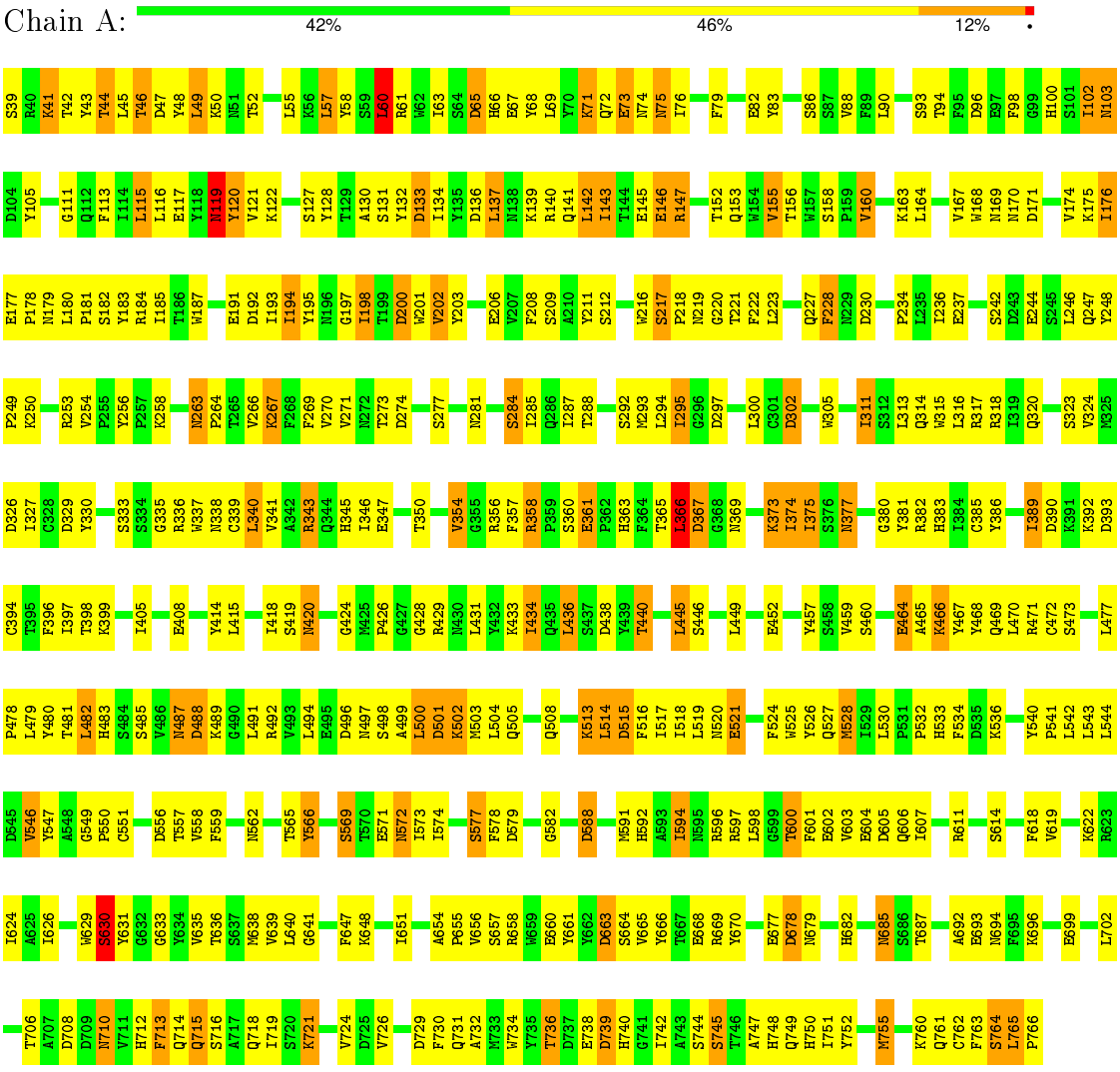
3 Residue-property plots (i)

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.

Note EDS was not executed.

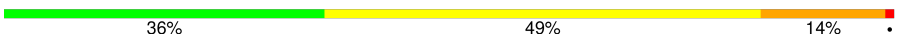
- Molecule 1: DIPEPTIDYL PEPTIDASE 4

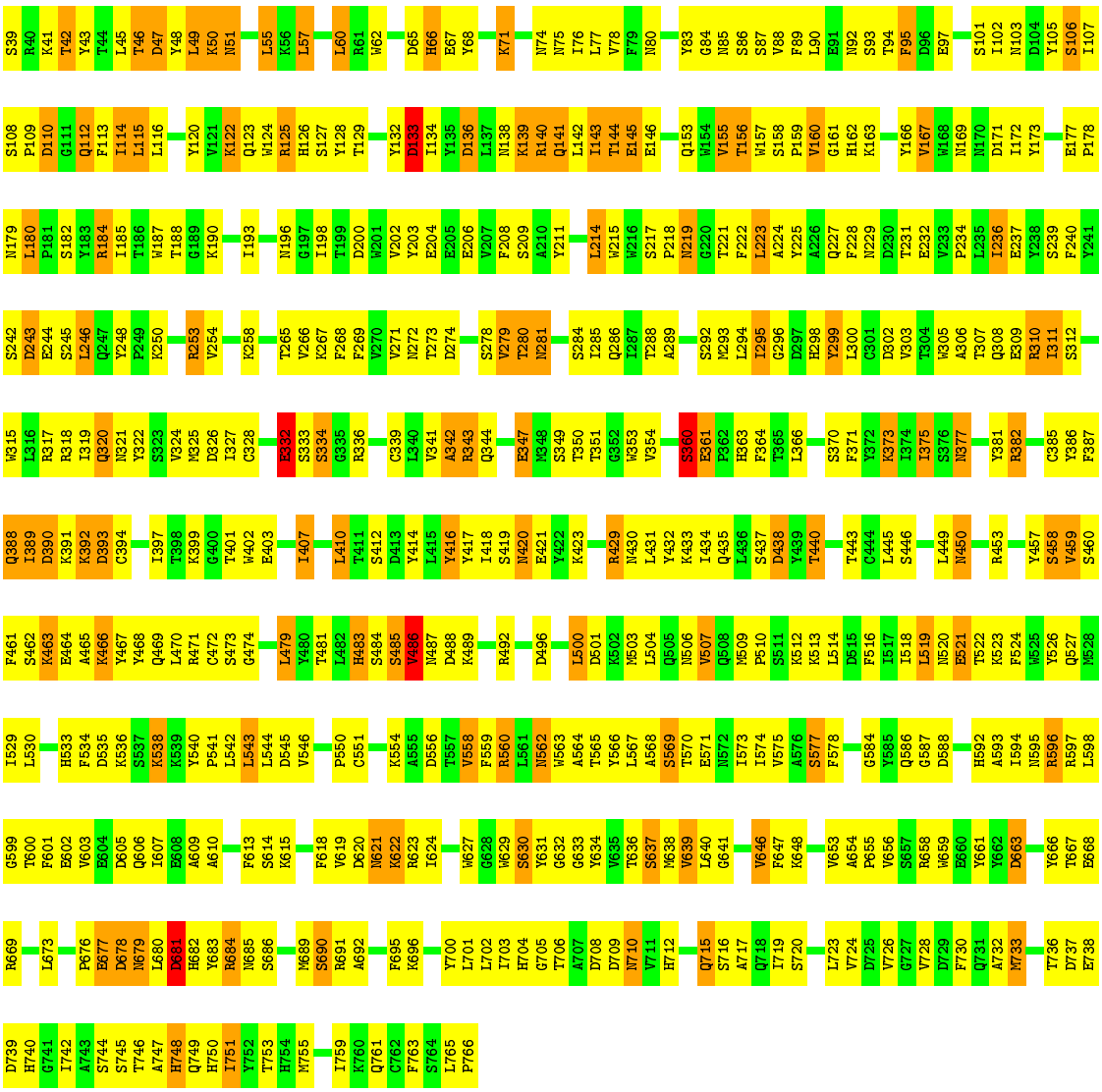
Chain A:



- Molecule 1: DIPEPTIDYL PEPTIDASE 4

Chain B:




S39 R40 K41 T42 Y43 T44 L45 T46 D47 Y48 L49 K50 N51 L55 K56 L57 L60 R61 W62 D65 H66 E67 Y68 K71 N74 N75 I76 L77 V78 F79 N80 Y83 G84 N85 S86 S87 V88 F89 L90 E91 N92 S93 T94 F95 D96 E97 S101 I102 N103 D104 Y105 S106 I107
S108 P109 D110 G111 Q112 F113 I114 L115 L116 Y120 V121 K122 Q123 W124 R125 H126 S127 Y128 T129 Y132 D133 I134 Y135 D136 L137 N138 K139 R140 Q141 L142 I143 T144 E145 E146 Q153 W154 V155 T156 W157 S158 P159 V160 G161 H162 K163 Y166 V167 W168 N169 N170 D171 I172 Y173 E177 P178
N179 L180 P181 S182 Y183 R184 I185 T186 W187 T188 G189 K190 I193 N196 G197 I198 T199 D200 W201 V202 Y203 E204 E205 E206 V207 F208 S209 A210 Y211 L214 W215 W216 S217 P218 N219 G220 T221 F222 L223 A224 Y225 A226 Q227 F228 N229 D230 T231 E232 V233 P234 L235 I236 E237 Y238 S239 F240 Y241
S242 D243 E244 S245 L246 Q247 Y248 P249 K250 R253 V254 K258 T265 V266 K267 F268 F269 V270 V271 M272 T273 D274 S278 V279 T280 M281 S284 I285 Q286 I287 T288 A289 S292 M293 L294 I295 G296 D297 H298 Y299 L300 C301 D302 V303 T304 W305 A306 T307 Q308 E309 R310 I311 S312
W315 L316 R317 R318 I319 Q320 N321 Y322 S323 V324 M325 D326 I327 G328 E332 S333 S334 G335 R336 C339 L340 V341 A342 R343 Q344 E347 M348 S349 T350 T351 G352 W353 V354 S360 E361 P362 H363 F364 T365 L366 S370 F371 Y372 K373 I374 I375 S376 N377 Y381 R382 C385 Y386 F387
Q388 I389 D390 K391 K392 D393 C394 I397 T398 K399 G400 T401 W402 E403 I407 L410 T411 S412 D413 Y414 L415 Y416 Y417 I418 S419 N420 E421 Y422 K423 R429 N430 L431 Y432 K433 I434 Q435 L436 S437 D438 Y439 T440 T443 C444 L445 S446 L449 N450 R453 Y457 S458 V459 S460
F461 S462 K463 E464 A465 K466 Y467 Y468 Q469 L470 R471 C472 S473 G474 L479 Y480 T481 L482 H483 S484 S485 V486 N487 D488 K489 R492 D496 L500 D501 K502 M503 L504 Q505 N506 V507 Q508 M509 P510 S511 K512 K513 L514 D515 F516 I517 I518 L519 N520 E521 T522 K523 F524 W525 Y526 Q527 M528
I529 L530 H533 F534 D535 K536 S537 K538 K539 Y540 P541 L542 L543 L544 D545 V546 P550 C551 K554 A555 D556 T557 V558 F559 R560 L561 N562 W563 A564 T565 Y566 L567 A568 S569 T570 E571 N572 I573 I574 V575 A576 S577 F578 G584 Y585 Q586 G587 D588 H592 A593 I594 N595 R596 R597 L598
G599 T600 F601 E602 V603 E604 D605 Q606 I607 E608 A609 A610 F613 S614 K615 F618 V619 D620 M621 K622 R623 I624 W627 G628 W629 S630 Y631 G632 G633 Y634 V635 T636 S637 M638 V639 L640 G641 V646 F647 K648 V653 A654 P655 V656 S657 R658 W659 E660 Y661 Y662 D663 Y666 T667 E668
R669 L673 P676 E677 D678 N679 L680 D681 H682 Y683 R684 N685 S686 M689 S690 R691 A692 F695 K696 Y700 L701 L702 I703 H704 G705 T706 A707 D708 D709 N710 V711 H712 Q715 S716 A717 Q718 I719 S720 L723 V724 D725 V726 G727 V728 D729 F730 Q731 A732 M733 T736 D737 E738
D739 H740 G741 I742 A743 S744 S745 T746 A747 H748 Q749 H750 I751 Y752 T753 H754 M755 I759 K760 Q761 C762 F763 S764 L765 P766

4 Data and refinement statistics (i)

Xtriage (Phenix) and EDS were not executed - this section will therefore be incomplete.

Property	Value	Source
Space group	P 21 21 21	Depositor
Cell constants a, b, c, α, β, γ	65.49Å 66.77Å 425.41Å 90.00° 90.00° 90.00°	Depositor
Resolution (Å)	20.00 – 2.66	Depositor
% Data completeness (in resolution range)	96.6 (20.00-2.66)	Depositor
R_{merge}	0.00	Depositor
R_{sym}	(Not available)	Depositor
Refinement program	REFMAC 5.1.24	Depositor
R, R_{free}	0.258 , 0.329	Depositor
Estimated twinning fraction	No twinning to report.	Xtriage
Total number of atoms	12414	wwPDB-VP
Average B, all atoms (Å^2)	38.0	wwPDB-VP

5 Model quality

5.1 Standard geometry

Bond lengths and bond angles in the following residue types are not validated in this section: NAG, NDG, FPB

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	0.62	0/6135	0.88	20/8344 (0.2%)
1	B	0.62	0/6135	0.85	21/8344 (0.3%)
All	All	0.62	0/12270	0.86	41/16688 (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	A	0	1

There are no bond length outliers.

The worst 5 of 41 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	B	133	ASP	CB-CG-OD2	8.18	125.66	118.30
1	B	47	ASP	CB-CG-OD2	7.81	125.33	118.30
1	A	200	ASP	CB-CG-OD2	7.54	125.08	118.30
1	B	545	ASP	CB-CG-OD2	7.41	124.97	118.30
1	A	678	ASP	CB-CG-OD2	6.80	124.42	118.30

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	491	LEU	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	5963	0	5681	467	0
1	B	5963	0	5681	507	0
2	A	42	0	39	1	0
2	B	42	0	39	1	0
3	A	14	0	13	0	0
3	B	14	0	13	2	0
4	A	28	0	26	3	0
4	B	28	0	26	4	0
5	A	161	0	0	29	0
5	B	159	0	0	41	0
All	All	12414	0	11518	969	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 41.

The worst 5 of 969 close contacts within the same asymmetric unit are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:136:ASP:OD2	1:B:139:LYS:HD2	1.35	1.26
1:A:75:ASN:HD22	1:A:75:ASN:N	1.27	1.23
1:B:594:ILE:CD1	1:B:601:PHE:HB2	1.67	1.23
1:B:600:THR:CG2	1:B:601:PHE:H	1.54	1.20
1:A:682:HIS:ND1	1:A:685:ASN:HB3	1.57	1.19

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	726/728 (100%)	626 (86%)	88 (12%)	12 (2%)	11 25
1	B	726/728 (100%)	624 (86%)	82 (11%)	20 (3%)	6 13
All	All	1452/1456 (100%)	1250 (86%)	170 (12%)	32 (2%)	8 19

5 of 32 Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	713	PHE
1	A	764	SER
1	B	140	ARG
1	B	219	ASN
1	B	279	VAL

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%)	532 (82%)	121 (18%)	2 4
1	B	653/653 (100%)	515 (79%)	138 (21%)	1 3
All	All	1306/1306 (100%)	1047 (80%)	259 (20%)	1 3

5 of 259 residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	745	SER
1	B	125	ARG
1	B	621	ASN
1	A	765	LEU
1	B	71	LYS

Some sidechains can be flipped to improve hydrogen bonding and reduce clashes. 5 of 46 such sidechains are listed below:

Mol	Chain	Res	Type
1	A	750	HIS
1	B	119	ASN
1	B	685	ASN
1	B	51	ASN
1	B	74	ASN

5.3.3 RNA

There are no RNA molecules in this entry.

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

There are no non-standard protein/DNA/RNA residues in this entry.

5.5 Carbohydrates

There are no carbohydrates in this entry.

5.6 Ligand geometry

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	$\#\|Z\| > 2$
2	NAG	A	1767	1	14,14,15	0.50	0	15,19,21	0.97	1 (6%)
2	NAG	A	1768	1	14,14,15	0.65	0	15,19,21	1.94	5 (33%)
3	NDG	A	1769	1	14,14,15	0.42	0	15,19,21	1.81	2 (13%)
2	NAG	A	1770	1	14,14,15	0.52	0	15,19,21	1.16	2 (13%)
4	FPB	A	1771	-	30,30,30	0.90	1 (3%)	34,40,40	1.07	2 (5%)
2	NAG	B	1767	1	14,14,15	0.48	0	15,19,21	1.16	2 (13%)
2	NAG	B	1768	1	14,14,15	0.61	0	15,19,21	1.83	3 (20%)
2	NAG	B	1769	1	14,14,15	0.46	0	15,19,21	1.12	0

Mol	Type	Chain	Res	Link	Bond lengths			Bond angles		
					Counts	RMSZ	$\#\|Z\|>2$	Counts	RMSZ	$\#\|Z\|>2$
3	NDG	B	1770	1	14,14,15	0.68	0	15,19,21	0.92	0
4	FPB	B	1771	-	30,30,30	0.72	1 (3%)	34,40,40	1.16	4 (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	NAG	A	1767	1	-	0/6/23/26	0/1/1/1
2	NAG	A	1768	1	-	0/6/23/26	0/1/1/1
3	NDG	A	1769	1	1/1/5/7	0/6/23/26	0/1/1/1
2	NAG	A	1770	1	-	0/6/23/26	0/1/1/1
4	FPB	A	1771	-	-	0/21/31/31	0/3/3/3
2	NAG	B	1767	1	-	2/6/23/26	0/1/1/1
2	NAG	B	1768	1	-	0/6/23/26	0/1/1/1
2	NAG	B	1769	1	-	0/6/23/26	0/1/1/1
3	NDG	B	1770	1	-	0/6/23/26	0/1/1/1
4	FPB	B	1771	-	-	0/21/31/31	0/3/3/3

All (2) bond length outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(Å)	Ideal(Å)
4	B	1771	FPB	C22-N17	2.08	1.38	1.33
4	A	1771	FPB	C22-N17	3.16	1.40	1.33

The worst 5 of 21 bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
2	B	1768	NAG	C2-N2-C7	-5.55	115.91	123.04
3	A	1769	NDG	C4-C3-C2	-4.25	104.61	111.23
2	A	1768	NAG	C4-C3-C2	-3.07	106.45	111.23
2	A	1768	NAG	C6-C5-C4	-3.04	105.53	113.02
4	A	1771	FPB	C9-C10-C11	-2.69	106.22	110.81

All (1) chirality outliers are listed below:

Mol	Chain	Res	Type	Atom
3	A	1769	NDG	C1

All (2) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	B	1767	NAG	O7-C7-N2-C2
2	B	1767	NAG	C8-C7-N2-C2

There are no ring outliers.

5 monomers are involved in 11 short contacts:

Mol	Chain	Res	Type	Clashes	Symm-Clashes
2	A	1767	NAG	1	0
4	A	1771	FPB	3	0
2	B	1769	NAG	1	0
3	B	1770	NDG	2	0
4	B	1771	FPB	4	0

5.7 Other polymers (i)

There are no such residues in this entry.

5.8 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Fit of model and data

6.1 Protein, DNA and RNA chains

EDS was not executed - this section will therefore be empty.

6.2 Non-standard residues in protein, DNA, RNA chains

EDS was not executed - this section will therefore be empty.

6.3 Carbohydrates

EDS was not executed - this section will therefore be empty.

6.4 Ligands

EDS was not executed - this section will therefore be empty.

6.5 Other polymers

EDS was not executed - this section will therefore be empty.