Full wwPDB X-ray Structure Validation Report (i)

Jan 3, 2024 – 04:27 am GMT

PDB ID	:	4YWT
Title	:	Crystal structure of full-length glypican-1 core protein after controlled crystal dehydration to 87% relative humidity
Authors	:	Awad, W.; Mani, K.; Logan, D.T.
Deposited on	:	2015-03-21
Resolution	:	2.38 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:

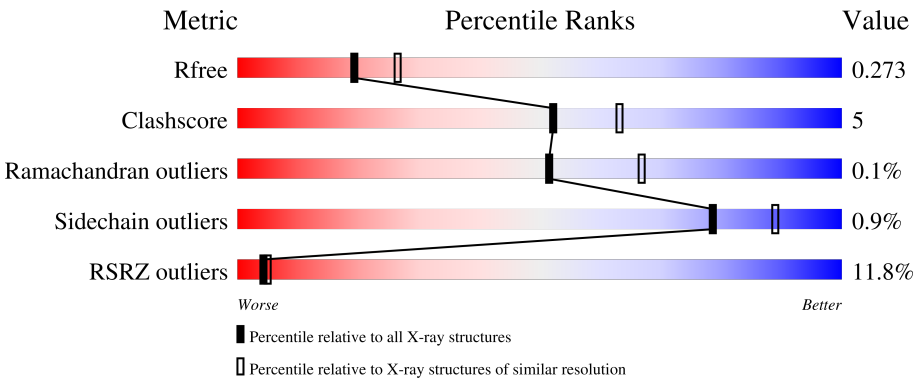
MolProbity	:	4.02b-467
Mogul	:	1.8.4, CSD as541be (2020)
Xtriage (Phenix)	:	1.13
EDS	:	2.36
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
Refmac	:	5.8.0158
CCP4	:	7.0.044 (Gargrove)
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.36

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 2.38 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	5509 (2.40-2.36)
Clashscore	141614	6082 (2.40-2.36)
Ramachandran outliers	138981	5973 (2.40-2.36)
Sidechain outliers	138945	5975 (2.40-2.36)
RSRZ outliers	127900	5397 (2.40-2.36)

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	526	11% 69% 9% 22%
1	B	526	9% 70% 12% 18%
1	C	526	6% 68% 8% 24%
1	D	526	11% 67% 15% 18%

2 Entry composition (i)

There are 4 unique types of molecules in this entry. The entry contains 13110 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 Glypican-1.

Mol	Chain	Residues	Atoms					ZeroOcc	AltConf	Trace
			Total	C	N	O	S			
1	A	411	3166	1981	568	594	23	0	4	0
1	B	431	3283	2062	581	616	24	0	3	0
1	C	400	3058	1918	544	575	21	0	2	0
1	D	429	3254	2043	571	613	27	0	4	0

There are 100 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	2	ALA	-	expression tag	UNP P35052
A	3	PRO	-	expression tag	UNP P35052
A	4	GLN	-	expression tag	UNP P35052
A	5	LEU	-	expression tag	UNP P35052
A	6	HIS	-	expression tag	UNP P35052
A	7	HIS	-	expression tag	UNP P35052
A	8	HIS	-	expression tag	UNP P35052
A	9	HIS	-	expression tag	UNP P35052
A	10	HIS	-	expression tag	UNP P35052
A	11	HIS	-	expression tag	UNP P35052
A	12	ASP	-	expression tag	UNP P35052
A	13	LEU	-	expression tag	UNP P35052
A	14	TYR	-	expression tag	UNP P35052
A	15	GLU	-	expression tag	UNP P35052
A	16	ASN	-	expression tag	UNP P35052
A	17	LEU	-	expression tag	UNP P35052
A	18	TYR	-	expression tag	UNP P35052
A	19	PHE	-	expression tag	UNP P35052
A	20	GLN	-	expression tag	UNP P35052
A	21	GLY	-	expression tag	UNP P35052
A	22	LYS	-	expression tag	UNP P35052

Continued on next page...

Continued from previous page...

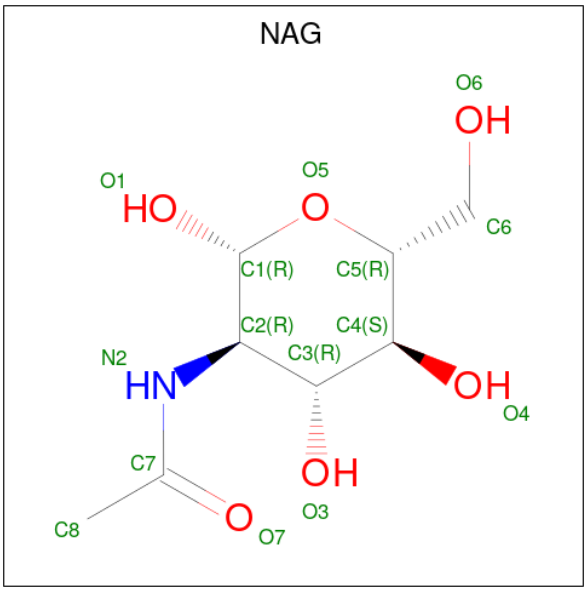
Chain	Residue	Modelled	Actual	Comment	Reference
A	23	LEU	-	expression tag	UNP P35052
A	486	ALA	SER	engineered mutation	UNP P35052
A	488	ALA	SER	engineered mutation	UNP P35052
A	490	ALA	SER	engineered mutation	UNP P35052
B	2	ALA	-	expression tag	UNP P35052
B	3	PRO	-	expression tag	UNP P35052
B	4	GLN	-	expression tag	UNP P35052
B	5	LEU	-	expression tag	UNP P35052
B	6	HIS	-	expression tag	UNP P35052
B	7	HIS	-	expression tag	UNP P35052
B	8	HIS	-	expression tag	UNP P35052
B	9	HIS	-	expression tag	UNP P35052
B	10	HIS	-	expression tag	UNP P35052
B	11	HIS	-	expression tag	UNP P35052
B	12	ASP	-	expression tag	UNP P35052
B	13	LEU	-	expression tag	UNP P35052
B	14	TYR	-	expression tag	UNP P35052
B	15	GLU	-	expression tag	UNP P35052
B	16	ASN	-	expression tag	UNP P35052
B	17	LEU	-	expression tag	UNP P35052
B	18	TYR	-	expression tag	UNP P35052
B	19	PHE	-	expression tag	UNP P35052
B	20	GLN	-	expression tag	UNP P35052
B	21	GLY	-	expression tag	UNP P35052
B	22	LYS	-	expression tag	UNP P35052
B	23	LEU	-	expression tag	UNP P35052
B	486	ALA	SER	engineered mutation	UNP P35052
B	488	ALA	SER	engineered mutation	UNP P35052
B	490	ALA	SER	engineered mutation	UNP P35052
C	2	ALA	-	expression tag	UNP P35052
C	3	PRO	-	expression tag	UNP P35052
C	4	GLN	-	expression tag	UNP P35052
C	5	LEU	-	expression tag	UNP P35052
C	6	HIS	-	expression tag	UNP P35052
C	7	HIS	-	expression tag	UNP P35052
C	8	HIS	-	expression tag	UNP P35052
C	9	HIS	-	expression tag	UNP P35052
C	10	HIS	-	expression tag	UNP P35052
C	11	HIS	-	expression tag	UNP P35052
C	12	ASP	-	expression tag	UNP P35052
C	13	LEU	-	expression tag	UNP P35052
C	14	TYR	-	expression tag	UNP P35052

Continued on next page...

Continued from previous page...

Chain	Residue	Modelled	Actual	Comment	Reference
C	15	GLU	-	expression tag	UNP P35052
C	16	ASN	-	expression tag	UNP P35052
C	17	LEU	-	expression tag	UNP P35052
C	18	TYR	-	expression tag	UNP P35052
C	19	PHE	-	expression tag	UNP P35052
C	20	GLN	-	expression tag	UNP P35052
C	21	GLY	-	expression tag	UNP P35052
C	22	LYS	-	expression tag	UNP P35052
C	23	LEU	-	expression tag	UNP P35052
C	486	ALA	SER	engineered mutation	UNP P35052
C	488	ALA	SER	engineered mutation	UNP P35052
C	490	ALA	SER	engineered mutation	UNP P35052
D	2	ALA	-	expression tag	UNP P35052
D	3	PRO	-	expression tag	UNP P35052
D	4	GLN	-	expression tag	UNP P35052
D	5	LEU	-	expression tag	UNP P35052
D	6	HIS	-	expression tag	UNP P35052
D	7	HIS	-	expression tag	UNP P35052
D	8	HIS	-	expression tag	UNP P35052
D	9	HIS	-	expression tag	UNP P35052
D	10	HIS	-	expression tag	UNP P35052
D	11	HIS	-	expression tag	UNP P35052
D	12	ASP	-	expression tag	UNP P35052
D	13	LEU	-	expression tag	UNP P35052
D	14	TYR	-	expression tag	UNP P35052
D	15	GLU	-	expression tag	UNP P35052
D	16	ASN	-	expression tag	UNP P35052
D	17	LEU	-	expression tag	UNP P35052
D	18	TYR	-	expression tag	UNP P35052
D	19	PHE	-	expression tag	UNP P35052
D	20	GLN	-	expression tag	UNP P35052
D	21	GLY	-	expression tag	UNP P35052
D	22	LYS	-	expression tag	UNP P35052
D	23	LEU	-	expression tag	UNP P35052
D	486	ALA	SER	engineered mutation	UNP P35052
D	488	ALA	SER	engineered mutation	UNP P35052
D	490	ALA	SER	engineered mutation	UNP P35052

- 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	B	1	14	8	1	5	0	0
2	C	1	14	8	1	5	0	0
2	D	1	14	8	1	5	0	0

- Molecule 3 is CALCIUM ION (three-letter code: CA) (formula: Ca).

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	Ca		
3	A	2	2	2	0	0
3	B	2	2	2	0	0
3	C	3	3	3	0	0
3	D	2	2	2	0	0

- Molecule 4 is water.

Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	A	102	102	102	0	0

Continued on next page...

Continued from previous page...

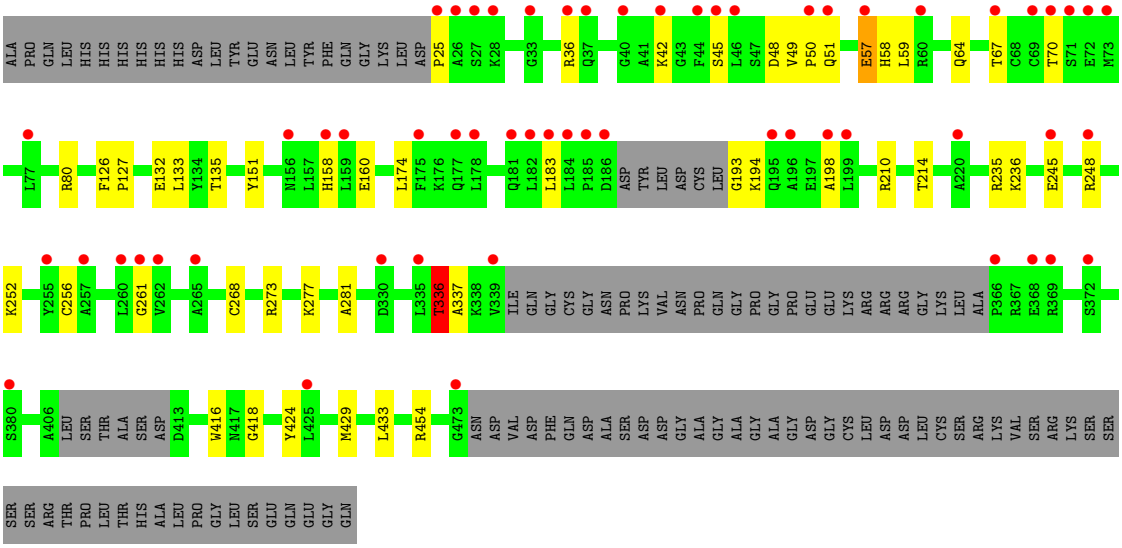
Mol	Chain	Residues	Atoms		ZeroOcc	AltConf
			Total	O		
4	B	86	86	86	0	0
			Total	O		
4	C	53	53	53	0	0
			Total	O		
4	D	43	43	43	0	0

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.

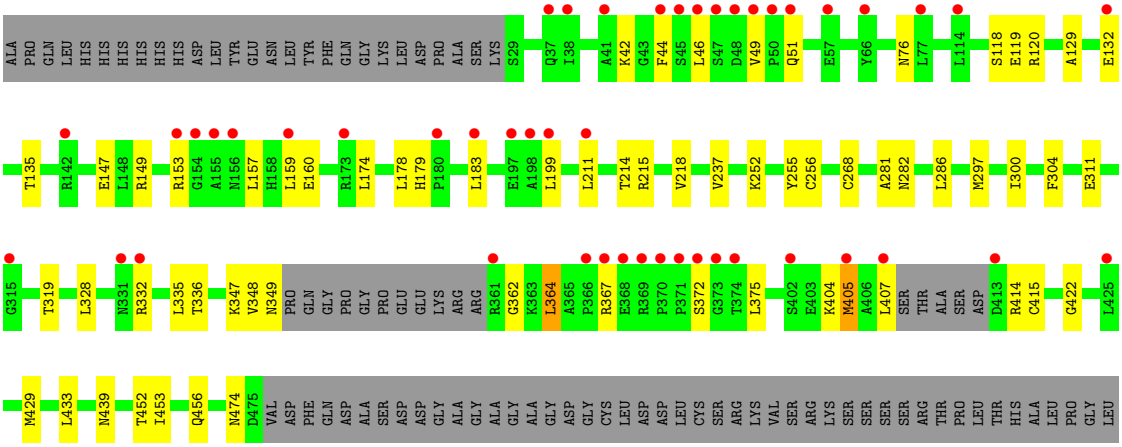
- Molecule 1: Glypican-1

Chain A:



- Molecule 1: Glypican-1

Chain B:



SER GLU GLN GLU GLY GLN

- Molecule 1: Glypican-1

Chain C:

ALA PRO GLN LEU HIS HIS HIS HIS HIS HIS ASP LEU TYR GLU ASN LEU TYR PHE GLN GLY LYS LEU ASP PRO ALA SER LYS S29 R36 Y39 K42 V49 M73 R94 Q97 E119 E132 T135 F141 R149 G154 A155 N156 L157 H158 L171

E172 R173 L174 F175 K176 Q177 L178 H179 P180 Q181 L182 L183 L184 P185 ASP ASP TYR TYR LEU ASP CYS LEU GLY LYS GLN ALA GLU A198 L199 A205 P206 R207 R221 V233 V234 R235 A238 K252 Y255 C256 D270 R273 A281 V298 T307 V310 E311 L328

R332 D333 T334 L335 T336 A337 K338 V339 ILE GLN GLY CYS GLY ASN PRO LYS VAL ASN PRO GLN GLY PRO GLY PRO GLU GLU LYS ARG ARG ARG GLY LYS LEU ALA PRO R367 S380 K383 R387 E403 K404 M405 A406 LEU SER THR ALA SER ASP ASP R414 G422 R423 Y424

L425 P426 E427 G430 D431 G432 L433 Q436 D445 I446 T447 Q456 N474 ASP VAL ASP PHE GLN ASP ALA SER ASP ASP GLY ALA GLY ALA GLY ALA GLY ASP GLY CYS LEU ASP ASP LEU CYS SER ARG LYS VAL SER ARG LYS SER SER SER ARG THR PRO LEU THR HIS

ALA LEU PRO GLY LEU SER GLU GLN GLU GLY GLN

- Molecule 1: Glypican-1

Chain D:

ALA PRO GLN LEU HIS HIS HIS HIS HIS HIS ASP LEU TYR GLU ASN LEU TYR PHE GLN GLY LYS LEU ASP PRO ALA SER LYS S29 I38 K42 G43 F44 S45 L46 S47 D48 A52 G56 E57 H58 G65 T70 N76 L89 Q103 L114 E119 A129

E132 L133 Y134 T135 Q136 R139 R142 R149 R153 G154 A155 N156 L157 H158 L159 E160 E161 E165 R169 L170 L171 E172 R173 L174 F175 K176 Q177 L178 Q181 L182 L183 L184 D187 Y188 L189 Q195 A196 E197 A198 L199 R200 P201 F202 L211 T214 R215 A216 F217

V218 R221 S222 G243 K252 C256 V262 A265 C268 A281 N282 R291 L294 D295 S296 M297 V298 L299 I300 T301 F304 V310 E311 S316 T319 A324 D333 T334 L335 T336 V339 G344 V348 N349 PRO GLN GLY PRO GLY PRO GLU

GLU LYS ARG ARG R361 G362 K363 L364 A365 P366 R369 P370 P371 S372 G373 T374 L375 E376 K377 S380 E381 R387 Q390 L396 P397 C401 S402 E403 K404 M405 A406 LEU SER THR ALA SER ASP D413 R414 C415 R421 G422 L425 V428 M429 G430 L433 Q436

I453 L467 G473 N474 ASP VAL ASP PHE GLN ASP ALA SER ASP ASP GLY ALA GLY ALA GLY ALA ASP GLY CYS LEU ASP ASP LEU CYS SER ARG LYS VAL SER ARG SER SER SER SER ARG THR PRO LEU THR HIS ALA LEU PRO GLY LEU SER GLU GLN GLU GLY GLN

4 Data and refinement statistics (i)

Property	Value	Source
Space group	P 1 21 1	Depositor
Cell constants a, b, c, α, β, γ	46.78Å 166.59Å 137.70Å 90.00° 90.38° 90.00°	Depositor
Resolution (Å)	43.22 – 2.38 43.22 – 2.38	Depositor EDS
% Data completeness (in resolution range)	97.8 (43.22-2.38) 97.9 (43.22-2.38)	Depositor EDS
R_{merge}	0.11	Depositor
R_{sym}	(Not available)	Depositor
$< I/\sigma(I) >$ [1]	1.35 (at 2.37Å)	Xtriage
Refinement program	PHENIX 1.9_1692	Depositor
R, R_{free}	0.239 , 0.273 0.241 , 0.273	Depositor DCC
R_{free} test set	1985 reflections (2.41%)	wwPDB-VP
Wilson B-factor (Å^2)	38.7	Xtriage
Anisotropy	0.569	Xtriage
Bulk solvent k_{sol}(e/Å^3), B_{sol}(Å^2)	0.34 , 59.8	EDS
L-test for twinning[2]	$< \|L\| > = 0.50$, $< L^2 > = 0.33$	Xtriage
Estimated twinning fraction	0.055 for h,-k,-l	Xtriage
F_o,F_c correlation	0.94	EDS
Total number of atoms	13110	wwPDB-VP
Average B, all atoms (Å^2)	59.0	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 46.82 % 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 1.0833e-04. The detected translational NCS is most likely also responsible for the elevated intensity ratio.*

[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: CA, 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	0.24	0/3238	0.45	1/4390 (0.0%)
1	B	0.29	0/3352	0.47	1/4550 (0.0%)
1	C	0.26	0/3120	0.44	0/4234
1	D	0.25	0/3326	0.44	0/4517
All	All	0.26	0/13036	0.45	2/17691 (0.0%)

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

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

There are no bond length outliers.

All (2) bond angle outliers are listed below:

Mol	Chain	Res	Type	Atoms	Z	Observed(^{o})	Ideal(^{o})
1	A	336	THR	N-CA-C	6.11	127.49	111.00
1	B	404	LYS	CA-CB-CG	5.15	124.74	113.40

There are no chirality outliers.

All (1) planarity outliers are listed below:

Mol	Chain	Res	Type	Group
1	A	198	ALA	Peptide

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	3166	0	3070	30	0
1	B	3283	0	3184	39	0
1	C	3058	0	2949	25	0
1	D	3254	0	3126	45	0
2	A	14	0	13	0	0
2	B	14	0	13	0	0
2	C	14	0	13	0	0
2	D	14	0	13	0	0
3	A	2	0	0	0	0
3	B	2	0	0	0	0
3	C	3	0	0	0	0
3	D	2	0	0	0	0
4	A	102	0	0	3	0
4	B	86	0	0	4	0
4	C	53	0	0	0	0
4	D	43	0	0	2	0
All	All	13110	0	12381	138	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 (138) 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:407:LEU:HD22	1:B:415:CYS:HB3	1.66	0.77
1:A:25:PRO:HD2	1:A:261:GLY:HA2	1.67	0.76
1:B:319:THR:HG21	1:B:367:ARG:HA	1.69	0.74
1:B:268[B]:CYS:SG	4:B:781:HOH:O	2.47	0.73
1:B:129:ALA:HB2	1:B:364:LEU:HD11	1.70	0.72
1:C:119:GLU:OE1	1:C:149:ARG:NH2	2.23	0.71
1:B:147:GLU:HG3	1:B:157:LEU:HD21	1.75	0.69
1:A:245:GLU:OE1	1:A:248:ARG:NH2	2.25	0.68
1:D:119:GLU:OE1	1:D:149:ARG:NH2	2.28	0.67
1:D:187:ASP:OD1	1:D:188:TYR:N	2.30	0.65
1:A:268[B]:CYS:SG	4:A:790:HOH:O	2.54	0.65

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:268[B]:CYS:SG	4:D:741:HOH:O	2.54	0.65
1:B:178:LEU:HD22	1:B:332:ARG:HD3	1.79	0.64
1:A:194:LYS:N	1:A:194:LYS:HD3	2.13	0.64
1:D:160:GLU:N	1:D:160:GLU:OE1	2.28	0.63
1:B:311:GLU:HG3	1:B:375:LEU:HD22	1.81	0.63
1:B:118:SER:OG	4:B:701:HOH:O	2.16	0.62
1:D:195:GLN:O	1:D:199:LEU:HB2	1.98	0.62
1:A:273:ARG:NH2	4:A:702:HOH:O	2.32	0.62
1:A:51:GLN:H	1:A:51:GLN:CD	2.02	0.62
1:D:129:ALA:HA	1:D:362:GLY:HA2	1.82	0.62
1:C:403:GLU:HG3	1:C:404:LYS:HD3	1.85	0.59
1:B:282:ASN:HB3	1:B:453:ILE:HD12	1.84	0.59
1:B:281:ALA:HB1	1:B:433:LEU:HA	1.85	0.59
1:D:333:ASP:HA	1:D:336:THR:HG22	1.84	0.58
1:C:36:ARG:HA	1:C:49:VAL:HG21	1.86	0.58
1:B:348:VAL:HG12	1:B:349:ASN:H	1.69	0.58
1:C:157:LEU:HD12	1:C:158:HIS:H	1.69	0.57
1:D:282:ASN:HB3	1:D:453:ILE:HD12	1.85	0.57
1:D:211:LEU:O	1:D:214:THR:OG1	2.21	0.56
1:B:179:HIS:NE2	1:B:336:THR:OG1	2.39	0.54
1:D:58:HIS:HE2	1:D:70:THR:HA	1.72	0.54
1:B:237:VAL:HG21	1:B:286:LEU:HD11	1.90	0.54
1:A:193:GLY:C	1:A:194:LYS:HD3	2.28	0.54
1:B:119:GLU:OE1	1:B:149:ARG:NH2	2.33	0.54
1:B:304:PHE:HB3	1:B:375:LEU:HD21	1.90	0.54
1:B:414:ARG:HD2	1:B:422:GLY:HA2	1.89	0.54
1:D:281:ALA:HB1	1:D:433:LEU:HA	1.90	0.53
1:A:57:GLU:HG3	1:A:58:HIS:N	2.24	0.53
1:C:270:ASP:OD1	1:C:273:ARG:NH2	2.39	0.53
1:B:429:MET:HG3	1:B:439:ASN:HD22	1.74	0.53
1:A:281:ALA:HB1	1:A:433:LEU:HA	1.91	0.52
1:D:421:ARG:NH2	4:D:703:HOH:O	2.42	0.52
1:C:430:GLY:O	1:C:436:GLN:NE2	2.41	0.52
1:D:160:GLU:H	1:D:160:GLU:CD	2.13	0.52
1:A:416:TRP:CE2	1:A:418:GLY:HA2	2.45	0.52
1:B:120:ARG:NH1	4:B:709:HOH:O	2.43	0.52
1:B:132:GLU:HA	1:B:135:THR:OG1	2.10	0.52
1:C:281:ALA:HB1	1:C:433:LEU:HA	1.90	0.52
1:A:59:LEU:HD21	1:A:67:THR:HG21	1.92	0.51
1:C:154:GLY:HA2	1:C:221:ARG:HD2	1.92	0.51
1:C:252:LYS:HA	1:C:256:CYS:SG	2.50	0.51

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:B:211:LEU:O	1:B:214:THR:OG1	2.27	0.51
1:D:297:MET:O	1:D:300:ILE:HG22	2.11	0.50
1:C:36:ARG:HB2	1:C:49:VAL:HG11	1.93	0.50
1:D:415:CYS:O	1:D:422:GLY:N	2.38	0.50
1:C:425:LEU:HG	1:C:426:PRO:HD3	1.94	0.49
1:D:252:LYS:HA	1:D:256:CYS:SG	2.53	0.49
1:B:46:LEU:O	1:B:49:VAL:HG12	2.12	0.49
1:D:430:GLY:O	1:D:436:GLN:NE2	2.45	0.49
1:A:252:LYS:HA	1:A:256:CYS:SG	2.52	0.49
1:C:173:ARG:O	1:C:177:GLN:HG2	2.13	0.49
1:D:298:VAL:O	1:D:301:THR:OG1	2.30	0.49
1:A:158:HIS:HE2	1:A:214:THR:HG23	1.77	0.49
1:D:175:PHE:HZ	1:D:336:THR:HG1	1.58	0.48
1:C:233:VAL:HG13	1:C:456:GLN:HB3	1.94	0.48
1:B:51:GLN:H	1:B:51:GLN:CD	2.16	0.48
1:D:172:GLU:O	1:D:176:LYS:HG3	2.14	0.48
1:D:291:ARG:HG2	1:D:390:GLN:HG3	1.96	0.48
1:D:169:ARG:HD3	1:D:169:ARG:C	2.35	0.47
1:A:210:ARG:NH1	4:A:711:HOH:O	2.40	0.47
1:D:222:SER:HB3	1:D:467:LEU:HD23	1.97	0.47
1:C:132:GLU:HA	1:C:135:THR:HB	1.97	0.47
1:D:311:GLU:HB2	1:D:375:LEU:HD22	1.96	0.46
1:D:304:PHE:HB3	1:D:375:LEU:HD21	1.97	0.46
1:A:158:HIS:CE1	1:A:160:GLU:H	2.34	0.46
1:C:383:LYS:HB3	1:C:387:ARG:HH12	1.79	0.46
1:A:42:LYS:O	1:A:80:ARG:NH2	2.38	0.46
1:D:154:GLY:HA2	1:D:221:ARG:HD2	1.98	0.46
1:C:383:LYS:HB3	1:C:387:ARG:NH1	2.30	0.46
1:D:38:ILE:O	1:D:42:LYS:HG3	2.15	0.45
1:B:252:LYS:HA	1:B:256:CYS:SG	2.56	0.45
1:A:36:ARG:HA	1:A:49:VAL:HG21	1.98	0.45
1:B:199:LEU:HD13	1:B:335:LEU:HD12	1.98	0.45
1:D:132:GLU:HA	1:D:135:THR:OG1	2.17	0.45
1:A:268[B]:CYS:SG	1:A:424:TYR:HB3	2.57	0.44
1:B:174:LEU:HD23	1:B:328:LEU:HD23	1.99	0.44
1:B:347:LYS:HG2	1:B:348:VAL:H	1.81	0.44
1:A:183:LEU:HD12	1:A:183:LEU:HA	1.81	0.44
1:C:387:ARG:NH1	1:D:376:GLU:OE1	2.50	0.44
1:D:295:ASP:OD1	1:D:387:ARG:NH1	2.49	0.44
1:A:45:SER:HB2	1:A:48:ASP:OD2	2.17	0.44
1:B:199:LEU:HD23	1:B:199:LEU:HA	1.76	0.44

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:202:PHE:HB3	1:D:324:ALA:HB1	2.00	0.43
1:D:216:ALA:HB1	1:D:310:VAL:HA	1.99	0.43
1:A:133:LEU:HD22	1:A:174:LEU:HD22	1.99	0.43
1:B:179:HIS:CE1	1:B:332:ARG:HH11	2.36	0.43
1:A:59:LEU:HD12	1:A:64:GLN:HA	2.00	0.43
1:B:297:MET:O	1:B:300:ILE:HG22	2.18	0.43
1:B:42:LYS:HE3	1:B:255:TYR:CE1	2.53	0.43
1:B:405:MET:H	1:B:405:MET:HG3	1.40	0.43
1:D:195:GLN:OE1	1:D:339:VAL:HA	2.18	0.43
1:C:97:GLN:NE2	1:C:235:ARG:HA	2.34	0.43
1:D:198:ALA:HA	1:D:200:ARG:HH11	1.84	0.43
1:A:336:THR:O	1:A:336:THR:CG2	2.67	0.43
1:B:129:ALA:HA	1:B:362:GLY:HA2	2.00	0.43
1:C:310:VAL:HG12	1:C:311:GLU:H	1.84	0.43
1:D:403:GLU:OE1	1:D:404:LYS:HB2	2.19	0.43
1:A:454:ARG:HD3	1:A:454:ARG:HA	1.81	0.42
1:B:452:THR:O	1:B:456:GLN:HG2	2.19	0.42
1:D:159:LEU:N	1:D:160:GLU:OE1	2.52	0.42
1:C:432:GLY:O	1:C:436:GLN:HG2	2.19	0.42
1:D:214:THR:O	1:D:218:VAL:HB	2.20	0.42
1:A:235:ARG:NH2	1:A:236:LYS:HE3	2.35	0.42
1:C:42:LYS:HE2	1:C:255:TYR:OH	2.19	0.42
1:C:175:PHE:O	1:C:179:HIS:CD2	2.73	0.42
1:C:39:TYR:CZ	1:C:73:MET:HG2	2.55	0.42
1:B:183:LEU:HD13	1:B:347:LYS:HD3	2.01	0.41
1:D:44:PHE:CE1	1:D:76:ASN:HB3	2.55	0.41
1:A:126:PHE:N	1:A:127:PRO:HD2	2.35	0.41
1:D:133:LEU:HA	1:D:177:GLN:OE1	2.20	0.41
1:D:316:SER:O	1:D:319:THR:OG1	2.23	0.41
1:B:414:ARG:HD3	1:B:414:ARG:HA	1.69	0.41
1:D:373:GLY:O	1:D:377:LYS:HG2	2.20	0.41
1:D:56:GLY:HA3	1:D:65:GLY:O	2.20	0.41
1:B:215:ARG:O	1:B:218:VAL:HG12	2.20	0.41
1:D:199:LEU:HD12	1:D:199:LEU:HA	1.85	0.41
1:A:277:LYS:NZ	1:A:429:MET:O	2.39	0.41
1:B:44:PHE:CE1	1:B:76:ASN:HB3	2.55	0.41
1:B:159:LEU:HD12	1:B:159:LEU:HA	1.91	0.41
1:B:372:SER:HB3	4:B:705:HOH:O	2.21	0.41
1:C:156:ASN:ND2	1:C:221:ARG:HH12	2.19	0.41
1:A:151:TYR:CE1	1:A:158:HIS:HA	2.57	0.40
1:A:50:PRO:HD3	1:A:70:THR:HG23	2.04	0.40

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Interatomic distance (Å)	Clash overlap (Å)
1:D:188:TYR:OH	1:D:344:GLY:O	2.30	0.40
1:A:132:GLU:HA	1:A:135:THR:OG1	2.21	0.40
1:C:445:ASP:OD2	1:C:447:THR:OG1	2.25	0.40
1:D:396:LEU:HB2	1:D:397:PRO:HD3	2.04	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	407/526 (77%)	401 (98%)	4 (1%)	2 (0%)	29	39
1	B	428/526 (81%)	412 (96%)	16 (4%)	0	100	100
1	C	394/526 (75%)	385 (98%)	9 (2%)	0	100	100
1	D	427/526 (81%)	418 (98%)	9 (2%)	0	100	100
All	All	1656/2104 (79%)	1616 (98%)	38 (2%)	2 (0%)	51	67

All (2) Ramachandran outliers are listed below:

Mol	Chain	Res	Type
1	A	336	THR
1	A	337	ALA

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	332/439 (76%)	331 (100%)	1 (0%)	92	97
1	B	341/439 (78%)	336 (98%)	5 (2%)	65	79
1	C	317/439 (72%)	315 (99%)	2 (1%)	86	93
1	D	336/439 (76%)	331 (98%)	5 (2%)	65	79
All	All	1326/1756 (76%)	1313 (99%)	13 (1%)	78	87

All (13) residues with a non-rotameric sidechain are listed below:

Mol	Chain	Res	Type
1	A	57	GLU
1	B	153	ARG
1	B	160	GLU
1	B	364	LEU
1	B	405	MET
1	B	474	ASN
1	C	94	ARG
1	C	307	THR
1	D	103	GLN
1	D	136	GLN
1	D	165	GLU
1	D	381[A]	GLU
1	D	381[B]	GLU

Sometimes sidechains can be flipped to improve hydrogen bonding and reduce clashes. All (1) such sidechains are listed below:

Mol	Chain	Res	Type
1	C	181	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)

Of 13 ligands modelled in this entry, 9 are monoatomic - leaving 4 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	$\#\|Z\| > 2$
2	NAG	C	601	1	14,14,15	0.34	0	17,19,21	0.54	0
2	NAG	B	601	1	14,14,15	0.31	0	17,19,21	0.38	0
2	NAG	A	601	1	14,14,15	0.40	0	17,19,21	0.50	0
2	NAG	D	601	1	14,14,15	0.24	0	17,19,21	0.55	0

In the following table, the Chirals column lists the number of chiral outliers, the number of chiral centers analysed, the number of these observed in the model and the number defined in the Chemical Component Dictionary. Similar counts are reported in the Torsion and Rings columns. '-' means no outliers of that kind were identified.

Mol	Type	Chain	Res	Link	Chirals	Torsions	Rings
2	NAG	C	601	1	-	4/6/23/26	0/1/1/1
2	NAG	B	601	1	-	4/6/23/26	0/1/1/1
2	NAG	A	601	1	-	2/6/23/26	0/1/1/1
2	NAG	D	601	1	-	3/6/23/26	0/1/1/1

There are no bond length outliers.

There are no bond angle outliers.

There are no chirality outliers.

All (13) torsion outliers are listed below:

Mol	Chain	Res	Type	Atoms
2	C	601	NAG	O5-C5-C6-O6
2	B	601	NAG	O5-C5-C6-O6
2	A	601	NAG	C8-C7-N2-C2
2	A	601	NAG	O7-C7-N2-C2
2	B	601	NAG	C8-C7-N2-C2
2	B	601	NAG	O7-C7-N2-C2
2	C	601	NAG	C8-C7-N2-C2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Atoms
2	C	601	NAG	O7-C7-N2-C2
2	D	601	NAG	C8-C7-N2-C2
2	D	601	NAG	O7-C7-N2-C2
2	B	601	NAG	C4-C5-C6-O6
2	C	601	NAG	C4-C5-C6-O6
2	D	601	NAG	O5-C5-C6-O6

There are no ring outliers.

No monomer is involved in short contacts.

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	411/526 (78%)	1.01	58 (14%)	2	3	25, 48, 96, 119	0
1	B	431/526 (81%)	0.93	47 (10%)	5	6	30, 53, 98, 134	0
1	C	400/526 (76%)	0.89	33 (8%)	11	12	40, 58, 107, 139	0
1	D	429/526 (81%)	0.99	59 (13%)	2	3	37, 60, 104, 135	0
All	All	1671/2104 (79%)	0.96	197 (11%)	4	5	25, 56, 102, 139	0

All (197) RSRZ outliers are listed below:

Mol	Chain	Res	Type	RSRZ
1	C	180	PRO	9.3
1	A	196	ALA	8.3
1	B	156	ASN	8.1
1	C	337	ALA	8.1
1	A	184	LEU	6.2
1	D	174	LEU	6.0
1	B	155	ALA	5.7
1	C	335	LEU	5.5
1	B	407	LEU	5.4
1	C	177	GLN	5.3
1	A	44	PHE	5.1
1	C	158	HIS	4.9
1	C	184	LEU	4.6
1	A	199	LEU	4.6
1	A	25	PRO	4.5
1	D	366	PRO	4.5
1	D	405	MET	4.4
1	A	45	SER	4.4
1	D	335	LEU	4.3
1	D	184	LEU	4.2
1	C	175	PHE	4.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	366	PRO	4.2
1	B	413	ASP	4.1
1	D	132	GLU	4.1
1	C	334	THR	4.1
1	A	366	PRO	4.1
1	D	370	PRO	4.1
1	A	262	VAL	4.0
1	C	199	LEU	4.0
1	A	26	ALA	4.0
1	A	261	GLY	4.0
1	D	170	LEU	4.0
1	A	158	HIS	4.0
1	A	156	ASN	4.0
1	C	179	HIS	3.9
1	D	372	SER	3.9
1	B	50	PRO	3.9
1	A	71	SER	3.8
1	D	156	ASN	3.8
1	A	198	ALA	3.8
1	A	335	LEU	3.8
1	A	181	GLN	3.7
1	B	371	PRO	3.7
1	B	361	ARG	3.6
1	D	155	ALA	3.6
1	C	157	LEU	3.6
1	B	211	LEU	3.6
1	C	185	PRO	3.5
1	D	182	LEU	3.5
1	A	185	PRO	3.5
1	D	189	LEU	3.4
1	B	402	SER	3.4
1	C	423	ARG	3.4
1	C	414	ARG	3.4
1	B	41	ALA	3.3
1	B	44	PHE	3.3
1	D	413	ASP	3.3
1	C	207[A]	ARG	3.3
1	B	45	SER	3.3
1	C	336	THR	3.3
1	B	77	LEU	3.3
1	B	369	ARG	3.3
1	C	474	ASN	3.2

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	D	175	PHE	3.2
1	B	370	PRO	3.2
1	D	403	GLU	3.2
1	D	361	ARG	3.2
1	A	50	PRO	3.2
1	C	328	LEU	3.2
1	A	265	ALA	3.1
1	B	405	MET	3.1
1	B	373	GLY	3.1
1	A	186	ASP	3.1
1	B	183	LEU	3.1
1	D	268[A]	CYS	3.1
1	A	195	GLN	3.0
1	B	37	GLN	3.0
1	D	195	GLN	3.0
1	B	154	GLY	3.0
1	D	371	PRO	3.0
1	D	89	LEU	2.9
1	D	157	LEU	2.9
1	B	38	ILE	2.9
1	C	182	LEU	2.9
1	A	372	SER	2.9
1	A	46	LEU	2.9
1	C	171	LEU	2.9
1	A	159	LEU	2.8
1	C	333	ASP	2.8
1	D	169	ARG	2.8
1	B	173	ARG	2.8
1	C	181	GLN	2.8
1	A	27	SER	2.8
1	B	368	GLU	2.8
1	B	48	ASP	2.7
1	B	51	GLN	2.7
1	D	402	SER	2.7
1	A	72	GLU	2.7
1	C	132	GLU	2.7
1	A	73	MET	2.7
1	B	332	ARG	2.7
1	A	178	LEU	2.7
1	C	178	LEU	2.7
1	C	298	VAL	2.7
1	C	427	GLU	2.7

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	372	SER	2.7
1	D	200	ARG	2.6
1	D	473	GLY	2.6
1	D	336	THR	2.6
1	B	114	LEU	2.6
1	C	380	SER	2.6
1	D	348	VAL	2.6
1	D	265	ALA	2.6
1	D	48	ASP	2.6
1	D	181	GLN	2.5
1	B	47	SER	2.5
1	A	67	THR	2.5
1	D	428	VAL	2.5
1	B	198	ALA	2.5
1	A	339	VAL	2.5
1	A	42	LYS	2.5
1	A	177	GLN	2.5
1	D	199	LEU	2.5
1	A	183	LEU	2.5
1	A	260	LEU	2.5
1	D	369	ARG	2.5
1	B	132	GLU	2.5
1	B	46	LEU	2.5
1	D	364	LEU	2.5
1	B	66	TYR	2.4
1	A	37	GLN	2.4
1	A	77	LEU	2.4
1	C	422	GLY	2.4
1	A	51	GLN	2.4
1	D	52	ALA	2.4
1	D	415	CYS	2.4
1	A	473	GLY	2.4
1	C	141	PHE	2.4
1	D	44	PHE	2.4
1	B	199	LEU	2.4
1	A	57	GLU	2.4
1	D	161	GLU	2.4
1	A	60	ARG	2.3
1	C	238	ALA	2.3
1	B	142	ARG	2.3
1	D	401	CYS	2.3
1	D	70	THR	2.3

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	A	28	LYS	2.3
1	D	262	VAL	2.3
1	A	36	ARG	2.3
1	B	331	ASN	2.3
1	A	220	ALA	2.3
1	D	406	ALA	2.3
1	B	153	ARG	2.2
1	B	159	LEU	2.2
1	A	69	CYS	2.2
1	B	49	VAL	2.2
1	A	175	PHE	2.2
1	C	332	ARG	2.2
1	D	380	SER	2.2
1	D	159	LEU	2.2
1	A	257	ALA	2.2
1	D	114	LEU	2.2
1	D	139	ARG	2.2
1	D	178	LEU	2.2
1	D	425	LEU	2.2
1	A	245	GLU	2.2
1	C	383	LYS	2.2
1	A	70	THR	2.2
1	D	142	ARG	2.2
1	A	330	ASP	2.1
1	B	425	LEU	2.1
1	A	255	TYR	2.1
1	A	380	SER	2.1
1	D	45	SER	2.1
1	A	368	GLU	2.1
1	D	165	GLU	2.1
1	D	243	GLY	2.1
1	A	182	LEU	2.1
1	D	211	LEU	2.1
1	A	33	GLY	2.1
1	B	180	PRO	2.1
1	B	197	GLU	2.1
1	D	153	ARG	2.1
1	B	374	THR	2.1
1	D	196	ALA	2.1
1	A	40	GLY	2.1
1	A	369	ARG	2.0
1	C	205	ALA	2.0

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	RSRZ
1	B	57	GLU	2.0
1	A	248	ARG	2.0
1	A	425	LEU	2.0
1	D	29	SER	2.0
1	D	47	SER	2.0
1	B	367	ARG	2.0
1	D	294	LEU	2.0
1	B	315	GLY	2.0

6.2 Non-standard residues in protein, DNA, RNA chains (i)

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

6.3 Carbohydrates (i)

There are no 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
3	CA	A	603	1/1	0.27	0.15	90,90,90,90	0
3	CA	B	603	1/1	0.28	0.10	43,43,43,43	1
3	CA	C	603	1/1	0.34	0.27	99,99,99,99	0
3	CA	D	602	1/1	0.39	0.11	93,93,93,93	0
3	CA	C	604	1/1	0.43	0.18	60,60,60,60	0
3	CA	A	602	1/1	0.70	0.15	47,47,47,47	0
3	CA	C	602	1/1	0.72	0.14	71,71,71,71	0
2	NAG	D	601	14/15	0.86	0.20	80,89,95,96	0
2	NAG	C	601	14/15	0.86	0.20	79,84,89,90	0
3	CA	D	603	1/1	0.86	0.11	82,82,82,82	0
2	NAG	B	601	14/15	0.87	0.19	65,72,80,84	0
2	NAG	A	601	14/15	0.90	0.16	56,68,78,81	0
3	CA	B	602	1/1	0.91	0.06	87,87,87,87	0

6.5 Other polymers (i)

There are no such residues in this entry.