WORLDWIDE PDB
PROTEIN DATA BANK

Full wwPDB NMR Structure Validation Report (i)

May 28, 2020 – 10:50 pm BST

PDB ID	:	2L2A
Title	:	Mutated Domain 11 of the Cytoplasmic region of the Cation-independent mannose-6-phosphate receptor
Authors	:	Williams, C.; Hoppe, H.; Rezgui, D.; Strickland, M.; Frago, S.; Ellis, R.Z.; Wattana-Amorn, P.; Prince, S.N.; Zaccheo, O.J.; Forbes, B.; Jones, E.Y.; Crump, M.P.; Hassan, A.B.
Deposited on	:	2010-08-13

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

|		
Cyrange	:	Kirchner and Güntert (2011)
NmrClust	:	Kelley et al. (1996)
MolProbity	:	4.02b-467
Percentile statistics	:	20191225.v01 (using entries in the PDB archive December 25th 2019)
RCI	:	v_1n_11_5_13_A (Berjanski et al., 2005)
PANAV	:	Wang et al. (2010)
ShiftChecker	:	2.11
Ideal geometry (proteins)	:	Engh & Huber (2001)
Ideal geometry (DNA, RNA)	:	Parkinson et al. (1996)
Validation Pipeline (wwPDB-VP)	:	2.11

1 Overall quality at a glance (i)

The following experimental techniques were used to determine the structure:
SOLUTION NMR

The overall completeness of chemical shifts assignment was not calculated.

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.
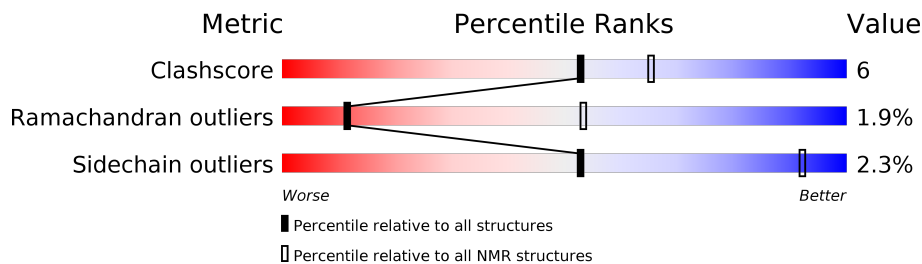
Metric	Whole archive (#Entries)	NMR archive (#Entries)
Clashscore	158937	12864
Ramachandran outliers	154571	11451
Sidechain outliers	154315	11428

The table below summarises the geometric issues observed across the polymeric chains and their fit to the experimental data. The red, orange, yellow and green segments indicate the fraction of residues that contain outliers for >=3, 2, 1 and 0 types of geometric quality criteria. A cyan segment indicates the fraction of residues that are not part of the well-defined cores, and 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%

Mol	Chain	Length	Quality of chain
1	A	142	80% 11% 9%

2 Ensemble composition and analysis (i)

This entry contains 20 models. Model 10 is the overall representative, medoid model (most similar to other models). The authors have identified model 1 as representative, based on the following criterion: *lowest energy*.

The following residues are included in the computation of the global validation metrics.

Well-defined (core) protein residues			
Well-defined core	Residue range (total)	Backbone RMSD (Å)	Medoid model
1	A:1515-A:1617, A:1622-A:1647 (129)	0.67	10

Ill-defined regions of proteins are excluded from the global statistics.

Ligands and non-protein polymers are included in the analysis.

The models can be grouped into 4 clusters and 1 single-model cluster was found.

Cluster number	Models
1	2, 5, 6, 7, 9, 10, 15, 16, 18, 19
2	3, 8, 11, 13, 20
3	4, 14
4	1, 12
Single-model clusters	17

3 Entry composition (i)

There is only 1 type of molecule in this entry. The entry contains 2126 atoms, of which 1049 are hydrogens and 0 are deuteriums.

- Molecule 1 is a protein called Insulin-like growth factor 2 receptor variant.

Mol	Chain	Residues	Atoms						Trace
			Total	C	H	N	O	S	
1	A	142	2126	672	1049	186	208	11	0

There are 5 discrepancies between the modelled and reference sequences:

Chain	Residue	Modelled	Actual	Comment	Reference
A	1544	LYS	GLU	ENGINEERED MUTATION	UNP Q59EZ3
A	1545	SER	LYS	ENGINEERED MUTATION	UNP Q59EZ3
A	1547	VAL	LEU	ENGINEERED MUTATION	UNP Q59EZ3
A	1648	PRO	-	EXPRESSION TAG	UNP Q59EZ3
A	1649	GLU	-	EXPRESSION TAG	UNP Q59EZ3

4 Residue-property plots (i)

4.1 Average score per residue in the NMR ensemble

These plots are provided for all protein, RNA and DNA chains in the entry. The first graphic is the same as shown in the summary in section 1 of this report. The second graphic shows the sequence where residues are colour-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. Stretches of 2 or more consecutive residues without any outliers are shown as green connectors. Residues which are classified as ill-defined in the NMR ensemble, are shown in cyan with an underline colour-coded according to the previous scheme. Residues which were present in the experimental sample, but not modelled in the final structure are shown in grey.

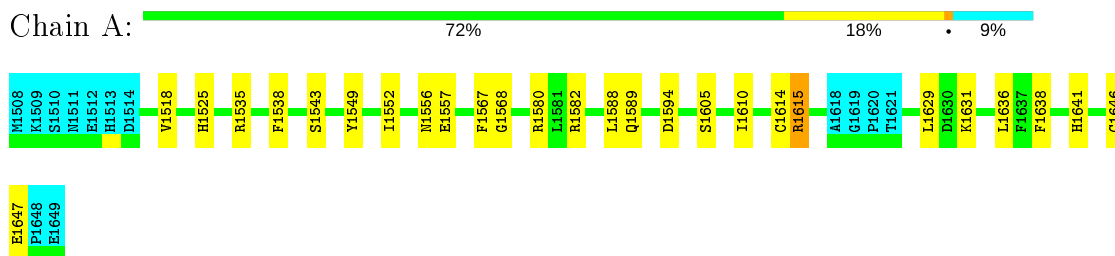• Molecule 1: Insulin-like growth factor 2 receptor variant
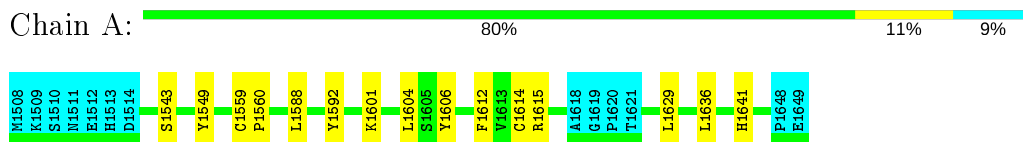
4.2 Scores per residue for each member of the ensemble
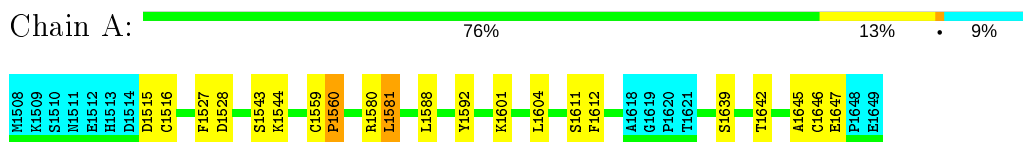
Colouring as in section 4.1 above.

4.2.1 Score per residue for model 1

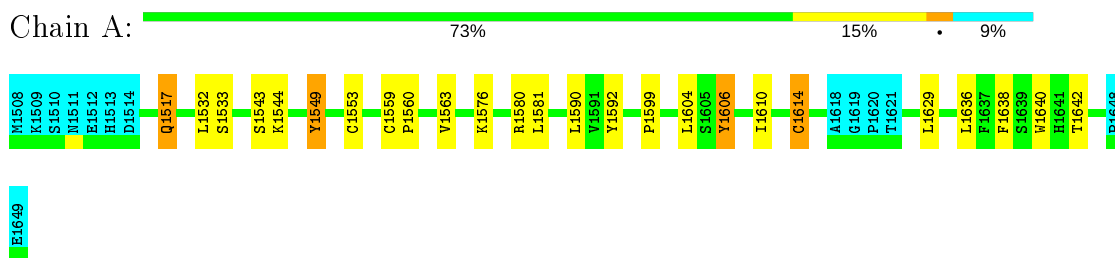• Molecule 1: Insulin-like growth factor 2 receptor variant

4.2.2 Score per residue for model 2

• Molecule 1: Insulin-like growth factor 2 receptor variant

4.2.3 Score per residue for model 3

• Molecule 1: Insulin-like growth factor 2 receptor variant

Chain A:

73%   15% · 9%

M1508 K1509 S1510 N1511 E1512 H1513 D1514 Q1517 L1532 S1533 S1543 K1544 Y1549 C1553 C1559 P1560 V1563 K1576 R1580 L1581 L1590 V1591 Y1592 P1599 L1604 S1605 Y1606 I1610 C1614 A1618 G1619 P1620 T1621 L1629 L1636 F1637 F1638 S1639 W1640 H1641 T1642 P1648

E1649

4.2.4 Score per residue for model 4

• Molecule 1: Insulin-like growth factor 2 receptor variant

Chain A:

75%   15% · 9%

M1508 K1509 S1510 N1511 E1512 H1513 D1514 L1529 L1532 A1540 S1543 C1559 P1560 F1567 G1568 Q1569 T1570 R1571 K1601 L1604 S1608 V1613 C1614 R1615 A1618 G1619 P1620 T1621 N1622 R1623 P1624 L1629 Q1632 L1636 W1640 H1641 T1642 P1648 E1649

4.2.5 Score per residue for model 5

• Molecule 1: Insulin-like growth factor 2 receptor variant

Chain A:

75%   14% ·· 9%

M1508 K1509 S1510 N1511 E1512 H1513 D1514 V1518 A1540 S1543 V1548 Y1549 C1559 P1560 G1568 Q1569 T1570 L1588 K1601 L1604 I1610 S1611 F1612 V1613 C1614 A1618 G1619 P1620 T1621 L1629 L1636 F1637 F1638 H1641 T1642 P1643 L1644 P1648 E1649

4.2.6 Score per residue for model 6

• Molecule 1: Insulin-like growth factor 2 receptor variant

Chain A:

77%   12% · 9%

M1508 K1509 S1510 N1511 E1512 H1513 D1514 D1515 T1519 L1526 S1533 S1543 K1544 I1552 C1559 P1560 L1588 Y1592 P1599 F1612 V1613 C1614 R1615 P1616 E1617 A1618 G1619 P1620 T1621 L1629 L1636 T1642 P1643 P1648 E1649

4.2.7 Score per residue for model 7

• Molecule 1: Insulin-like growth factor 2 receptor variant

Chain A:

80%   8% ·· 9%

4.2.8 Score per residue for model 8

- Molecule 1: Insulin-like growth factor 2 receptor variant

4.2.9 Score per residue for model 9

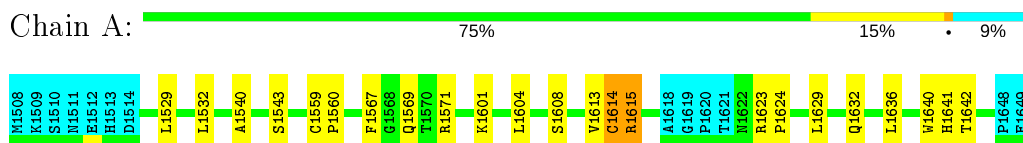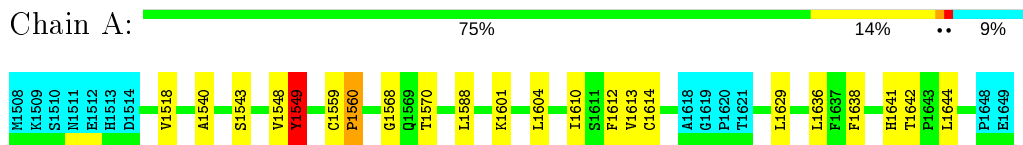- Molecule 1: Insulin-like growth factor 2 receptor variant

4.2.10 Score per residue for model 10 (medoid)

- Molecule 1: Insulin-like growth factor 2 receptor variant

4.2.11 Score per residue for model 11

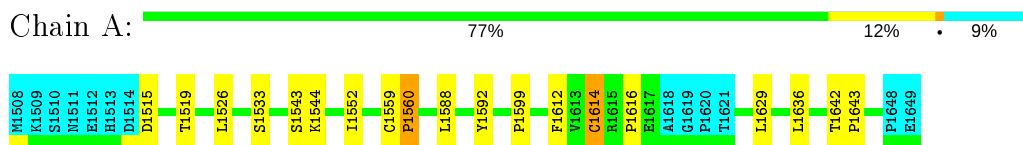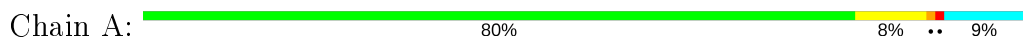- Molecule 1: Insulin-like growth factor 2 receptor variant

4.2.12 Score per residue for model 12

• Molecule 1: Insulin-like growth factor 2 receptor variant
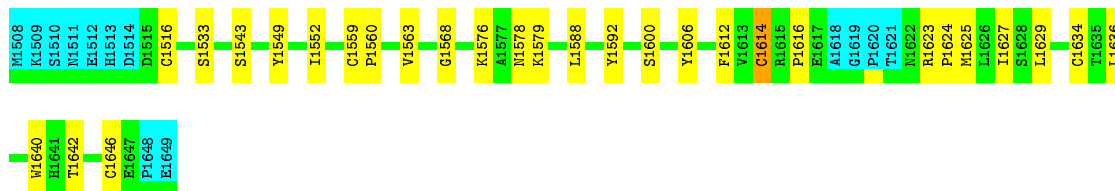
Chain A: 73% 15% 9%

4.2.13 Score per residue for model 13

• Molecule 1: Insulin-like growth factor 2 receptor variant

Chain A: 73% 15% 9%

4.2.14 Score per residue for model 14

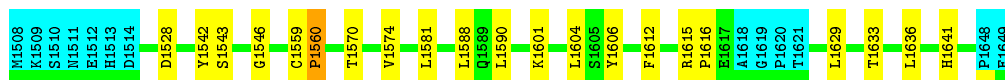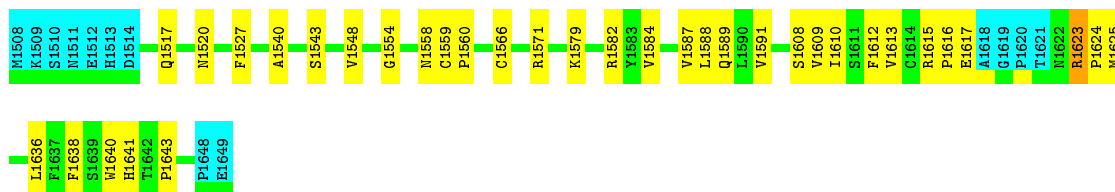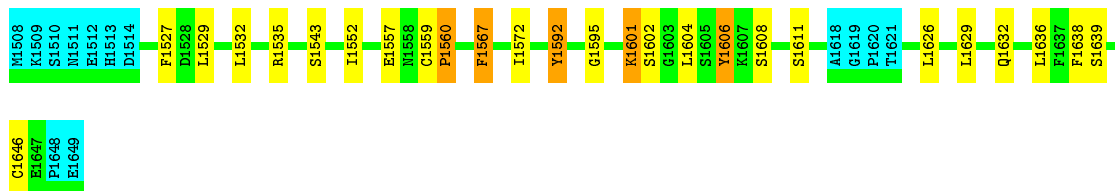• Molecule 1: Insulin-like growth factor 2 receptor variant

Chain A: 74% 16% 9%

4.2.15 Score per residue for model 15

• Molecule 1: Insulin-like growth factor 2 receptor variant

Chain A: 73% 16% 9%

T1642 P1648 E1649

4.2.16 Score per residue for model 16

• Molecule 1: Insulin-like growth factor 2 receptor variant

Chain A: 70%   20%   9%

M1508 K1509 S1510 N1511 E1512 H1513 D1514 T1539 S1543 V1548 Y1549 P1560 G1568 Q1569 T1570 K1576 A1577 N1578 V1584 V1587 L1588 Y1592 G1595 C1598 K1601 L1604 S1605 Y1606 S1611 F1612 V1613 C1614 R1615 P1616 E1617 A1618 G1619 P1620 T1621 N1622 R1623 P1624 L1629 L1636

S1639 W1640 H1641 T1642 P1648 E1649
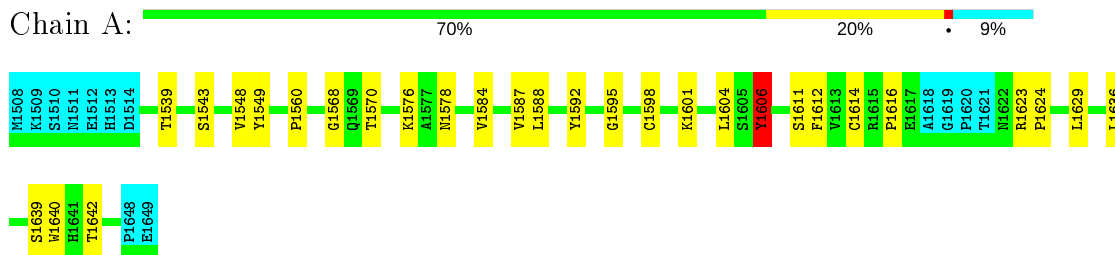
4.2.17 Score per residue for model 17

• Molecule 1: Insulin-like growth factor 2 receptor variant

Chain A: 75%   15%   9%

M1508 K1509 S1510 N1511 E1512 H1513 D1514 D1515 C1516 Q1517 V1518 F1527 S1543 Y1549 E1555 N1556 G1568 Q1569 T1570 V1574 K1579 L1588 G1595 S1596 K1601 L1604 S1605 Y1606 F1612 R1615 A1618 G1619 P1620 T1621 L1626 L1629 C1634 P1648 E1649
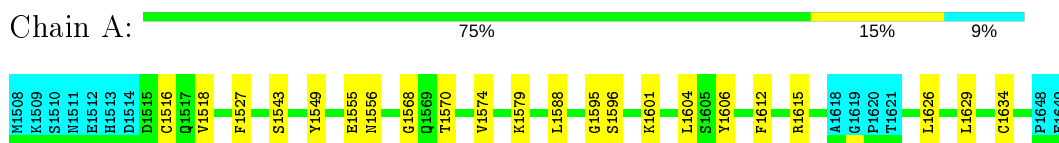
4.2.18 Score per residue for model 18

• Molecule 1: Insulin-like growth factor 2 receptor variant

Chain A: 77%   13%   9%

M1508 K1509 S1510 N1511 E1512 H1513 D1514 T1519 L1526 S1543 C1559 P1560 G1564 T1570 V1574 G1575 Y1592 Y1606 I1610 A1618 G1619 P1620 T1621 L1629 C1634 T1635 L1636 F1637 F1638 T1642 P1643 L1644 E1647 P1648 E1649
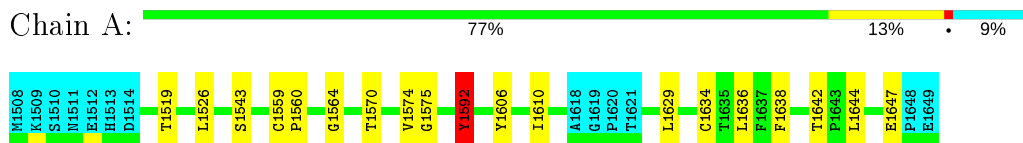
4.2.19 Score per residue for model 19

• Molecule 1: Insulin-like growth factor 2 receptor variant

Chain A: 69%   22%   9%

M1508 K1509 S1510 N1511 E1512 H1513 D1514 Q1517 F1527 A1536 G1537 F1538 T1539 Y1542 S1543 K1544 S1545 V1548 Y1549 N1556 E1557 N1558 C1559 P1560 V1563 G1568 Q1569 T1570 K1576 L1581 K1593 D1594 G1595 K1601 Y1606 I1610 C1614 A1618 G1619 P1620 T1621 M1625

L1629 Q1632 T1633 C1634 T1635 L1636 F1637 F1638 P1648 E1649
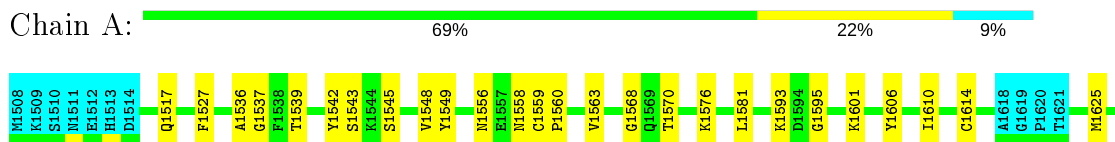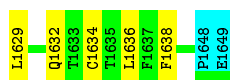
4.2.20 Score per residue for model 20

• Molecule 1: Insulin-like growth factor 2 receptor variant

Chain A:

76% 14% • 9%

M1508 K1509 S1510 N1511 E1512 H1513 D1514 D1515 C1516 Q1517 V1518 T1519 L1526 L1529 L1532 S1543 S1551 N1556 P1560 T1570 A1577 N1578 V1587 Y1592 L1604 I1610 S1611 F1612 A1618 G1619 P1620 T1621 F1638 T1642 P1643 L1644 P1648 E1649
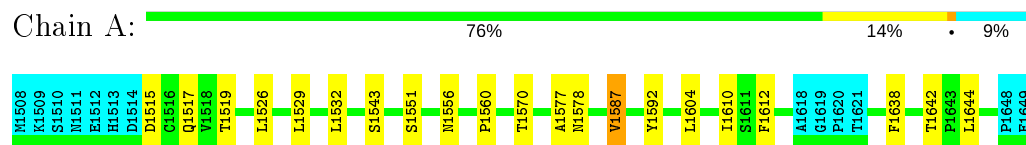
5 Refinement protocol and experimental data overview (i)

The models were refined using the following method: *simulated annealing, simulated annealing*.

Of the 200 calculated structures, 20 were deposited, based on the following criterion: *structures with the lowest energy*.

The following table shows the software used for structure solution, optimisation and refinement.

Software name	Classification	Version
CNS	refinement	
CNS	structure solution	
TALOS	geometry optimization	
iCing	refinement	r765
ARIA	refinement	2.2
ARIA	structure solution	2.2

No chemical shift data was provided. No validations of the models with respect to experimental NMR restraints is performed at this time.

COVALENT-GEOMETRY INFOmissingINFO

5.1 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 each 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 averaged over the ensemble.

Mol	Chain	Non-H	H(model)	H(added)	Clashes
1	A	979	961	959	11±3
All	All	19580	19220	19180	216

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 6.

All unique clashes are listed below, sorted by their clash magnitude.

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models Worst	Models Total
1:A:1574:VAL:HA	1:A:1596:SER:HB3	0.80	1.51	17	1
1:A:1629:LEU:HG	1:A:1636:LEU:HG	0.79	1.54	15	8
1:A:1629:LEU:HD13	1:A:1636:LEU:HG	0.76	1.56	5	4
1:A:1601:LYS:HB3	1:A:1604:LEU:HD12	0.73	1.60	16	6
1:A:1533:SER:HB2	1:A:1552:ILE:HG22	0.72	1.62	6	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:1588:LEU:HB2	1:A:1612:PHE:HB2	0.68	1.64	7	12
1:A:1554:GLY:HA2	1:A:1579:LYS:HE2	0.68	1.65	10	1
1:A:1615:ARG:HB3	1:A:1641:HIS:HB3	0.67	1.64	10	3
1:A:1559:CYS:HB3	1:A:1560:PRO:HD2	0.64	1.70	3	13
1:A:1535:ARG:HG3	1:A:1557:GLU:HG3	0.64	1.68	12	1
1:A:1601:LYS:HD2	1:A:1632:GLN:HA	0.63	1.71	12	1
1:A:1601:LYS:HG2	1:A:1604:LEU:HD12	0.63	1.70	15	1
1:A:1610:ILE:HG13	1:A:1638:PHE:HB2	0.62	1.71	3	8
1:A:1535:ARG:HD3	1:A:1557:GLU:HG3	0.62	1.70	1	1
1:A:1542:TYR:HB3	1:A:1546:GLY:HA3	0.61	1.72	9	1
1:A:1602:SER:HB3	1:A:1604:LEU:HD22	0.61	1.73	12	1
1:A:1553:CYS:SG	1:A:1579:LYS:HE3	0.60	2.36	15	1
1:A:1615:ARG:HB2	1:A:1641:HIS:HB3	0.60	1.72	15	1
1:A:1518:VAL:HG21	1:A:1588:LEU:HD11	0.59	1.74	1	1
1:A:1529:LEU:HA	1:A:1532:LEU:HD13	0.58	1.75	4	2
1:A:1629:LEU:HD13	1:A:1636:LEU:HD13	0.58	1.76	4	1
1:A:1615:ARG:HB2	1:A:1641:HIS:HB2	0.58	1.74	14	1
1:A:1575:GLY:HA3	1:A:1592:TYR:HB3	0.57	1.75	18	1
1:A:1582:ARG:HB2	1:A:1589:GLN:HB3	0.57	1.74	10	1
1:A:1628:SER:HB2	1:A:1637:PHE:HB2	0.57	1.75	14	1
1:A:1540:ALA:HB3	1:A:1548:VAL:HG13	0.57	1.77	10	2
1:A:1519:THR:HA	1:A:1526:LEU:HA	0.57	1.77	6	3
1:A:1533:SER:HB2	1:A:1552:ILE:HB	0.56	1.76	8	1
1:A:1608:SER:HA	1:A:1636:LEU:HB2	0.56	1.77	10	2
1:A:1567:PHE:O	1:A:1571:ARG:HA	0.56	2.01	4	1
1:A:1577:ALA:HA	1:A:1592:TYR:CE2	0.56	2.36	15	1
1:A:1516:CYS:SG	1:A:1579:LYS:HE3	0.55	2.40	15	2
1:A:1614:CYS:HA	1:A:1642:THR:O	0.55	2.01	8	6
1:A:1615:ARG:HG3	1:A:1617:GLU:HG2	0.55	1.79	10	1
1:A:1642:THR:HG22	1:A:1644:LEU:HG	0.54	1.79	5	1
1:A:1529:LEU:HG	1:A:1532:LEU:HD12	0.54	1.79	12	1
1:A:1613:VAL:O	1:A:1641:HIS:HA	0.54	2.02	5	5
1:A:1642:THR:OG1	1:A:1644:LEU:HG	0.54	2.03	18	2
1:A:1642:THR:HG23	1:A:1645:ALA:HB2	0.54	1.78	2	1
1:A:1615:ARG:HB3	1:A:1641:HIS:HB2	0.54	1.78	13	2
1:A:1549:TYR:CD1	1:A:1568:GLY:HA3	0.53	2.39	1	1
1:A:1598:CYS:O	1:A:1602:SER:HA	0.53	2.04	11	1
1:A:1559:CYS:SG	1:A:1560:PRO:HD2	0.52	2.44	5	1
1:A:1516:CYS:HA	1:A:1581:LEU:HD23	0.52	1.81	2	1
1:A:1527:PHE:CE1	1:A:1646:CYS:HA	0.51	2.40	12	2
1:A:1551:SER:CB	1:A:1556:ASN:HB3	0.51	2.34	13	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:1584:VAL:O	1:A:1587:VAL:HG12	0.51	2.06	16	2
1:A:1590:LEU:HB2	1:A:1612:PHE:HE1	0.51	1.66	15	1
1:A:1560:PRO:O	1:A:1563:VAL:HG22	0.51	2.05	19	1
1:A:1598:CYS:HA	1:A:1606:TYR:HE1	0.51	1.66	16	1
1:A:1529:LEU:HA	1:A:1532:LEU:HD23	0.51	1.81	14	1
1:A:1538:PHE:CD1	1:A:1552:ILE:HD13	0.50	2.41	1	1
1:A:1564:GLY:HA3	1:A:1592:TYR:CE2	0.50	2.42	18	1
1:A:1518:VAL:HB	1:A:1581:LEU:HD11	0.50	1.82	15	1
1:A:1540:ALA:HA	1:A:1623:ARG:HD3	0.49	1.85	4	1
1:A:1542:TYR:HD2	1:A:1545:SER:HB2	0.49	1.67	19	1
1:A:1516:CYS:HB3	1:A:1529:LEU:HB2	0.49	1.85	13	1
1:A:1549:TYR:CE1	1:A:1568:GLY:HA3	0.48	2.43	1	2
1:A:1549:TYR:HB3	1:A:1556:ASN:HD21	0.48	1.67	1	1
1:A:1542:TYR:CD2	1:A:1545:SER:HB2	0.48	2.44	19	1
1:A:1564:GLY:HA3	1:A:1592:TYR:CD2	0.48	2.43	13	1
1:A:1624:PRO:HB3	1:A:1640:TRP:HB2	0.48	1.85	4	2
1:A:1576:LYS:HG3	1:A:1595:GLY:HA2	0.48	1.85	19	1
1:A:1626:LEU:HD21	1:A:1629:LEU:HD23	0.48	1.84	17	1
1:A:1601:LYS:HD2	1:A:1604:LEU:HG	0.48	1.86	2	1
1:A:1527:PHE:HE1	1:A:1646:CYS:HA	0.47	1.69	12	1
1:A:1636:LEU:HB3	1:A:1638:PHE:CE2	0.47	2.44	10	1
1:A:1543:SER:HB3	1:A:1546:GLY:N	0.47	2.24	11	1
1:A:1551:SER:HB2	1:A:1556:ASN:HB3	0.47	1.86	20	1
1:A:1574:VAL:HB	1:A:1606:TYR:HB2	0.47	1.87	9	2
1:A:1629:LEU:HG	1:A:1636:LEU:HD13	0.47	1.86	3	1
1:A:1552:ILE:O	1:A:1592:TYR:OH	0.46	2.30	12	1
1:A:1515:ASP:HB3	1:A:1517:GLN:HG2	0.46	1.86	20	1
1:A:1549:TYR:CD2	1:A:1568:GLY:HA2	0.46	2.45	7	2
1:A:1577:ALA:HA	1:A:1592:TYR:CE1	0.46	2.45	20	1
1:A:1549:TYR:HD2	1:A:1568:GLY:HA3	0.46	1.70	19	1
1:A:1594:ASP:HA	1:A:1605:SER:CB	0.46	2.41	1	1
1:A:1520:ASN:HB2	1:A:1527:PHE:CE2	0.46	2.46	10	1
1:A:1533:SER:HA	1:A:1553:CYS:O	0.46	2.11	14	1
1:A:1549:TYR:HE1	1:A:1568:GLY:HA3	0.46	1.70	15	1
1:A:1574:VAL:HA	1:A:1596:SER:CB	0.45	2.34	17	1
1:A:1582:ARG:HB3	1:A:1589:GLN:HB3	0.45	1.88	1	1
1:A:1566:CYS:SG	1:A:1571:ARG:HA	0.45	2.51	10	1
1:A:1613:VAL:HG12	1:A:1615:ARG:HG2	0.45	1.87	14	1
1:A:1537:GLY:HA3	1:A:1556:ASN:ND2	0.45	2.26	19	1
1:A:1611:SER:O	1:A:1639:SER:HA	0.44	2.12	12	3
1:A:1563:VAL:HA	1:A:1576:LYS:HA	0.44	1.89	8	2

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:1629:LEU:CD2	1:A:1636:LEU:HG	0.44	2.41	14	1
1:A:1549:TYR:CE2	1:A:1568:GLY:HA2	0.44	2.46	17	1
1:A:1606:TYR:HA	1:A:1634:CYS:O	0.44	2.13	17	2
1:A:1600:SER:HB3	1:A:1634:CYS:SG	0.44	2.52	8	1
1:A:1557:GLU:CD	1:A:1558:ASN:H	0.44	2.15	13	1
1:A:1629:LEU:HD22	1:A:1631:LYS:HG2	0.44	1.89	1	1
1:A:1532:LEU:HD13	1:A:1640:TRP:CZ3	0.44	2.48	3	1
1:A:1580:ARG:O	1:A:1590:LEU:HA	0.44	2.12	3	1
1:A:1629:LEU:CD1	1:A:1636:LEU:HG	0.44	2.43	9	1
1:A:1591:VAL:HG12	1:A:1609:VAL:HG22	0.44	1.90	10	1
1:A:1539:THR:HA	1:A:1548:VAL:O	0.44	2.13	19	3
1:A:1612:PHE:HE1	1:A:1640:TRP:HE3	0.44	1.56	10	1
1:A:1535:ARG:O	1:A:1557:GLU:HB2	0.43	2.12	1	1
1:A:1601:LYS:HG3	1:A:1633:THR:HG22	0.43	1.89	9	1
1:A:1518:VAL:HB	1:A:1527:PHE:HB2	0.43	1.90	17	1
1:A:1615:ARG:HD2	1:A:1617:GLU:OE2	0.43	2.13	11	1
1:A:1629:LEU:HD22	1:A:1636:LEU:HG	0.43	1.90	14	1
1:A:1616:PRO:HA	1:A:1643:PRO:HD3	0.43	1.90	10	1
1:A:1592:TYR:HB2	1:A:1608:SER:HB3	0.43	1.91	12	1
1:A:1517:GLN:HA	1:A:1527:PHE:O	0.43	2.14	19	2
1:A:1588:LEU:CB	1:A:1612:PHE:HB2	0.42	2.41	7	1
1:A:1587:VAL:HA	1:A:1612:PHE:O	0.42	2.14	20	1
1:A:1581:LEU:HD13	1:A:1590:LEU:HB2	0.42	1.91	9	1
1:A:1590:LEU:HG	1:A:1612:PHE:HE1	0.42	1.75	13	1
1:A:1616:PRO:HB3	1:A:1643:PRO:HB3	0.41	1.91	6	1
1:A:1601:LYS:HG3	1:A:1632:GLN:O	0.41	2.15	19	1
1:A:1599:PRO:HD2	1:A:1606:TYR:OH	0.41	2.14	15	2
1:A:1590:LEU:HG	1:A:1612:PHE:CE1	0.41	2.50	13	1
1:A:1549:TYR:CD2	1:A:1568:GLY:HA3	0.41	2.49	19	1
1:A:1590:LEU:HB2	1:A:1610:ILE:HB	0.41	1.92	13	1
1:A:1614:CYS:HB3	1:A:1646:CYS:HB2	0.41	1.62	1	1
1:A:1589:GLN:HA	1:A:1610:ILE:O	0.41	2.15	13	1
1:A:1516:CYS:HB2	1:A:1579:LYS:HD2	0.41	1.93	8	1
1:A:1525:HIS:CE1	1:A:1647:GLU:HB2	0.41	2.51	1	1
1:A:1527:PHE:CE2	1:A:1646:CYS:HA	0.41	2.51	2	1
1:A:1542:TYR:HB3	1:A:1546:GLY:CA	0.41	2.43	9	1
1:A:1551:SER:HB3	1:A:1556:ASN:HB3	0.41	1.91	13	1
1:A:1549:TYR:CE2	1:A:1568:GLY:HA3	0.41	2.51	8	1
1:A:1533:SER:OG	1:A:1553:CYS:HB2	0.40	2.16	3	1
1:A:1615:ARG:HA	1:A:1615:ARG:NE	0.40	2.31	1	1
1:A:1525:HIS:HE1	1:A:1647:GLU:HB2	0.40	1.76	1	1

Continued on next page...

Continued from previous page...

Atom-1	Atom-2	Clash(Å)	Distance(Å)	Models	
				Worst	Total
1:A:1549:TYR:CZ	1:A:1568:GLY:HA2	0.40	2.51	5	1
1:A:1626:LEU:HG	1:A:1638:PHE:CE1	0.40	2.52	12	1
1:A:1623:ARG:HB3	1:A:1640:TRP:CG	0.40	2.51	16	1
1:A:1606:TYR:HA	1:A:1634:CYS:HB3	0.40	1.92	18	1
1:A:1629:LEU:HG	1:A:1636:LEU:CD1	0.40	2.47	3	1
1:A:1623:ARG:HB3	1:A:1640:TRP:CD1	0.40	2.52	8	1
1:A:1625:MET:HG3	1:A:1627:ILE:HG23	0.40	1.93	8	1
1:A:1543:SER:HB3	1:A:1546:GLY:H	0.40	1.74	11	1
1:A:1576:LYS:HG2	1:A:1595:GLY:HA2	0.40	1.93	16	1
1:A:1536:ALA:HB1	1:A:1558:ASN:HD21	0.40	1.76	19	1

5.2 Torsion angles (i)

5.2.1 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 PDB entries followed by that with respect to all NMR entries. 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	129/142 (91%)	116±2 (90±2%)	11±3 (8±2%)	2±1 (2±1%)	11	53
All	All	2580/2840 (91%)	2315 (90%)	216 (8%)	49 (2%)	11	53

All 15 unique Ramachandran outliers are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	1543	SER	19
1	A	1560	PRO	9
1	A	1544	LYS	3
1	A	1556	ASN	2
1	A	1624	PRO	2
1	A	1587	VAL	2
1	A	1595	GLY	2
1	A	1515	ASP	2
1	A	1616	PRO	2
1	A	1614	CYS	1
1	A	1601	LYS	1
1	A	1585	ASP	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	1572	ILE	1
1	A	1557	GLU	1
1	A	1558	ASN	1

5.2.2   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 PDB entries followed by that with respect to all NMR entries. 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	112/123 (91%)	109±1 (98±1%)	3±1 (2±1%)	53	92
All	All	2240/2460 (91%)	2189 (98%)	51 (2%)	53	92

All 25 unique residues with a non-rotameric sidechain are listed below. They are sorted by the frequency of occurrence in the ensemble.

Mol	Chain	Res	Type	Models (Total)
1	A	1570	THR	9
1	A	1614	CYS	5
1	A	1578	ASN	5
1	A	1517	GLN	3
1	A	1604	LEU	3
1	A	1641	HIS	2
1	A	1581	LEU	2
1	A	1528	ASP	2
1	A	1615	ARG	2
1	A	1555	GLU	2
1	A	1647	GLU	2
1	A	1646	CYS	1
1	A	1530	SER	1
1	A	1599	PRO	1
1	A	1560	PRO	1
1	A	1557	GLU	1
1	A	1625	MET	1
1	A	1642	THR	1
1	A	1588	LEU	1
1	A	1623	ARG	1
1	A	1569	GLN	1
1	A	1559	CYS	1

Continued on next page...

Continued from previous page...

Mol	Chain	Res	Type	Models (Total)
1	A	1593	LYS	1
1	A	1632	GLN	1
1	A	1549	TYR	1

5.2.3 RNA (i)

There are no RNA molecules in this entry.

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

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

5.4 Carbohydrates (i)

There are no carbohydrates in this entry.

5.5 Ligand geometry (i)

There are no ligands in this entry.

5.6 Other polymers (i)

There are no such molecules in this entry.

5.7 Polymer linkage issues (i)

There are no chain breaks in this entry.

6 Chemical shift validation (i)

No chemical shift data were provided