Supplementary material

Peptide bond planarity constrains hydrogen bond geometry and influences secondary structure conformations

Kuan Pern Tan[1,2], Khushboo Singh[3], Anirban Hazra[3] and M. S. Madhusudhan [1,3]*

[1]Bioinformatics Institute, 30 Biopolis Street, #07-01, Matrix, Singapore 138671,

[2]School of Computer Engineering, Nanyang Technological University, Singapore 639798,

[3]Indian Institute of Science Education and Research, Pune, India

* madhusudhan@iiserpune.ac.in

Supplementary data 1: List of structure codes (PDB and Open Crystallography Database)

(A) High resolution single domain proteins

This high-resolution data set was used to derive the geometric parameters of hydrogen bond in proteins. Protein Data Bank search criteria: resolution 1.7 Å or better; R-free 0.2 or better, non-redundant at 30% sequence identity, chain lengths between 80-180 amino acids.The first 4 letters specify the PDB code and the fifth letter specifies the chain identifier.

1A6MA, 1AQZA, 1B0BA, 1B9OA, 1BQKA, 1C7KA, 1CCWA, 1CXQA, 1DK8A, 1DWKA, 1DY5A, 1E7LA, 1EAJA, 1EAQA, 1EB6A, 1EUWA, 1EXRA, 1F4PA, 1FJJA, 1FM0E, 1G8KB, 1GK8I, 1GNYA, 1GQVA, 1GU2A, 1GWMA, 1H4AX, 1H97A, 1HFES, 1I76A, 1J98A, 1J9BA, 1JBEA, 1JLTA, 1JLTB, 1K2XA, 1K2XB, 1KMTA, 1KNGA, 1KNMA, 1KQ6A, 1KQRA, 1L2HA, 1LO7A, 1MC2A, 1MQKL, 1N62A, 1NC7A, 1NKIA, 1NWWA, 1NWZA, 1OALA, 1OAQH, 1OAQL, 1OD3A, 1OH4A, 1OKOA, 1OOHA, 1OUWA, 1P0ZA, 1P28A, 1P6OA, 1PM1X, 1PQHA, 1PWAA, 1QD9A, 1QDDA, 1QJPA, 1QTNA, 1QVEA, 1R26A, 1R29A, 1R2QA, 1R8SA, 1RG8A, 1RM6C, 1S69A, 1SENA, 1SJWA, 1SMBA, 1SX7A, 1SXVA, 1T0AA, 1T3YA, 1TO4A, 1TT8A, 1TU9A, 1TYJA, 1U69A, 1U7IA, 1UFYA, 1UGWA, 1UNQA, 1UOWA, 1UUYA, 1UXZA, 1UY4A, 1VJFA, 1VKEA, 1VKKA, 1VL7A, 1VL9A, 1VMFA, 1VMHA, 1VMJA, 1VPMA, 1VR7A, 1W0NA, 1W9SA, 1WCUA, 1WDDS, 1WKQA, 1WLUA, 1WMSA, 1WMZA, 1WWIA, 1X6ZA, 1XBIA, 1XG0C, 1XJUA, 1XO7A, 1XPHA, 1XRKA, 1XSOA, 1XT5A, 1Y55X, 1Y93A, 1YN9A, 1YPHC, 1YPQA, 1Z2UA, 1Z6MA, 1ZKIA, 1ZKKA, 1ZLBA, 1ZV9A, 2A2KA, 2A3MA, 2A50B, 2AENA, 2AJ7A, 2ANXA, 2AQ6A, 2AQMA, 2AQPA, 2AU7A, 2AXWA, 2BEMA, 2BK9A, 2BKMA, 2BWKA, 2C4BA, 2C4WA, 2C6UA, 2CALA, 2CE2X, 2CHCA, 2CJ4A, 2CJTA, 2CKKA, 2CM5A, 2D37A, 2D4PA, 2DKOA, 2DT4A, 2DXEA, 2ERXA, 2F01A, 2F22A, 2F62A, 2F9SA, 2FCLA, 2FG1A, 2FLHA, 2FR5A, 2FUFA, 2FWHA, 2G1UA, 2G2SB, 2G7BA, 2GJ3A, 2GKGA, 2GLZA, 2GMYA, 2GUDA, 2GYQA, 2H2RA, 2H8EA, 2H8FA, 2H8FB, 2HEWF, 2HUHA, 2HW2A, 2HX5A, 2HXSA, 2I3HA, 2I6CA, 2I8DA, 2I8TA, 2IA1A, 2IBLA, 2IMJA, 2IONA, 2J1SA, 2J23A, 2J8WA, 2JCQA, 2JDAA, 2JDCA, 2JJUA, 2JLIA, 2LISA, 2NRLA, 2NSZA, 2NWFA, 2O7AA, 2O8QA, 2O90A, 2OA2A, 2OFCA, 2OFZA, 2OH1A, 2OIZD, 2OKFA, 2OLMA, 2OSSA, 2OXGY, 2OZHA, 2P45B, 2P6OA, 2P7OA, 2P8GA, 2PBDP, 2PIEA, 2PR7A, 2PRVA, 2Q2HA, 2Q3TA, 2Q4NA, 2QCBA, 2QEUA, 2QF4A, 2QF9A, 2QIMA, 2QIPA, 2QJZA, 2QL8A, 2QLWA, 2QNLA, 2QNTA, 2QUDA, 2R0XA, 2R4IA, 2R6QA, 2R99A, 2RA9A, 2RBDA, 2RE2A, 2RFRA, 2RI7A, 2RKQA, 2UYKA, 2UYWA, 2UYZA, 2V0UA, 2V1MA, 2V4XA, 2V7FA, 2V9VA, 2VB1A, 2VIFA, 2VLQB, 2VMHA, 2VUVA, 2VVPA, 2VXTI, 2VY7A, 2VZCA, 2VZPA, 2W1WA, 2W3JA, 2W72A, 2W72B, 2W87A, 2WCWA, 2WFIA, 2WKKA, 2WLVA, 2WNXA, 2WQ4A, 2WRAA, 2WTPA, 2WVFA, 2WWEA, 2WYTA, 2WZ8A, 2WZ9A, 2X2SA, 2X32A, 2X46A, 2X52A, 2X5PA, 2X5YA, 2X7KA, 2XFGB, 2XG3A, 2XHFA, 2XOLA, 2XOMA, 2XR6A, 2XS4A, 2XSTA, 2Y0OA, 2Y1QA, 2Y71A, 2Y78A, 2Y8EA, 2Y8GA, 2Y9FA, 2YFUA, 2YH5A, 2YKZA, 2YZ1A, 2Z0XA, 2Z3HA, 2Z51A, 2Z6WA, 2ZEXA, 2ZFDB, 2ZHNA, 2ZIBA, 2ZNRA, 2ZS0A, 2ZS0B, 2ZS0C, 2ZS0D, 3A57A, 3A6RA, 3AP9A, 3AQJA, 3AWUB, 3B79A, 3BA3A, 3BEDA, 3BFQG, 3BLNA, 3BS2A, 3BT5A, 3C3MA, 3C8LA, 3CE1A, 3CHMA, 3CI6A, 3CIPG, 3CJEA, 3CM3A, 3CP5A, 3CT6A, 3CTGA, 3CXKA, 3D0JA, 3D1PA, 3D22A, 3D5PA, 3D9XA, 3DAUA, 3DB7A, 3DD7A, 3E10A, 3E4GA, 3E8MA, 3EBTA, 3EC9A, 3EDOA, 3EERA, 3EF4A, 3EF8A, 3EJVA, 3EURA, 3EVOA, 3EY6A, 3EY8A, 3F1PB, 3F2ZA, 3F37A, 3F6QA, 3F6VA, 3F7EA, 3F7LA, 3F7XA, 3F8XA, 3FAJA, 3FCNA, 3FIAA, 3FIQA, 3FK8A, 3FQ4A, 3FSAA, 3FYMA, 3FZWA, 3G0KA, 3G16A, 3G39A, 3G46A, 3G9AB, 3GA4A, 3GAXA, 3GBWA, 3GHJA, 3GIXA, 3GKMA, 3GMXB, 3GPGA, 3GRDA, 3GXHA, 3GY9A, 3GZBA, 3GZRA, 3H08A, 3H5JA, 3H6QA, 3H87A, 3HM4A, 3HT1A, 3HTNA, 3HUPA, 3HV2A, 3HWUA, 3HX8A, 3HZAA, 3HZPA, 3I3FA, 3I7MA, 3IA4A, 3IISM, 3IJMA, 3IMKA, 3IP0A, 3IQ1A, 3IQ2A, 3IQTA, 3ITFA, 3IVVA, 3JRVA, 3JUDA, 3JVLA, 3K67A, 3KA8A, 3KBEA, 3KE7A, 3KFFA, 3KHQA, 3KKGA, 3KLRA, 3KUUA, 3KWKA, 3KWUA, 3KYZA, 3L0FA, 3L0FB, 3L4RA, 3L51A, 3L51B, 3L8HA, 3L91A, 3LASA, 3LB2A, 3LBEA, 3LHNA, 3LJDA, 3LJWA, 3LLOA, 3LQWA, 3LWZA, 3LX3A, 3LYDA, 3LYGA, 3LZLA, 3M1HA, 3M1OA, 3M1XA, 3M2OA, 3M7AA, 3MC3A, 3MDXA, 3ML1B, 3MMHA, 3MNGA, 3MPZA, 3MR0A, 3N08A, 3N1FA, 3N4IB, 3N4JA, 3N5AA, 3NBCA, 3NBKA, 3NECA, 3NJKA, 3NKVA, 3NOHA, 3O1CA, 3O2RA, 3O5QA, 3OBLA, 3OD9A, 3OGNA, 3OHEA, 3OJ0A, 3OJ6A, 3OMDA, 3ON9A, 3ONHA, 3OR5A, 3OSEA, 3OSTA, 3OXPA, 3P1GA, 3P3VA, 3P48A, 3P6DA, 3P6IA, 3PG6A, 3PIWA, 3PJYA, 3PNAA, 3PP2A, 3PWFA, 3Q62A, 3Q64A, 3QDSA, 3QHBA, 3QL9A, 3QOOA, 3QU3A, 3QZBA, 3QZMA, 3R0LD, 3R0NA, 3R2RA, 3R3QA, 3RDOA, 3ROBA, 3ROFA, 3RRIA, 3RWNA, 3S4KA, 3S6FA, 3S6MA, 3S8IA, 3S9XA, 3SB1A, 3SK2A, 3SM1A, 3SXHA, 3T3LA, 3TQ5A, 3ZW5A, 3ZXHA, 4A2VA, 4UBPB

(B) Small organic molecules in Open Crystallography Database

The following lists the small organic molecules deposited in Open Crystallography Database (Grazulis et al., 2009; Grazulis et al., 2012)

(http://www.crystallography.net/cod/) that are used in our study. Each structure contains at least one hydrogen bond with a nitrogen donor, oxygen acceptor and carbon acceptor-antecedent.

1_00_00_1000001, 1_00_00_1000001, 1_10_01_1100181, 1_10_02_1100269, 1_10_02_1100296, 1_10_03_1100369, 1_10_03_1100376, 1_10_10_1101093, 1_10_11_1101103, 1_10_11_1101134,
1_50_00_1500006, 1_50_00_1500006, 1_50_00_1500023, 1_50_00_1500023, 1_50_00_1500029, 1_50_15_1501546, 1_50_15_1501549, 1_50_16_1501600, 1_50_16_1501614, 1_50_16_1501644,
1_50_18_1501851, 1_50_18_1501853, 1_50_18_1501881, 1_50_19_1501965, 1_50_22_1502224, 1_50_23_1502300, 1_50_23_1502327, 1_50_24_1502438, 1_50_25_1502573, 1_50_26_1502669,
1_50_27_1502728, 1_50_27_1502749, 1_50_27_1502751, 1_50_27_1502769, 1_50_27_1502769, 1_50_29_1502911, 1_50_31_1503122, 1_50_31_1503124, 1_50_31_1503125, 1_50_31_1503158,
1_50_31_1503187, 1_50_32_1503239, 1_50_32_1503267, 1_50_32_1503268, 1_50_32_1503269, 1_50_33_1503384, 1_50_33_1503384, 1_50_34_1503456, 1_50_36_1503605, 1_50_37_1503717,
1_50_37_1503720, 1_50_37_1503730, 1_50_37_1503758, 1_50_37_1503793, 1_50_37_1503797, 1_50_37_1503797, 1_50_38_1503850, 1_50_39_1503968, 1_50_40_1504026, 1_50_41_1504113,
1_50_41_1504149, 1_50_41_1504173, 1_50_41_1504174, 1_50_42_1504273, 1_50_42_1504278, 1_50_43_1504321, 1_50_43_1504362, 1_50_43_1504362, 1_50_44_1504435, 1_50_44_1504498,
1_50_45_1504587, 1_50_47_1504720, 1_50_48_1504852, 1_50_49_1504928, 1_50_51_1505171, 1_50_51_1505172, 1_50_51_1505173, 1_50_51_1505174, 1_50_51_1505177, 1_50_51_1505178,
1_50_51_1505181, 1_50_52_1505216, 1_50_52_1505286, 1_50_54_1505408, 1_50_54_1505409, 1_50_54_1505424, 1_50_55_1505583, 1_50_56_1505625, 1_50_56_1505647, 1_50_56_1505668,
1_50_57_1505767, 1_50_57_1505769, 1_50_57_1505769, 1_50_59_1505927, 1_50_60_1506013, 1_50_61_1506107, 1_50_61_1506121, 1_50_61_1506132, 1_50_62_1506286, 1_50_63_1506357,
1_50_63_1506358, 1_50_65_1506527, 1_50_66_1506650, 1_50_66_1506674, 1_50_66_1506674, 1_50_67_1506747, 1_50_67_1506748, 1_50_67_1506749, 1_50_67_1506750, 1_50_67_1506753,
1_50_67_1506755, 1_50_68_1506821, 1_50_69_1506968, 1_50_70_1507005, 1_50_70_1507066, 1_50_71_1507124, 1_50_71_1507128, 1_50_72_1507215, 1_50_73_1507388, 1_50_75_1507504,
1_50_75_1507547, 1_50_75_1507559, 1_50_75_1507571, 1_50_76_1507609, 1_50_76_1507678, 1_50_77_1507721, 1_50_77_1507723, 1_50_78_1507868, 1_50_78_1507873, 1_50_78_1507874,
1_50_79_1507994, 1_50_81_1508161, 1_50_81_1508162, 1_50_83_1508316, 1_50_83_1508337, 1_50_83_1508373, 1_50_86_1508621, 1_50_86_1508659, 1_50_88_1508808, 1_50_89_1508934,
1_50_89_1508956, 1_51_17_1511773, 1_51_18_1511849, 1_51_18_1511850, 1_51_18_1511851, 1_51_18_1511870, 1_51_20_1512023, 1_51_20_1512038, 1_51_20_1512098, 1_51_21_1512139,
1_51_22_1512263, 1_51_23_1512329, 1_51_26_1512633, 1_51_26_1512636, 1_51_28_1512851, 1_51_29_1512987, 1_51_30_1513028, 1_51_30_1513028, 1_51_30_1513071, 1_51_30_1513098,
1_51_31_1513100, 1_51_33_1513315, 1_51_34_1513401, 1_51_34_1513401, 1_51_34_1513402, 1_51_34_1513402, 1_51_34_1513452, 1_51_34_1513452, 1_51_34_1513496, 1_51_35_1513520,
1_51_36_1513606, 1_51_38_1513822, 1_51_38_1513831, 1_51_38_1513832, 1_51_38_1513838, 1_51_39_1513921, 1_51_41_1514127, 1_51_41_1514128, 1_51_41_1514131, 1_51_41_1514132,
1_51_41_1514136, 1_51_43_1514334, 1_51_44_1514433, 1_51_45_1514524, 1_51_45_1514581, 1_51_46_1514645, 1_51_46_1514680, 1_51_48_1514827, 1_51_48_1514827, 1_51_48_1514876,
1_51_51_1515119, 1_51_51_1515119, 1_51_51_1515120, 1_51_51_1515120, 1_51_51_1515121, 1_51_51_1515121, 1_51_52_1515200, 1_51_52_1515200, 1_51_52_1515241, 1_51_52_1515243,
1_51_52_1515244, 1_51_52_1515245, 1_51_52_1515293, 1_51_53_1515316, 1_51_53_1515345, 1_51_54_1515423, 1_51_54_1515481, 1_51_56_1515622, 1_51_56_1515622, 1_51_56_1515623,
1_51_56_1515623, 1_51_57_1515785, 1_51_57_1515785, 1_51_58_1515861, 1_51_60_1516080, 1_51_63_1516392, 1_51_64_1516411, 1_51_64_1516464, 1_51_65_1516519, 1_51_66_1516623,
1_51_67_1516708, 1_51_67_1516717, 1_51_67_1516720, 1_51_67_1516781, 1_51_67_1516794, 1_51_70_1517025, 1_51_72_1517246, 1_51_72_1517247, 1_51_75_1517530, 1_51_75_1517533,
1_51_78_1517816, 1_51_78_1517817, 1_51_81_1518163, 1_51_81_1518188, 1_51_82_1518236, 1_51_82_1518242, 1_51_82_1518266, 1_51_83_1518362, 1_51_83_1518362, 1_51_83_1518378,
1_51_83_1518383, 1_51_83_1518387, 1_51_84_1518454, 1_51_85_1518568, 1_51_87_1518700, 1_51_87_1518785, 2_00_02_2000202, 2_00_02_2000222, 2_00_02_2000223, 2_00_02_2000270,
2_00_03_2000305, 2_00_04_2000437, 2_00_07_2000768, 2_00_09_2000931, 2_00_11_2001120, 2_00_12_2001216, 2_00_14_2001406, 2_00_14_2001468, 2_00_18_2001815, 2_00_29_2002920,
2_00_29_2002921, 2_00_36_2003612, 2_00_38_2003839, 2_00_38_2003861, 2_00_39_2003937, 2_00_39_2003975, 2_00_40_2004089, 2_00_41_2004136, 2_00_41_2004187, 2_00_42_2004227,
2_00_43_2004347, 2_00_45_2004546, 2_00_48_2004890, 2_00_51_2005192, 2_00_54_2005476, 2_00_55_2005550, 2_00_58_2005815, 2_00_58_2005871, 2_00_58_2005891, 2_00_61_2006122,
2_00_62_2006298, 2_00_63_2006392, 2_00_64_2006423, 2_00_65_2006502, 2_00_65_2006540, 2_00_67_2006747, 2_00_67_2006755, 2_00_70_2007091, 2_00_73_2007308, 2_00_74_2007425,
2_00_74_2007480, 2_00_75_2007515, 2_00_76_2007693, 2_00_77_2007758, 2_00_79_2007943, 2_00_79_2007974, 2_00_79_2007991, 2_00_79_2007991, 2_00_80_2008053, 2_00_80_2008081,
2_00_82_2008207, 2_00_82_2008209, 2_00_82_2008210, 2_00_82_2008266, 2_00_83_2008366, 2_00_83_2008366, 2_00_85_2008519, 2_00_85_2008527, 2_00_85_2008589, 2_00_86_2008686,
2_00_86_2008698, 2_00_87_2008708, 2_00_87_2008712, 2_00_88_2008867, 2_00_89_2008956, 2_00_89_2008969, 2_00_90_2009047, 2_00_90_2009047, 2_00_91_2009196, 2_00_92_2009202,
2_00_92_2009202, 2_00_93_2009317, 2_00_93_2009331, 2_00_93_2009351, 2_00_96_2009613, 2_00_96_2009692, 2_00_98_2009809, 2_00_99_2009966, 2_00_99_2009968, 2_01_00_2010080,
2_01_01_2010109, 2_01_01_2010140, 2_01_02_2010206, 2_01_05_2010528, 2_01_06_2010673, 2_01_06_2010689, 2_01_06_2010695, 2_01_06_2010699, 2_01_08_2010827, 2_01_08_2010846,

2_01_08_2010847, 2_01_08_2010848, 2_01_08_2010859, 2_01_09_2010950, 2_01_10_2011017, 2_01_10_2011024, 2_01_11_2011125, 2_01_11_2011154, 2_01_11_2011155, 2_01_11_2011194,
2_01_13_2011305, 2_01_13_2011379, 2_01_13_2011385, 2_01_15_2011504, 2_01_15_2011541, 2_01_15_2011542, 2_01_15_2011545, 2_01_16_2011603, 2_01_16_2011613, 2_01_18_2011824,
2_01_18_2011847, 2_01_18_2011865, 2_01_18_2011866, 2_01_18_2011867, 2_01_18_2011896, 2_01_18_2011896, 2_01_19_2011913, 2_01_19_2011938, 2_01_19_2011939, 2_01_19_2011968,
2_01_19_2011969, 2_01_20_2012057, 2_01_20_2012068, 2_01_21_2012163, 2_01_21_2012163, 2_01_21_2012193, 2_01_22_2012211, 2_01_22_2012219, 2_01_22_2012220, 2_01_22_2012222,
2_01_22_2012223, 2_01_23_2012335, 2_01_24_2012458, 2_01_24_2012497, 2_01_24_2012497, 2_01_26_2012627, 2_01_26_2012628, 2_01_26_2012629, 2_01_26_2012630, 2_01_26_2012678,
2_01_26_2012679, 2_01_27_2012768, 2_01_28_2012810, 2_01_28_2012847, 2_01_29_2012910, 2_01_30_2013013, 2_01_30_2013017, 2_01_30_2013017, 2_01_30_2013092, 2_01_30_2013092,
2_01_30_2013093, 2_01_30_2013093, 2_01_30_2013097, 2_01_31_2013124, 2_01_31_2013147, 2_01_31_2013147, 2_01_31_2013151, 2_01_31_2013151, 2_01_31_2013158, 2_01_31_2013161,
2_01_32_2013202, 2_01_32_2013216, 2_01_32_2013232, 2_01_32_2013281, 2_01_33_2013336, 2_01_33_2013336, 2_01_33_2013348, 2_01_34_2013415, 2_01_34_2013465, 2_01_34_2013495,
2_01_35_2013502, 2_01_35_2013554, 2_01_35_2013579, 2_01_36_2013630, 2_01_36_2013679, 2_01_36_2013679, 2_01_38_2013823, 2_01_38_2013823, 2_01_38_2013874, 2_01_38_2013889,
2_01_38_2013890, 2_01_39_2013910, 2_01_39_2013949, 2_01_39_2013959, 2_01_40_2014074, 2_01_41_2014115, 2_01_41_2014144, 2_01_41_2014144, 2_01_41_2014162, 2_01_42_2014235,
2_01_42_2014235, 2_01_42_2014281, 2_01_43_2014333, 2_01_43_2014354, 2_01_43_2014399, 2_01_44_2014400, 2_01_44_2014441, 2_01_44_2014441, 2_01_44_2014474, 2_01_44_2014475,
2_01_45_2014536, 2_01_46_2014605, 2_01_46_2014678, 2_01_47_2014739, 2_01_47_2014798, 2_01_48_2014855, 2_01_49_2014902, 2_01_49_2014965, 2_01_50_2015042, 2_01_50_2015068,
2_01_51_2015153, 2_01_52_2015204, 2_01_52_2015217, 2_01_52_2015236, 2_01_53_2015314, 2_01_54_2015407, 2_01_54_2015409, 2_01_54_2015413, 2_01_54_2015413, 2_01_54_2015414,
2_01_54_2015434, 2_01_54_2015435, 2_01_54_2015436, 2_01_54_2015459, 2_01_54_2015479, 2_01_55_2015525, 2_01_55_2015538, 2_01_55_2015543, 2_01_55_2015564, 2_01_55_2015564,
2_01_55_2015577, 2_01_55_2015596, 2_01_56_2015614, 2_01_56_2015614, 2_01_56_2015625, 2_01_56_2015639, 2_01_56_2015688, 2_01_57_2015702, 2_01_57_2015703, 2_01_57_2015729,
2_01_57_2015771, 2_01_58_2015859, 2_01_58_2015867, 2_01_58_2015878, 2_01_58_2015881, 2_01_59_2015912, 2_01_59_2015912, 2_01_59_2015914, 2_01_59_2015929, 2_01_59_2015942,
2_01_59_2015943, 2_01_59_2015952, 2_01_59_2015955, 2_01_59_2015961, 2_01_59_2015961, 2_01_59_2015962, 2_01_59_2015962, 2_01_59_2015963, 2_01_59_2015963, 2_01_59_2015969,
2_01_60_2016032, 2_01_60_2016032, 2_01_60_2016033, 2_01_60_2016033, 2_01_61_2016104, 2_01_61_2016136, 2_01_62_2016233, 2_01_62_2016234, 2_01_62_2016253, 2_01_62_2016256,
2_01_62_2016282, 2_01_63_2016360, 2_01_64_2016422, 2_01_64_2016482, 2_01_64_2016482, 2_01_64_2016483, 2_01_64_2016484, 2_01_64_2016485, 2_01_64_2016486, 2_01_64_2016486,
2_01_66_2016637, 2_01_66_2016639, 2_01_66_2016645, 2_01_66_2016646, 2_01_67_2016730, 2_01_67_2016794, 2_01_67_2016794, 2_01_67_2016795, 2_01_68_2016807, 2_01_68_2016808,
2_01_68_2016835, 2_01_68_2016849, 2_01_68_2016850, 2_01_69_2016909, 2_01_69_2016922, 2_01_69_2016922, 2_01_69_2016968, 2_01_69_2016969, 2_01_70_2017005, 2_01_70_2017005,
2_01_70_2017006, 2_01_70_2017007, 2_01_70_2017035, 2_01_70_2017043, 2_01_70_2017099, 2_01_71_2017104, 2_01_71_2017115, 2_01_71_2017116, 2_01_71_2017117, 2_01_71_2017132,
2_01_71_2017133, 2_01_71_2017134, 2_01_71_2017185, 2_01_71_2017195, 2_01_71_2017195, 2_01_71_2017198, 2_01_72_2017231, 2_01_72_2017257, 2_01_72_2017290, 2_01_72_2017291,
2_01_73_2017329, 2_01_73_2017330, 2_01_73_2017344, 2_01_73_2017345, 2_01_73_2017346, 2_01_73_2017359, 2_01_73_2017379, 2_01_73_2017380, 2_01_74_2017404, 2_01_74_2017404,
2_01_74_2017405, 2_01_74_2017436, 2_01_74_2017437, 2_01_74_2017443, 2_01_74_2017486, 2_01_74_2017487, 2_01_75_2017514, 2_01_75_2017556, 2_01_75_2017558, 2_01_75_2017558,
2_01_75_2017562, 2_01_75_2017564, 2_01_75_2017565, 2_01_75_2017570, 2_01_75_2017574, 2_01_75_2017576, 2_01_75_2017584, 2_01_75_2017589, 2_01_75_2017597, 2_01_75_2017598,
2_01_75_2017598, 2_01_76_2017666, 2_01_76_2017667, 2_01_76_2017668, 2_01_76_2017669, 2_01_76_2017692, 2_01_76_2017694, 2_01_76_2017698, 2_01_76_2017698, 2_01_76_2017699,
2_01_76_2017699, 2_01_77_2017700, 2_01_77_2017700, 2_01_77_2017702, 2_01_77_2017746, 2_01_77_2017747, 2_01_77_2017747, 2_01_77_2017758, 2_01_77_2017758, 2_01_77_2017759,
2_01_77_2017793, 2_01_77_2017798, 2_01_78_2017841, 2_01_78_2017884, 2_01_78_2017889, 2_01_78_2017890, 2_01_78_2017890, 2_01_78_2017892, 2_01_79_2017924, 2_01_79_2017925,
2_01_79_2017926, 2_01_79_2017927, 2_01_79_2017928, 2_01_79_2017941, 2_01_79_2017948, 2_01_79_2017998, 2_01_80_2018015, 2_01_80_2018017, 2_01_80_2018080, 2_01_80_2018081,
2_01_80_2018082, 2_01_80_2018083, 2_01_80_2018084, 2_01_81_2018104, 2_01_81_2018105, 2_01_81_2018121, 2_01_81_2018128, 2_01_81_2018132, 2_01_81_2018132, 2_01_81_2018133,
2_01_81_2018133, 2_01_81_2018145, 2_01_81_2018171, 2_01_81_2018179, 2_01_81_2018190, 2_01_82_2018205, 2_01_82_2018242, 2_01_82_2018279, 2_01_82_2018280, 2_01_82_2018298,
2_01_82_2018299, 2_01_83_2018302, 2_01_83_2018307, 2_01_83_2018314, 2_01_83_2018366, 2_01_84_2018401, 2_01_84_2018406, 2_01_84_2018426, 2_01_84_2018427, 2_01_84_2018439,
2_01_84_2018483, 2_01_85_2018523, 2_01_85_2018546, 2_01_85_2018585, 2_01_85_2018592, 2_01_86_2018635, 2_01_86_2018652, 2_01_86_2018655, 2_01_86_2018671, 2_01_86_2018674,
2_01_86_2018674, 2_01_86_2018675, 2_01_87_2018731, 2_01_87_2018745, 2_01_87_2018746, 2_01_87_2018747, 2_01_87_2018747, 2_01_87_2018748, 2_01_87_2018748, 2_01_87_2018749,
2_01_87_2018781, 2_01_88_2018838, 2_01_88_2018897, 2_01_88_2018898, 2_01_89_2018954, 2_01_89_2018966, 2_01_89_2018966, 2_01_89_2018967, 2_01_90_2019040, 2_01_90_2019055,
2_01_90_2019074, 2_01_90_2019074, 2_01_91_2019105, 2_01_91_2019107, 2_01_91_2019120, 2_01_91_2019121, 2_01_91_2019126, 2_01_91_2019127, 2_01_91_2019127, 2_01_91_2019141,
2_01_91_2019163, 2_01_91_2019185, 2_01_91_2019185, 2_01_91_2019186, 2_01_91_2019186, 2_01_91_2019187, 2_01_91_2019187, 2_01_92_2019225, 2_01_92_2019226, 2_01_92_2019227,
2_01_92_2019229, 2_01_92_2019240, 2_01_92_2019254, 2_01_92_2019255, 2_01_92_2019255, 2_01_92_2019256, 2_01_93_2019318, 2_01_93_2019342, 2_01_93_2019361, 2_01_93_2019362,

2_01_93_2019374, 2_01_93_2019376, 2_01_93_2019377, 2_01_93_2019378, 2_01_93_2019378, 2_01_93_2019379, 2_01_93_2019381, 2_01_93_2019382, 2_01_93_2019383, 2_01_93_2019383,
2_01_94_2019414, 2_01_94_2019416, 2_01_96_2019655, 2_01_96_2019655, 2_01_96_2019657, 2_01_96_2019658, 2_01_97_2019749, 2_01_97_2019749, 2_01_97_2019763, 2_01_97_2019771,
2_01_97_2019772, 2_01_97_2019773, 2_01_97_2019774, 2_01_98_2019800, 2_01_98_2019800, 2_01_98_2019801, 2_01_98_2019801, 2_01_98_2019803, 2_01_98_2019846, 2_01_98_2019872,
2_01_98_2019876, 2_01_98_2019878, 2_01_98_2019878, 2_01_98_2019879, 2_01_98_2019880, 2_01_98_2019881, 2_01_98_2019882, 2_10_01_2100131, 2_10_02_2100287, 2_10_02_2100290,
2_10_03_2100351, 2_10_03_2100352, 2_10_03_2100353, 2_10_03_2100354, 2_10_03_2100355, 2_10_03_2100356, 2_10_03_2100357, 2_10_03_2100358, 2_10_03_2100359, 2_10_03_2100360,
2_10_03_2100361, 2_10_03_2100362, 2_10_05_2100503, 2_10_05_2100505, 2_10_05_2100533, 2_10_05_2100542, 2_10_05_2100554, 2_10_05_2100555, 2_10_05_2100556, 2_10_05_2100596,
2_10_08_2100881, 2_10_09_2100981, 2_10_10_2101074, 2_10_10_2101075, 2_10_10_2101076, 2_10_10_2101087, 2_10_10_2101094, 2_10_11_2101127, 2_10_11_2101128, 2_10_11_2101162,
2_10_13_2101327, 2_10_13_2101337, 2_10_14_2101407, 2_10_14_2101410, 2_10_14_2101413, 2_10_15_2101500, 2_10_15_2101501, 2_10_15_2101502, 2_10_15_2101522, 2_10_15_2101547,
2_10_15_2101548, 2_10_16_2101616, 2_10_16_2101617, 2_10_16_2101646, 2_10_16_2101647, 2_10_16_2101648, 2_10_18_2101819, 2_10_18_2101820, 2_10_18_2101820, 2_10_18_2101821,
2_10_18_2101821, 2_10_18_2101822, 2_10_19_2101919, 2_10_19_2101920, 2_10_19_2101957, 2_10_19_2101965, 2_10_19_2101965, 2_10_19_2101966, 2_10_19_2101966, 2_10_19_2101967,
2_10_19_2101967, 2_10_19_2101968, 2_10_19_2101976, 2_10_19_2101978, 2_10_19_2101979, 2_10_19_2101980, 2_10_19_2101983, 2_10_20_2102033, 2_10_20_2102097, 2_10_22_2102227,
2_10_22_2102229, 2_10_22_2102229, 2_10_24_2102479, 2_10_25_2102581, 2_10_26_2102682, 2_10_28_2102810, 2_10_28_2102814, 2_10_29_2102969, 2_10_29_2102970, 2_10_29_2102971,
2_10_29_2102975, 2_10_29_2102982, 2_10_29_2102985, 2_10_31_2103127, 2_10_31_2103129, 2_10_31_2103130, 2_10_31_2103138, 2_10_31_2103138, 2_10_32_2103240, 2_10_32_2103251,
2_10_32_2103253, 2_10_32_2103264, 2_10_32_2103265, 2_10_32_2103266, 2_10_33_2103324, 2_10_33_2103325, 2_10_34_2103455, 2_10_34_2103455, 2_10_36_2103612, 2_10_37_2103718,
2_10_37_2103719, 2_10_37_2103720, 2_10_37_2103721, 2_10_37_2103747, 2_10_38_2103832, 2_10_38_2103833, 2_10_38_2103834, 2_10_38_2103840, 2_10_38_2103846, 2_10_38_2103846,
2_10_38_2103860, 2_10_38_2103862, 2_10_40_2104013, 2_10_42_2104209, 2_10_42_2104251, 2_10_42_2104253, 2_10_42_2104254, 2_10_48_2104838, 2_10_48_2104877, 2_10_49_2104908,
2_10_49_2104909, 2_10_51_2105158, 2_10_51_2105159, 2_10_51_2105160, 2_10_51_2105161, 2_10_51_2105162, 2_10_51_2105163, 2_10_51_2105165, 2_10_51_2105166, 2_10_51_2105168,
2_10_51_2105169, 2_10_51_2105170, 2_10_51_2105171, 2_10_51_2105171, 2_10_53_2105324, 2_10_53_2105325, 2_10_55_2105511, 2_10_55_2105512, 2_10_55_2105513, 2_10_55_2105579,
2_10_55_2105580, 2_10_55_2105582, 2_10_55_2105582, 2_10_55_2105583, 2_10_55_2105583, 2_10_55_2105592, 2_10_56_2105605, 2_10_56_2105608, 2_10_56_2105652, 2_10_56_2105653,
2_10_56_2105654, 2_10_56_2105655, 2_10_56_2105656, 2_10_57_2105705, 2_10_57_2105755, 2_10_57_2105756, 2_10_57_2105770, 2_10_59_2105989, 2_10_59_2105990, 2_20_00_2200001,
2_20_01_2200182, 2_20_03_2200304, 2_20_03_2200394, 2_20_04_2200412, 2_20_04_2200439, 2_20_05_2200547, 2_20_05_2200556, 2_20_05_2200578, 2_20_06_2200645, 2_20_07_2200709,
2_20_08_2200811, 2_20_08_2200876, 2_20_09_2200925, 2_20_09_2200925, 2_20_09_2200960, 2_20_10_2201047, 2_20_11_2201144, 2_20_13_2201373, 2_20_14_2201497, 2_20_15_2201578,
2_20_16_2201647, 2_20_16_2201658, 2_20_16_2201659, 2_20_16_2201676, 2_20_17_2201705, 2_20_17_2201713, 2_20_17_2201713, 2_20_17_2201717, 2_20_18_2201845, 2_20_18_2201850,
2_20_20_2202058, 2_20_20_2202068, 2_20_20_2202076, 2_20_21_2202116, 2_20_21_2202124, 2_20_21_2202124, 2_20_21_2202187, 2_20_22_2202241, 2_20_22_2202241, 2_20_23_2202352,
2_20_23_2202396, 2_20_24_2202458, 2_20_25_2202537, 2_20_26_2202683, 2_20_27_2202728, 2_20_27_2202797, 2_20_28_2202817, 2_20_28_2202895, 2_20_29_2202901, 2_20_29_2202937,
2_20_29_2202964, 2_20_29_2202966, 2_20_29_2202996, 2_20_32_2203245, 2_20_33_2203360, 2_20_33_2203360, 2_20_33_2203361, 2_20_33_2203384, 2_20_34_2203489, 2_20_35_2203517,
2_20_35_2203526, 2_20_37_2203709, 2_20_37_2203790, 2_20_38_2203837, 2_20_38_2203894, 2_20_39_2203978, 2_20_39_2203999, 2_20_41_2204123, 2_20_41_2204148, 2_20_41_2204151,
2_20_41_2204171, 2_20_42_2204284, 2_20_42_2204285, 2_20_42_2204297, 2_20_42_2204298, 2_20_43_2204372, 2_20_43_2204382, 2_20_43_2204383, 2_20_44_2204405, 2_20_45_2204524,
2_20_45_2204531, 2_20_45_2204532, 2_20_45_2204545, 2_20_45_2204545, 2_20_45_2204550, 2_20_46_2204612, 2_20_46_2204642, 2_20_46_2204690, 2_20_46_2204690, 2_20_46_2204697,
2_20_47_2204713, 2_20_47_2204759, 2_20_47_2204798, 2_20_48_2204829, 2_20_48_2204831, 2_20_48_2204857, 2_20_49_2204925, 2_20_49_2204925, 2_20_49_2204948, 2_20_49_2204953,
2_20_49_2204954, 2_20_50_2205042, 2_20_50_2205042, 2_20_50_2205072, 2_20_50_2205088, 2_20_50_2205096, 2_20_51_2205116, 2_20_51_2205155, 2_20_52_2205292, 2_20_53_2205391,
2_20_53_2205391, 2_20_54_2205475, 2_20_55_2205589, 2_20_56_2205656, 2_20_58_2205821, 2_20_58_2205841, 2_20_59_2205956, 2_20_59_2205976, 2_20_60_2206032, 2_20_60_2206033,
2_20_60_2206038, 2_20_60_2206039, 2_20_61_2206102, 2_20_61_2206182, 2_20_62_2206236, 2_20_62_2206242, 2_20_62_2206254, 2_20_63_2206325, 2_20_63_2206362, 2_20_64_2206405,
2_20_64_2206412, 2_20_64_2206414, 2_20_64_2206414, 2_20_64_2206445, 2_20_64_2206479, 2_20_65_2206527, 2_20_65_2206552, 2_20_66_2206620, 2_20_67_2206753, 2_20_67_2206753,
2_20_69_2206930, 2_20_69_2206968, 2_20_70_2207034, 2_20_70_2207034, 2_20_70_2207076, 2_20_70_2207081, 2_20_70_2207081, 2_20_70_2207084, 2_20_70_2207084, 2_20_70_2207098,
2_20_70_2207098, 2_20_71_2207116, 2_20_71_2207144, 2_20_71_2207144, 2_20_71_2207162, 2_20_74_2207422, 2_20_75_2207511, 2_20_75_2207511, 2_20_75_2207551, 2_20_75_2207574,
2_20_75_2207597, 2_20_76_2207659, 2_20_76_2207660, 2_20_76_2207661, 2_20_77_2207754, 2_20_77_2207789, 2_20_77_2207790, 2_20_77_2207793, 2_20_78_2207838, 2_20_78_2207884,
2_20_79_2207940, 2_20_80_2208010, 2_20_80_2208015, 2_20_80_2208023, 2_20_81_2208160, 2_20_82_2208291, 2_20_84_2208430, 2_20_84_2208431, 2_20_84_2208432, 2_20_85_2208551,
2_20_87_2208761, 2_20_87_2208771, 2_20_88_2208817, 2_20_88_2208817, 2_20_88_2208864, 2_20_89_2208935, 2_20_89_2208938, 2_20_89_2208999, 2_20_90_2209013, 2_20_90_2209047,

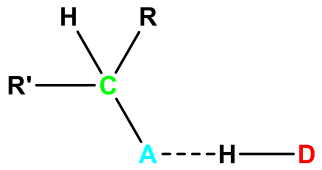
2_20_92_2209209, 2_20_95_2209506, 2_20_95_2209536, 2_20_96_2209631, 2_20_97_2209725, 2_20_98_2209849, 2_20_98_2209875, 2_20_98_2209884, 2_20_98_2209889, 2_20_99_2209900,
2_20_99_2209927, 2_20_99_2209979, 2_21_00_2210053, 2_21_00_2210055, 2_21_00_2210062, 2_21_00_2210062, 2_21_01_2210167, 2_21_01_2210169, 2_21_01_2210182, 2_21_02_2210226,
2_21_04_2210427, 2_21_04_2210439, 2_21_04_2210439, 2_21_05_2210561, 2_21_06_2210615, 2_21_06_2210672, 2_21_07_2210723, 2_21_07_2210741, 2_21_07_2210791, 2_21_07_2210795,
2_21_08_2210803, 2_21_08_2210827, 2_21_08_2210855, 2_21_09_2210929, 2_21_09_2210929, 2_21_10_2211088, 2_21_10_2211089, 2_21_11_2211137, 2_21_12_2211256, 2_21_15_2211556,
2_21_15_2211557, 2_21_17_2211753, 2_21_18_2211821, 2_21_18_2211875, 2_21_18_2211879, 2_21_20_2212039, 2_21_20_2212039, 2_21_20_2212085, 2_21_21_2212140, 2_21_21_2212162,
2_21_21_2212193, 2_21_21_2212193, 2_21_22_2212224, 2_21_23_2212320, 2_21_23_2212371, 2_21_24_2212411, 2_21_24_2212452, 2_21_25_2212534, 2_21_26_2212698, 2_21_27_2212751,
2_21_27_2212772, 2_21_28_2212814, 2_21_28_2212837, 2_21_28_2212837, 2_21_29_2212953, 2_21_29_2212985, 2_21_30_2213014, 2_21_30_2213052, 2_21_30_2213094, 2_21_31_2213107,
2_21_31_2213119, 2_21_31_2213119, 2_21_31_2213152, 2_21_31_2213170, 2_21_31_2213175, 2_21_32_2213261, 2_21_32_2213286, 2_21_33_2213353, 2_21_33_2213358, 2_21_33_2213363,
2_21_34_2213473, 2_21_34_2213495, 2_21_35_2213515, 2_21_35_2213580, 2_21_36_2213649, 2_21_36_2213675, 2_21_36_2213690, 2_21_36_2213690, 2_21_36_2213695, 2_21_37_2213731,
2_21_37_2213745, 2_21_37_2213774, 2_21_38_2213837, 2_21_38_2213840, 2_21_38_2213897, 2_21_39_2213945, 2_21_39_2213945, 2_21_39_2213964, 2_21_41_2214101, 2_21_41_2214110,
2_21_41_2214184, 2_21_44_2214405, 2_21_44_2214417, 2_21_45_2214561, 2_21_45_2214599, 2_21_46_2214646, 2_21_46_2214685, 2_21_49_2214951, 2_21_49_2214990, 2_21_51_2215178,
2_21_51_2215178, 2_21_51_2215188, 2_21_51_2215199, 2_21_52_2215217, 2_21_52_2215217, 2_21_52_2215218, 2_21_53_2215377, 2_21_53_2215377, 2_21_54_2215442, 2_21_54_2215493,
2_21_54_2215493, 2_21_55_2215517, 2_21_55_2215520, 2_21_55_2215524, 2_21_55_2215563, 2_21_55_2215586, 2_21_55_2215588, 2_21_56_2215697, 2_21_57_2215773, 2_21_58_2215817,
2_21_58_2215817, 2_21_58_2215819, 2_21_58_2215867, 2_21_58_2215880, 2_21_59_2215905, 2_21_59_2215934, 2_21_59_2215968, 2_21_59_2215968, 2_21_60_2216054, 2_21_61_2216147,
2_21_62_2216289, 2_21_62_2216289, 2_21_63_2216322, 2_21_64_2216459, 2_21_65_2216578, 2_21_67_2216708, 2_21_67_2216716, 2_21_67_2216756, 2_21_68_2216872, 2_21_69_2216913,
2_21_69_2216944, 2_21_70_2217016, 2_21_70_2217079, 2_21_70_2217082, 2_21_70_2217093, 2_21_71_2217161, 2_21_71_2217196, 2_21_72_2217233, 2_21_72_2217248, 2_21_72_2217274,
2_21_72_2217282, 2_21_73_2217366, 2_21_73_2217366, 2_21_73_2217377, 2_21_74_2217431, 2_21_74_2217482, 2_21_76_2217614, 2_21_76_2217647, 2_21_76_2217647, 2_21_77_2217774,
2_21_78_2217890, 2_21_79_2217929, 2_21_79_2217935, 2_21_80_2218070, 2_21_81_2218138, 2_21_82_2218208, 2_21_82_2218233, 2_21_82_2218242, 2_21_83_2218393, 2_21_84_2218485,
2_21_84_2218485, 2_21_85_2218518, 2_21_85_2218524, 2_21_86_2218603, 2_21_86_2218635, 2_21_86_2218676, 2_21_87_2218765, 2_21_87_2218766, 2_21_87_2218767, 2_21_88_2218837,
2_21_89_2218946, 2_21_90_2219012, 2_21_90_2219021, 2_21_90_2219023, 2_21_91_2219128, 2_21_91_2219137, 2_21_91_2219144, 2_21_92_2219271, 2_21_93_2219343, 2_21_94_2219441,
2_21_94_2219471, 2_21_95_2219503, 2_21_95_2219511, 2_21_95_2219514, 2_21_95_2219571, 2_21_95_2219578, 2_21_95_2219583, 2_21_95_2219595, 2_21_96_2219626, 2_21_96_2219674,
2_21_97_2219715, 2_21_97_2219764, 2_21_97_2219784, 2_21_98_2219847, 2_21_98_2219851, 2_21_98_2219865, 2_21_98_2219865, 2_21_99_2219924, 2_21_99_2219924, 2_21_99_2219936,
2_22_00_2220067, 2_22_00_2220086, 2_22_01_2220107, 2_22_01_2220197, 2_22_03_2220349, 2_22_04_2220435, 2_22_04_2220455, 2_22_05_2220528, 2_22_06_2220605, 2_22_07_2220783,
2_22_09_2220924, 2_22_09_2220975, 2_22_09_2220975, 2_22_10_2221078, 2_22_11_2221132, 2_22_11_2221138, 2_22_11_2221141, 2_22_11_2221141, 2_22_11_2221152, 2_22_12_2221223,
2_22_12_2221248, 2_22_12_2221248, 2_22_12_2221281, 2_22_13_2221330, 2_22_14_2221417, 2_22_14_2221417, 2_22_14_2221420, 2_22_14_2221427, 2_22_14_2221449, 2_22_15_2221515,
2_22_16_2221621, 2_22_16_2221644, 2_22_16_2221667, 2_22_16_2221686, 2_22_17_2221722, 2_22_17_2221742, 2_22_17_2221783, 2_22_17_2221799, 2_22_18_2221835, 2_22_18_2221840,
2_22_18_2221849, 2_22_18_2221883, 2_22_19_2221913, 2_22_19_2221945, 2_22_19_2221961, 2_22_20_2222083, 2_22_21_2222199, 2_22_22_2222233, 2_22_22_2222250, 2_22_22_2222257,
2_22_22_2222261, 2_22_22_2222264, 2_22_22_2222264, 2_22_22_2222272, 2_22_22_2222272, 2_22_23_2222353, 2_22_23_2222359, 2_22_23_2222363, 2_22_23_2222366, 2_22_24_2222410,
2_22_24_2222472, 2_22_24_2222472, 2_22_25_2222520, 2_22_25_2222574, 2_22_25_2222574, 2_22_27_2222742, 2_22_27_2222774, 2_22_28_2222847, 2_22_28_2222847, 2_22_28_2222890,
2_22_28_2222890, 2_22_29_2222929, 2_22_29_2222971, 2_22_29_2222997, 2_22_30_2223030, 2_22_31_2223131, 2_22_31_2223163, 2_22_32_2223250, 2_22_32_2223250, 2_22_33_2223330,
2_22_33_2223375, 2_22_33_2223399, 2_22_34_2223404, 2_22_34_2223416, 2_22_34_2223455, 2_22_34_2223470, 2_22_35_2223520, 2_22_35_2223521, 2_22_36_2223647, 2_22_36_2223653,
2_22_37_2223748, 2_22_38_2223816, 2_22_38_2223816, 2_22_38_2223844, 2_22_38_2223847, 2_22_38_2223871, 2_22_38_2223871, 2_22_38_2223884, 2_22_38_2223899, 2_22_39_2223983,
2_22_40_2224065, 2_22_40_2224065, 2_22_41_2224145, 2_22_41_2224145, 2_22_41_2224152, 2_22_41_2224185, 2_22_42_2224236, 2_22_42_2224269, 2_22_42_2224280, 2_22_42_2224284,
2_22_42_2224285, 2_22_42_2224288, 2_22_44_2224492, 2_22_44_2224492, 2_22_45_2224566, 2_22_45_2224567, 2_22_45_2224579, 2_22_45_2224595, 2_22_46_2224613, 2_22_47_2224726,
2_22_47_2224781, 2_22_49_2224915, 2_22_49_2224915, 2_22_50_2225053, 2_22_50_2225093, 2_22_51_2225113, 2_22_51_2225128, 2_22_52_2225277, 2_22_52_2225286, 2_22_53_2225300,
2_22_53_2225300, 2_22_53_2225308, 2_22_53_2225313, 2_22_53_2225313, 2_22_53_2225348, 2_22_53_2225356, 2_22_54_2225433, 2_22_54_2225465, 2_22_55_2225568, 2_22_55_2225587,
2_22_55_2225587, 2_22_56_2225685, 2_22_56_2225685, 2_22_57_2225721, 2_22_57_2225724, 2_22_57_2225779, 2_22_58_2225843, 2_22_58_2225889, 2_22_59_2225907, 2_22_59_2225957,
2_22_60_2226065, 2_22_60_2226079, 2_22_60_2226080, 2_22_60_2226082, 2_22_61_2226158, 2_22_61_2226158, 2_22_61_2226170, 2_22_62_2226289, 2_22_62_2226289, 2_22_62_2226293,
2_22_62_2226296, 2_22_63_2226369, 2_22_64_2226434, 2_22_64_2226434, 2_22_64_2226468, 2_22_65_2226573, 2_22_65_2226596, 2_22_65_2226596, 2_22_66_2226667, 2_22_66_2226675,

2_22_67_2226781, 2_22_67_2226784, 2_22_67_2226786, 2_22_68_2226803, 2_22_68_2226804, 2_22_68_2226818, 2_22_68_2226818, 2_22_68_2226823, 2_22_68_2226825, 2_22_68_2226826,
2_22_68_2226826, 2_22_68_2226844, 2_22_68_2226848, 2_22_68_2226854, 2_22_69_2226920, 2_22_69_2226920, 2_22_69_2226952, 2_22_69_2226952, 2_22_69_2226987, 2_22_69_2226992,
2_22_70_2227003, 2_22_71_2227121, 2_22_71_2227122, 2_22_71_2227126, 2_22_71_2227168, 2_22_71_2227199, 2_22_71_2227199, 2_22_72_2227200, 2_22_72_2227200, 2_22_72_2227202,
2_22_72_2227202, 2_22_72_2227212, 2_22_72_2227236, 2_22_72_2227237, 2_22_72_2227262, 2_22_72_2227281, 2_22_73_2227331, 2_22_73_2227335, 2_22_73_2227339, 2_22_73_2227339,
2_22_73_2227351, 2_22_73_2227352, 2_22_73_2227368, 2_22_74_2227415, 2_22_74_2227415, 2_22_74_2227471, 2_22_74_2227471, 2_22_74_2227494, 2_22_74_2227494, 2_22_75_2227537,
2_22_75_2227598, 2_22_76_2227649, 2_22_76_2227674, 2_22_76_2227674, 2_22_77_2227727, 2_22_77_2227775, 2_22_77_2227775, 2_22_77_2227781, 2_22_77_2227784, 2_22_78_2227801,
2_22_78_2227801, 2_22_78_2227805, 2_22_78_2227805, 2_22_78_2227814, 2_22_78_2227832, 2_22_78_2227846, 2_22_78_2227882, 2_22_80_2228041, 2_22_80_2228042, 2_22_80_2228094,
2_22_81_2228121, 2_22_82_2228259, 2_22_82_2228265, 2_22_82_2228295, 2_22_83_2228317, 2_22_83_2228391, 2_22_84_2228489, 2_22_84_2228489, 2_22_84_2228492, 2_22_85_2228519,
2_22_85_2228524, 2_22_85_2228524, 2_22_85_2228567, 2_22_85_2228574, 2_22_86_2228695, 2_22_87_2228722, 2_22_87_2228722, 2_22_87_2228758, 2_22_87_2228759, 2_22_88_2228825,
2_22_88_2228851, 2_22_88_2228885, 2_22_88_2228886, 2_22_89_2228929, 2_22_89_2228967, 2_22_89_2228967, 2_22_89_2228971, 2_22_90_2229010, 2_22_90_2229077, 2_22_91_2229106,
2_22_91_2229140, 2_22_91_2229171, 2_22_91_2229177, 2_22_92_2229219, 2_22_94_2229408, 2_22_94_2229459, 2_22_94_2229472, 2_22_94_2229473, 2_22_95_2229509, 2_22_95_2229533,
2_22_95_2229558, 2_22_95_2229558, 2_22_96_2229647, 2_22_96_2229665, 2_22_98_2229872, 2_22_98_2229872, 2_22_99_2229951, 2_22_99_2229951, 2_22_99_2229993, 2_23_00_2230004,
2_23_00_2230023, 2_23_00_2230095, 2_23_00_2230098, 2_23_01_2230163, 2_23_02_2230227, 2_23_02_2230249, 2_23_02_2230257, 2_23_03_2230396, 2_23_04_2230422, 2_23_04_2230489,
2_23_05_2230565, 2_23_06_2230642, 2_23_07_2230745, 2_23_08_2230813, 2_23_08_2230834, 2_23_10_2231011, 2_23_10_2231064, 2_23_10_2231083, 2_23_12_2231211, 2_23_12_2231271,
2_23_13_2231327, 2_23_13_2231350, 2_23_13_2231350, 2_23_13_2231353, 2_23_13_2231359, 2_23_13_2231363, 2_23_13_2231385, 2_23_14_2231405, 2_23_14_2231478, 2_23_15_2231506,
2_23_15_2231506, 2_23_15_2231537, 2_23_15_2231591, 2_23_15_2231591, 2_23_16_2231627, 2_23_16_2231633, 2_23_16_2231659, 2_23_17_2231714, 2_23_17_2231739, 2_23_17_2231751,
2_23_17_2231791, 2_23_18_2231836, 2_23_18_2231868, 2_23_18_2231885, 2_23_19_2231972, 2_23_19_2231979, 2_23_20_2232015, 2_23_20_2232031, 2_23_20_2232069, 2_23_20_2232080,
2_23_20_2232082, 2_23_21_2232128, 2_23_21_2232194, 2_23_22_2232209, 2_23_22_2232220, 2_23_22_2232267, 2_23_22_2232297, 2_23_22_2232297, 2_23_23_2232316, 2_23_23_2232316,
2_23_23_2232350, 2_23_24_2232464, 2_23_25_2232579, 2_23_25_2232592, 2_23_26_2232610, 2_23_27_2232730, 2_23_27_2232753, 2_23_27_2232753, 2_23_29_2232963, 2_23_29_2232995,
2_23_31_2233127, 2_23_31_2233142, 2_23_31_2233142, 2_23_31_2233178, 2_23_32_2233239, 2_23_32_2233239, 2_23_32_2233275, 2_23_33_2233368, 2_23_34_2233418, 2_23_34_2233418,
2_23_34_2233456, 2_23_34_2233466, 2_23_34_2233477, 2_23_34_2233496, 2_23_35_2233519, 2_23_35_2233579, 2_23_35_2233587, 2_23_36_2233607, 2_23_36_2233619, 2_23_37_2233769,
2_23_39_2233954, 2_23_40_2234060, 2_23_40_2234060, 2_23_41_2234120, 2_23_41_2234133, 2_23_41_2234135, 2_23_43_2234323, 2_23_43_2234334, 2_23_43_2234393, 2_23_43_2234399,
2_23_44_2234400, 2_23_44_2234401, 2_23_44_2234459, 2_23_44_2234475, 2_23_45_2234529, 2_23_45_2234540, 2_23_45_2234540, 2_23_45_2234584, 2_23_47_2234795, 2_23_47_2234799,
2_23_47_2234799, 2_23_48_2234809, 2_23_48_2234864, 2_23_49_2234932, 2_23_49_2234959, 2_23_49_2234959, 2_23_50_2235004, 2_23_50_2235065, 2_23_51_2235145, 2_23_52_2235201,
2_23_52_2235206, 2_23_52_2235252, 2_23_53_2235304, 2_23_53_2235311, 2_23_53_2235367, 2_23_53_2235393, 2_23_54_2235436, 2_23_55_2235503, 2_23_56_2235618, 2_23_57_2235701,
2_23_57_2235774, 2_23_57_2235775, 2_23_57_2235775, 2_23_60_2236031, 2_23_61_2236149, 2_23_62_2236211, 2_23_62_2236252, 2_23_62_2236261, 2_23_62_2236288, 2_23_64_2236400,
2_23_64_2236400, 2_23_64_2236445, 2_23_64_2236478, 2_23_65_2236515, 2_23_65_2236515, 2_23_67_2236722, 2_23_67_2236739, 2_23_67_2236754, 2_23_67_2236793, 2_23_68_2236870,
2_23_69_2236911, 2_23_69_2236931, 2_23_69_2236934, 2_23_69_2236946, 2_23_69_2236954, 2_23_69_2236954, 2_23_69_2236988, 2_23_70_2237047, 2_23_71_2237133, 2_23_71_2237179,
2_23_72_2237213, 2_23_72_2237236, 2_23_72_2237254, 2_23_72_2237254, 2_23_72_2237257, 2_23_72_2237286, 2_23_73_2237309, 2_23_74_2237493, 2_23_75_2237507, 2_23_75_2237508,
2_23_75_2237508, 2_23_75_2237549, 2_23_76_2237651, 2_23_76_2237656, 2_23_76_2237659, 2_23_76_2237668, 2_23_76_2237677, 2_23_76_2237677, 2_23_77_2237788, 2_23_79_2237912,
2_23_79_2237930, 2_23_80_2238050, 2_23_80_2238057, 2_23_81_2238122, 2_23_81_2238156, 2_23_81_2238187, 2_23_81_2238196, 2_23_82_2238225, 2_23_82_2238286, 2_23_83_2238331,
2_23_83_2238396, 2_23_84_2238481, 2_23_85_2238505, 2_23_86_2238687, 2_23_87_2238738, 2_23_88_2238828, 2_23_88_2238828, 2_23_88_2238872, 2_23_89_2238930, 2_23_89_2238940,
2_23_89_2238947, 2_23_89_2238997, 2_23_90_2239003, 2_23_90_2239083, 2_23_90_2239083, 2_23_91_2239134, 2_23_91_2239148, 2_23_92_2239252, 2_23_92_2239252, 2_23_92_2239274,
2_23_92_2239287, 2_23_94_2239428, 2_23_95_2239576, 2_23_95_2239576, 2_23_95_2239577, 2_23_95_2239577, 2_23_96_2239623, 2_23_97_2239783, 2_23_98_2239870, 2_23_99_2239928,
2_23_99_2239952, 2_23_99_2239952, 2_24_00_2240035, 2_24_00_2240036, 2_24_00_2240036, 2_24_00_2240071, 2_24_00_2240071, 2_24_01_2240175, 2_24_02_2240238, 2_24_03_2240301,
2_24_03_2240308, 2_24_03_2240309, 2_24_03_2240391, 3_50_00_3500006, 3_50_00_3500007, 5_00_00_5000078, 5_00_01_5000125, 5_00_01_5000129, 5_00_02_5000250, 7_00_00_7000085,
7_00_06_7000660, 7_00_06_7000662, 7_00_07_7000720, 7_00_11_7001176, 7_00_11_7001180, 7_00_11_7001182, 7_00_11_7001183, 7_00_11_7001184, 7_00_18_7001822, 8_00_00_8000048,
8_10_03_8100374, 8_10_04_8100475, 8_10_05_8100556, 8_10_05_8100557, 8_10_05_8100567, 8_10_06_8100624, 8_10_07_8100722, 8_10_07_8100763, 8_10_13_8101353, 8_10_14_8101437,
8_10_14_8101481, 8_10_14_8101497, 8_10_15_8101544, 8_10_15_8101544, 8_10_16_8101641, 8_10_17_8101721, 8_10_17_8101725, 8_10_17_8101778, 8_10_17_8101778, 8_10_18_8101800,

8_10_18_8101840, 8_10_19_8101994, 8_10_20_8102070, 8_10_20_8102070, 8_10_21_8102159, 8_10_21_8102184, 8_10_21_8102187, 8_10_22_8102228, 8_10_23_8102330, 8_10_24_8102411, 8_10_24_8102447, 8_10_24_8102481, 8_10_25_8102515, 8_10_25_8102515, 8_10_26_8102647, 8_10_27_8102791, 8_10_28_8102804, 8_10_28_8102866, 8_10_30_8103073, 8_10_31_8103163, 8_10_32_8103270, 9_01_12_9011247

Supplementary data 2: Quantum calculation methods and results

(A) Model systems used



A schematic of the hydrogen bond, where the Donor, Acceptor, Acceptor antecedent and hydrogen are labeled as D (red), A(cyan), C(green) and H respectively. Solids lines represent covalent bonds and the dashed line represents the hydrogen bonds. The $\hat{H}$ angle is the angle formed between atoms D, A and C.

The four different kinds (depending on whether the donor/acceptor belong to the main-chain or side-chain) of hydrogen bonds in proteins and the following model systems have been employed to study these bonds *in silico*. The minimum energy geometries of these models were obtained by carrying out optimization using Møller-Plesset second order perturbation theory (MP2) and 6-31G(d,p)++ basis set.

1. Side Chain-Side Chain (SC-SC): To model hydrogen bond interaction between two side chain residues of a protein, hydrogen form a donor, either an amine or an alcohol, is donated to an acceptor alcohol molecule (Fig. 1). Since an amine rarely accepts hydrogen, amines have not been used to model an acceptor. The two models for a SC-SC interaction are (A) Methanol donating to Methanol; (B) Methylamine donating to Methanol.

Note: Our model systems of amino side chains are representative of most side chains, with the exception of side chains such as Asparagine or Glutamine that resemble the main chain.

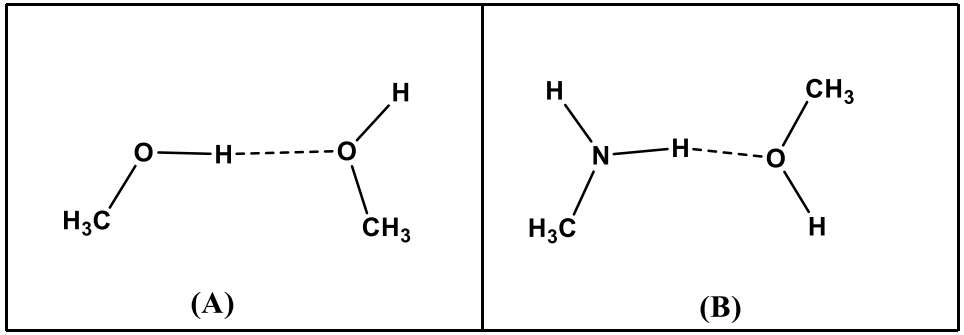


Figure S2.1. Model system used for studying side chain-side chain interactions

2. Main Chain-Side Chain (MC-SC): To mimic the backbone of a protein, N-Methylacetamide (NMA) is employed as a model system. The hydrogen bond interaction between a main chain as a donor and a side chain as an acceptor is studied by N-H donation from NMA to the oxygen of acceptor methanol.

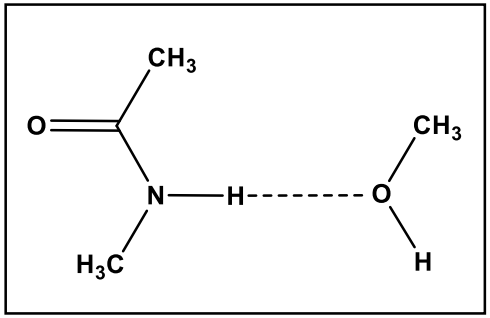


Figure S2.2. Hydrogen bond between a donor main chain (N-Methylacetamide) and an acceptor side chain (Alcohol)

3. Side Chain- Main Chain (SC-MC): The hydrogen bond interaction between a side chain as donor and the main chain as an acceptor is studied by hydrogen donation of methylamine to the carbonyl of NMA.

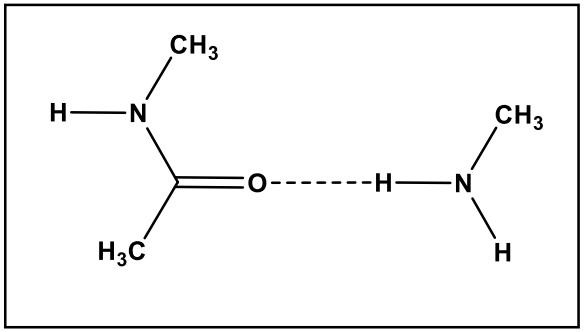


Figure S2.3. Hydrogen bond between a donor side chain (methylamine) and an acceptor main chain (N-Methylacetamide)

4. Main Chain-Main Chain (MC-MC): The hydrogen bond interaction between backbones of the proteins is studied by N-H donation from NMA to carbonyl acceptor of NMA.

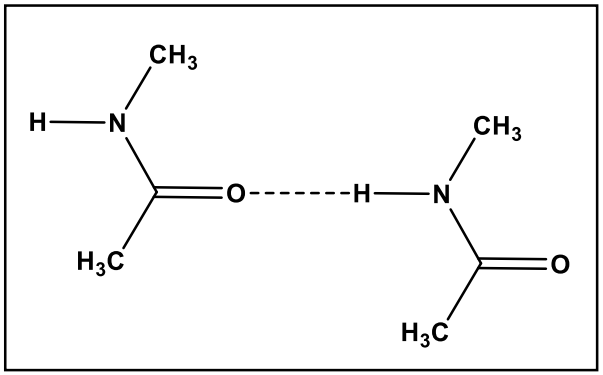


Figure S2.4. Hydrogen bond interaction between two main chain residues

(B) Sampling coordinate space

As stated in the main article, the minimum energy, optimized structure obtained is dependent on the initial structure that one begins with. We therefore began with optimizations of randomly chosen initial geometries to arrive at an optimized geometry. The optimized structure that had (1) hydrogen bond as a primary interaction, measured by its DHA angle being greater than 150°, (2) minimal secondary interactions between the dimers (the distance between any other pair of atoms between donor and acceptor was greater than 2.3 Å) and (3) trans amide bonds was chosen for further calculations. This *native* structure served as the basic unit for further calculations. The italicized entries of DHA angle and $\hat{H}$ angle in the table refer to this native initial geometry upon which sampling is carried out.

To ensure that the whole co-ordinate space has been sampled, the $\hat{H}$ angle was varied from 85° to 160° in steps of 15° and optimized as follows: (a) the $\hat{H}$ angle was constrained to a particular angle (Constrained $\hat{H}$ angle in Table S2.1) and the structure optimized. (b) The constraints on the optimized structure from step (a) were then released and the geometry fully optimized. The relaxed $\hat{H}$ angle and relaxed DHA in column 3 and 4 are angles of this final optimized geometry.

The two step sampling was carried out for all the four kinds of hydrogen bond interactions. It is noted that irrespective of constrained $\hat{H}$ angle that one begins with, the relaxed angle generally converges to one value representative of the minimum energy geometry of the molecule. For instance, in an interaction where main chain is the donor and side chain is the acceptor, sampling the space, from 85° to 160° yields one structure with $\hat{H}$ angle and DHA angle of about 112° and 168°, respectively.

Case 3 (main chain as an acceptor and side chain as a donor) has in fact two minimum energy structures, both of which satisfy the criterion of DHA angle being greater than 150°. Since both model 1 and 2 have DHA angles varying by less than 2°, sampling was carried out for both these geometries. Model 1 in the table has $\hat{H}$ angle of 101.5° while model 2 in the table has $\hat{H}$ angle of 124°. Model 1 and 2 structurally are different. Model 2 has the N-CH_3 group of the amide close to the CH_3 group of the amine like in Figure S2.3 while model 1 has C-CH_3 group of the amide closer to the CH_3 group of the amine. The average DHA angle and $\hat{H}$ angle for SC-MC is 157.2° and 112.7° respectively.

Case 4 (main chain as acceptor and as donor) yielded various values for $\hat{H}$ angle as listed in the table. However, only those structures were taken for further calculations that had the most linear hydrogen bond and therefore the largest DHA angle (170.2°). The $\hat{H}$ angle corresponding to this structure is 138°.

Table S2.1 Sampling of co-ordinate space for initial structure in the geometry optimization calculation

	Constrained $\hat{H}$ angle (°)	Relaxed $\hat{H}$ angle (°)	Relaxed DHA angle (°)
1. Side chain as donor and acceptor			
Alcohol- Alcohol			
	*Native $\hat{H}$ angle: **107.6**; DHA Angle: **172.4***		
	85	107.5	171.9
	100	107.3	172.0
	115	107.6	172.6
	130	107.7	172.5
	145	107.2	172.1
	160	107.3	172.2
Alcohol(A)- Amine(D)			
	*Native $\hat{H}$ angle: **92.4**; DHA Angle: **139.6***		
	85	92.5	139.2
	100	92.4	138.9
	115	92.3	138.9
	130	92.4	139.1
	145	92.0	140.5
	160	91.7	142.1

2. Main chain as donor and side chain as acceptor			
	*Native $\hat{H}$ angle: **112.5**; DHA Angle: **168.3***		
	85	112.9	169.0
	100	112.5	168.1
	115	112.6	168.0
	130	112.5	168.1
	145	112.7	168.2
	160	112.4	168.0
3. Main chain as acceptor and side chain as donor			
Model 1	*Native $\hat{H}$ angle: **101.5**; DHA Angle: **156.5***		
	85	101.3	156.2
	100	101.5	156.5
	115	101.6	156.5
	130	101.8	156.4
	145	89.9	139.4
	160	124.3	157.1
Model 2	*Native $\hat{H}$ angle: **124**; DHA Angle: **157.9***		
	85	123.8	158.2
	100	123.6	158.1
	115	123.6	157.9

	130	124.2	156.8
	145	123.8	157.0
	160	123.8	157.1
4. Main chain as donor and acceptor			
	Native $\hat{H}$ angle: ***138****; DHA Angle:* ***170.2***		
	85	103.8	149.1
	100	103.8	149.1
	115	103.8	149.1
	130	136.4	168.1
	145	114.8	156.1
	160	136.1	169.6

SUMMARY

Tabulated below are the details on the minimum energy geometries of the four kinds of hydrogen bond. D-H is hydrogen bond donor. HA Distance refers to the distance between the hydrogen donated and the acceptor atom A. Distance of closest secondary interaction is the smallest distance between an atom on donor in closest proximity to any atom on the acceptor unit.

	Donor	Acceptor	$\hat{H}$ angle (°)	DHA angle (°)	HA distance (Å)	Distance of closest secondary interaction (Å)

Side chain-Side Chain						
	Alcohol	Alcohol	107.6	172.4	1.89	2.49
	Amine	Alcohol	92.4	139.1	2.18	2.71
		Average	100.0	155.7	2.03	2.60
Main Chain- Side Chain						
	NMA	Alcohol	112.5	168.3	1.97	2.74
Side Chain- Main Chain						
Model 1	Amine	NMA	124	157.9	2.13	2.34
Model 2	Amine	NMA	101.5	156.5	2.14	2.43
		Average	112.8	157.2	2.13	2.38
Main Chain- Main Chain						
	NMA	NMA	138	170.2	1.96	2.48

(C) Explaining the trend in $\hat{H}$ angle

Quantum mechanical calculations were performed on two sets of molecular systems, namely (i) N-methyl acetamide (NMA) and (ii) di-Glycine peptides as proxy to protein main-chain main-chain hydrogen bonds. Figure S2.5 shows the original, saturated planar and saturated optimized structures of NMA system, while Figure S2.6 shows the Gly-Gly dipeptide dimers (original, saturated planar and saturated optimized structures with the electron density mapped). It can be seen that in original structure the electron density is uniformly

distributed around donor N-atom, while in the saturated structure the distribution is skewed toward one side of the peptide bond. Interestingly, in the saturated planar structure, the electron density around the donor N-atom is similar to the original structure, while it is distinctly different from the saturated optimized structure. This suggests that the breaking of planarity of the peptide bond qualitatively affects the distribution of electron density on the N atom. The geometrical parameters of the simulated system are given in Table S2.2.

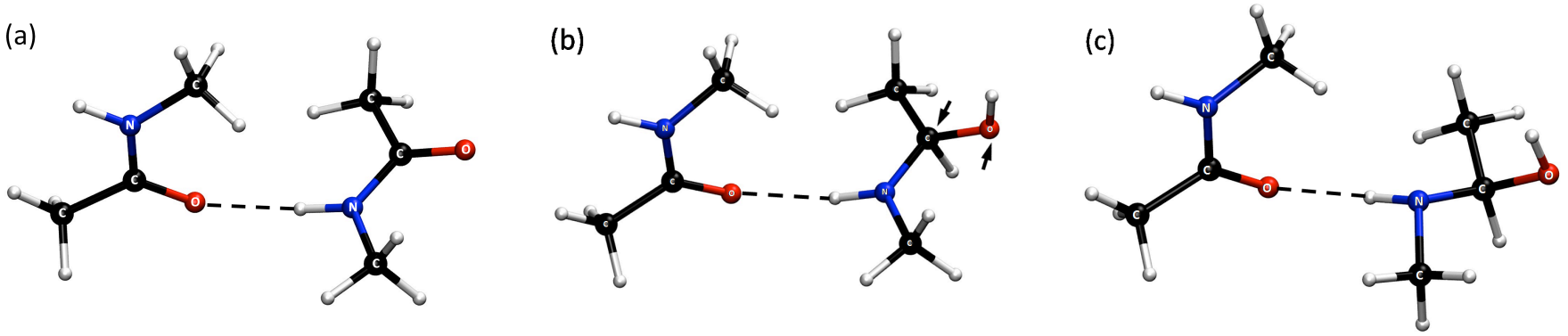


Figure S2.5 (a) Original structure of NMA dimer that is a model for MC-MC type of hydrogen bond, (b) Saturated planar structure and (c) the Saturated optimized structure which is a model for the SC-MC type of hydrogen bond. All models are in ball and stick representation with Carbons in black, Oxygens in red, Nitrogens in blue and Hydrogens in white.

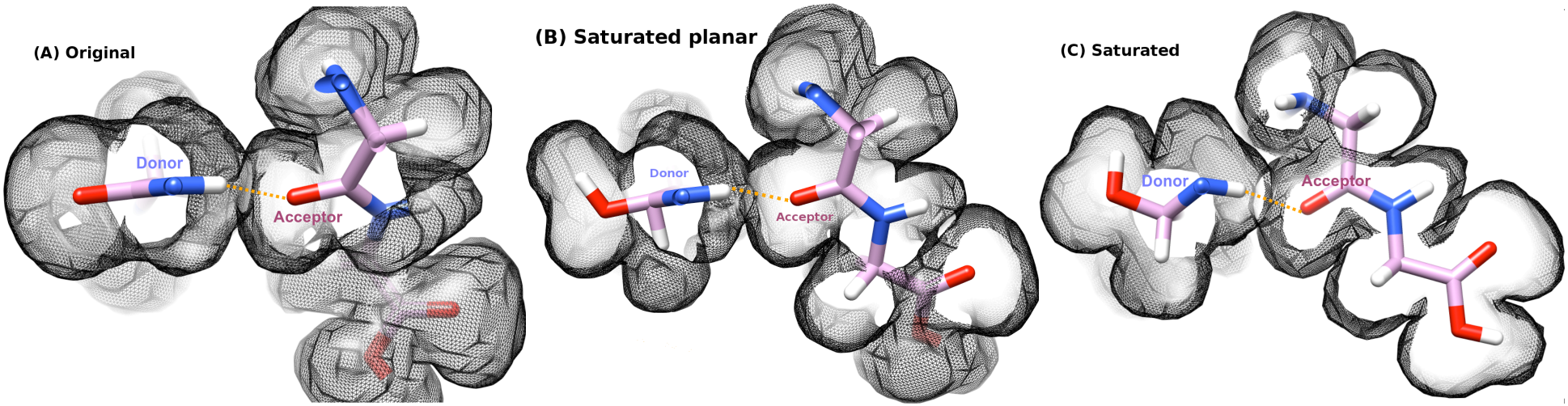


Figure S2.6. Electron density distribution in GG-dipeptide dimers of (a) Original structure (b) Saturated planar structure and (c) Saturated optimized structure. The dipeptides are shown in stick representation where the oxygen, nitrogen, carbon and hydrogen atoms are coloured res, blue, pink and white respectively. The orange dotted line represents the hydrogen bond. Electron density map is calculated using MOLDEN (Schaftenaar and Noordik, 2000), and the figures were rendered using UCSF Chimera (Pettersen et al, 2004).

Table S2.2. Hydrogen bond angle (DHA), donor-acceptor distance, $\hat{H}$ angle, Nitrogen-Carbon distance, hydrogen-acceptor distance and partial charges in the original optimized and saturated optimized structures (angles are reported in (°) and distances in (Å)).

	Original structure (Model for MC-MC hydrogen bond)	Saturated optimized structure (Model for SC-MC hydrogen bond)
_N_-Methyl Acetamide		

Hydrogen bond angle (DHA)	170.2	173.4
$\hat{H}$	138.0	120.7
NO distance	2.97	3.04
NC distance	3.98	3.83
Hydrogen bond length	1.96	2.03
Mulliken partial charges		
N of donor	-0.42	-0.44
H of donor	0.38	0.29
O of acceptor	-0.54	-0.53
C of acceptor	0.46	0.38
Glycine di-peptide dimer Anti-parallel beta-sheet model 1 (Ψ=140º, ϕ=-120º)		
DHA	170.2	174.1
$\hat{H}$	150.3	131.2
NO distance	2.93	3.04
NC distance	4.05	3.96
Hydrogen bond length	1.92	2.02
Mulliken partial charges		
N of donor	-0.27	-0.33

H of donor	0.45	0.32
O of acceptor	-0.61	-0.57
C of acceptor	0.19	0.25
Glycine di-peptide dimer Anti-parallel beta-sheet model 2 (Ψ=140º, ϕ=-100º)		
DHA	175.7	170.6
$\hat{H}$	147.8	133.7
NO distance	2.93	3.13
NC distance	4.03	4.08
Hydrogen bond length	1.91	2.12
Mulliken partial charges		
N of donor	-0.29	-0.29
H of donor	0.46	0.34
O of acceptor	-0.60	-0.56
C of acceptor	0.19	0.17

Supplementary data 3: Aggregated geometrical distribution of hydrogen bonds in proteins

The aggregated distribution of hydrogen bond geometry, represented by donor-acceptor distance and donor-acceptor-antecedent angle ($\hat{H}$ angle) geometry of hydrogen bonds in proteins were investigated. The preferred geometry of hydrogen bonds are shown as the log-odd ratio as compared to an ideal gas background model (Supplementary figure S3).

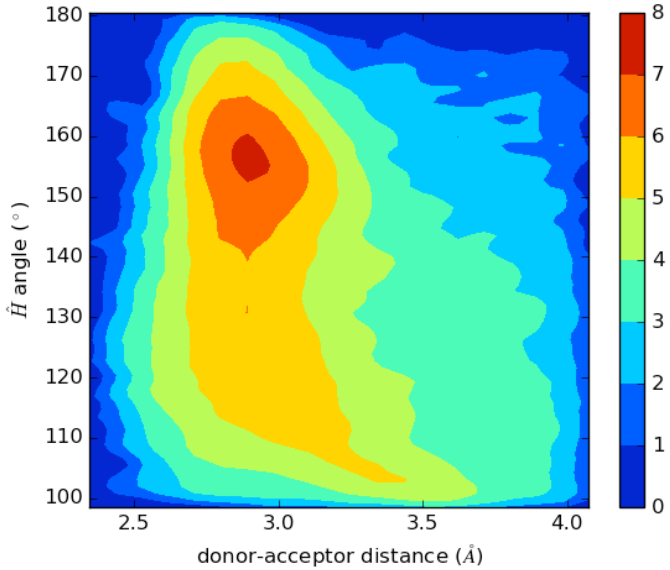


Figure S3.The average geometry ($\hat{H}$ angle – donor-acceptor distance) of hydrogen bond in proteins is depicted with a heat-map of normalized log-odd of occurrence compared to background distribution $\log(N^{obs}/N^{exp})$. A 0.05 Å and 1° standard deviation for bond distance and $\hat{H}$ angle was applied for Gaussian blurring of the frequency counting for plotting purpose. The heat map is coloured red to blue for the higher and lower log odd ratios respectively.

Supplementary data 4: Case study on Hen Egg White Lysozyme

Both high-resolution X-ray structure (PDB:193L) and hydrogen-deuterium exchange (HDX) experiments under physiological conditions (Radford et al, 1992) are available for Hen Egg White Lysozyme (HEWL). It is widely known that HDX exchange rates of amides depend on their solvation and hydrogen bonding state. It has been shown that residue depth correlates better with hydrogen-deuterium exchange rates as compared to solvent accessibility (Chakravarty and Varadarajan, 1999), and hence the measure was used to gauge solvation. A depth value of 5.5 Å (~ 2 hydration shells) approximately separates the protected (exchange rates of 0.008 s^{-1} or slower) and unprotected amides (Figures S4) and we have used this cutoff to label amides as either exposed or buried.

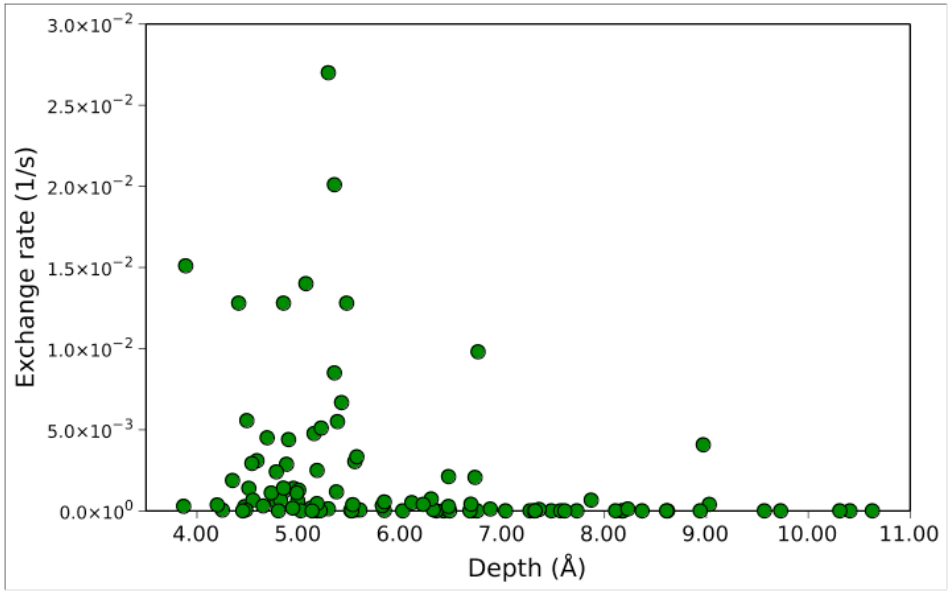


Figure S4. Scatter-plot of HDX exchange rate as a function of residue depth in the dataset.

We tested the effect on relaxation of $\hat{H}$ angle criteria from 120° to 100°. In the HEWL high-resolution structure, we found that 8 amide donors would have been categorized as non-hydrogen bonded if 120° was used as $\hat{H}$ angle cut-off (Table S4). 4 out of 8 amide donors are buried in protein interior (depth > 5.5 Å), while the other 4 were exposed on protein surface. These exposed amide groups are all protected from hydrogen exchange, which strongly suggest that they are hydrogen bonded.

Residue	Exchange rate	Depth (Å)	Exposed	Bond-type	donor-acceptor distance (Å)	$\hat{H}$ angle
78	3.62E-04	4.20	TRUE	MC-SC	2.99	111.7
42	2.56E-04	4.47	TRUE	MC-MC	3.27	118.5
23	3.09E-03	4.59	TRUE	MC-MC	3.04	119.3

41	1.28E-02	5.47	TRUE	MC-SC	2.97	117.9
39	5.50E-05	5.60	FALSE	MC-SC	3.42	101.5
111	4.07E-04	6.69	FALSE	MC-MC	3.04	118.1
57	1.79E-06	7.03	FALSE	MC-MC	3.36	113.7
28	4.39E-07	8.62	FALSE	MC-MC	3.05	119.3

Table S4. List of hydrogen bonds with $\hat{H}$ angle less than 120° in HEWL.

Supplementary data 5: Statistical significance of hydrogen bond donor-acceptor distance and $\hat{H}$ angle among different classes of hydrogen bonds

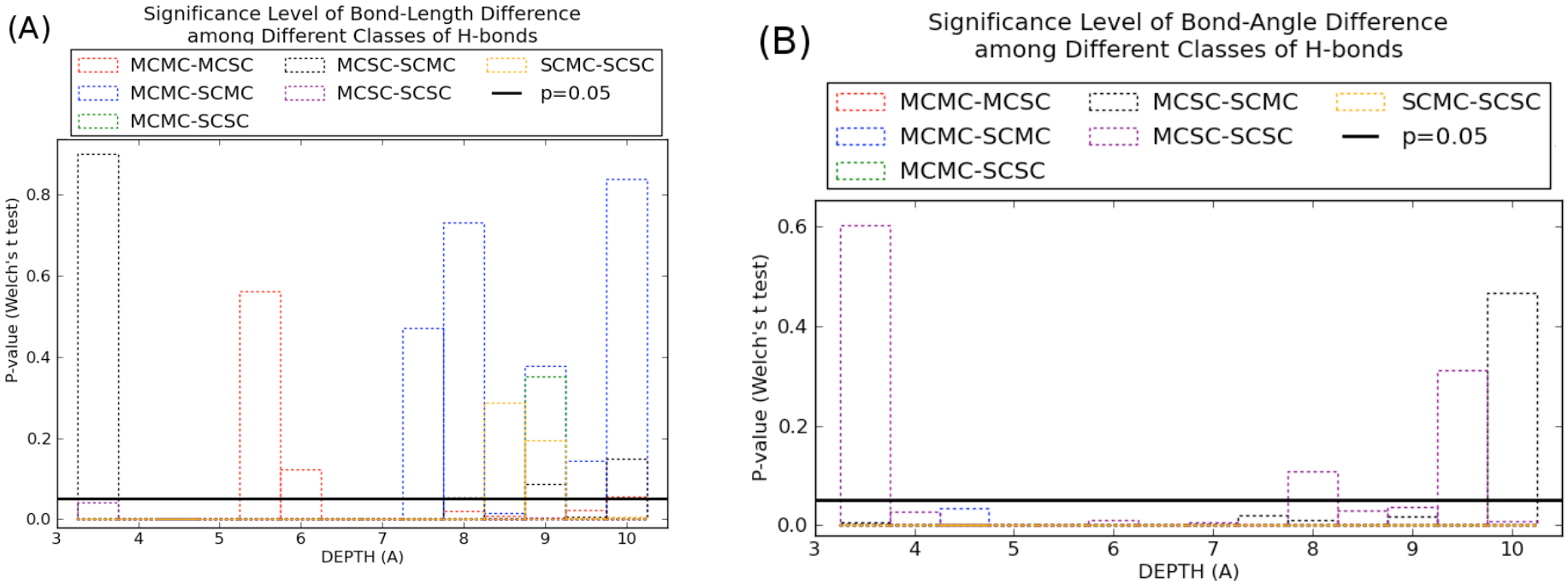


Figure S5. Statistical significance of difference as measured in p-value in (A) donor-acceptor distance and (B) $\hat{H}$ angle in difference classes of hydrogen bonds, at different levels of depth.

Supplementary data 6: Statistical significance of donor-acceptor distance of hydrogen bond in alpha-helix and beta-strand

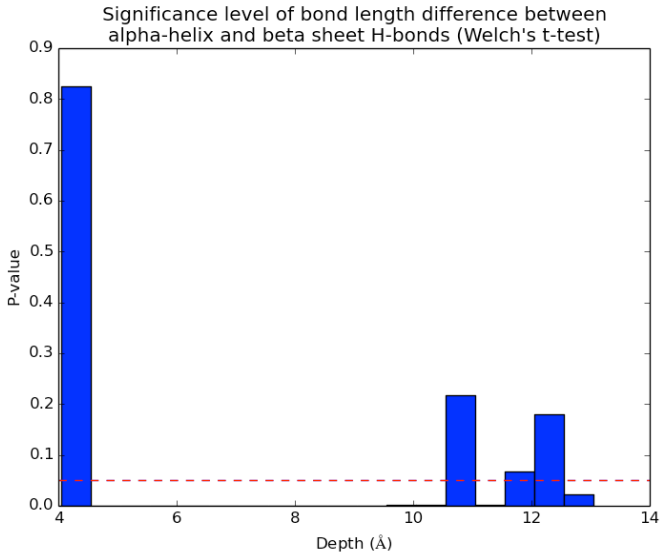


Figure S6. Statistical significance of donor-acceptor distance between hydrogen bonds in alpha-helix and beta-strand secondary structure, as measured by p-value at various levels of depth.

Supplementary data 7: Hydrogen bond $\hat{H}$ angle in small peptides

This following table lists 22 hydrogen bonds found in 16 structures deposited in Cambridge structure database. The first 4 hydrogen bonds (in structure TEHJAR, GESLUL) involves GLY residues.

Table S7. Donor-acceptor distance and $\hat{H}$ angles of intermolecular hydrogen bond in small peptides deposited in Cambridge structure database.

No	CODE	DONOR	ACCEPTOR	ANTECEDENT	donor-acceptor distance (Å)	$\hat{H}$ angle (°)
1	TEHJAR	37	1	42	3.05	124.49

2	TEHJAR	29	7	78	2.97	140.91
3	GESLUL	37	12	11	3.03	138.18
4	GESLUL	5	44	43	2.86	139.51
5	FUDGAM	55	25	6	3.02	141.00
6	FUDGAM	24	56	37	2.78	146.79
7	RELWIO	32	1	13	2.91	151.37
8	ZZZIFQ01	26	63	40	3.33	152.74
9	ZZZIFQ01	61	31	14	3.16	155.71
10	BEVYIL	6	24	39	2.94	152.89
11	BEVYIL	28	2	17	2.90	158.78
12	XEPZAU	58	32	31	2.96	156.72
13	VISSAR	39	22	21	3.03	156.95
14	VOHFAA	56	44	16	2.89	157.92
15	COCGIK	8	40	61	3.10	158.55
16	RELWIO	7	26	38	2.81	158.68
17	NAFZID	81	115	125	2.89	156.71
18	VOHFAA	41	87	86	3.00	161.30
19	VUNFEQ	35	59	60	3.10	141.87
20	VUNFEQ	89	5	8	2.83	161.42
21	XEPZAU	10	67	66	2.95	162.03
22	VISSAR	10	51	50	2.99	163.22

COD_Id	Donor	Acceptor	Donor-Acceptor distance (Å)	HB_angle (degrees)
1546040	N_12	O_23	2.8	148.9
1546040	N_2	O_13	2.8	156.8
1546040	N_4	O_1	3.0	128.6
2012623	N3	O0P	3.2	147.1
7115368	N213	O106	3.1	120.7
7115368	N407	O316	2.8	158.0
7115369	N110	O219	2.8	142.2
7150102	N27A	O7B	3.2	164.6
7150102	N8A	O26B	3.0	138.5
7150102	N8C	O26A	3.1	134.4
7150405	N(17)	O(3A)	3.0	155.7
7150405	N(7)	O(13A)	2.9	172.6
7207943	N3	O2	2.9	148.3
7207943	N5	O3	3.0	166.3
7207943	N6	O8	2.9	149.2
7207944	N2	O4	2.9	150.1
7216422	N10	O24	3.0	136.4
7216422	N2	O15	2.9	160.4
7216422	N3	O23	3.0	163.5
7216422	N6	O12	3.0	165.8
7216422	N8	O4	2.9	152.6
7216423	N2	O30	2.9	160.3
7216423	N3	O16	2.9	148.4
7216423	N4	O18	2.9	142.6
7216423	N5	O12	2.9	144.0
7216423	N6	O11	2.9	157.0

Table S7A. Donor-acceptor distance and $\hat{H}$ angles of intermolecular hydrogen bond in tri-peptides deposited in Crystallography Open Database.

In selecting hydrogen bonds between peptides from the databases, we manually removed those that involved modified amino acids, additional links to close proximity to amino acids and/or close to the capping groups of the peptide N and C termini. The hydrogen bonds in this list are almost always between two amide N and carbonyl O of a pair of (the 20) naturally occurring amino acids.

Supplementary data 8: Comparison of melting temperature between all-alpha proteins and all-beta proteins

Experimentally determined melting temperature data were collected from the Protherm database (Kumar et al, 2006). Only wild-type proteins (*i.e.* no mutations) were chosen. As the pH value of buffer has a substantial effect on melting temperature, we choose only thermal denaturation data from pH range of 6 - 8. These data corresponds to a total of 326 structures with 3D structure solved and deposited in the PDB. The dataset consists of 53 all-alpha proteins, 46 all-beta proteins and 227 others.

The melting temperature of all-alpha proteins is 58.11°C (±13.04°C) and all-beta proteins have a slightly higher mean value at 62.47°C (±14.18°C). Independent t-test result showed that the difference is not significant (p ~=0.12). However from histogram it was observed that all-beta protein has a heavy-tailed distribution at the high temperature and a secondary peak at ~= 70°C (Figure S8). Fisher's exact test showed the association of all-alpha/all-beta and heat resistance (using threshold temperature of 70°C) is statistically significant ($p < 0.05$). In this test, we have also shown that the density of hydrogen bonds (hydrogen bond per residue) is not correlated with melting temperature, for both all-alpha (correlation coefficient ~ 0.03) and all-beta proteins (correlation coefficient ~ -0.03).

	<70°C	>=70°C	Total
All-alpha	45	8	53
All-beta	29	17	46
Total	74	25	99

Table S8. Fisher's exact test on heat-resistant (melting temperature >70°C) with all-alpha / all-beta protein. p-value of the test ~= 0.02.

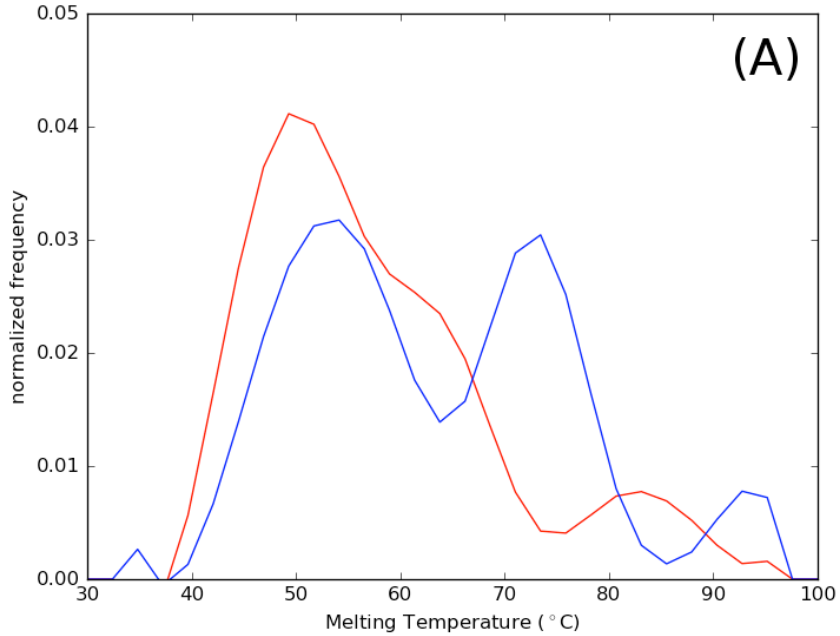


Figure S8. Comparison of distribution of melting temperatures of all-alpha (red) and all-beta (blue) proteins in Protherm database.

Supplementary data 9: Hydrogen bond donor-acceptor distance and secondary structure preference

We investigated into the relationship between the hydrogen bond donor-acceptor distance and donor amino acid residue secondary structure preference. Mean donor-acceptor distance of each type of amino acid was computed from our high resolution data training set, while alpha-helix and beta-sheet propensities of all amino acid types were obtained from Chou-Fasman table (Chou and Fasman, 1974). We found that helical propensity (as compared to strand propensity) has a correlation coefficient of ~0.55 (Figure S9).

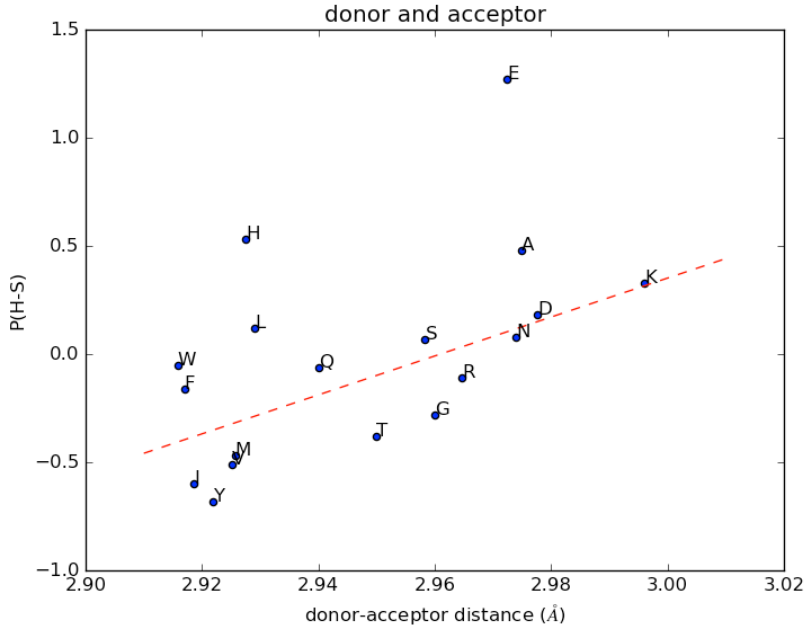


Figure S9. MC-MC donor-acceptor distances involving different amino acid types (labeled by their single letter codes) and their secondary structure preferences. The y-axis is the difference in propensities between helix and sheet of a particular amino acid residue. The red dashed line shows the best-fit line to the data set (His and Glu are excluded from the fitting).

Supplementary data 10: Modeling alpha helices at different $\hat{H}$ angles

A 24mer poly-Alanine was attempted to fold into an alpha-helix starting from an initial extended conformation using different value of $\hat{H}$ angles. In the resulting structures, regardless of initial restraining angles, in 7 simulations the $\hat{H}$ angle values were able to converge to ~150° and form stable helices. In the other 4 cases helices are not formed during the optimization.

The Ramachandran map used in this study was produced with the dihedral angles statistics of a high resolution (resolution < 2 Å, R-factor < 0.25), non-redundant (sequence similarity < 30%), single domain (chain length 100 – 200 residues) training set of 1175 protein structures. The figure is rendered using Gnuplot with Gaussian kernel density with bandwidth of 3°.

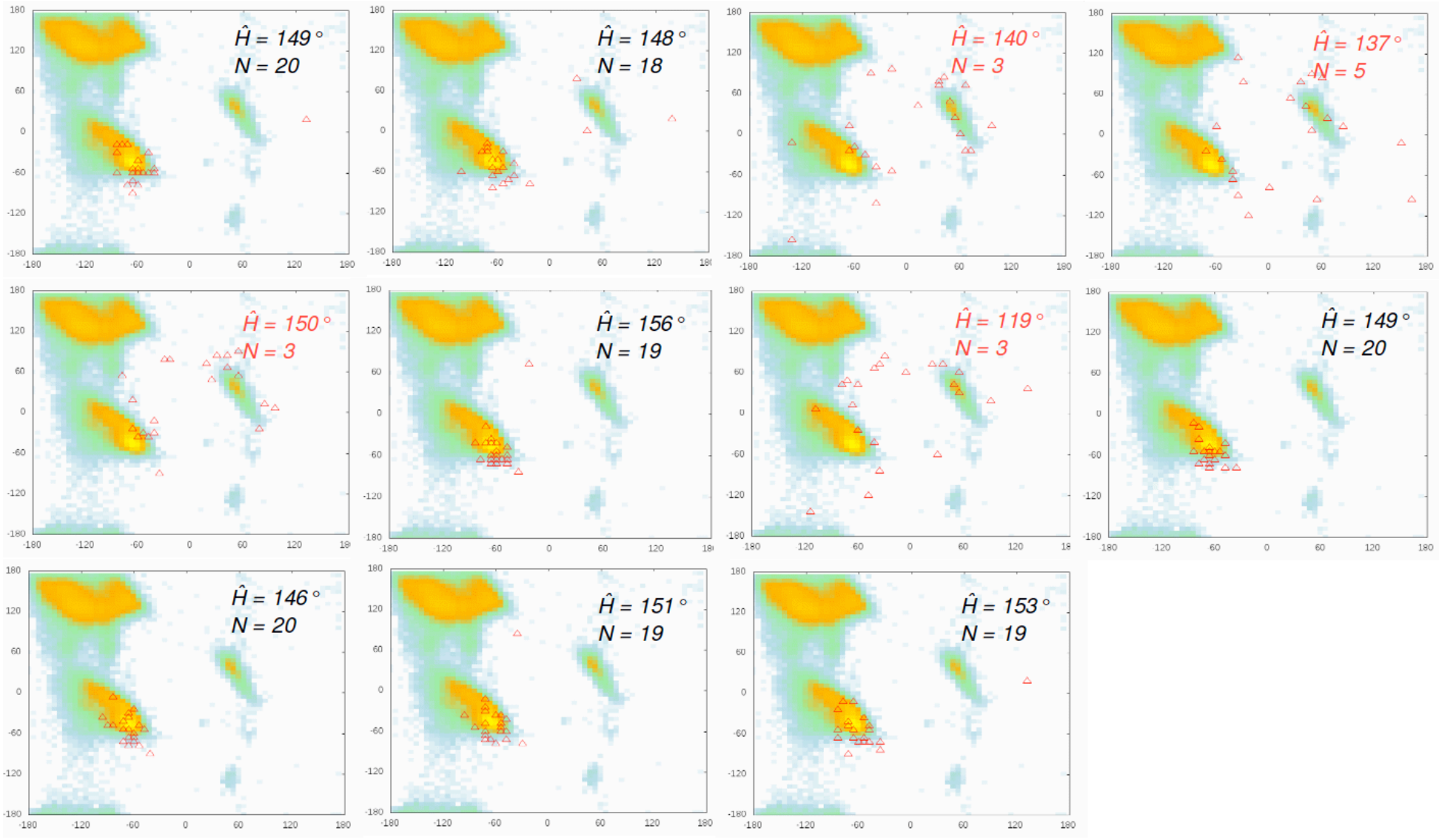

Ĥ = 149 °
N = 20
Ĥ = 148 °
N = 18
Ĥ = 140 °
N = 3
Ĥ = 137 °
N = 5
Ĥ = 150 °
N = 3
Ĥ = 156 °
N = 19
Ĥ = 119 °
N = 3
Ĥ = 149 °
N = 20
Ĥ = 146 °
N = 20
Ĥ = 151 °
N = 19
Ĥ = 153 °
N = 19

Figure S10. Ramachandran plots (red triangles) of all 11 helices folding simulation experiments. The density of the map is color-coded from blue (low density region) to yellow (high density region). Legends indicate the mean $\hat{H}$ angle of hydrogen bond forming helix (i.e. N_{i+4} -> O_i if the two atoms fall within bonding distance and angle) and number of hydrogen bond formed in the system. Legend texts of the 4 cases that failed to form an alpha-helix were colored red.

Supplementary data 11: Aggregated geometrical distribution of hydrogen bonds in proteins

The allowed $\hat{H}$ angle in beta-sheets were simulated by sliding a pair of parallel and anti-parallel along the two axes (x- and y- axes) (Figure S11.1). The allowed geometries (w/o atomic clashes) are shown in Figure S11.2.

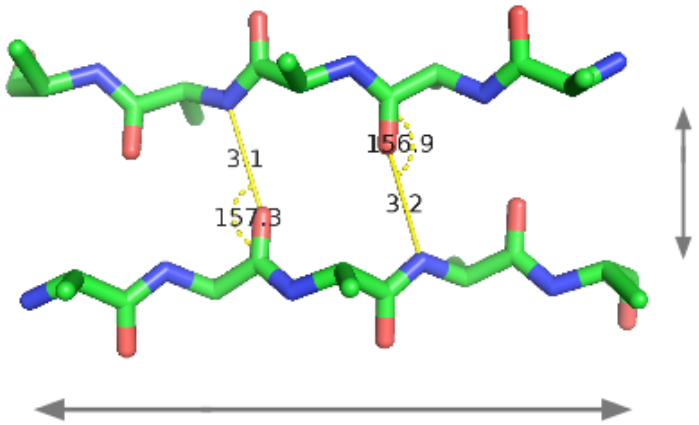


Figure S11.1. Schematic diagram showing the empirical simulation of sliding two parallel beta-sheets alongside one another. The atoms of the sheet are shown in stick representation where the C, N and O atoms are coloured green, blue and red respectively. A couple of the central hydrogen bonds are shown in yellow dotted lines connecting the donor and acceptor atoms. The corresponding $\hat{H}$ angle is also indicated.

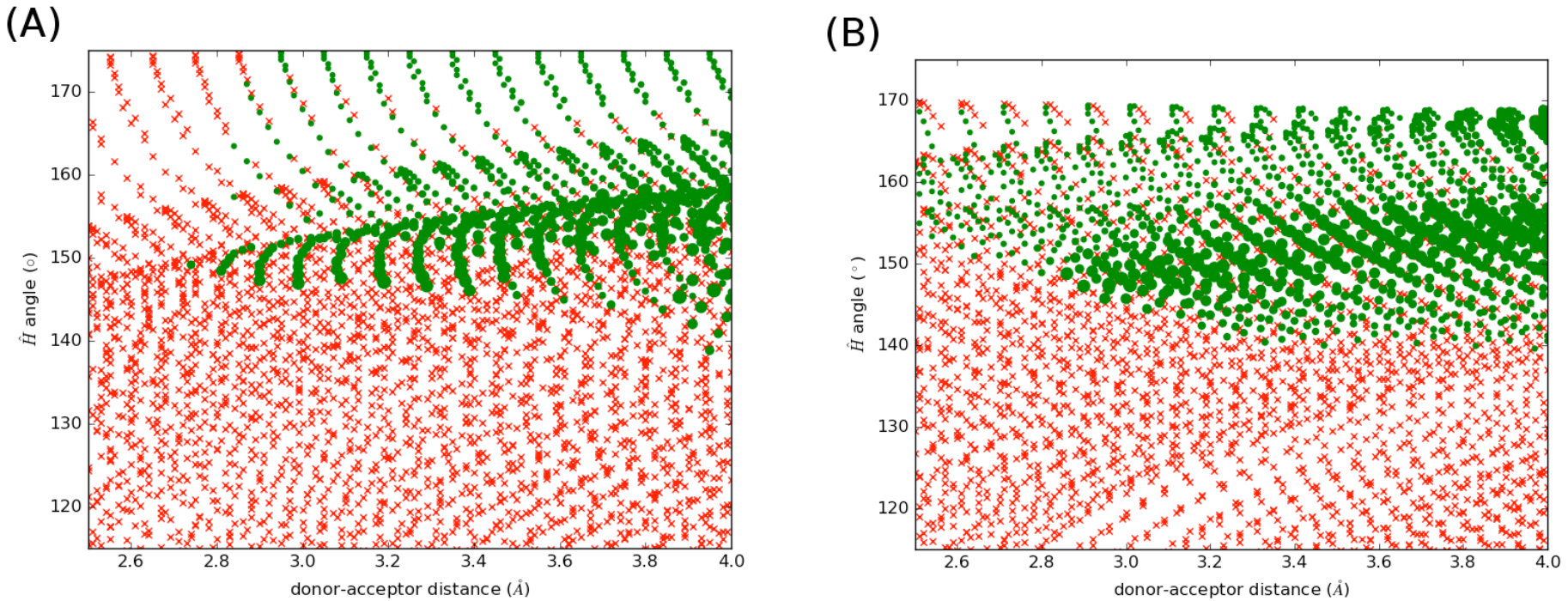


Figure S11.2. Allowed donor-acceptor distance and $\hat{H}$ angle configuration in a planar (A) anti-parallel beta-sheet and (B) parallel beta-sheet generated from simulations. Structures with fully formed hydrogen bonds and steric clashes are indicated with green and red points respectively.

Supplementary data 12: Molecular dynamics simulation on alpha-helix

Three sets of molecular dynamics simulations were performed on a 16-mer poly Alanine to probe into the relationship between hydrogen bond $\hat{H}$ angle and backbone dihedral angle of an alpha-helix. The three set of simulations sampled $\hat{H}$ angle and dihedral angle at different value regions, as shown in Supplement figure S12.1 ($\hat{H}$ angle distribution) and S12.2 (dihedral angle distribution).

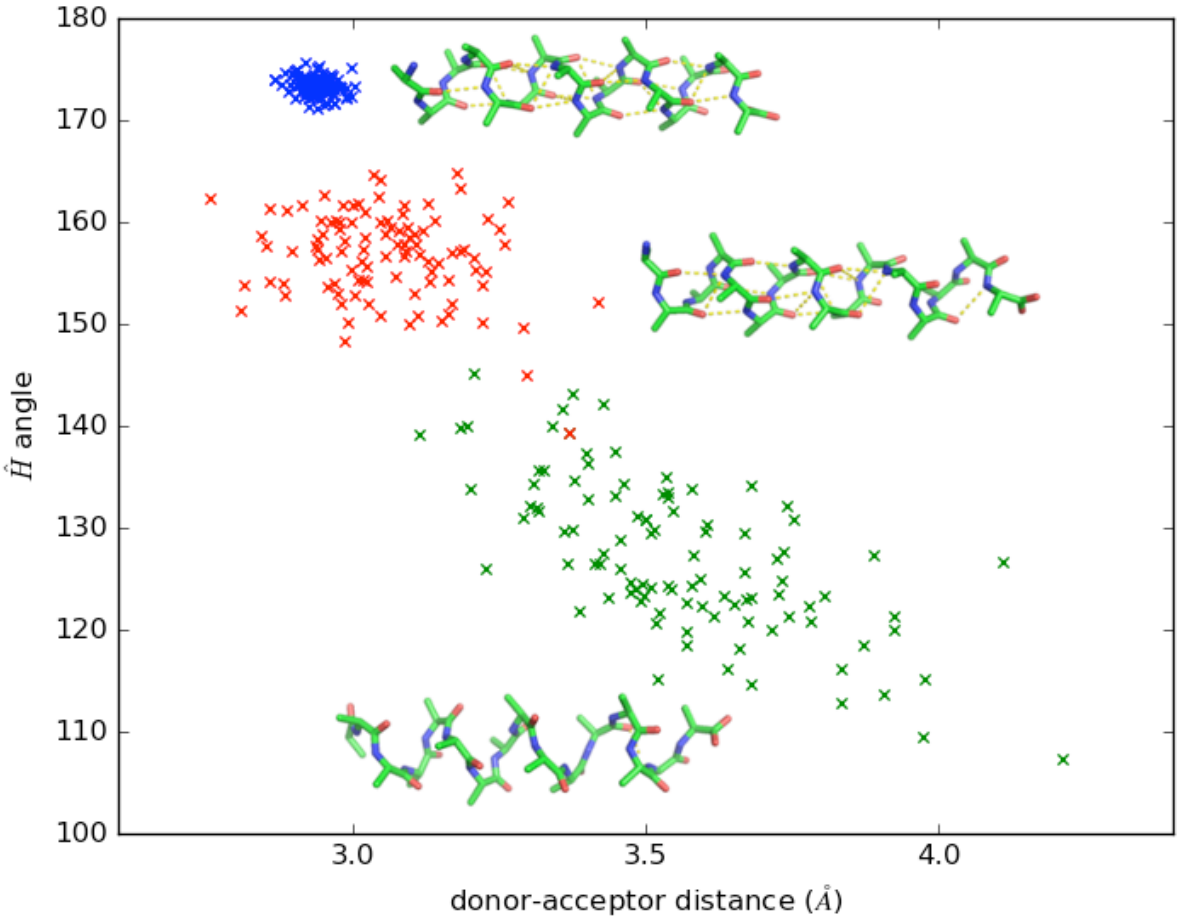


Figure S12.1. Figure shows the $\hat{H}$ angle and donor-acceptor distance sampled in all three simulations (c.f. main text) on a 16-mer alpha-helix. A snapshot structure was taken every 100 ps along the simulation trajectory, and the average of $\hat{H}$ angle and donor-acceptor distance is represented as a data point. The data points from the first, second and third simulations are colored blue, green and red respectively. Snapshots of the alpha-helix during the (top) first, (center) third and (bottom) second simulation are also shown in the figure. Hydrogen bonds are represented in yellow dashed lines.

Supplementary Data 13

GROMACS input files to harmonically restrain bond angles http://manual.gromacs.org/current/reference-manual/functions/restraints.html

1. To make the bond angles explore angles away from 155

atom1	atom2	atom2	atom3	angle	energy	multiplicity
43	12	12	11	1	155	40	1
53	22	22	21	1	155	40	1
63	32	32	31	1	155	40	1
73	42	42	41	1	155	40	1
83	52	52	51	1	155	40	1
93	62	62	61	1	155	40	1
103	72	72	71	1	155	40	1
113	82	82	81	1	155	40	1
123	92	92	91	1	155	40	1
133	102	102	101	1	155	40	1
143	112	112	111	1	155	40	1

2. To restrain the bond angles close to155

atom1	atom2	atom2	atom3	angle	energy	multiplicity
43	12	12	11	1	155	-40	1
53	22	22	21	1	155	-40	1
63	32	32	31	1	155	-40	1
73	42	42	41	1	155	-40	1
83	52	52	51	1	155	-40	1
93	62	62	61	1	155	-40	1
103	72	72	71	1	155	-40	1
113	82	82	81	1	155	-40	1
123	92	92	91	1	155	-40	1
133	102	102	101	1	155	-40	1
143	112	112	111	1	155	-40	1

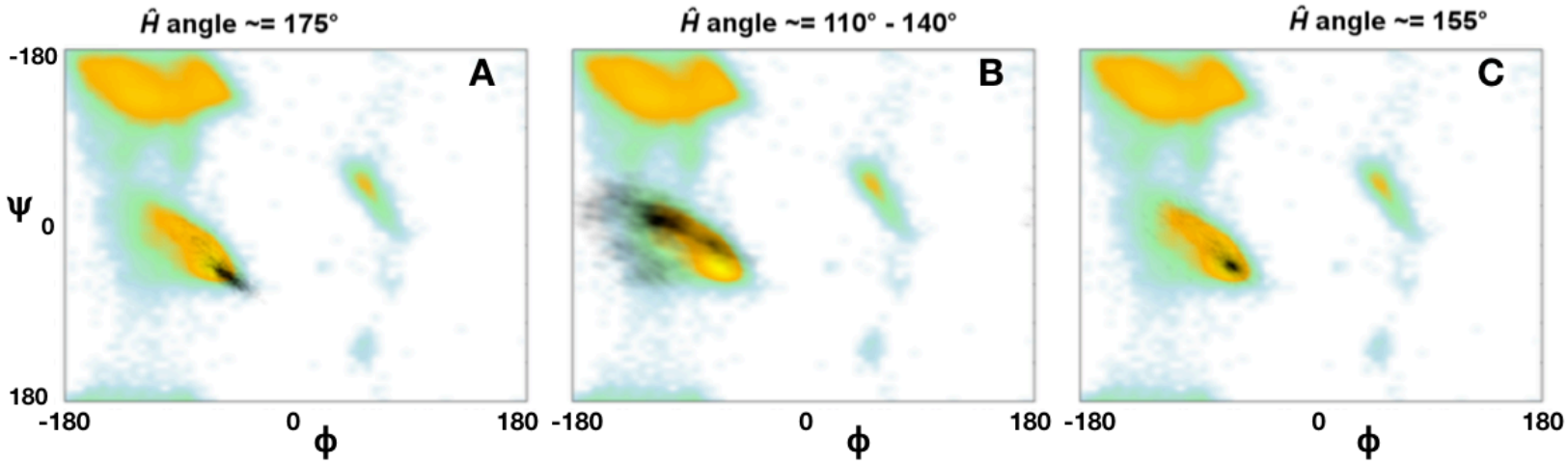


Figure S12.2. Dihedral angle propensity aggregated from the first (A), second (B) and third (C) set of simulations. A snapshot structure was taken every 100 ps along the simulation trajectory, and the phi and psi angles of all residues in the structure are plotted. The density of dihedral angles in the Ramachandran map is represented as black dots. The orange, green, blue and white colours represent favoured, allowed, generously allowed and disallowed regions on the map. A Gaussian kernel density estimation with bandwidth of 0.1° is applied to the data set for plotting purposes.

References

1. Radford, S.E., Buck, M., Topping, K.D., Dobson, C.M., and Evans, P.A. (1992). Hydrogen exchange in native and denatured states of hen egg-white lysozyme. Proteins: Structure, Function and Bioinformatics *14*, 237-248
2. Chakravarty, S., and Varadarajan, R. (1999). Residue depth: a novel parameter for the analysis of protein structure and stability. Structure *7*, 723-732

3. Kumar, M.D., Bava, K.A., Gromiha, M.M., Prabakaran, P., Kitajima, K., Uedaira, H., and Sarai, A. (2006). ProTherm and ProNIT: thermodynamic databases for proteins and protein-nucleic acid interactions. Nucleic Acids Res 34, D204-206.
4. Kabsch, W., and Sander, C. (1983). Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22, 2577-2637.
5. Chou PY, Fasman GD (1974). "Prediction of protein conformation".Biochemistry **13** (2): 222–245.
6. G.Schaftenaar and J.H. Noordik (2000), "Molden: a pre- and post-processing program for molecular and electronic structures", J. Comput.-Aided Mol. Design, 14 123-134
7. Pettersen, E. F. et al (2004). UCSF Chimera–a visualization system for exploratory research and analysis. eng. J Comput Chem 25, 1605–1612.