The BRCT domain from the large subunit of human Replication Factor C

Kobayashi, Masakazu

Chapter 4

^{1}H, ^{15}N and ^{13}C resonance assignments and secondary structure determination of the BRCT Region of the large subunit of human Replication Factor C, bound to DNA

Abstract

Essentially complete ^{1}H, ^{15}N and ^{13}C resonances of the protein moiety of the 19 kDa p140(375-480)-DNA complex are presented. Secondary structure prediction based on the chemical shifts of Cα, Cβ and Hα and on the pattern of backbone NOEs indicate the presence of a consensus BRCT domain with an extra alpha helix in sequences N-terminal to the BRCT domain.

Parts of this chapter have been published as:

Kobayashi, M. and Siegal, G. (2005) *J.Biomol.NMR* **31,** 183-184

Introduction

Replication Factor C (RFC) is a complex of five proteins required for replication and repair of chromosomal DNA (1). The primary function of RFC appears to be to open the toroidally shaped, "sliding clamp" protein PCNA and "load" it onto DNA where it serves as a binding platform for a multitude of enzymes and regulatory proteins involved in the replication and repair of DNA. RFC consists of five subunits, four homologous proteins with molecular mass between 35 and 40 kDa, and a fifth, which has a molecular mass of 140 kDa in mammals (referred to as p140). The N-terminal half of RFC p140 contains sequences unique to RFC, including a region shown to have DNA binding activity(2) (3-6), but that is not required for the clamp loading activity (7;8).

The DNA binding region, which is homologous to bacterial DNA ligases (5), has an unusual specificity for the 5' phosphorylated terminus of dsDNA (4). Interestingly, binding is independent of the sequence of the double-stranded region, so long as there are at least 10 fully paired bases (Figure 2.3). Since the ligase homology was initially noted, it has been recognized that sequences between residues 403 and 486 of human RFC p140 form part of a distinct class of BRCT domains (9;10). However, it has been shown that residues outside of this region are additionally required for DNA binding ((4) and Figure 2.2) thus bringing into question whether residues 403-486 actually fold similarly to known BRCT domains. The DNA binding region of human p140 subunit has been mapped between 375 and 480, which consists of an N-terminus of 28 amino acids followed by a BRCT domain (403-480), and hereafter referred to as the "BRCT region". As part of a project to determine the biological role of the N-terminal half of RFC p140 and to further understand the molecular mechanism of the unusual DNA recognition, we have determined the solution structure of a protein-DNA complex that consists of residues 375-480 and a 10 bp, hairpin oligonucleotide with a recessed, 5' phosphorylated terminus using NMR spectroscopy. Presumably due to dynamics that occur intermediate on the NMR timescale, the spectra of some residues of the protein are of poor quality. Despite this fact, we report here essentially complete ^{1}H, ^{15}N and ^{13}C resonances of the protein moiety of the 19 kDa protein-DNA complex.

Methods and experiments

The gene coding for human RFC p140 residues 375-480 was cloned into pET-20b (Novagen) to allow its expression in fusion with an C-terminal His_6-tag. The recombinant gene codes for Met and Asn prior to residue 375 and Asn, Leu and Glu before the C-terminal His_6-tag. Recombinant protein was produced in *E. coli* BL21(DE3), purified by immobilized metal affinity chromatography on HisBind resin (Novagen) charged with Ni^{2+} ions and subsequently gel filtration using Superose 12 resin (1.6 x 75 cm, Amersham Biosciences). Isotopically labeled proteins were prepared from cells grown in M9-based minimal medium supplemented with $^{15}NH_4Cl$ as the sole nitrogen source, and either $^{13}C_6$-glucose or unlabeled glucose. To form the complex, RFC p140(375-480) was diluted to 10 μM in 25 mM Tris-HCl pH 7.5, 5 mM NaCl, 1 mM DTT and 1.2 equivalents of the oligonucleotide (pCTCGAGGTCGTCATCGACCTCGAGATCA) were added. The complex was concentrated to 0.5 mM using vacuum dialysis (Spectrum Labs) and the buffer was exchanged to 25 mM D_{11}-Tris-HCl pH 7.5, 5mM NaCl in 95/5 H_2O/D_2O.
All NMR data were acquired at 25 °C on a Bruker DMX600 spectrometer. Most of the sequential assignments for the backbone were obtained using 3D HNCACB, CBCA(CO)NH and HBHA(CO)NH spectra. Aliphatic side-chain resonances were derived from 3D HCCH-TOCSY and CCH-TOCSY spectra. Additional data provided by 2D [^{1}H,^{1}H] NOESY, 3D [^{15}N,^{1}H] NOESY-HSQC and [^{13}C,^{1}H] NOESY-HSQC experiments were used for further assignment as well as confirmation of the through-bond data.

Results and Discussion

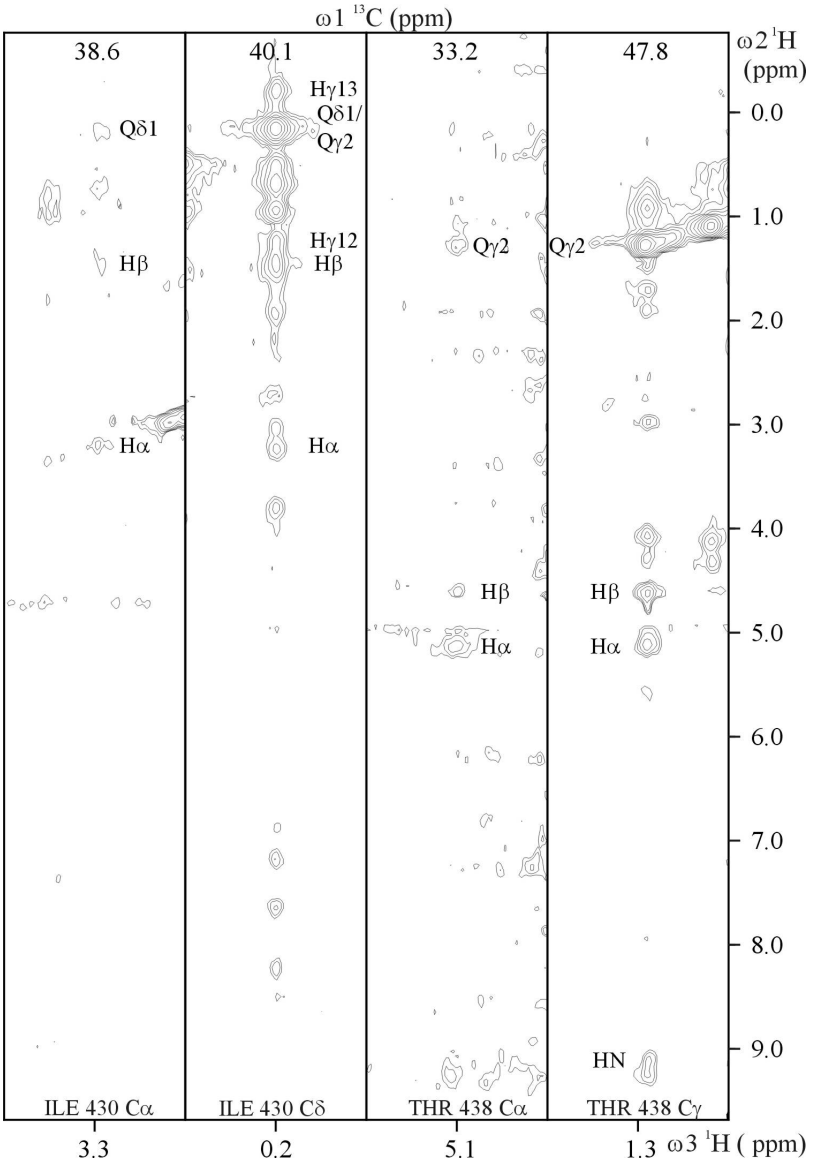
In general, the spectra of the protein-DNA complex are of moderate quality (Figure 4.1) displaying linewidths that are broader than expected for a 19 kDa complex and with missing correlations. However, we have been able to assign over 95 % of the non-labile aliphatic ^{1}H's and 96 % of the protonated aliphatic ^{13}C resonances. More than 99 % of the amide ^{1}H and ^{15}N resonances of the backbone have been assigned with chemical shifts missing only for Tyr 379. ^{1}H-^{15}N correlations have been found for all glutamine and asparagine side chains. RFC p140(375-480) has a rather low aromatic content (5 Tyr and 1 Phe). Of these, we have unambiguously assigned the ^{1}H aromatic resonances of Tyr 382, 385 433 and 447, and Phe 412 using a combination of homonuclear and heteronuclear

edited NOESY spectra. The chemical shift values (Supplementary materials Table 1S) have been deposited in the BioMagResBank database under the accession number 6353.

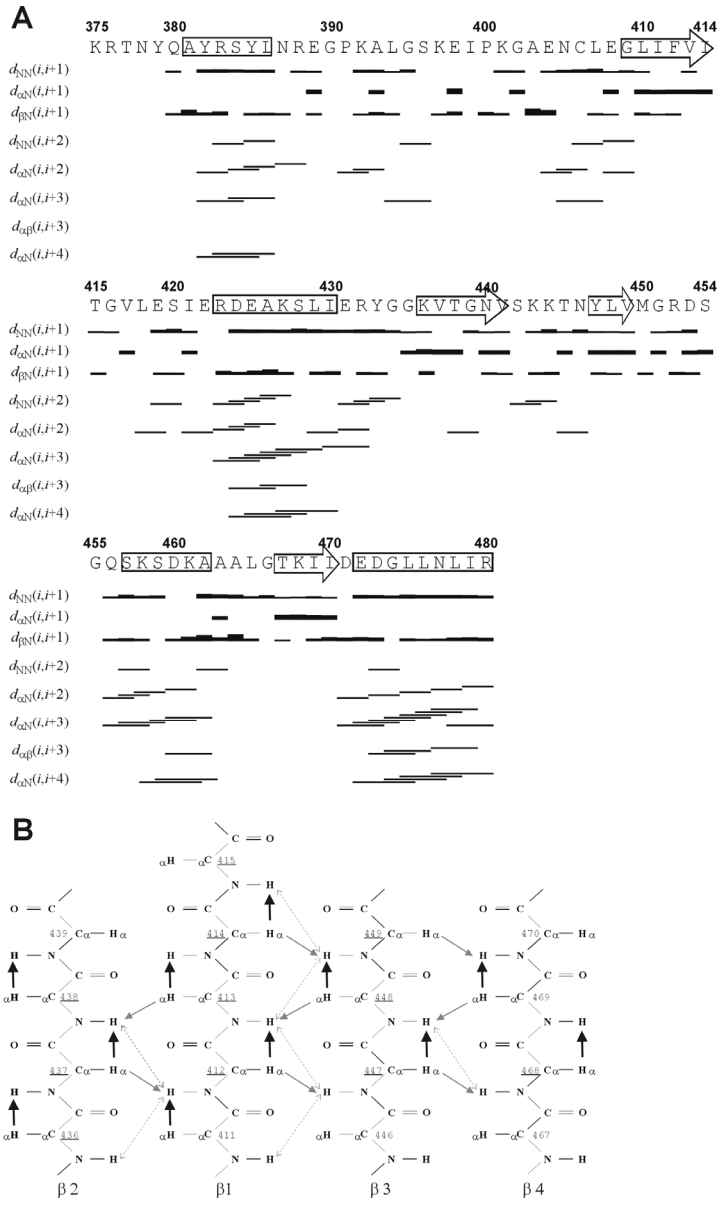
Secondary structure analysis

Regular secondary structures in a polypeptide can be distinguished by the presence of medium and long-range H^1- H^1 distances that are readily observed by NOE's (11). Using the essentially complete chemical shift assignment, the presence of such NOE patterns in the 3D [^{15}N,^{1}H] NOESY-HSQC spectrum was analyzed by manual assignment of NOESY crosspeaks. The resulting sequential and medium range H^1- H^1 NOE's observed are summarized in Figure 4.2A, which indicates that the BRCT region, p140(375-480) consists of four α helices spanning residues 381-386, 423-431, 458-462 and 471-480 and four β - strands spanning residues 410-414, 436-438, 447-449 and 467-470.

In regular β – sheets, the long range H^1- H^1 distances involving the polypeptide backbone are sufficiently short to be observable by NOEs. The pattern of such NOEs is unique for anti-parallel and parallel β – sheets (11). Analysis of the 3D [^{15}N,^{1}H] NOESY-HSQC spectrum reveals a pattern of cross-strand NOE's that includes medium $d\alpha_N(i,j)$ weak d_{NN} (i,j) and strong sequential $d\alpha_N$ $(i, i+1)$ NOE's, the signature of a parallel stranded β -sheet (Figure 4.2B) (11). Due to the absence of Hα-Cα correlations in the 3D [^{13}C,^{1}H] NOESY-HSQC, $d\alpha\alpha(i,j)$ NOE's could not be found. The pattern of secondary structure is in good agreement with all known BRCT domains which have a four stranded parallel β – sheet that comprises the core of the domain (12;13). In general the length of β - strands in BRCT domains (12;13) are rather short, typically between 2 and 4 residues, as also observed here.

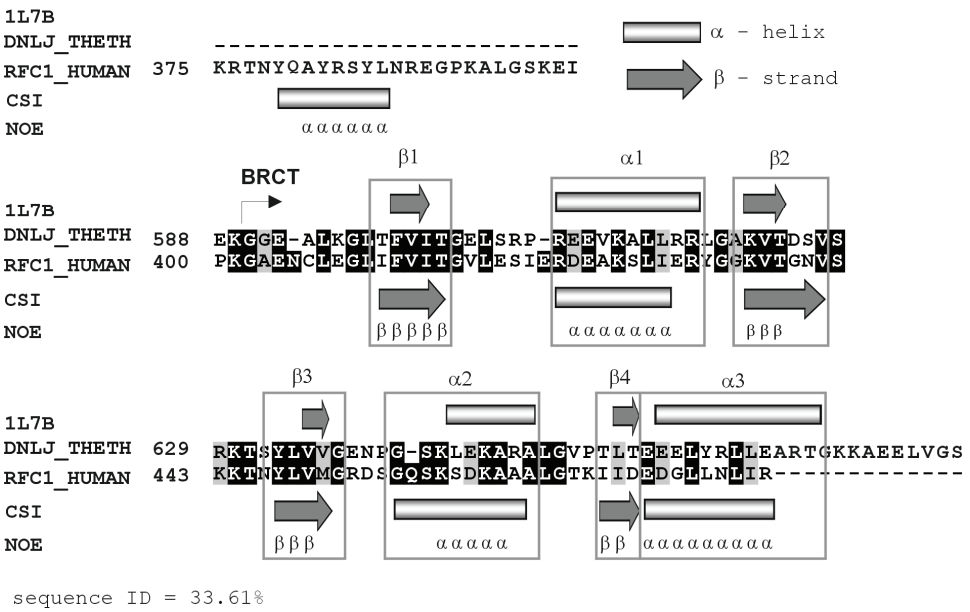


(Figure 4.1) Selected strips from the 3D [^{13}C,1H]-NOESY HSQC of the RFC p140(375-480)-DNA complex. Strips from the Hα-Cα and Qδ1-Cδ1 correlation of Ile 430 and the Hα-Cα and Qγ2-Cγ2 correlation of Thr 438 are shown. The ^{13}C chemical shift is shown above each strip. ^{1}H-^{13}C correlations are significantly weaker in the backbone than in the sidechain. This pattern is observed 83% of the residues in the protein.



(Figure 4.2) The secondary structure prediction generated from NOE connectivity.
(A) Short and medium range upper-distance limits were identified by manual assignments of NOEY crosspeaks in [^{15}N, ^{1}H] NOESY-HSQC spectrum, and plotted against the sequence using the DYANA (14). The resulting secondary structures determined by the NOE patterns were drawn on the sequence where the arrow and the square indicate the presence of a β strand and α helix respectively. **(B)** The backbone amide representing the parallel β-strands. The cross-strand NOE's particular to the parallel β-strands observed in [^{15}N, ^{1}H] NOESY-HSQC spectrum are indicated with arrows. The arrows represent the observed strong sequential $d\alpha_N$ (*i, i+1*) (in black), the medium cross-strand NOE's $d\alpha_N(i,j)$ (in gray), and weak d_{NN} (*i,j*) (dotted) correlations. The underlined residue numbers are conserved amino acid residues, which are also found to participate in forming the parallel β-strands of the homologous BRCT domain from the bacterial NAD+ dependent DNA ligase (see Figure 4.3 for details).

The secondary structure determined from the NOE connectivity is further supported by prediction based on the chemical shift index (CSI) using the $^{13}C\alpha$, $^{13}C\beta$ and $^{1}H\alpha$ resonance assignment of DNA bound p140(375-480) (15). The deviation of chemical shifts of those nuclei from their random coil values can be used to predict secondary structure in proteins. The CSI plot for each of the $^{13}C\alpha$, $^{13}C\beta$ and $^{1}H\alpha$ nuclei can be found as a supplement data (Figure S1) at the end of this chapter. Comparison of the secondary structure derived from the CSI plot and the NOE pattern shows slight differences, especially in the size of strand β2 and helix α3, however in general they are in good agreement with each other (Figure 4.3). Comparison with one member of the distinct class of BRCT superfamily indicates that the secondary structure of the BRCT domain (403-480) is indeed consistent with the known structures of BRCT domains from the NAD+ dependent DNA ligase (Figure 4.3). In addition, the BRCT region includes an extra α-helix (379-386) at the N-terminus which is followed by a long loop (387-410) before the BRCT domain (Figure 4.3).



(Figure 4.3) Amino acid sequence and secondary structure determination of the DNA bound RFC p140(375-480). The secondary structure as determined by CSI analysis (in Supplementary materials) or from the pattern of NOE's is indicated schematically below the sequence. The start of the conserved BRCT domain is indicated. For comparison, the secondary structure of the BRCT domain from the NAD+ dependent ligase from *Th. filiformus* determined by NMR, is also shown .

The complete chemical shift assignment of the free protein is not currently available. A considerable number of the expected peaks were missing from the HNCACB spectrum and 40 % of the backbone amide correlations were missing from the [^{15}N,^{1}H] NOESY-HSQC spectrum. Presumably the amide protons are in rapid exchange with those of water. Accordingly poor dispersion and non-uniform linewidth of resonances were observed in the [^{15}N,^{1}H]-HSQC spectrum of the free p140(375-480) indicating that the protein may undergo conformational exchange intermediate on the NMR time scale. However partial backbone assignments and circular dichroism spectroscopic data (Figure 2.1) suggest the existence of secondary structure in the free protein. Upon DNA binding, the conformational dynamics of the p140(375-480) apparently become restricted resulting in a more defined conformation. The similarity of the secondary structure content to the BRCT domains of known structure suggests that the BRCT domain of the p140(375-480) folds similarly to that of the NAD+ dependent DNA ligase upon DNA binding. The position of the α-helix and the loop at the N-terminus relative to the BRCT domain awaits elucidation of the three-dimensional structure of p140(375-480).

Reference list

1. Waga, S. and Stillman, B. (1998) *Annual Review of Biochemistry* **67,** 721-751

2. Allen, B. L., Uhlmann, F., Gaur, L. K., Mulder, B. A., Posey, K. L., Jones, L. B., and Hardin, S. H. (1998) *Nucleic Acids Res.* **26,** 3877-3882

3. Fotedar, R., Mossi, R., Fitzgerald, P., Rousselle, T., Maga, G., Brickner, H., Messier, H., Kasibhatla, S., Hubscher, U., and Fotedar, A. (1996) *EMBO Journal* **15,** 4423-4433

4. Allen, B. L., Uhlmann, F., Gaur, L. K., Mulder, B. A., Posey, K. L., Jones, L. B., and Hardin, S. H. (1998) *Nucleic Acids Research* **26,** 3877-3882

5. Burbelo, P. D., Utani, A., Pan, Z. Q., and Yamada, Y. (1993) *Proceedings Of The National Academy Of Sciences Of The United States Of America* **90,** 11543-11547

6. Tsurimoto, T. and Stillman, B. (1991) *Journal of Biological Chemistry* **266,** 1950-1960

7. Gomes, X. V., Gary, S. L., and Burgers, P. M. J. (2000) *Journal of Biological Chemistry* **275,** 14541-14549

8. Uhlmann, F., Cai, J. S., Gibbs, E., ODonnell, M., and Hurwitz, J. (1997) *Journal of Biological Chemistry* **272,** 10058-10064

9. Bork, P., Hofmann, K., Bucher, P., Neuwald, A. F., Altschul, S. F., and Koonin, E. V. (1997) *FASEB Journal* **11,** 68-76

10. Callebaut, I. and Mornon, J. P. (1997) *FEBS Letters* **400,** 25-30

11. Wuthrich, K. (1986) *NMR of Proteins and Nucleic acids*, Wiley, New York,

12. Krishnan, V. V., Thornton, K. H., Thelen, M. P., and Cosman, M. (2001) *Biochemistry* **40,** 13158-13166

13. Zhang, X., Morera, S., Bates, P. A., Whitehead, P. C., Coffer, A. I., Hainbucher, K., Nash, R. A., Sternberg, M. J., Lindahl, T., and Freemont, P. S. (1998) *EMBO J.* **17,** 6404-6411

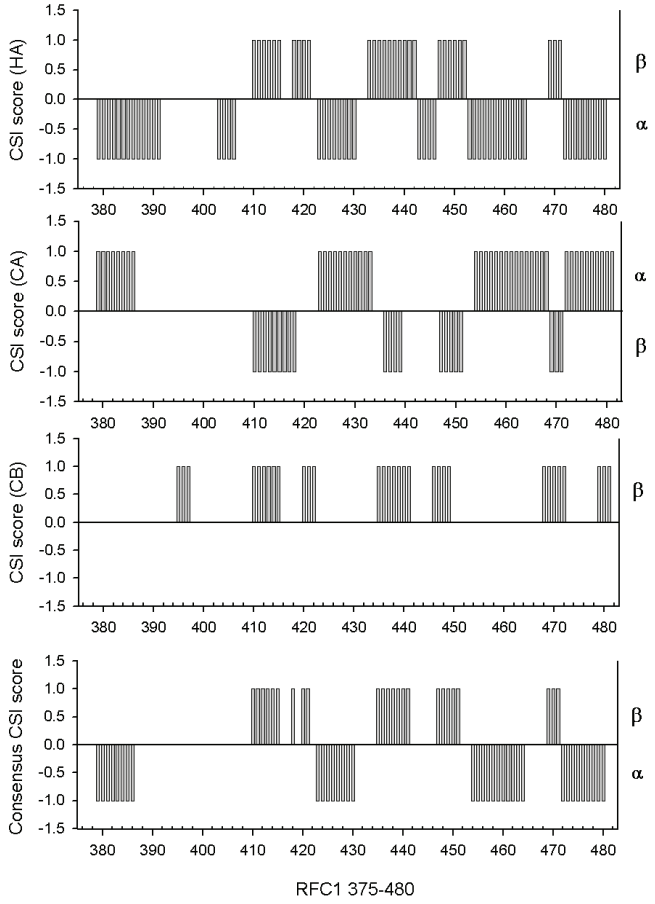
14. Guntert, P., Mumenthaler, C., and Wuthrich, K. (1997) *J.Mol.Biol.* **273,** 283-298

15. Wishart, D. S. and Sykes, B. D. (1994) *Journal of Biomolecular NMR* **4,** 171-180

16. Lee, J. Y., Chang, C., Song, H. K., Moon, J., Yang, J. K., Kim, H. K., Kwon, S. T., and Suh, S. W. (2000) *EMBO J.* **19,** 1119-1129

Supplementary materials

Figure S1. CSI calculaton.



CSI scoring and secondary structures generated based on the CSI consensus. Chemical shift data of residue number 375-480 were used as input for CSI calculation. The secondary structures were assigned to the residues based on their CSI score (left side the plots). The consensus score was derived from the CSI scre of the nuclei CA, CB and HA and was used to generate the final secondary structures for the Figure 3.

Table 1S : RFC p140 375-480 BRCT region : assignments; pH 7.5, 298 ° K **(proton chemical shifts are given in parentheses).**

residue	N	C^{α}	C^{β}	other
K375	122.3 (8.35)	56.8 (4.33)	33.2 (1.79, 1.86)	C^{γ}, 25.0 (1.46, 1.46); C^{δ}, 29.1 (1.71, 1.71); C^{ε}, 30.7 (3.05, 3.05)
R376	123.4 (8.60)	56.6 (4.38)	30.8 (1.80, 1.87)	C^{γ}, 27.2 (1.68, 1.52); C^{δ}, 43.6 (3.17, 3.17)
T377	115.7 (8.28)	62.2 (4.32)	70.1 (4.13)	$C^{\gamma 2}$, 21.6 (1.11)
N378	122.4 (8.57)	53.0 (4.70)	38.2 (2.99, 2.77)	$N^{\delta 2}$, 112.9 (7.58, 6.89)
Y379		59.4 (4.46)	38.4 (3.18, 3.00)	$C^{\delta 1}$, * (7.00); $C^{\delta 2}$, * (7.00); $C^{\varepsilon 1}$, * (6.77); $C^{\varepsilon 2}$, * (6.77)
Q380	119.4 (8.46)	59.2 (3.92)	28.1 (2.10, 2.16)	C^{γ}, 33.9 (2.44, 2.44); $N^{\varepsilon 2}$, 112.3 (6.94, 7.54)
A381	122.0 (8.21)	54.4 (4.16)	18.3 (1.53)	
Y382	121.0 (7.94)	58.8 (4.42)	39.4 (2.75, 3.14)	$C^{\delta 1}$, * (6.53); $C^{\delta 2}$, * (6.53); $C^{\varepsilon 1}$, * (5.89); $C^{\varepsilon 2}$, * (5.89)
R383	117.9 (8.16)	59.3 (3.33)	29.2 (1.64, 1.64)	C^{γ}, 27.0 (1.38, 1.56); C^{δ}, 43.1 (3.04, 3.04)
S384	113.5 (7.94)	61.6 (3.94)	63.1 (4.09, 4.09)	
Y385	124.0 (7.52)	60.8 (4.06)	37.3 (3.28, 3.03)	$C^{\delta 1}$, * (6.66); $C^{\delta 2}$, * (6.66)
L386	119.3 (8.00)	56.7 (3.33)	42.1 (1.31, 0.91)	C^{γ}, 25.8 (0.82); $C^{\delta 1}$, 24.3 (-0.24); $C^{\delta 2}$, 21.5 (0.41)
N387	113.8 (7.22)	52.8 (4.61)	39.8 (2.81, 2.54)	$N^{\delta 2}$, 112.5 (6.89, 7.43)
R388	122.1 (7.22)	56.7 (4.24)	30.8 (1.88, 1.88)	C^{γ}, * (1.37, 1.37); C^{δ}, * (3.33, 3.33)
E389	125.8 (9.00)	56.6 (4.23)	30.8 (2.11, 1.95)	C^{γ}, 36.9 (2.38, 2.38)
G390	109.8 (8.55)	45.3 (3.85, 4.25)		
P391		64.4 (4.09)	32.8 (1.56, 2.49)	C^{γ}, 28.0 (1.80, 2.15); C^{δ}, 49.6 (3.62, 3.62)
K392	123.0 (9.22)	57.5 (4.31)	33.8 (1.99, 1.53)	C^{γ}, 25.5 (1.66, 1.66); C^{δ}, 29.3 (1.63, 1.63); C^{ε}, 42.3 (2.91, 2.91)
A393	127.3 (9.00)	50.4 (4.80)	19.0 (1.08)	
L394	123.2 (8.46)	56.9 (3.99)	41.4 (1.68, 1.46)	C^{γ}, * (1.55); $C^{\delta 1}$, 23.8 (0.77); $C^{\delta 2}$, 25.4 (0.94)
G395	111.8 (8.74)	46.3 (4.16, 4.00)		
S396	113.3 (7.95)	60.0 (4.19)	63.7 (3.96, 3.80)	
K397	121.6 (7.89)	55.3 (4.62)	35.1 (1.47, 1.60)	C^{γ}, 25.6 (1.39, 1.39); C^{δ}, 29.4 (1.43, 1.43)
E398	124.3 (8.52)	56.3 (4.04)	30.2 (1.89, 1.83)	C^{γ}, 36.2 (1.97, 2.23)
I399	128.2 (8.63)	56.7 (4.40)	36.6 (2.00)	$C^{\gamma 1}$, 26.7 (1.52, 1.45); $C^{\gamma 2}$, 17.5 (0.96); $C^{\delta 1}$, 10.4 (0.73)
P400		68.2 (4.42)	32.7 (1.63, 2.12)	C^{γ}, 26.6 (1.91, 2.02); C^{δ}, 51.6 (4.01, 3.56)
K401	120.1 (8.26)	55.7 (4.35)	32.7 (1.85, 1.74)	C^{γ}, 25.0 (1.53, 1.46); C^{δ}, 29.1 (1.73, 1.73); C^{ε}, 42.3 (3.06, 3.06)
G402	110.9 (7.42)	44.6 (3.07, 4.34)		
A403	123.4 (8.58)	51.7 (4.35)	19.2 (1.38)	
E404	120.7 (8.55)	58.0 (4.15)	29.6 (2.02, 2.02)	C^{γ}, 36.2 (2.38, 2.38)
N405	117.7 (9.35)	54.3 (4.52)	38.1 (3.17, 2.84)	$N^{\delta 2}$, 112.5 (7.58, 6.87)
C406	114.6 (7.93)	60.8 (4.35)	27.9 (3.05, 3.05)	
L407	118.5 (8.58)	53.4 (4.77)	40.3 (1.85, 1.58)	C^{γ}, 26.2 (0.82); $C^{\delta 1}$, 22.6 (0.72); $C^{\delta 2}$, * (0.72)
E408	119.0 (7.36)	58.5 (3.99)	29.4 (2.19, 1.98)	C^{γ}, * (2.31, 2.31)
G409	113.9 (8.28)	45.5 (3.72, 4.26)		
L410	120.7 (8.10)	54.1 (4.71)	45.4 (2.38, 2.38)	C^{γ}, * (1.54); $C^{\delta 1}$, 22.9 (0.80); $C^{\delta 2}$, 26.3 (1.04)
I411	123.3 (9.35)	61.0 (5.02)	39.3 (1.90)	$C^{\gamma 1}$, 28.2 (1.05, 1.70); $C^{\gamma 2}$, 19.3 (0.94); $C^{\delta 1}$, 13.5 (0.90)
F412	127.8 (9.77)	56.4 (5.71)	43.7 (3.08, 2.75)	$C^{\delta 1}$, * (7.18); $C^{\delta 2}$, * (7.18); $C^{\varepsilon 1}$, * (6.89); $C^{\varepsilon 2}$, * (6.89)
V413	119.6 (8.32)	61.3 (4.60)	35.6 (1.64)	$C^{\gamma 1}$, 21.2 (0.76); $C^{\gamma 2}$, 22.7 (0.80)
I414	127.7 (9.18)	60.7 (4.95)	40.2 (1.55)	$C^{\gamma 1}$, 27.1 (1.60, 1.60); $C^{\gamma 2}$, 14.3 (0.77); $C^{\delta 1}$, 18.2 (0.84)
T415	120.7 (9.04)	59.4 (5.17)	70.5 (3.88)	$C^{\gamma 2}$, 20.0 (1.22)
G416	114.6 (9.94)	44.0 (3.81, 3.65)		
V417	120.0 (9.59)	63.2 (3.84)	34.5 (1.74)	$C^{\gamma 1}$, 20.3 (0.88); $C^{\gamma 2}$, 21.1 (0.83)
L418	131.1 (8.62)	53.8 (4.53)	37.9 (2.19, 1.77)	C^{γ}, 26.9 (1.70); $C^{\delta 1}$, 21.8 (0.60); $C^{\delta 2}$, 25.1 (0.67)
E419	118.6 (8.44)	57.5 (4.57)	29.9 (2.00, 2.26)	C^{γ}, 34.8 (2.48, 2.38)
S420	111.6 (10.16)	57.4 (4.94)	65.3 (3.80, 3.69)	
I421	119.8 (7.17)	60.1 (4.41)	42.3 (1.66)	$C^{\gamma 1}$, 26.6 (1.17, 1.58); $C^{\gamma 2}$, 16.8 (0.88); $C^{\delta 1}$, 13.6 (0.66)
E422	123.8 (8.88)	56.7 (4.23)	30.8 (1.94, 1.94)	C^{γ}, 36.8 (2.56, 2.35)
R423	124.0 (8.92)	60.1 (3.84)	(1.93, 1.93)	C^{γ}, * (1.53, 1.53); C^{δ}, * (3.26, 3.26); N^{ε}, 81.2 (7.81)
D424	116.3 (8.98)	57.6 (4.37)	39.8 (2.68, 2.68)	
E425	121.0 (7.05)	51.1 (4.18)	29.7 (2.08, 2.24)	C^{γ}, 36.3 (2.39, 2.25)
A426	124.3 (8.55)	55.4 (3.86)	17.3 (1.28)	
K427	117.2 (8.23)	60.2 (3.85)	32.7 (1.99, 1.80)	C^{γ}, 24.9 (1.53, 1.53); C^{δ}, 30.0 (1.75, 1.75); C^{ε}, 42.0 (3.00, 3.00)

S428	113.0 (7.90)	61.6 (4.23)	62.8 (3.98, 3.92)	
L429	124.7 (8.13)	58.3 (4.00)	42.2 (2.00, 1.62)	C^{γ}, 26.9 (1.49); $C^{\delta 1}$, 25.5 (0.74); $C^{\delta 2}$, 24.3 (1.01)
I430	117.2 (7.62)	65.1 (3.27)	38.2 (1.55)	$C^{\gamma 1}$, 29.9 (1.46, -0.13); $C^{\gamma 2}$, 19.1 (0.30); $C^{\delta 1}$, 13.6 (0.21)
E431	118.4 (8.21)	52.1 (4.36)	31.2 (2.05, 2.05)	C^{γ}, 36.8 (2.35, 2.56)
R432	123.2 (8.49)	58.7 (4.03)	29.1 (1.96, 1.83)	C^{γ}, 26.7 (1.22, 1.22); C^{δ}, 43.7 (2.92, 2.86)
Y433	116.0 (7.13)	58.9 (4.76)	38.7 (3.94, 2.50)	$C^{\delta 1}$, * (7.39); $C^{\delta 2}$, * (7.39); $C^{\varepsilon 1}$, * (6.69); $C^{\varepsilon 2}$, * (6.69)
G434	107.0 (7.91)	45.9 (3.85, 4.31)		
G435	109.8 (8.24)	44.9 (4.37, 3.34)		
K436	119.5 (8.53)	54.6 (4.98)	36.1 (1.87, 1.87)	C^{γ}, 25.2 (1.51, 1.41); C^{δ}, 29.3 (1.73, 1.73); C^{ε}, 42.4 (2.99, 2.99)
V437	125.6 (9.27)	60.4 (5.56)	33.5 (1.94)	$C^{\gamma 1}$, 23.4 (0.98); $C^{\gamma 2}$, 22.2 (1.03)
T438	120.1 (9.10)	59.7 (5.11)	71.4 (4.61)	$C^{\gamma 2}$, 21.3 (1.29)
G439	106.1 (9.23)	45.2 (3.81, 4.62)		
N440	118.2 (7.91)	52.2 (4.94)	42.6 (2.80, 2.18)	$N^{\delta 2}$, 117.1 (7.42, 8.71)
V441	125.3 (8.69)	63.2 (4.07)	31.6 (1.92)	$C^{\gamma 1}$, 22.8 (1.01); $C^{\gamma 2}$, 22.0 (0.82)
S442	125.6 (9.37)	57.3 (4.67)	66.6 (5.01, 5.01)	
K443	120.9 (9.22)	59.5 (4.05)	32.1 (2.08, 1.97)	C^{γ}, 25.8 (1.64, 1.64); C^{δ}, 29.4 (1.80, 1.80); C^{ε}, 41.9 (3.06, 3.06)
K444	116.2 (8.09)	56.2 (4.09)	32.8 (1.84, 1.69)	C^{γ}, 24.8 (1.46, 1.46); C^{δ}, 29.2 (1.70, 1.70); C^{ε}, 42.3 (3.00, 3.00)
T445	114.6 (7.40)	65.1 (3.40)	68.9 (3.92)	$C^{\gamma 2}$, 25.3 (1.07)
N446	126.0 (8.99)	57.4 (4.56)	41.4 (2.45, 2.37)	$N^{\delta 2}$, 109.0 (6.87, 7.67)
Y447	114.1 (7.42)	57.0 (5.22)	44.0 (2.45, 2.26)	$C^{\delta 1}$, * (6.63); $C^{\delta 2}$, * (6.63); $C^{\varepsilon 1}$, * (6.29); $C^{\varepsilon 2}$, * (6.29)
L448	124.5 (9.03)	53.1 (5.15)	44.6 (2.02, 0.99)	C^{γ}, 27.5 (1.27); $C^{\delta 1}$, 24.8 (0.80); $C^{\delta 2}$, 26.7 (0.76)
V449	128.6 (9.61)	61.2 (4.55)	31.3 (2.29)	$C^{\gamma 1}$, 21.4 (1.03); $C^{\gamma 2}$, 21.4 (0.76)
M450	127.4 (9.15)	55.1 (4.98)	35.9 (2.07, 1.62)	C^{γ}, 32.2 (2.39, 2.39)
G451	114.4 (8.96)	43.6 (4.50, 3.45)		
R452	118.4 (8.05)	55.7 (4.51)	29.8 (1.76, 1.76)	C^{γ}, 25.9 (1.65, 1.65); C^{δ}, * (2.98, 2.98); N^{ε}, 84.9 (9.30)
D453	119.8 (9.21)	56.7 (4.41)	40.1 (2.84, 2.84)	
S454	110.6 (7.70)	58.2 (3.95)	65.3 (3.94, 3.84)	
G455	108.5 (9.10)	47.9 (3.81, 4.09)		
Q456	128.6 (9.13)	58.3 (4.01)	29.4 (2.16, 2.07)	C^{γ}, 33.9 (2.44, 2.44); $N^{\varepsilon 2}$, 114.8 (7.78, 6.84)
S457	115.7 (8.86)	61.8 (4.11)	62.5 (3.96, 3.96)	
K458	121.4 (8.23)	61.6 (4.00)	33.8 (2.15, 2.15)	C^{γ}, * (1.51, 1.36); C^{δ}, * (1.93, 1.93); C^{ε}, * (2.94, 2.94)
S459	112.5 (7.91)	62.4 (3.90)	68.0 (3.86, 3.86)	
D460	123.8 (8.40)	57.5 (4.45)	40.2 (2.95, 2.69)	
K461	122.5 (8.39)	59.5 (4.08)	32.5 (1.93, 1.71)	C^{γ}, 25.5 (1.46, 1.46); C^{δ}, 29.4 (1.65, 1.65); C^{ε}, 42.3 (2.92, 2.92)
A462	120.4 (8.28)	54.3 (4.21)	19.4 (1.43)	
A463	120.1 (8.03)	54.8 (4.13)	18.0 (1.56)	
A464	120.7 (7.76)	54.7 (4.15)	18.4 (1.56)	
L465	115.7 (8.14)	55.4 (4.25)	42.9 (1.31, 1.88)	C^{γ}, 26.4 (1.82); $C^{\delta 1}$, 22.7 (0.82); $C^{\delta 2}$, 25.8 (0.56)
G466	108.7 (7.92)	45.9 (4.11, 3.83)		
T467	120.1 (8.11)	64.0 (3.66)	69.4 (3.39)	$C^{\gamma 2}$, 22.9 (1.00)
K468	126.7 (7.56)	57.8 (4.14)	32.3 (1.77, 1.77)	C^{γ}, 25.3 (1.49, 1.49); C^{δ}, 29.1 (1.59, 1.59); C^{ε}, 41.8 (2.96, 2.96)
I469	125.4 (8.30)	60.5 (5.13)	39.4 (1.76)	$C^{\gamma 1}$, 28.0 (1.03, 1.03); $C^{\gamma 2}$, 18.6 (0.96); $C^{\delta 1}$, 13.5 (0.87)
I470	122.5 (9.12)	59.0 (5.04)	43.2 (2.01)	$C^{\gamma 1}$, 25.3 (1.56, 1.56); $C^{\gamma 2}$, 18.2 (1.00); $C^{\delta 1}$, 13.5 (0.54)
D471	120.0 (7.87)	50.9 (5.04)	41.8 (2.84, 3.46)	
E472	119.0 (8.68)	60.8 (3.64)	31.1 (2.14, 2.14)	C^{γ}, 36.7 (2.01, 2.01)
D473	117.0 (7.97)	57.6 (4.39)	39.7 (2.65, 2.65)	
G474	110.5 (8.70)	47.3 (3.87, 3.87)		
L475	124.6 (8.47)	58.4 (3.95)	41.0 (2.17, 2.17)	C^{γ}, 27.5 (1.28); $C^{\delta 1}$, 22.2 (0.56); $C^{\delta 2}$, 24.6 (0.69)
L476	117.0 (7.85)	58.2 (3.82)	39.9 (1.28, 2.00)	C^{γ}, 26.5 (2.04); $C^{\delta 1}$, 21.9 (0.72); $C^{\delta 2}$, 25.1 (0.86)
N477	116.2 (8.57)	55.5 (4.45)	38.2 (2.75, 2.87)	$N^{\delta 2}$, 113.1 (7.64, 7.00)
L478	122.3 (7.73)	57.9 (4.06)	42.3 (1.85, 1.57)	C^{γ}, 25.8 (1.65); $C^{\delta 1}$, 25.3 (0.54); $C^{\delta 2}$, 23.9 (0.61)
I479	116.6 (7.03)	64.0 (3.48)	38.3 (1.59)	$C^{\gamma 1}$, 29.5 (0.59, 1.57); $C^{\gamma 2}$, 17.2 (0.17); $C^{\delta 1}$, 13.3 (0.60)
R480	116.3 (7.64)	58.6 (3.91)	31.3 (1.80, 1.80)	C^{γ}, 28.0 (1.68, 1.60); C^{δ}, 43.5 (3.20, 3.20)