BMB Reports - Manuscript Submission

Manuscript Draft

Manuscript Number: BMB-17-152

Title: An NMR Study on the Intrinsically Disordered Core Transactivation Domain of Human Glucocorticoid Receptor

Article Type: Article

Keywords: Human glucocorticoid receptor (hGR); Transcriptional activation domains (TADs); Nuclear magnetic resonance (NMR); Intrinsically disordered protein (IDP); Pre-structured motif (PreSMo)

Corresponding Author: Kyou-Hoon Han

Authors: Kyou-Hoon Han[1,*,#], Do-Hyoung Kim[1], Anthony Wright[1]

Institution: [1]Genome Editing Research Center,, Korea Research Institute of Bioscience and Biotechnology, Daejeon 34141, Korea,
[2]Department of Nano and Bioinformatics,, University of Science and Technology, Daejeon 34113, Korea,
[3]Department of Laboratory Medicine,, Clinical Research Center, Karolinska Institutet, Stockholm, Sweden,

An NMR Study on the Intrinsically Disordered Core Transactivation Domain of Human Glucocorticoid Receptor

Do-Hyoung Kim[1], Anthony Wright[2], and Kyou-Hoon Han[1,3*]

[1]*Genome Editing Research Center, Korea Research Institute of Bioscience and Biotechnology, Daejeon 34141, Korea*

[2]*Clinical Research Center, Department of Laboratory Medicine, Karolinska Institutet, Stockholm, Sweden*

[3]*Department of Nano and Bioinformatics, University of Science and Technology, Daejeon 34113, Korea*

*Corresponding author

Tel : +82-42-860-4250

Fax : +82-42-860-4259

Email : khhan600@kribb.re.kr

Running Title: structural characterization of hGR tau1c

Keywords:

Human glucocorticoid receptor (hGR), Nuclear magnetic resonance (NMR), intrinsically disordered protein (IDP), Pre-structured motif (PreSMo),

ABSTRACT

A large number of transcriptional activation domains (TADs) are intrinsically unstructured, meaning they are devoid of a three-dimensional structure. The fact that these TADs are transcriptionally active without forming a 3-D structure raises the question of what features in these domains enable them to function. One of two TADs in human glucocorticoid receptor (hGR) is located at its N-terminus and is responsible for ~70% of the transcriptional activity of hGR. This 58-residue intrinsically-disordered TAD, named tau1c in an earlier study, was shown to form three helices under trifluoroethanol, which might be important for its activity. We carried out heteronuclear multi-dimensional NMR experiments on hGR tau1c in a more physiological aqueous buffer solution and found that it forms three helices that are ~30% pre-populated. Since pre-populated helices in several TADs were shown to be key elements for transcriptional activity, the three pre-formed helices in hGR tau1c delineated in this study should be critical determinants of the transcriptional activity of hGR. The presence of pre-structured helices in hGR tau1c strongly suggests that the existence of pre-structured motifs in target-unbound TADs is a very broad phenomenon.

INTRODUCTION

Although many globular proteins are three-dimensionally structured, they contain short flexible/disordered loops composed of less than 20 amino acid residues [1]. This phenomenon of so-called protein disorder has been known for decades, but since the late 1990s, we have begun encountering some peculiar proteins that contain long unfolded/disordered regions (more than 40 and up to hundreds of residues) that do not form 3-D structures [2, 3]. These proteins are now named as intrinsically disordered proteins (IDPs) [4, 5] and represent a special case of protein disorder. IDPs are highly unorthodox because even without three-dimensional structures, they are capable of performing specific biological functions (including transcription, translation, chaperoning, and cell cycle regulation) or are responsible for many fatal diseases (such as cancers, prion diseases, neurodegenerative diseases including Alzheimer's and Parkinson's, and so on) [6-9]. Many viral proteins in HIV-1, HBV, HCV, SARS virus, and AI virus are also IDPs or contain intrinsically disordered regions (IDRs) [4, 10-13].

The unexpected correlation between the "unstructured" nature of IDPs/IDRs and their functionality ended up nullifying the decades-old structure-function paradigm, 3-D structure = function, in protein science and structural biology. Approximately 40% of the entire protein kingdom is predicted to consist of IDPs/IDRs [14], and the proportion of IDPs/IDRs is much higher (~60%) in transcription factors [15]. In the case of globular proteins, the collective structural features (secondary, tertiary and quaternary) provide a reasonable explanation of function. However, the absence of tertiary structures in IDPs makes it quite challenging to come up with an explanation on why and how they should function at all. For example, we still do not have a clear understanding on how IDPs bind to their targets.

Initially, IDPs/IDRs were erroneously thought to be completely unstructured (CU) without any trace of secondary structures [16, 17]. In contrast to this early view, a more quantitative structural picture on IDPs/IDR has emerged from many high-resolution multi-dimensional NMR investigations conducted over the last two decades. These studies have revealed that at least ~70% of IDPs/IDRs are not fully "unstructured", but contain transient local structural elements in their free state that mediate binding of IDPs to targets [4]. The IDPs/IDRs

containing transient local structural elements are therefore described to be in a mostly unstructured (MU) state [3] since they are not completely unstructured in terms of secondary structure. Although these transient secondary structures in IDPs, which were recently named in 2012 as pre-structured motifs (PreSMos) [4], were first noticed in the late 1990s [2, 3, 18, 19], they were not given the general name of PreSMos until structural details for a statistically significant number (~ 4 dozens) of MU type IDPs became available. Most well-known PreSMos are the amphipathic helix and two turns found in the 73-residue long transactivation domain (TAD) of tumor suppressor p53 [3]. These PreSMos are the key determinants that enable binding of p53TAD to mdm2 [20], p62 [21], RPA [22] as well as the NCBD of CBP [23] and Bcl-2 [24]. In other words, the PreSMos of p53 TAD are the "active sites" that are pre-populated transient secondary structures primed for target-binding. These PreSMos found in the p53 TAD, and other PreSMos observed in various IDPs/IDRs [4, 25-27] are likely to be at least part of the long-sought answer to the question of how IDPs/IDRs bind to various targets including proteins, nucleic acids, lipids and metals.

The glucocorticoid receptor (GR), a member of the nuclear receptor family, contains independent domains for DNA-binding and transactivation [28]. Human GR contains two transactivation domains (TADs), one at its N-terminus (tau1) and the other at its C-terminus (tau2) [29]. The former consists of residues 77-262 (186-residues) and is responsible for the major transcriptional activity of hGR. We have previously shown that the minimal ("core") domain of tau1 containing residues 187-244 (called tau1c hereafter) retained ~70% of the transactivation activity of the intact tau1 [30]. hGR tau1c was one of the IDPs that was studied in the early days before the PreSMo concept was introduced. Since many transcription factors and TADs contain PreSMos, we wanted to learn whether tau1c also contains PreSMos. A previous NMR study on hGR tau1c indicated that it was largely unstructured in an aqueous solution, but it formed three helices under a helix-inducing solvent, trifluoroethanol (TFE) [31]. To determine whether the helices observed under TFE may be considered as PreSMos, we performed multi-dimensional NMR experiments on hGR tau1c in aqueous solvents using a $^{15}N/^{13}C$-double labeled hGR tau1c.

RESULTS

Figure 1 shows a fingerprint region of an ^{15}N-^{1}H heteronuclear single quantum coherence (HSQC) spectrum of hGR tau1c with resonance assignment. Based on the narrow chemical-shift dispersion in both ^{15}N and ^{1}H dimensions, we can confirm its overall unfolded/disordered nature in agreement with the bioinformatics prediction (Suppl. Fig. S1). Using the standard triple-resonance assignment procedure, we achieved a full NMR resonance assignment for the backbone ^{15}N and the amide protons of the 64-residue hGR tau1c excluding the first two N-terminal residues, Met, and His, that originated from the N-terminal glutathione-S-transferase fusion linker. For 3 prolines which do not have backbone amide NH protons, their aliphatic protons were fully assigned. The resonance assignment of hGR tau1c summarized in Table S1 was sufficient for structural characterization of the hGR tau1c, i.e., delineation of the PreSMo-forming residues.

Figure 2 shows the amino acid sequence of hGR tau1c along with the associated interproton NOEs and chemical shift indices (CSI). Continuous d_{NN} interproton NOEs from a NOESY spectrum (a mixing time of 150 ms) and CSIs are observed for three regions of the hGR tau1c, indicating the pre-structured (non-random) nature of hGR tau1c with several helix-forming residues. A quick examination of Figure 2 reveals that the location of the three helices detected in the current study mostly overlaps with those previously detected under TFE. Figure 3 shows several NMR parameters measured for hGR tau1c. Whether a PreSMo exists or not is usually determined by a combination of all available NMR parameters [26, 32]. The left panel in Fig. 3 shows a summary of the chemical shift and deviations from random chemical shift values and SSP (secondary structure propensity) scores. The SSP scores were obtained by combining various chemical shifts (Hα, Cα, Cβ) [33] to provide the degree of pre-population of a PreSMo. The SSP values are often more conclusive about the location of a PreSMo than individual chemical shift deviations. Positive SSP scores over 4 residues or more indicate the formation of a helix, whereas negative values suggest non-helical (β-type) secondary structures. Fig. 3d shows that hGR tau1c contains three helical PreSMos, the first (Helix 1) formed by residues 185-202, the second (Helix 2) by residues 206-225, and the third (Helix 3) by residues 232-244. All helices are ~20-30% pre-populated. For most PreSMos positive (0.0–0.5), heteronuclear NOEs were observed. However, some reported

heteronuclear NOEs for PreSMos were non-positive, being zero or slightly negative [3, 32].

DISCUSSION

By definition, PreSMos are *transient or nascent* secondary structures (average population of ~30%) detected by NMR in *aqueous* solution [4]. With the discovery of PreSMos, we can classify IDPs/IDRs into two subclasses, a CU type and an MU type. Although several NMR parameters are used in combination to accurately delineate the residues that form a PreSMo, the SSP scores alone allow one to quickly judge if any IDP/IDR is a CU type or an MU type [3, 26, 34, 35]. The SSP scores of hGR tau1c clearly indicate that it is an MU type, and the combination of all NMR parameters show that it forms three helical PreSMos: the first helix (Helix 1) formed by residues 185-202, the second (Helix 2) by residues 206-225, and the third (Helix 3) by residues 232-244, respectively. The three helices previously observed under TFE were Helix I (189-201), Helix II (215-226) and Helix III (234-239), respectively [31]. Even though the three helices delineated in this study reasonably overlap with the three helices observed under TFE, notable differences exist at the N-termini of Helix 1 and Helix 2 and at the C-terminus of Helix 3.

For example, the first residue of Helix 1 is 185T whereas the first one in Helix I is 189S. This difference can be ascribed to the fact that the starting residue of hGR tau1c is 181V while that in the previous work was 188Q. Thus, it was not possible to observe the potential helix formation by residues 181-187 in the previous work. Another significant difference is noted at the N-terminus of Helix 2. Helix 2 determined in the current study has a 9-residue longer N-terminal portion than Helix II, and its N-terminus starts at residue 206K instead of 215S. The previous work used only interproton NOEs since the SSP scores were introduced only in 2006. The usage of TFE must have generated helices in a higher population so that the interproton NOEs were stronger than those observed in the current study when we compare Figure 2 of this report with Figure 3 in the previous report [31]. Previous data indicated that 213W is important for transcriptional activity [36]. Even though this residue was not a part of Helix II, a thorough reexamination of the data insinuates that 213W may belong to Helix II. It would be interesting to see if the inclusion of this bulky hydrophobic residue in a helical conformation would better explain the activity of hGR tau1c. Since the residue at 212 is a

proline in Helix 2, it might introduce a kink just prior to 213W. Another difference is that the C-terminus of Helix 3 is longer by 5 residues than Helix III. It also is very interesting to learn that two prolines, 204P and 226P, flank the C-termini of the two helices, Helix 1 and Helix 2. The helix-flanking prolines of helix PreSMos that might act as a subtle activity switch were described earlier [37].

Tau1c of hGR is one of the IDRs that was studied very early in IDP research when controversy was keen regarding the question of whether transactivation domains should contain some sort of specificity determinants that mediate target binding [3]. Although hGR tau1c was shown to form three helices that might be important for transcriptional activity [31], this pioneering data was not considered when the PreSMo concept was formulated [4] since the observation of helices was made in TFE that may artificially induce helix formation. The current study was initiated to re-investigate this IDR in aqueous solvents in order to determine if it forms PreSMos and if it should be classified as an MU type. We confirmed that hGR tau1c is an MU type IDR. Our work demonstrated that the PreSMo concept is further expandable to other IDPs/IDRs, including the IDPs studied before the introduction of the PreSMos concept. Recent mutation studies on p53 TAD and CNBR [38, 39] have elegantley shown that the degree of pre-population of a helix PreSMo is critical for target binding, suggesting the degree of PreSMo pre-population is a variable inherent in the MU type IDPs/IDRs that is quite subtly tuned for target binding. It would be interesting to see if one can also establish such a correlation between the degree of helix pre-population and the activity of hGR tau1c.

MATERIALS AND METHODS

Detailed experimental and computational procedures are described in the Supplementary Information.

ACKNOWLEDGMENTS

This work was supported by a Korea-Hungary and Pan EU collaborative project from the National Research Council of Science and Technology (NST) (NTC2251422) of Korea (to KH & AW) as well as grants from the Swedish Research Council and Swedish Cancer Society (to AW).

REFERENCES

1. The RCSB Protein Data Bank (https://www.rcsb.org)
2. Daughdrill GW, Chadsey M, Karlinsey JE, Hughes KT, Dahlquist FW et al (1997) The C-terminal half of the anti-sigma factor, FlgM, becomes structured when bound to its target, σ^{28}. Nat Struct Mol Biol. 4, 285-291.
3. Lee H, Mok KH, Muhandiram R et al. (2000) Local structural elements in the mostly unstructured transcriptional activation domain of human p53. J Biol Chem. 275, 29426-29432
4. Lee S-H, Kim D-H, Han JJ et al. (2012) Understanding pre-structured motifs (PreSMos) in intrinsically unfolded proteins. Curr Protein Pept Sci. 13, 34-54
5. Dunker AK, Babu MM, Barbar E et al. (2013) What's in a name? Why these proteins are intrinsically disordered. Intrinsically Disordered Proteins. 1, e24157
6. Uversky VN, Oldfield CJ, Dunker AK et al (2008) Intrinsically disordered proteins in human diseases: Introducing the D2 concept. Annual Review of Biophysics. 37, 215-246
7. James TL, Liu H, Ulyanov NB et al. (1997) Solution structure of a 142-residue recombinant prion protein corresponding to the infectious fragment of the scrapie□isoform. Proc Natl Acad Sci USA. 94, 10086-10091
8. Eliezer D, Kutluay E, Bussell JrR, Browne G et al (2001) Conformational properties of α-synuclein in its free and lipid-associated states. J Mol Biol. 307, 1061-1073.
9. Mukrasch MD, Bibow S, Korukottu J, Jeganathan S, Biernat J, Griesinger C, Mandelkow E, Zweckstetter M et al (2009) Structural polymorphism of 441-residue tau at single residue resolution. PLoS Biol. 7, e1000034
10. Chi S-W, Kim D-H, Lee S-H, Chang I, Han K-H et al (2007) Pre-structured motifs in the natively unstructured preS1 surface antigen of hepatitis B virus. Protein Sci. 16, 2108-2117
11. Kim D-H, Lee S-H, Nam KH, Chi S-W, Chang, I, and Han K-H (2009) Multiple hTAFII31-binding motifs in the intrinsically unfolded transcriptional activation domain of VP16. BMB Reports. 42, 411-417.
12. Lee SH, Cha EJ, Lim JE et al. (2012) Structural characterization of an intrinsically

unfolded mini-HBX protein from hepatitis B virus. Mol Cells. 34, 165-169

13. Xue B, Blocquel D, Habchi J et al. (2014) Structural disorder in viral proteins. Chem Rev. 114, 6880-6911
14. Dunker AK, Obradovic Z, Romero P, Gamer EC, Brown CJ et al (2000) Intrinsic protein disorder in complete genomes. Gen. Inform. 11, 161-171
15. Yoshiaki Minezaki, Keiichi Homma, Akira R. Kinjo, Ken Nishikawa et al (2006) J Mol Biol. 359, 1137–1149
16. Radhakrishnan I, Pérez-Alvarado GC, Parker D, Dyson HJ, Montminy MR, Wright PE et al (1997) Solution structure of the KIX domain of CBP bound to the transactivation domain of CREB: A model for activator:coactivator interactions. Cell. 91, 741-752
17. Fletcher CM and Wagner G (1998) The interaction of eIF4E with 4E-BP1 is an induced fit to a completely disordered protein. Protein Sci. 7, 1639-1642
18. Ramelot TA, Gentile LN, Nicholson LK et al (2000) Transient structure of the amyloid precursor protein cytoplasmic tail indicates preordering of structure for binding to cytosolic factors. Biochemistry. 39, 2714-2725.
19. Sayers EW, Gerstner RB, Draper DE, Torchia DA et al (2000) Structural preordering in the N-terminal region of ribosomal protein S4 revealed by heteronuclear NMR spectroscopy. Biochemistry. 39, 13602-13613.
20. Chi S-W, Lee S-H, Kim D-H et al. (2005) Structural details on mdm2-p53 interaction. J Biol Chem. 280, 38795-38802
21. Di Lello P, Jenkins LMM, Jones TN et al. (2006) Structure of the Tfb1/p53 complex: Insights into the interaction between the p62/Tfb1 subunit of TFIIH and the activation domain of p53. Mol cell. 22, 731-740
22. Bochkareva E, Kaustov L, Ayed A et al. (2005) Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A. Proc Natl Acad Sci USA. 102, 15412-15417
23. Lee CW, Martinez-Yamout MA, Dyson HJ, Wright PE et al (2010) Structure of the p53 transactivation domain in complex with the nuclear receptor coactivator binding domain of CREB binding protein. Biochemistry. 49, 9964-9971
24. Ha J-H, Shin J-S, Yoon M-K et al. (2013) Dual-site interactions of p53 protein transactivation domain with anti-apoptotic Bcl-2 family proteins reveal a highly

convergent mechanism of divergent p53 pathways. J Biol Chem. 288, 7387-7398

25. Andresen C, Helander S, Lemak A, Farès C, Csizmok V, Carlsson J, Penn LZ, Forman-Kay JD, Arrowsmith CH, Lundström P, Sunnerhagen M et al (2012) Transient structure and dynamics in the disordered c-Myc transactivation domain affect Bin1 binding. Nucleic Acids Research. 40, 6353–6366
26. Kim D-H, Lee C, Cho Y-J et al. (2015) A pre-structured helix in the intrinsically disordered 4EBP1. Mol BioSyst. 11, 366-369
27. Berlow RB, Dyson HJ, Wright PE et al (2017) Hypersensitive termination of the hypoxic response by a disordered protein switch. Nature. 543, 447-451
28. Evans RM. (1988) The steroid and thyroid hormone receptor superfamily. Science. 240, 889-895
29. Hollenberg SM, and Evans RM. (1988) Multiple and cooperative trans-activation domains of the human glucocorticoid receptor. Cell. 55, 899-906
30. Dahlman-Wright K, Almlöf T, McEwan IJ, Gustafsson JA, Wright AP et al (1994) Delineation of a small region within the major transactivation domain of the human glucocorticoid receptor that mediates transactivation of gene expression. Proc Natl Acad Sci USA. 91, 1619-1623
31. Dahlman-Wright K, Baumann H, McEwan IJ, Almlöf T, Wright AP, Gustafsson JA, Härd T et al (1995) Structural characterization of a minimal functional transactivation domain from the human glucocorticoid receptor. Proc Natl Acad Sci USA. 92, 1699-1703
32. Kim DH, Lee C, Lee SH, Kim KT, Han JJ, Cha EJ, Lim JE, Cho YJ, Hong SH, Han KH et al (2017) The Mechanism of p53 Rescue by SUSP4. Angew Chem Int Ed Engl. 56, 1278–1282
33. Marsh JA, Singh VK, Jia Z, Forman-Kay JD et al (2006) Sensitivity of secondary structure propensities to sequence differences between α- and γ-synuclein: Implications for fibrillation. Protein Sci. 15, 2795-2804
34. Baker JM, Hudson RP, Kanelis V, Choy WY, Thibodeau PH, Thomas PJ, Forman-Kay JD et al (2007) CFTR regulatory region interacts with NBD1 predominantly via multiple transient helices. Nat Struct Mol Biol. 14, 738-745
35. Zhang X, Perugini MA, Yao S, Adda CG, Murphy VJ, Low A, Anders RF, Norton RS

et al (2008) Solution conformation, backbone dynamics and lipid interactions of the intrinsically unstructured malaria surface protein MSP2. J Mol Biol. 379, 105-121.

36. Almlöf T, Gustafsson JA, Wright AP et al (1997) Role of hydrophobic amino acid clusters in the transactivation activity of the human glucocorticoid receptor. Mol Cell Biol. 17, 934-945

37. Lee C, Kalmar L, Xue B, Tompa P, Daughdrill GW, Uversky VN, Han KH et al (2014) Contribution of proline to the pre-structuring tendency of transient helical secondary structure elements in intrinsically disordered proteins. Biochim Biophys Acta. 1840, 993-1003

38. Borcherds W, Theillet FX, Katzer A, Finzel A, Mishall KM, Powell AT, Wu H, Manieri W, Dieterich C, Selenko P, Loewer A, Daughdrill GW et al (2014) Disorder and residual helicity alter p53-Mdm2 binding affinity and signaling in cells. Nat Chem Biol. 10, 1000-1002.

39. Iešmantavičius V, Dogan J, Jemth P, Teilum K, Kjaergaard M et al (2014) Helical propensity in an intrinsically disordered protein accelerates ligand binding. Angew Chem Int Ed. 53, 1548-1551

FIGURE LEGEND

Figure 1.

A fingerprint region in an ^{15}N-^{1}H HSQC spectrum of the N-terminal region of hGR tau1c (residues 181-244) obtained at 10°C and pH 6.5 on 90% H_2O/10% D_2O.

Figure 2.

The amino acid sequence of hGR tau1c. tau1c is shown along with the associated interproton NOEs and chemical shift indices (CSI). Continuous d_{NN} interproton NOEs and CSIs are observed for three PreSMo regions of hGR tau1c, indicating their pre-structured (non-random) nature with transient helices. See the text for details. The three helix PreSMos identified in this study and those reported previously are shown as open and hatched boxes.

Figure 3. Left panel: deviation of (a) $^{1}H\alpha$, (b) $^{13}C\alpha$, (c) carbonyl chemical shifts from random coil values and (d) the SSP (secondary structure propensity) scores of hGR tau1c (181-244). In (d), positive scores indicate helical propensity while negative values suggest formation of non-helical type PreSMos. Right panel: ^{1}H–^{15}N heteronuclear NOEs (e), backbone ^{15}N relaxation times, T_1 (f), T_2 (g), and temperature coefficients of the backbone amide hydrogens (h). The horizontal lines in (f) and (g) indicate an average value. In (h), temperature coefficients less than 5 ppb/K suggests formation of a helix.

Figure 1.

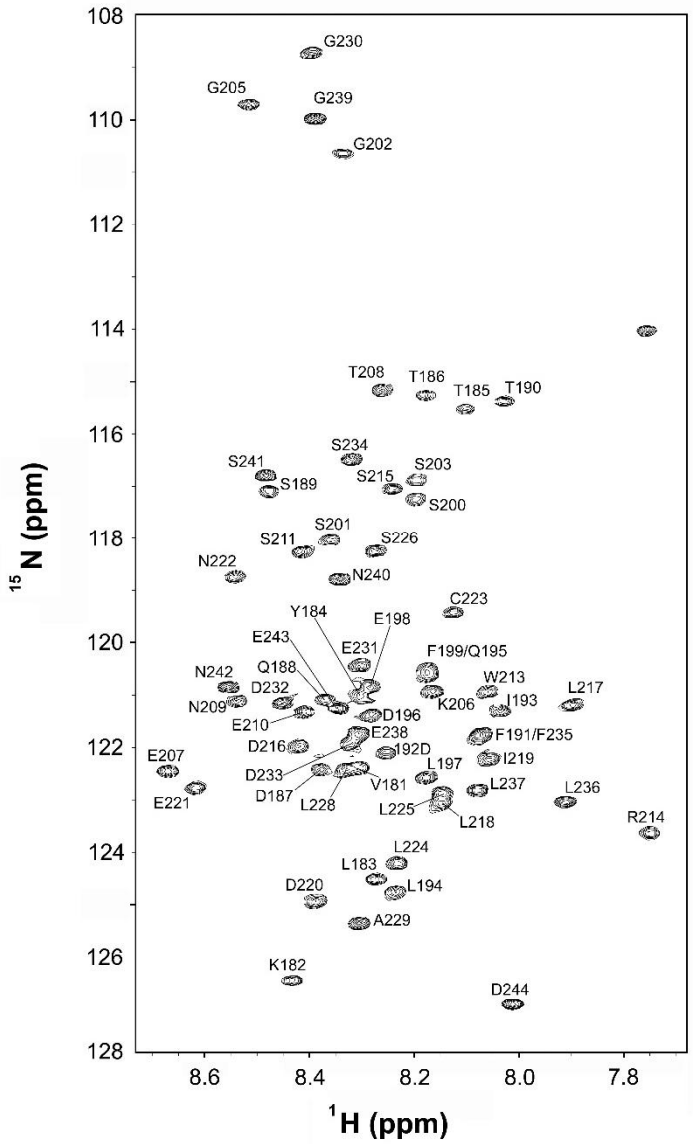

108
110
112
114
116
118
120
122
124
126
128
15N (ppm)
8.6
8.4
8.2
8.0
7.8
1H (ppm)
G230
G205
G239
G202
T208
T186
T185
T190
S234
S241
S203
S215
S189
S200
S201
S211
S226
N222
N240
C223
Y184
E198
E243
E231
F199/Q195
N242
Q188
W213
L217
D232
K206
I193
N209
D196
E210
E238
F191/F235
D216
192D
I219
E207
L197
D233
V181
L237
L236
D187
L228
L225
L218
R214
E221
L224
L183
D220
L194
A229
K182
D244

Figure 2.

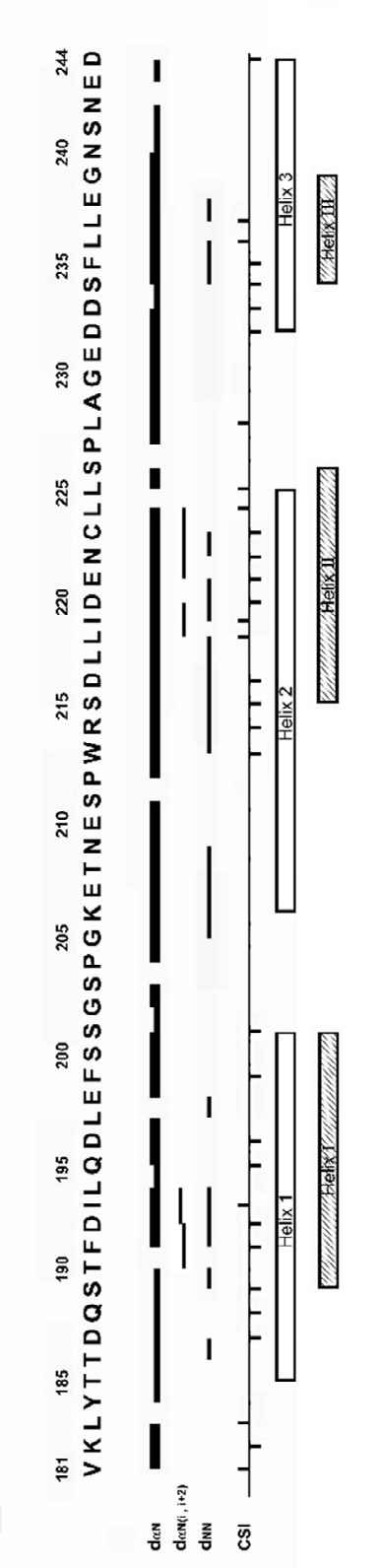
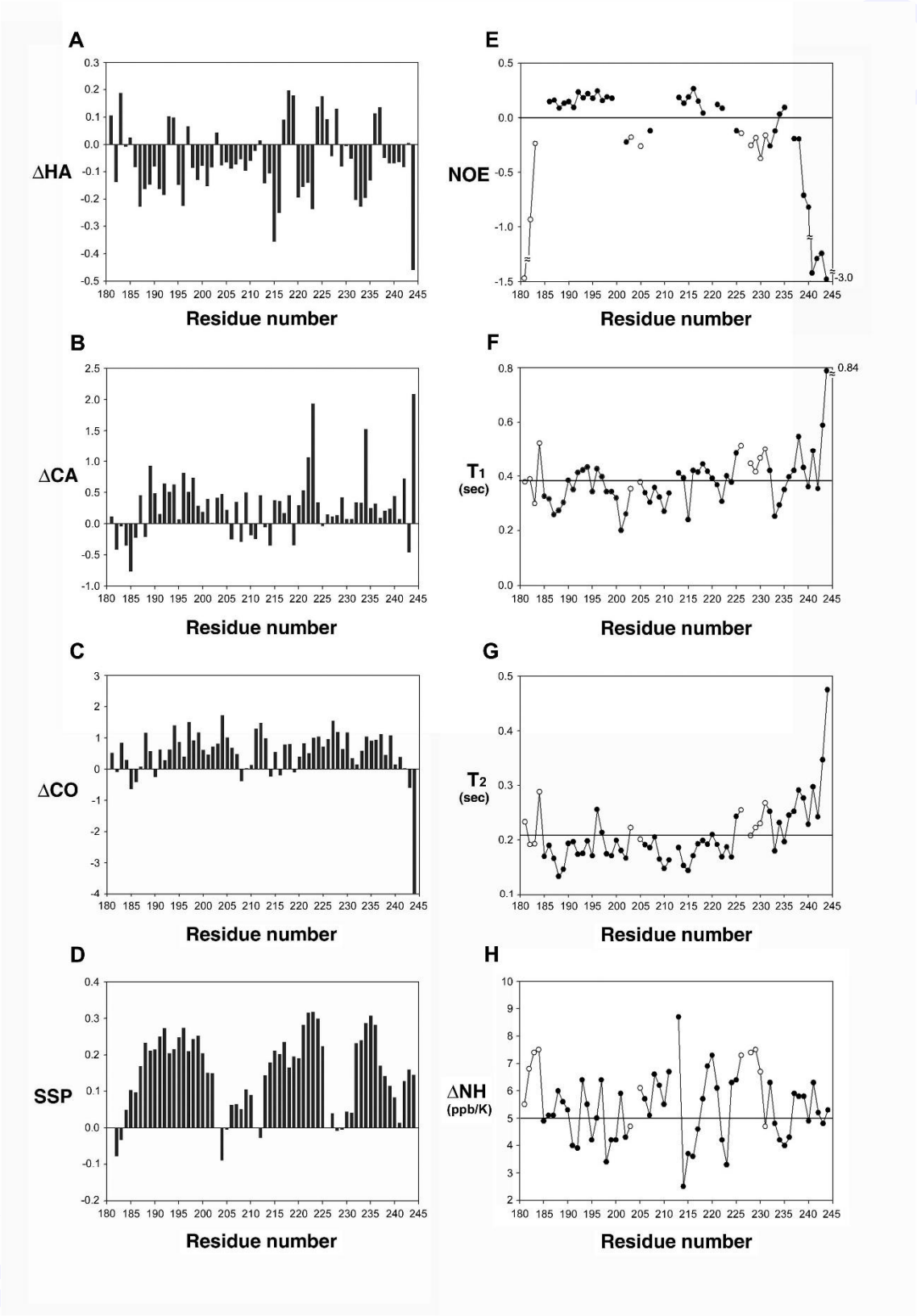


Figure 3.


A
ΔHA
Residue number
E
NOE
-3.0
Residue number
B
ΔCA
Residue number
F
T1
(sec)
0.84
Residue number
C
ΔCO
Residue number
G
T2
(sec)
Residue number
D
SSP
Residue number
H
ΔNH
(ppb/K)
Residue number

Supplementary Information

An NMR Study on the Intrinsically Disordered Core Transactivation Domain of Human Glucocorticoid Receptor

Do-Hyoung Kim[1], Anthony Wright[2], and Kyou-Hoon Han[1,3*]

MATERIALS AND METHODS

Preparation of GR tau1core for NMR experiments

DNA encoding amino acid residues 181-244 of the human alpha glucocorticoid receptor (acc. no. X03225.1) was amplified by PCR and cloned into pGEX4T-3 (Sigma-Aldrich Sweden AB, Stockholm, Sweden) cleaved with BamHI and XhoI. The insert lies downstream of the thrombin cleavage site encoded by pGEX4T-3, allowing cleavage of the free GR tau1core protein from the expressed GST fusion protein. The resulting protein consists of GR residues V181 to D244, preceded by Gly-Ser at the N-terminus. Cells transformed with the pGEX4T-3 vector containing the free GR tau1core were grown overnight on LB agar plates supplemented with 50 µg/mL ampicillin. A single colony was transferred into LB media containing ampicillin and the cells were grown until it reached an OD_{600} of 0.6. Protein expression was induced with 0.5 mM isopropyl-β-D-thiogalactopyranoside (IPTG). The cells were then further cultivated at 20°C for 16 hrs. The same procedure was used to produce the ^{15}N-labeled or $^{13}C/^{15}N$-labeled hGR tau1c except that cells were grown in M9 minimal media where the sole nitrogen source is ^{15}N labeled ammonium chloride and the sole carbon source is ^{13}C glucose, respectively. The cells were harvested by centrifugation and the bacterial pellet was resuspended in 50 mM Tris-HCl (pH 7.5), 0.4 M NaCl, 1 mM phenylmethylsulfonyl fluoride (PMSF), 10mM β–mercaptoethanol. The cells were then sonicated using an Ultra Cell TM (Sonics and Materials) and the pellet was separated by centrifugation (25,000 g for 30 min at 4°C). The supernatant was collected and filtered through a 0.45 µm syringe filter (Corning, SFCA membrane). The fraction containing hGR tau1c was applied to an SP-Sepharose column, a Q-Sepharose column, and a Sephacryl S-200 column (2.5 × 120 cm; GE Healthcare). Protein concentration was determined by measuring the absorbance using a UV spectrophotometer (GE Healthcare) in a 1 cm path length cell. The first two N-terminal residues of the recombinant protein, glycine and serine, originated from the glutathione-S-transferase fusion linker. The molecular masses of the purified proteins were confirmed by MALDI-TOF mass spectrometry.

NMR spectroscopy

For NMR experiments stable-isotope labeled (^{15}N and $^{15}N/^{13}C$) protein samples with a concentration of ~0.4 mM were prepared in 90% H_2O/10% ^2H_2O containing 20 mM sodium

acetate (pH 6.5), 50 mM NaCl, 1 mM DTT, 0.1 mM PMSF, and 0.01 mM EDTA. NMR spectra were acquired using a Varian Unity INOVA 600 MHz or a Bruker Avance II 800 MHz equipped with a cryogenic probe. Sequence-specific resonance assignment of hGR tau1c was obtained using standard multidimensional double- and triple-resonance NMR techniques as previously described (1). To achieve sequence-specific backbone and side chain assignment of all aliphatic residues of SUSP4, 2D ^{1}H-^{15}N HSQC, 3D HNCACB, 3D CBCA(CO)NH, 3D CC(CO)NH, 3D HNCO, 3D HN(CA)CO, 3D HNCA, 3D HBHA(CO)NH, 3D ^{15}N-edited TOCSY-HSQC, and ^{15}N-edited NOESY-HSQC (τ_{mix} = 80-150 ms) were obtained at 5°C. The sequence-corrected random coil chemical-shift values of Schwarzinger *et al*., (2) were used to calculate the secondary structure-related chemical shift deviations of Hα and Cα. To measure the interproton NOEs a NOESY spectrum with a mixing time of 150 ms was obtained. The secondary structure propensity program (SSP) (3) was used to calculate the propensity for secondary structures in hGR tau1c. Temperature coefficients for the backbone amide protons (Δ_{NH}) were calculated using the ^{1}H resonance assignments obtained at three temperatures (5°C, 10°C, and 15°C). All peaks were referenced to a residual water signal (4.76 ppm at 25 °C and 4.96 ppm at 5°C). The ^{15}N T_1 relaxation times were measured from spectra recorded with eight relaxation delays (20, 40, 80, 160, 320, 640, 1,280, and 2,560 ms) and the ^{15}N T_2 relaxation times were measured from spectra recorded using a CPMG sequence with eight relaxation delays (10, 30, 50, 70, 90, 130, 190, and 250 ms) according to the published procedure (4). The ^{1}H-^{15}N heteronuclear steady-state NOEs were measured from a pair of spectra recorded with and without a proton pre-saturation. All data were processed and analyzed on a Red Hat Linux version using a Varian VnmrJ, nmrPipe (5) and Sparky software.

Figure S1. Disorder prediction for the N-terminal half (~380 residues) of hGR by an IUPred program.[6] The residues with a disorder value of < 0.5 may form a globular structure. The hGR tau1c (residues 181-244) studied in this investigation is indicated by a thick box. The prediction result provides only a rough guidance to the overall disordered nature of a protein or a protein region, and does not tell if there should be any secondary structural features at all.

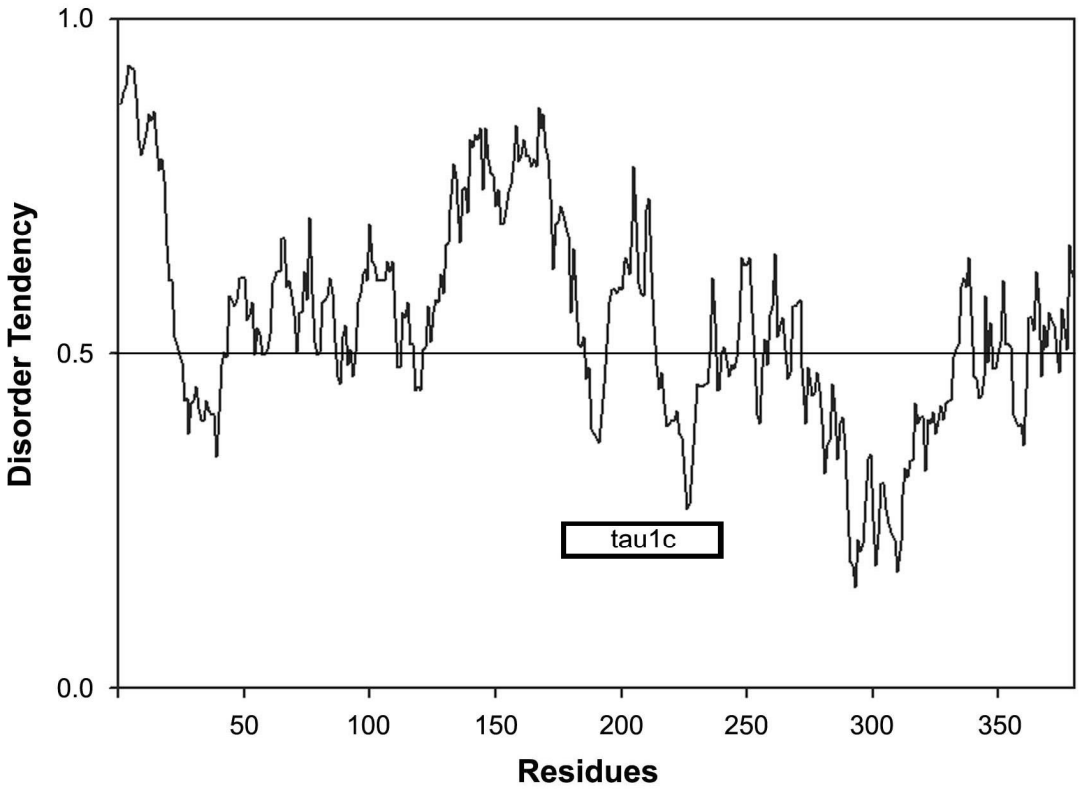


Table S1. Chemical shift assignments for the 64-residue hGR tau1c at pH 6.5 and 10 °C

Residue	^{15}N	NH	H^{α}	H^{β}	H^{γ}	H^{δ} (etc)	$^{13}C^{\alpha}$	$^{13}C^{\beta}$	^{13}CO
Val^{181}	122.35	8.35	4.03	2.08	0.95		62.33	32.85	175.86
Lys^{182}	126.36	8.48	4.17	1.69	1.34	2.96	55.96	32.93	175.73
Leu^{183}	124.47	5.31	4.24	1.55	1.44		54.71	42.61	176.71
Tyr^{184}	120.99	8.34	4.59	3.11	2.96		57.63	38.93	175.87
Thr^{185}	115.46	8.15	4.34	4.28	1.20		61.25	69.90	174.44
Thr^{186}	115.25	8.23	4.25	4.25	1.25		61.89	69.66	174.39
Asp^{187}	122.41	8.43	4.51	2.73			54.41	40.90	176.26
Gln^{188}	121.04	8.42	4.20	2.02	2.36		55.96	29.27	176.22
Ser^{189}	117.05	8.52	4.39	3.90			59.02	63.76	174.89
Thr^{190}	115.35	8.07	4.18	4.19	1.16		62.38	69.54	174.28
Phe^{191}	121.82	8.13	4.48	3.16/3.03			58.17	39.19	175.40
Asp^{192}	122.07	8.31	4.47	2.69/2.59			54.39	41.06	176.25
Ile^{193}	121.22	8.08	4.00	1.91	0.94		61.85	38.61	176.48
Leu^{194}	124.77	8.29	4.22	1.73	1.58		55.55	41.75	177.63
Gln^{195}	120.63	8.23	4.16	2.10/1.98	2.33		56.09	29.55	175.75
Asp^{196}	121.36	8.33	4.49	2.64/2.77			54.65	40.93	176.52
Leu^{197}	122.55	8.22	4.19	1.68	1.58		55.67	42.17	177.79
Glu^{198}	120.85	8.34	4.06	2.05/1.92	2.19		57.02	29.96	176.55
Phe^{199}	120.50	8.22	4.53	3.20/3.07			58.15	39.42	176.07
Ser^{200}	117.22	8.25	4.36	3.89			58.26	63.84	174.66
Ser^{201}	118.01	8.41	4.34	3.94			58.61	63.75	175.01
Gly^{202}	110.64	8.39	3.90				45.06		173.77
Ser^{203}	116.87	8.25	4.65	3.87			56.51	63.26	172.67
Pro^{204}			4.36				63.59	31.98	
Gly^{205}	109.68	8.56	3.85				45.12		174.02
Lys^{206}	120.93	8.21	4.25	1.77	1.87	1.43	56.16	33.06	176.69
Glu^{207}	122.43	8.72	4.25	2.11	1.99	2.32	56.78	29.88	176.71
Thr^{208}	115.11	8.31	4.26	4.26	1.22		61.73	69.79	174.25
Asn^{209}	121.12	8.58	4.63	2.85/2.79			53.24	38.89	175.01
Glu^{210}	121.32	8.46	4.23	2.09/1.98	2.29		56.46	30.31	176.22
Ser^{211}	118.22	8.46	4.56	3.78			56.35	63.33	172.94
Pro^{212}			4.35				63.40	31.81	
Trp^{213}	120.84	8.09	4.50	3.28			57.49	29.13	176.00
Arg^{214}	123.54	7.81	4.15	1.77/1.62	3.14	1.47	55.60	31.24	175.67
Ser^{215}	117.05	8.30	4.11	3.92/3.80			58.46	63.68	174.46
Asp^{216}	121.94	8.47	4.48	2.69			54.28	40.61	175.96
Leu^{217}	121.18	7.95	4.21	1.61			55.31	42.20	177.04
Leu^{218}	123.03	8.19	4.29	1.65			55.28	42.21	176.99
Ile^{219}	122.15	8.09	4.06	1.87/1.20	0.90		60.88	38.82	175.68
Asp^{220}	124.89	8.43	4.52	2.79/2.64			54.10	41.22	176.54
Glu^{221}	122.71	8.66	4.10	2.28/2.10	1.99		57.25	29.82	176.63
Asn^{222}	118.75	8.59	4.61	2.86			53.73	38.75	175.39
Cys^{223}	119.40	8.18	4.36	2.97			58.86	27.66	174.49
Leu^{224}	124.18	8.28	4.27	1.67			55.34	42.07	177.16
Leu^{225}	122.86	8.19	4.31	1.65			54.91	42.21	177.08
Ser^{226}	118.2	8.32	4.65	3.87			56.27	63.33	172.68
Pro^{227}			4.36				63.13	32.08	
Leu^{228}	122.36	8.37	4.24	1.63			54.98	42.10	177.22
Ala^{229}	125.31	8.35	4.22	1.43			52.64	19.26	178.10
Gly^{230}	108.68	8.44	3.96/3.91				45.05		174.22
Glu^{231}	120.44	8.35	4.22	2.27/2.09	1.96		56.56	30.25	176.36
Asp^{232}	121.15	8.50	4.52	2.75/2.70			54.29	41.17	176.15
Asp^{233}	121.91	8.37	4.53	2.69			54.39	40.67	176.85
Ser^{234}	116.47	8.37	4.19	3.85			59.64	63.47	174.77
Phe^{235}	121.74	8.12	4.49	3.03			58.17	39.11	175.75
Leu^{236}	123.03	7.96	4.17	1.62			55.28	42.21	177.07
Leu^{237}	122.81	8.13	4.23	1.66			55.04	42.13	177.41
Glu^{238}	121.73	8.35	4.19	2.31/2.08	2.00		56.73	30.30	176.91
Gly^{239}	109.96	8.44	3.87				45.25		173.82
Asn^{240}	118.80	8.39	4.70	2.86/2.78			53.05	39.22	175.34
Ser^{241}	116.78	8.53	4.40	3.91			58.37	63.74	174.32
Asn^{242}	120.86	8.60	4.46	2.87/2.80			53.47	39.03	174.99
Glu^{243}	121.27	8.40	4.26	2.15	2.29		56.27	30.44	175.32
Asp^{244}	126.92	8.06	4.28	2.70/2.59			56.03	41.91	170

REFERENCES

1. Bax A and Grzesiek S (1993) Methodological advances in protein NMR. Acc Chem Res 26, 131-138
2. Schwarzinger S, Kroon GJA, Foss TR, Chung J, Wright PE and Dyson HJ (2001) Sequence-dependent correction of random coil NMR chemical shifts. J Am Chem Soc 123, 2970-2978
3. Marsh JA, Singh VK, Jia Z and Forman-Kay JD (2006) Sensitivity of secondary structure propensities to sequence differences between α- and γ-synuclein: Implications for fibrillation. Protein Sci 15, 2795-2804
4. Lee H, Mok KH, Muhandiram R et al. (2000) Local structural elements in the mostly unstructured transcriptional activation domain of human p53. J Biol Chem 275, 29426-29432
5. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J and Bax A (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. J Biomol NMR 6, 277-293
6. Dosztányi, Z.; Csizmok, V.; Tompa, P.; Simon, I. *Bioinformatics* **2005**, *21*, 3433.