Journal of Biomolecular NMR, **24:** 73–74, 2002.
KLUWER/ESCOM

Letter to the Editor: ^{1}H, ^{13}C, and ^{15}N assignment of the N-terminal, catalytic domain of the replication initiation protein from the geminivirus TYLCV

Ramón Campos-Olivas[a,b], John M. Louis[a], Danielle Clérot[c], Bruno Gronenborn[c] & Angela M. Gronenborn[a,*]

[a]*Laboratory of Chemical Physics, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bethesda, MD 20892, U.S.A;* [b]*Structural and Computational Biology Program, Centro Nacional de Investigaciones Oncológicas, E-28029 Madrid, Spain;* [c]*Institut des Sciences Végétales, Centre National de la Recherche Scientifique, F-91198 Gif-sur-Yvette Cedex, France*

Received 17 June 2002; Accepted 30 July 2002

Key words: DNA-binding, geminivirus, Rep, replication initiation, rolling circle

Biological context

Geminiviruses are small single-stranded DNA plant viruses that replicate using a rolling circle mechanism (Gutierrez, 2000). Replication initiation and termination is accomplished by the viral Rep protein (∼360 residues), containing a N-terminal catalytic (∼130 residues), an oligomerization, and a C-terminal ATPase domain (Hanley-Bowdoin et al., 2000). As a first step in investigating the structural basis for DNA binding and catalytic activity of Rep, we report the NMR assignment for two different lengths polypeptides containing the N-terminal catalytic domain (Laufs et al., 1995) of *Tomato yellow leaf curl virus* (TYLCV). This geminivirus, a member of the *begomovirus* genus within the *geminiviridae*, causes devastating crop infections worldwide (Moffat, 1999).

Methods and results

Uniformly ^{15}N-, ^{13}C,^{15}N (100% and 10% ^{13}C) labeled proteins (H_{Tag}·Rep_{1-136}) containing a 23-residue N-terminal histidine tag (H_{Tag}, $MRGSH_6GIPGS$ $GSGD_4K$) preceding the first 136 residues of the Rep protein of TYLCV (Genbank CAA43466) were expressed using plasmid pQE-32 (Qiagen, Courtaboeuf, France) and *E. coli* strain BL21, and purified by affinity chromatography. NMR experiments on H_{Tag}·Rep_{1-136} were carried out on samples containing $\leq$0.6 mM protein in 20 mM sodium phosphate buffer pH 6.6 in the presence of 0.3 M NaCl due to limited solubility and aggregation in low salt buffer. The spectra of H_{Tag}·Rep_{1-136} suffer from low sensitivity and overlap. In addition, noise caused by the strong signals arising from both N- and C-terminal flexible regions result in spectral artifacts (see below). Consequently, assignment could only progress to ∼70% of the backbone resonances. Proteolytic cleavage of H_{Tag}·Rep_{1-136} by factor Xa resulted in the production of a protein containing residues 4 to 121 (Rep_{4-121}). As illustrated in Figure 1, Rep_{4-121} exhibited high-quality spectra and displayed significantly better solubility even at 0.1 M NaCl. Only very few minor resonances probably due to additional cleavage were noted. Consequently, all subsequent NMR data were acquired at 25 °C on samples containing 0.8–1.0 mM Rep_{4-121} in 20 mM sodium phosphate pH 6.6, 0.1 M NaCl, 1 mM DTT, 0.01% NaN_3.

Assignment of Rep_{4-121} was obtained via analysis of the following NMR experiments: 2D ^{1}H-^{1}H TOCSY and NOESY, ^{1}H-^{15}N HSQC, ^{1}H-^{13}C HMQC, HNCACB, CBCA(CO)NH, HNCO, HNHA, HNHB, 3D ^{15}N-edited NOESY, and 4D ^{13}C/^{15}N- and ^{13}C/^{13}C-edited NOESY (Bax and Grzesiek, 1993; Clore and Gronenborn, 1991). Acquisition parameters and pulse sequences were similar to those we commonly use for comparable proteins (Campos-Olivas et al., 2001). Spectra were processed with NMRPipe (Delaglio et al., 1995) and analyzed using NMRView (Johnson and Blevins, 1994). Ex-

*To whom correspondence should be addressed. E-mail: gronenborn@nih.gov

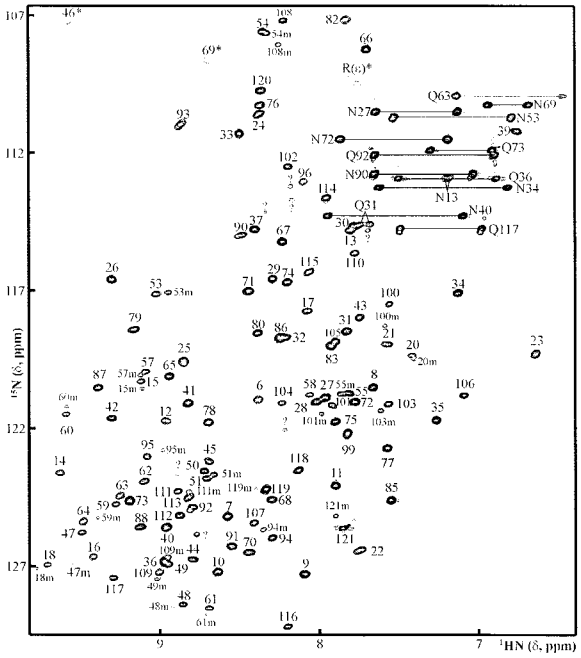

Figure 1. 2D ^{1}H-^{15}N HSQC spectrum of Rep_{4-121} recorded at 600 MHz. Backbone amide resonances are labeled according to residue position in the protein sequence and the side chain NH_2 signals are connected and labeled by residue type and number. Folded signals in the ^{15}N dimension are negative (dashed contours) and are marked with asterisks. Non-assigned signals are indicated with by question marks. Minor signals are marked with 'm'.

periments employing ^{13}C-^{13}C isotropic mixing (3D CC(CO)NH and HCCH-TOCSY) were only recorded for $H_{Tag}{\cdot}Rep_{1-136}$, since backbone assignments for this protein were incomplete. Side chain assignments for Rep_{4-121} were obtained via NOE correlations and comparison between the assignments for $H_{Tag}{\cdot}Rep_{1-136}$ and Rep_{4-121} revealed that they are essentially identical. Some side chain resonances revealed features of potential functional significance. For example, the aromatic ring of tyrosine 18 is restricted in ring flipping, such that four distinct resonances (δ_1, δ_2, ε_1, ε_2) are observed. This implies that Y_{18} is an essential component of the protein hydrophobic core and tightly packed. Interestingly, this amino acid is one of those highly conserved in the sequences of related proteins (Ilyina and Koonin, 1992). In addition, the presence of a minor set of imidazol ring resonances for H_{51}, H_{57}, and H_{59}, was revealed in a long range ^{1}H-^{15}N correlation spectrum. At pH 6.6, 4 histidines (H_{51}, H_{57}, H_{59} and H_{86}) are the neutral tautomer while H_{88} is protonated. All three cysteine residues (C_{21}, C_{47}, C_{70}) display $^{13}C_\beta$ chemical shifts indicative of being in the reduced state.

After completion of the assignments for Rep_{4-121}, backbone assignments were transferred to $H_{Tag}{\cdot}Rep_{1-136}$ using ^{1}H-^{15}N HSQC, HNCO, HNCA, and HNCOCA spectra. Comparison of the chemical shifts for equivalent residues revealed that the two different lengths proteins were essentially conformationally indistinguishable.

Extent of assignments and data deposition

Complete backbone (Figure 1) and side chain assignments (exceptions are backbone amides of E_{52}, S_{84}, S_{97}, and S_{98}) are available for Rep_{4-121}. Similarly, complete backbone assignments (^{1}HN, ^{15}N, $^{13}C_\alpha$, and $^{13}C'$ nuclei) were obtained for $H_{Tag}{\cdot}Rep_{1-136}$, including those for the N-terminal residues in the affinity tag (exceptions are residues M_{-23}, R_{-22}, H_{-19} to H_{-14}, and D_{-3}), N-terminal wild type residues M_1, P_2, R_3, and S_4, as well as residues S_{122} to K_{136} in the C-terminal extension. For Rep_{4-121}, stereospecific assignments for 24 of the 26 isopropyl methyl groups of the 3 Val and 10 Leu residues were obtained using CT-^{13}C-HSQC spectra recorded on the 10% ^{13}C-enriched sample. In addition, 42 pairs of methylene protons (5 out of the 7 Gly $H_{\alpha 1,2}$/ and 37 $H_{\beta 1,2}$ resonances) were stereospecifically assigned using HNHA/HNHB and NOE information. Complete assignments for Rep_{4-121} and backbone assignments for $H_{Tag}{\cdot}Rep_{1-136}$ have been deposited in BMRB under accession numbers 5297 and 5341, respectively.

Acknowledgements

This work was supported in part by the Intramural AIDS Targeted Antiviral Program of the Office of the Director of the National Institutes of Health to AMG, and by grant BIO2001-2287 from *Ministerio de Ciencia y Tecnología* to RCO.

References

Bax, A. and Gzesiek, S. (1993) *Acc. Chem. Res.*, **26**, 131–138.
Campos-Olivas, R., Hörr, I., Bormann, C., Jung, G. and Gronenborn, A.M. (2001) *J. Mol. Biol.*, **308**, 765–782.
Clore, G.M. and Gronenborn, A.M. (1998) *Trends Biotechnol.* **16**, 22–34.
Delaglio, F., Grzesiek, S., Vuister, G.W., Zhu, G., Pfeifer, J. and Bax, A. (1995) *J. Biomol. NMR*, **6**, 277–293.
Gutierrez, C. (2000) *EMBO J.*, **19**, 792–799.
Hanley-Bowdoin, L., Settlage, S.B., Orozco, B. M., Nagar, S. and Robertson, D. (2000) *Crit. Rev. Biochem. Mol. Biol.*, **35**, 105–140.
Ilyina, T.V. and Koonin, E.V. (1992) *Nucleic Acids Res.*, **20**, 3279.
Johnson, B. and Blevins, R.A. (1994) *J. Biomol. NMR*, **4**, 603–614.
Laufs, J., Jupin, I., David, C., Schumacher, S., Heyraud-Nitschke, F. and Gronenborn, B. (1995) *Biochimie*, **77**, 765–773.
Moffat, A.S. (1999) *Science*, **286**, 1835–1837.