

Full paper / Mémoire

# Optimizing HSQC experiment for the observation of exchange broadened signals in RNA–protein complexes

Pierre Barraud, Frédéric Dardel, Carine Tisné\*

*Laboratoire de cristallographie et RMN biologiques, Université Paris Descartes, CNRS UMR 8015,  
4, avenue de l'Observatoire, 75006 Paris, France*

Received 6 April 2007; accepted after revision 27 June 2007  
Available online 30 October 2007

## Abstract

Sites for interaction in protein–RNA complexes are often regions of conformational exchange. Although the study of exchange processes could bring valuable information about the recognition mode between protein and RNA, chemical exchange can be detrimental to the NMR spectra quality, resulting in broad, very weak or even unobservable signals. In the present report, we used CPMG-like experiments to improve HSQC spectra of an RNA–protein complex in fast exchange on the chemical shift time scale. The use of such improvement will allow us to handle the resolution of the three-dimensional structure of the complex by NMR. **To cite this article:** P. Barraud et al., *C. R. Chimie X 11 (2008)*.

© 2007 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

## Résumé

Les sites d'interaction des complexes ARN–protéine sont souvent des régions d'échange conformationnel. Même si l'étude de ces échanges peut apporter des informations importantes sur le mode de reconnaissance ARN–protéine, l'échange chimique est souvent une cause de mauvaise qualité des spectres RMN en donnant lieu à des signaux larges, de faible intensité et parfois non observables. Dans ce travail, nous avons utilisé des blocs CPMG pour améliorer l'intensité des signaux des spectres HSQC d'un complexe ARN–protéine en équilibre intermédiaire–rapide à l'échelle de temps spectrale. Cette amélioration nous permet maintenant d'envisager de résoudre la structure de ce complexe par RMN. **Pour citer cet article :** P. Barraud et al., *C. R. Chimie X 11 (2008)*.

© 2007 Académie des sciences. Published by Elsevier Masson SAS. All rights reserved.

**Keywords:** NMR; Chemical exchange; CPMG; HIV; Reverse transcription; Nucleocapsid

**Mots-clés :** RMN ; échange chimique ; CPMG ; VIH ; transcription inverse ; nucléocapside

**Abbreviations:** CPMG, Carr–Purcell–Meiboom–Gill; HIV, human immunodeficiency virus; HSQC, heteronuclear single quantum coherence; NC, nucleocapsid; PBS, primer binding site; SE, sensitivity enhanced; tRNA, transfer RNA.

\* Corresponding author.

*E-mail address:* [carine.tisne@univ-paris5.fr](mailto:carine.tisne@univ-paris5.fr) (C. Tisné).

## 1. Introduction

In the absence of radiofrequency radiation, the contribution of chemical exchange to relaxation is dependent on the rate of the exchange process, the chemical shift differences between involved sites and the population

of each site. Near coalescence, the rate of exchange is of the order of the chemical shift difference, and the contribution of chemical exchange to relaxation is therefore maximal. Regions of conformational exchange often correspond to some of the functionally most important sites in proteins, for example, sites for interactions with RNA. Although the study of exchange processes could bring valuable information about the recognition process between protein and RNA, chemical exchange is often detrimental to the NMR spectra quality resulting in broad, very weak or even unobservable signals. Several methods have been developed to characterize chemical exchange processes (for reviews, see Refs. [1–3]) and to improve NMR signal intensities in the presence of such exchange, i.e. to preserve spin coherence [4–6].

In the present report, we used previously described CPMG-like experiment [5,6] to improve HSQC spectrum of an RNA–protein complex. In this kind of experiment, the intensity of the signals involved in chemical exchange process is increased by refocusing spins during coherence transfer period. The HSQC experiment modified with a CPMG pulse train during INEPT period allowed us to improve the overall signal-to-noise ratio and to tear some signals off the noise. These NMR experiments were carried out on a complex between the HIV-1 nucleocapsid protein (NC) and an RNA hairpin that mimicks the D-arm of tRNA<sub>3</sub><sup>Lys</sup>. The HIV-1 nucleocapsid protein (NC) is a short, basic, nucleic acid binding protein with two zinc finger domains. NC is a versatile protein which is involved in multiple functions during the virus replication cycle, many of these functions relying on the nucleic acid chaperone activity of NC (for recent reviews see Refs. [7–9]). In previous studies [10], we have highlighted a strong and specific interaction between the NC protein and the D-arm of tRNA<sub>3</sub><sup>Lys</sup>. This interaction could be the key factor driving the unfolding process of tRNA<sub>3</sub><sup>Lys</sup>, thereby allowing complete annealing of tRNA<sub>3</sub><sup>Lys</sup> to the primer binding site (PBS), the first event in the HIV-1 reverse transcription. However, the NMR study of the complex between the NC protein and the D-arm of tRNA<sub>3</sub><sup>Lys</sup> is rendered tricky by chemical exchange process in the microsecond to millisecond time scale, leading to unobservable or broadened signals.

## 2. Materials and methods

### 2.1. NMR sample

The recombinant NC<sub>55</sub> protein was overexpressed from a recombinant plasmid (kind gift of M.F. Summers) and purified as previously described [11,12].

The RNA oligonucleotide mimicking the tRNA<sub>3</sub><sup>Lys</sup> D-arm (24 nucleotides length) was purchased from Dharmacon Research with 2'-*o*-bis (acetoxymethoxy)-methyl (ACE) protection. The sample was deprotected by following the manufacturer's recommendations and dialyzed several times against deionised water. The pH was adjusted between 6 and 6.5 before freeze-drying.

The <sup>15</sup>N-labelled NC<sub>55</sub> and the unlabelled RNA D-arm were prepared at about 2 mM and 10 mM, respectively, in an NMR buffer (25 mM Na-*d*<sub>3</sub>-acetate pH 6.5, 25 mM NaCl, 0.1 mM BME, 0.1 mM ZnCl<sub>2</sub>, 10% <sup>2</sup>H<sub>2</sub>O). The titration experiment is achieved by adding increased amount of RNA D-arm in a 0.5 mM sample of <sup>15</sup>N-labelled NC<sub>55</sub>. After RNA addition, the sample is immediately vortexed and kept at 30 °C during 30 min before NMR data acquisition.

### 2.2. NMR experiments

NMR experiments were carried out on a Bruker Avance DRX spectrometer operating at 600 MHz and equipped with a 5-mm triple-resonance gradient probe with *z*-axis gradients. All sequences used sine-bell shaped pulsed field gradients (PFG) along the *z*-axis. Acquisitions of data were performed at 303 K. All experiments were recorded with 512 points in the <sup>1</sup>H dimension with a spectral width of 4.6 ppm and 256 points in the <sup>15</sup>N dimension with a spectral width of 30.0 ppm. NMR data were processed with NMRPipe/NMRDraw software [13]. A pure squared sine bell and a pure squared cosine bell were applied along *t*<sub>2</sub> and *t*<sub>1</sub>, respectively. The data were zero-filled to 1024 points along *t*<sub>1</sub> and to 4096 along *t*<sub>2</sub> prior to Fourier transform. Spectra were analyzed with the software SPARKY (T.D. Goddard and D.G. Kneller, SPARKY 3, University of California, San Francisco).

The SE-HSQC experiment is a standard sensitivity-enhanced 2D <sup>1</sup>H/<sup>15</sup>N correlation via double INEPT transfer [14]. Frequency discrimination with retention of absorption mode lineshape is achieved by the echo anti-echo method [15]. Sensitivity enhancement is optimized for amide cross peaks by setting both reverse transfer periods to (1/2*J*). Heteronuclear decoupling during acquisition was achieved by GARP sequence.

The SE-CPMG-HSQC experiment (Fig. 1) is adapted from the SE-HSQC experiment in which the INEPT and reverse-INEPT periods are modified by the inclusion of a phase-modulated pulse train. This experiment is based on previously described HSQC experiments for the observation of exchange broadened signals [6]. The sequence was modified to optimize sensitivity enhancement for amide cross peaks by setting both reverse

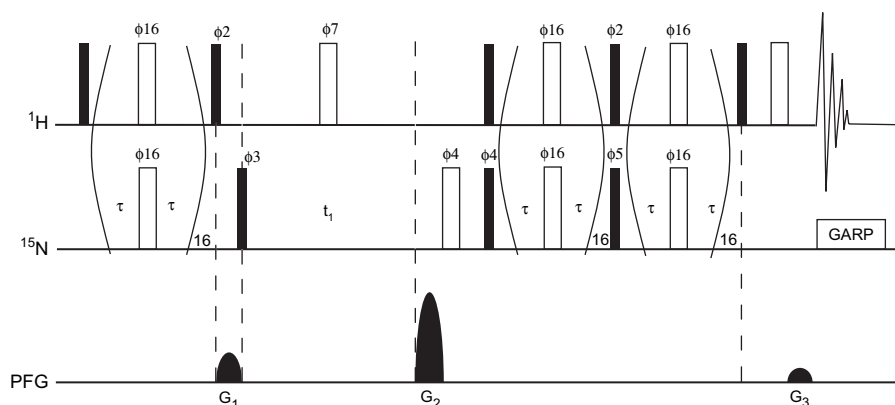


Fig. 1. Pulse scheme of the  $^{15}\text{N}$ – $^1\text{H}$  SE-CPMG-HSQC experiment. Narrow (filled) and wide (open) rectangles represent  $90^\circ$  and  $180^\circ$  pulses, respectively. Unless indicated, the pulses are applied along the  $x$ -axis. The delay  $\tau$  was set to  $140\ \mu\text{s}$  such that the total length of the spin-echo pulse train equalled  $5.55\ \text{ms}$ , i.e.  $1/2J$  with  $J = 90\ \text{Hz}$  for  $\text{N-H } ^1J$  coupling.  $^{15}\text{N}$  pulses were achieved using a field strength  $\gamma B_1 = 47\ \text{kHz}$ . The phase modulation scheme for  $\phi_{16}$  is  $x, y, x, y, x, y, x, y, x, -x, -y, -x, -y, -x, -y, -x$ . The following phase cycling was used:  $\phi_2 = y$ ;  $\phi_3 = x, -x$ ;  $\phi_7 = 2(x), 2(-x)$ ;  $\phi_4 = 2(x), 2(-x)$ ;  $\phi_5 = 2(y), 2(-y)$ ;  $\phi_{\text{rec}} = x, -x, -x, x$ . Heteronuclear decoupling during acquisition was achieved by GARP. For each  $t_1$  increment, axial peaks were shifted to the sides of the spectrum by inversion of  $\phi_3$  in concert with the receiver phase, similar as in States-TPPI [23]. The sequence uses sine-bell shaped pulsed field gradients (PFG) along the  $z$ -axis. To yield absorption mode spectra P- and N-type coherence selection is achieved by inversion of  $G_2$  in concert with inversion of  $\phi_5$ . The duration of all the gradients was  $1\ \text{ms}$ , followed by a  $400\ \mu\text{s}$  recovery delay. Gradient relative strengths were  $G_1 = 50\%$ ;  $G_2 = 80\%$ ;  $G_3 = 8.1\%$ .

transfer periods to  $(1/2J)$ . See caption of Fig. 1 for more details on NMR pulse sequence.

### 3. Results and discussion

#### 3.1. Biological context

The HIV-1 virus uses a cellular tRNA, the human  $\text{tRNA}_{3}^{\text{Lys}}$  as a primer to initiate the reverse transcription of the viral genome. In the virus, the  $3'$  end of the primer tRNA is strongly bound by 18 Watson–Crick base pairs to a complementary sequence of the viral RNA called the primer binding site (PBS). The first event in viral replication is thus the annealing of  $\text{tRNA}_{3}^{\text{Lys}}$  to the PBS, this step being promoted by the viral NC. Most of the known functions of the HIV-1 NC involve interactions with nucleic acids (for recent reviews, see Refs. [7–9]). The structure of the free NC consists in two zinc knuckle domains, connected by a flexible linker attached to flexible N- and C-terminal tails [12,16,17]. By NMR, it was shown that NC has an inherent conformational flexibility that could be crucial to allow its binding to different nucleic acid sequences and structures along the HIV-1 replication cycle [12,18]. In the currently available structures [11,19,20] of NC with nucleic acids, most of the specific contacts between the protein and the nucleic acids

occur via the zinc knuckles, notably residues  $\text{F}_{16}$  and  $\text{W}_{37}$ , interacting with the guanine-rich loop of RNA hairpin. We have shown by NMR of  $^{15}\text{N}$ -labelled  $\text{tRNA}_{3}^{\text{Lys}}$  that NC interacts preferentially with the acceptor and D stems and with the nucleotides involved in tertiary interactions between the D and T loops [21]. We also performed the reciprocal experiment, i.e. the NMR footprint of  $\text{tRNA}_{3}^{\text{Lys}}$  binding on  $^{15}\text{N}$ -labelled NC (data not shown). The signals are so broadened that roughly only the C-terminal amide group and the amino groups of side chains remain observable. Indeed, the multiplicity of NC binding sites on  $\text{tRNA}_{3}^{\text{Lys}}$  rendered analyses using the full-length tRNA intractable, even when stringency was raised by increasing salt concentration (data not shown). Reducing the target RNA length could decrease the number of non-specific sites and still have a biological interest. Indeed, studying separate hairpins cannot obviously mimic the entire tRNA, in particular at the level of the T $\Psi$ C/D loops interaction. However, in the annealing process, there is a step where the T $\Psi$ C/D loops interaction are broken and where the NC protein plays a crucial role [22]. Therefore, we have previously demonstrated that NC makes a specific interaction with the D hairpin mimicking the D-arm of  $\text{tRNA}_{3}^{\text{Lys}}$  (dissociation constant  $K_d = 28\ \text{nM}$  at  $100\ \text{mM}$  NaCl). This interaction could be a key factor driving the unfolding process,

thereby allowing complete annealing of tRNA<sup>Lys</sup><sub>3</sub> to the PBS [10]. Fig. 2a shows the HSQC experiment of <sup>15</sup>N-labelled NC with the tRNA<sup>Lys</sup><sub>3</sub> D hairpin. The NMR signals are quite broad for the size of the complex (14.2 kDa) and may indicate that the NC in the context of the complex is involved in a chemical exchange process in the microsecond to millisecond time scale. This chemical exchange is attributable to the equilibrium between the free and bound forms of the protein.

### 3.2. Analysis of the exchange process and NMR optimization of HSQC spectra

In the presence of such chemical exchange, multidimensional NMR experiments suffer from dephasing of spin coherence during periods of evolution and coherence transfer by scalar coupling [2,3]. The result of chemical exchange during periods of evolution is a broadening of the resonances, while it leads to loss of magnetization during periods of coherence transfer and therewith of signal intensity. Therefore, resonances are broad, very weak or even unobservable (Fig. 2a). The loss of magnetization during the period of coherence transfer can be minimized by using a modified INEPT block, i.e. the CPMG-INEPT block first described by Mueller and co-workers [5]. The intensity of the signals involved in the chemical exchange process is increased by refocusing them during the coherence transfer period. We then compared CPMG-HSQC [6] and SE-CPMG-HSQC described in Fig. 1, since the gain of intensity from the sensitivity-enhanced HSQC could be counter-balanced by the increase of the pulse sequence length. As a result, the SE-CPMG-HSQC spectrum shows more intense signals than those from the CPMG-HSQC (data not shown).

The comparison of the performance of the SE-HSQC and SE-CPMG-HSQC experiment is given in Fig. 2. The CPMG-derived sequence shows improved sensitivity for almost all resonances, resulting in an average enhancement of about 40%. This result confirms that the exchange process lies in the microsecond to millisecond time scale. Moreover, some unobservable signals in the classic SE-HSQC experiment grow up in the SE-CPMG-HSQC and become observable (see signals 1, 2, 3 in Fig. 2). This could be of significant importance, as these signals, which show considerably higher intensity, belong to residues directly involved in RNA binding.

Fig. 3 shows the titration of <sup>15</sup>N-labelled NC by increased amounts of D hairpin until reaching the equivalence. Two kinds of behaviour are observed: (1) a progressive variation of chemical shifts as the amount

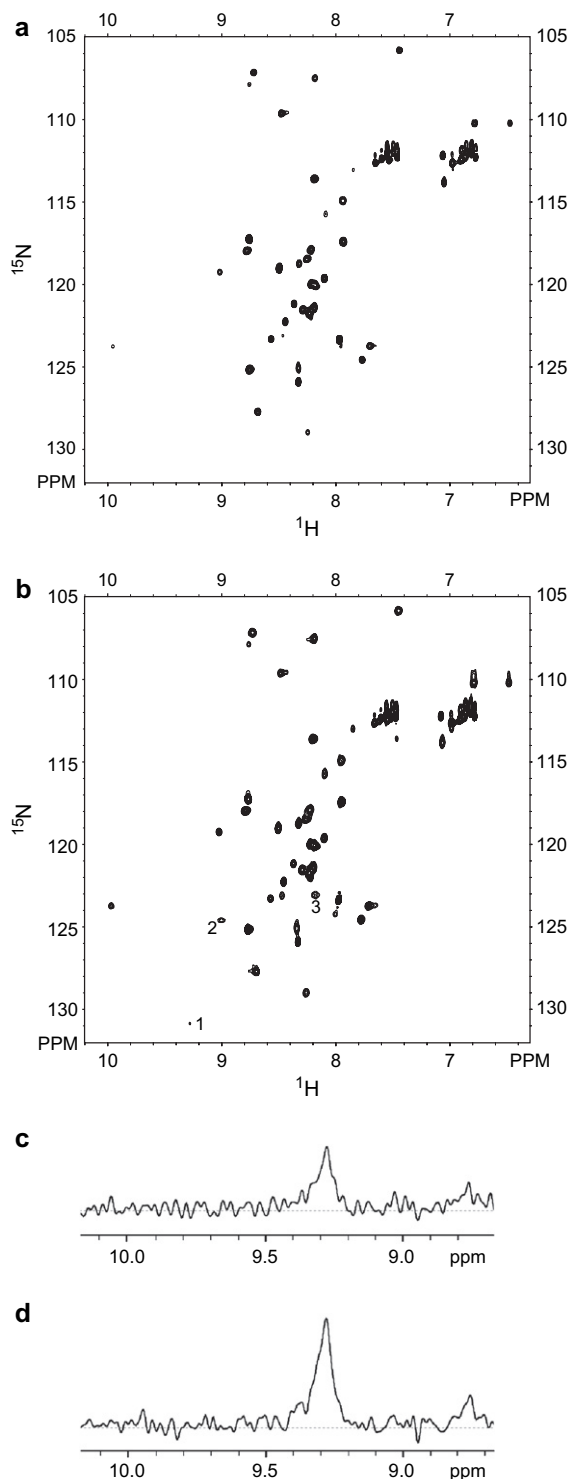


Fig. 2. Contour plot of the 2D HSQC-type spectra of the complex between uniformly <sup>15</sup>N-labelled NC and (a) unlabelled tRNA<sup>Lys</sup><sub>3</sub> D-arm (SE-HSQC experiment), (b) unlabelled tRNA<sup>Lys</sup><sub>3</sub> D-arm (SE-CPMG-HSQC experiment), (c) 1D SE-HSQC trace of signal 1 and (d) 1D SE-CPMG-HSQC trace of signal 1.

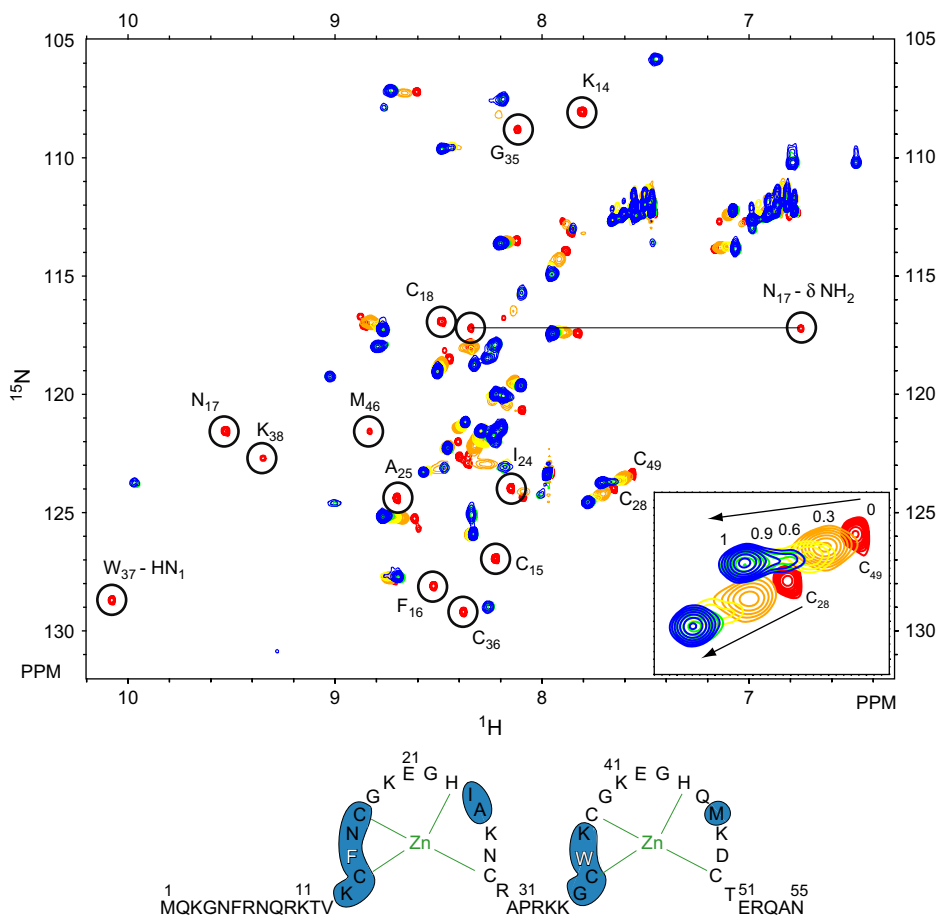


Fig. 3. Titration experiment (SE-CPMG-HSQC) of  $^{15}\text{N}$ -labelled NC protein by unlabelled  $\text{tRNA}_{3^{\text{Lys}}}$  D hairpin. In red: 1/0 protein/RNA ratio; in orange: 1/0.3; in yellow: 1/0.6, in green: 1/0.9, in blue: 1/1. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.) The circled signals disappear during the titration. The corresponding residues are highlighted on the NC sequence. The major aromatic residue of each zinc knuckle is written in open letter.

of RNA increases (see for instance the NMR signals of  $\text{C}_{28}$  and  $\text{C}_{49}$  – Fig. 3); (2) a disappearance at the substoichiometric level of NC signals and their reappearance close to the 1:1 ratio of NC:RNA (Fig. 3). These results are consistent with an exchange relaxation constant ( $R_{\text{ex}}$ ) behaviour in the fast exchange regime. Indeed, for a simple model of two-site exchange process in the fast exchange limit, the exchange relaxation constant is given by  $R_{\text{ex}} = p_a p_b (\Delta\omega^2 / k_{\text{ex}})$ , where  $k_{\text{ex}}$  is the exchange rate constant,  $\Delta\omega$  is the chemical shift difference between the two sites, and  $p_a$  and  $p_b$  the relative population of each site. For a complex with a dissociation constant more than 3 orders of magnitude smaller than the species concentrations, we can assume that the concentration of the complex in bound form is rather equal to the partner added in substoichiometric molarity. Thus, during the titration, for a fraction  $p$  of added

RNA in the protein ( $p \in [0, 1]$ ), the fraction of the bound protein is  $p$  and the fraction of the free one is  $(1 - p)$  and so  $R_{\text{ex}}$  becomes  $R_{\text{ex}} = p(1 - p)(\Delta\omega^2 / k_{\text{ex}})$ . Even if these equations apply only to the fast exchange limit ( $k_{\text{ex}} > \Delta\omega$ ), they provide useful qualitative information for the interpretation of the relaxation behaviour of a complex in a fast exchange regime. For instance, the  $R_{\text{ex}}$  maximum during the titration occurs for  $p = 0.5$ , and depends quadratically on the chemical shift difference between the free and the bound forms. Therefore, the NMR signals that experience the largest variation of chemical shifts between the free and bound forms of NC are also those that will experience the largest broadening during the titration due to the exchange regime.

We can reasonably think that these signals which present the largest differences in their chemical shifts

between the free and bound forms of NC are those involved in the interaction with the RNA partner. Actually, they belong to amino acids surrounding the major aromatic residue of each zinc knuckle, i.e. F<sub>16</sub> and W<sub>37</sub>, and to residues with hydrophobic side chain, i.e. I<sub>24</sub>, A<sub>25</sub> and M<sub>46</sub> (Fig. 3). These residues have been identified to largely participate in the recognition by the NC of RNA encapsidation signals SL2 and SL3 [11,20]. Indeed, in these structures, each zinc knuckle forms a hydrophobic cleft by means of the hydrophobic side chains listed above and binds to an exposed guanine nucleobase of the RNA loop. Thus the NMR footprint of the RNA D-arm on the NC (Fig. 3) suggests that the recognition mode between the NC and the tRNA<sub>3<sup>Lys</sup></sub> D-arm is similar to that of the encapsidation signals.

#### 4. Conclusion

In this paper, the use of modified HSQC experiment allowed us to significantly improve the signal to noise ratio in the spectra of the complex between the NC and the tRNA<sub>3<sup>Lys</sup></sub> D-arm. This experiment also allowed us to tear some signals off the noise and could be of pronounced importance in the way of the determination of the NC/D-arm complex structure by NMR. The determination of this structure is under investigation and should confirm the recognition mode similarity between the encapsidation signals and the tRNA<sub>3<sup>Lys</sup></sub> D-arm.

#### Acknowledgments

P. Barraud is supported by a studentship from Ministère de la Recherche. The French AIDS National Agency (ANRS) and 'Ensemble contre le sida (Sida-tion) supported this work.

#### References

- [1] M. Akke, *Curr. Opin. Struct. Biol.* 12 (2002) 642.
- [2] A.G. Palmer, C.D. Kroenke, J.P. Loria, Part B, *Nucl. Magn. Reson. Biol. Macromol.* 339 (2001) 204.
- [3] J.G. Kempf, J.P. Loria, *Methods Mol. Biol.* 278 (2004) 185.
- [4] H. Desvaux, P. Berthault, N. Birlirakis, M. Goldman, M. Piotto, *J. Magn. Reson.* 113 (1995) 47.
- [5] L. Mueller, P. Legault, A. Pardi, *J. Am. Chem. Soc.* 117 (1995) 11043.
- [6] F.A.A. Mulder, C.A.E.M. Spronk, M. Slijper, R. Kaptein, R. Boelens, *J. Biomol. NMR* 8 (1996) 223.
- [7] L. Kleiman, R. Halwani, H. Javanbakht, *Curr. HIV Res.* 2 (2004) 163.
- [8] J.G. Levin, J.H. Guo, I. Rouzina, K. Musier-Forsyth, *Prog. Nucleic Acid Res. Mol. Biol.* 80 (80) (2005) 217.
- [9] C. Tisné, *Curr. HIV Res.* 3 (2005) 147.
- [10] P. Barraud, C. Gaudin, F. Dardel, C. Tisné, *Biochimie* 89 (2007) 1204.
- [11] R.N. de Guzman, Z.R. Wu, C.C. Stalling, L. Pappalardo, P.N. Borer, M.F. Summers, *Science* 279 (1998) 384.
- [12] B.M. Lee, R.N. de Guzman, B.G. Turner, N. Tjandra, M.F. Summers, *J. Mol. Biol.* 279 (1998) 633.
- [13] F. Delaglio, S. Grzesiek, G.W. Vuister, G. Zhu, J. Pfeifer, A. Bax, *J. Biomol. NMR* 6 (1995) 277.
- [14] A.G. Palmer, J. Cavanagh, P.E. Wright, M. Rance, *J. Magn. Reson.* 93 (1991) 151.
- [15] L.E. Kay, P. Keifer, T. Saarinen, *J. Am. Chem. Soc.* 114 (1992) 10663.
- [16] M.F. Summers, L.E. Henderson, M.R. Chance, J.W. Bess Jr., T.L. South, P.R. Blake, I. Sagi, G. Perez-Alvarado, R.C. Sowder 3rd, D.R. Hare, et al., *Protein Sci.* 1 (1992) 563.
- [17] N. Morellet, N. Jullian, H. de Rocquigny, B. Maigret, J.-L. Darlix, B.P. Roques, *EMBO J.* 11 (1992) 3058.
- [18] S. Ramboarina, N. Srividya, R.A. Atkinson, N. Morellet, B.P. Roques, J.F. Lefevre, Y. Mely, B. Kieffer, *J. Mol. Biol.* 316 (2002) 611.
- [19] N. Morellet, H. Déméné, V. Teilleux, T. Huynh-Dinh, H. de Rocquigny, M.-C. Fournié-Zaluski, B.P. Roques, *J. Mol. Biol.* 283 (1998) 419.
- [20] G.K. Amarasinghe, R.N. de Guzman, R.B. Turner, K.J. Chancellor, Z.R. Wu, M.F. Summers, *J. Mol. Biol.* 301 (2000) 491.
- [21] C. Tisné, B.P. Roques, F. Dardel, *J. Mol. Biol.* 306 (2001) 443.
- [22] C. Tisné, B.P. Roques, F. Dardel, *J. Biol. Chem.* 279 (2004) 3588.
- [23] D. Marion, M. Ikura, R. Tschudin, A. Bax, *J. Magn. Reson.* 85 (1989) 393.