



# Assignment of the N-terminal domain of mouse cGAS

Hanna Aucharova<sup>1</sup> · Rasmus Linser<sup>1</sup>

Received: 1 November 2024 / Accepted: 18 December 2024 / Published online: 4 January 2025  
© The Author(s) 2024

## Abstract

Cyclic GMP-AMP synthase (cGAS) is a DNA-sensing enzyme that is a member of the nucleotidyltransferase (NTase) family and functions as a DNA sensor. The protein is comprised of a catalytic NTase core domain and an unstructured hypervariable N-terminal domain (NTD) that was reported to increase protein activity by providing an additional DNA-binding surface. We report nearly complete <sup>1</sup>H, <sup>15</sup>N, and <sup>13</sup>C backbone chemical-shift assignments of mouse cGAS NTD (residues 5–146), obtained with a set of 3D and 4D solution NMR experiments. Analysis of the chemical-shift values confirms that the NTD is intrinsically disordered. These resonance assignments can provide the basis for further studies such as activation by DNA and protein-protein interactions.

**Keywords** Cyclic GMP-AMP synthase · cGAS · Intrinsically disordered proteins · IDPs · NMR spectroscopy

## Biological context

cGAS is a major DNA sensor in humans and other animals that binds to double-stranded DNA and activates the STING-TBK1-IRF3 signaling axis to induce production of type I interferons and trigger innate immunity or apoptosis (Gao et al. 2013). Activation of cGAS by DNA makes it a promising drug target for autoimmune diseases and cancer (Decout et al. 2021; Gan et al. 2022). cGAS is a Mg<sup>2+</sup>- and Zn<sup>2+</sup>-binding enzyme belonging to the Mab-21-like family of nucleotidyltransferases (NTases). It was initially reported as a cytosolic DNA sensor, however, recent findings showed that the major fraction of cGAS resides in the nucleus bound to chromatin (Song et al. 2022; Jiang et al. 2019; Michalski et al. 2020), and cytoplasmic cGAS is anchored to the plasma membrane (Barnett et al. 2019). The NTD is critical for sensing nuclear chromatin, and this interaction is abrogated by multiple hyperphosphorylation at various serine and threonine residues. Chromatin-bound cGAS is kept inhibited, with its activity being maintained at a level 200-fold lower than that of cGAS stimulated with exogenous DNA (Gentili et al. 2019). Truncation of cGAS NTD leads

to cGAS mislocalization in cytoplasm and nucleus and hyper-response to genotoxic stress but a weaker response to viral infection (Barnett et al. 2019). Solution NMR studies on the cGAS N-terminal domain will enable a better understanding of its biophysical properties and associated intermolecular interactions. As a basis for such studies, we report the assignments of chemical shifts for its residues 5–146.

## Methods and experiments

### Expression and purification of cGAS NTD

The cDNA encoding *M. musculus* cGAS N-terminal domain (residues 1–146) was cloned into a pOPIN-8xHis-eGFP plasmid and overexpressed in *E. coli* RIL-X cells as an eGFP fusion protein with a PreScission protease cleavage site. To produce a uniformly <sup>15</sup>N- and <sup>13</sup>C-labeled protein, a single colony of *E. coli* RIL-X cells was inoculated in 1 L of a standard M9 minimal medium supplemented with 100 µg/ml ampicillin, 1 g/L <sup>15</sup>NH<sub>4</sub>Cl, 2 g/L <sup>13</sup>C-D-glucose, and 2 g/L <sup>15</sup>N/<sup>13</sup>C-labeled ISOGRO algal extract, and cells were grown at 37 °C and 110 rpm until an OD<sub>600</sub> of 0.8 was reached. The culture was chilled on ice for 20 min and expression was induced by the addition of isopropyl β-D-thiogalactopyranoside (IPTG) to a final concentration of 0.5 mM. Cells continued to grow upon shaking at 120 rpm and 20 °C for another 20 h, followed by harvesting by

✉ Rasmus Linser  
rasmus.linser@tu-dortmund.de

<sup>1</sup> Department of Chemistry and Chemical Biology, TU Dortmund University, Dortmund, Germany

centrifugation at 4 °C and 4,000 g for 20 min; pellets were frozen at -80 °C until further use. Frozen cell pellets were thawed at room temperature and resuspended in lysis buffer (50 mM HEPES pH 8.0, 0.3 M NaCl, 1 mM TCEP, 1 mM PMSF, 10% (w/v) glycerol, 20 mM imidazole). After vortexing, cells were lysed with an Avestin Emulsiflex C3. Lysates were aliquoted into 45 ml tubes and centrifuged at 100,000 g and 4 °C for 20 min, then incubated at 100 °C for 10 min to precipitate cellular proteases and centrifuged again as described above. The cleared lysate was applied onto a 5 ml HisTrap HP column (GE Healthcare) using an ÄKTA Pure chromatography system (GE Healthcare) in equilibration buffer. eGFP-mNTD fusion protein was eluted with a gradient of elution buffer (50 mM HEPES pH 8.0, 0.3 M NaCl, 1 mM TCEP, 1 mM PMSF, 10% (w/v) glycerol, 0.5 M imidazole) at 150–200 mM imidazole concentration. Elution fractions were then pooled together and concentrated to 2 ml using Amicon Ultra centrifugal filters with a molecular weight cut-off (MWCO) of 30 kDa (Sigma-Aldrich). The protein was buffer-exchanged into cleavage buffer (50 mM HEPES pH 8.0, 0.3 M NaCl, 1 mM TCEP) using a PD 10 desalting column (Cytiva). The protein was then mixed with 12 mg of PreScission protease (prepared in-house) and cleavage was performed overnight at 4 °C on a rocking shaker. After cleavage, the protein mixture was mixed 1:5 (v/v) with 6 M guanidine hydrochloride (Gu-HCl) to cancel electrostatic interactions of the eGFP tag and the protein. The protein was then concentrated to 1–2 ml using Amicon Ultra centrifugal filters (MWCO 3 kDa). The protein mixture was applied onto a manually packed Ni-NTA gravity flow column (5 ml) equilibrated with equilibration buffer. A total volume of 70 ml of a flow-through fraction that contained the protein was collected. The protein was then concentrated to 1–2 ml volume and renatured by rapid dilution with renaturing buffer at a ratio of 1:10 (v/v). The protein was then concentrated again to a final volume of 2 ml and exchanged into NMR buffer (20 mM MES pH 6.0, 0.1 M NaCl, 5 mM TCEP, 1 mM PMSF). For preparation of NMR samples, 270 µl of protein solution was mixed with D<sub>2</sub>O to a final concentration of 10% (v/v) and 0.1 M DSS to a final concentration of 1–3 mM.

## NMR spectroscopy

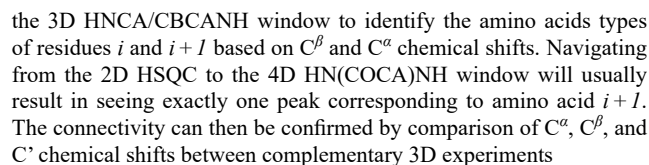
An approximately 1 mM uniformly <sup>15</sup>N/<sup>13</sup>C-labeled mNTD sample in NMR buffer was used for resonance assignments. DSS was used for direct <sup>1</sup>H chemical shift referencing, whereas <sup>13</sup>C and <sup>15</sup>N chemical shifts were indirectly referenced. Protein backbone resonances were assigned via standard 2D HSQC, 3D HNCO, 3D HNCA, 3D HN(CO)CA, 3D HN(CA)CO, and 3D CBCANH experiments, as well as a 4D HN(COCA)NH (Bracken et al. 1997) experiment.

All NMR experiments were recorded at 25 °C on a Bruker Avance III HD 800 MHz spectrometer equipped with a <sup>2</sup>H, <sup>1</sup>H, <sup>13</sup>C, <sup>15</sup>N four-channel cryogenic TCI probe. The 4D HN(COCA)NH experiment was recorded using non-uniform sampling with 0.1% sampling density and reconstructed using the SSA algorithm (Stanek and Koźmiński 2010). NMR data were processed using TopSpin (Bruker) and nmrPipe (NIST IBBR) and analyzed using CCPNMR Analysis v. 3.2.2 (Skinner et al. 2016). Automated sequence assignment was performed in FLYA (Schmidt and Güntert 2012), and prediction of chemical shifts was done in POTENCI (Nielsen and Mulder 2018).

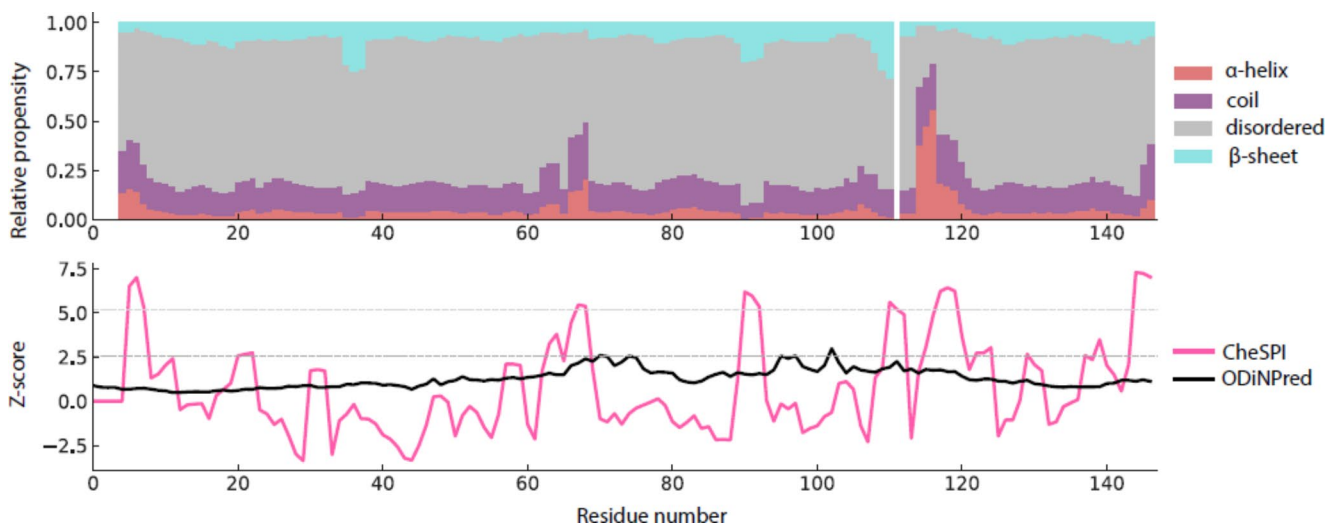
## Extent of assignments and data deposition

The mNTD <sup>1</sup>H-<sup>15</sup>N HSQC spectrum displays limited peak dispersion, with <sup>1</sup>H<sup>N</sup> chemical shifts clustered between 7.7 and 8.7 ppm, indicating a predominantly disordered structure (Fig. 1). An initial set of <sup>1</sup>H-detected 3D experiments suffered from signal overlap in the carbon dimension, and the resonances could hardly be assigned unambiguously without the 4D HN(COCA)NH spectrum. This experiment benefits from the long *T*<sub>2</sub> time of IDPs and allows to obtain unambiguous sequential assignments in a reasonably short time. Automated sequence assignment in FLYA resulted in 77% assignments of non-proline residues, which was then improved to 92% by manual assignment in combination with chemical shift prediction from POTENCI (Fig. 1a/b). Figure 1c represents the sequential connectivity through the 4D HN(COCA)NH experiment as well as the carbon chemical shifts for residues *i* and *i* + 1 for an exemplary amide peak.

The backbone N/H resonances are completely assigned other than residues Met1, Glu2, Asp3, Gln111, His127, Arg128, and Arg138. The protein primary sequence possesses two repeats, bearing the amino acid motif RGARS, in positions 39–43 and 129–133. These repeats show overlapping signals for residues Ala41 and Ala131, as well as for residues Arg42 and Arg132, which complicates unambiguous assignment. The number of assignments made for individual types of nuclei is: <sup>1</sup>H<sup>N</sup>: 113, <sup>15</sup>N<sup>H</sup>: 113, <sup>13</sup>C<sup>α</sup>: 131, <sup>13</sup>C<sup>β</sup>: 120, and <sup>13</sup>C<sup>γ</sup>: 134. The protein sequence contains 146 amino acids, of which ten are Gly (having no <sup>13</sup>C<sup>β</sup>) and 22 are Pro (lacking <sup>1</sup>H<sup>N</sup> and yielding no <sup>15</sup>N<sup>H</sup> assignments in the triple-resonance experiments). Consequently, 92% of backbone <sup>15</sup>N and <sup>1</sup>H resonances could be assigned and, excluding proline residues, 89% of <sup>13</sup>C<sup>α</sup>, 96% of <sup>13</sup>C<sup>β</sup>, and 91% of <sup>13</sup>C<sup>γ</sup> resonances could be assigned. By using C<sup>α</sup> and C<sup>β</sup> NMR chemical shifts (Δδ(C<sup>α</sup>) – Δδ(C<sup>β</sup>)), we determined the deviation from random-coil chemical shift values to identify potential α-helices, β-sheets, and disordered loops. Analysis via CheSPI (Nielsen and Mulder 2021) showed



peaks hinting to alternative proline conformations were confidently detected either. On the basis of the signal-to-noise ratio obtained, we can hence exclude major secondary (cis) proline populations (> around 10% or higher). The assigned proline shifts (major conformer) all have a  $^{13}\text{C}^\beta$  chemical shift of around 32 ppm and agree very well with (trans)



**Fig. 2** Secondary structural propensities of mouse cGAS NTD according to CheSPI, in comparison with secondary-structure Z-score calculation of the primary sequence obtained via ODINPred (Dass et al. 2020)

chemical shifts predicted by ncIDP (Tamiola et al. 2010). Minor *cis* proline populations, however, are still likely.

The profile of Z-scores derived from the experimental NMR data suggests the presence of five regions of decreased mobility within the NTD: Arg5-Arg8, Ala66-Ala68, Ala90-Arg92, Pro110-Arg119, and Lys145-Pro147 (Fig. 2, bottom). Although the NTD is largely disordered, our analysis suggests some residual order spanning the entire protein. The lower part of Fig. 2 also includes assessment of the Z-score purely on the basis of sequence, taking the ODINPred framework (Dass et al. 2020) into account.

**Acknowledgements** We are immensely grateful for continuous biochemical and strategic support, including plasmids for various cGAS constructs, by Sebastian Michalski and Karl-Peter Hopfner, Gene Center of the LMU Munich. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy - EXC 2033 - 390677874 - RESOLV, and EXC-114 - 24286268 - CiPS-M. Funded by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) - 27112786, 325871075 and the Emmy Noether program. Funded/Co-funded by the European Union (ERC, 101082494 bypassNMR). Views and opinions expressed are however those of the author(s) only and do not necessarily reflect those of the European Union or the European Research Council. Neither the European Union nor the granting authority can be held responsible for them.

**Author contributions** HA conceptualized the project, prepared the NMR samples, performed the NMR experiments, processed and analysed the data, produced the figures, and wrote and reviewed the manuscript. RL provided project supervision, reviewed the data analysis, and reviewed the manuscript.

**Funding** Open Access funding enabled and organized by Projekt DEAL.

**Data availability** The reported  $^1\text{H}$ ,  $^{13}\text{C}$  and  $^{15}\text{N}$  chemical-shift assignments of the cGAS N-terminus from *mus musculus* (residues 5 - 146)

at pH 6.0 and 25 °C have been deposited in the Biological Magnetic Resonance Data Bank (BMRB, <http://www.bmrbl.wisc.edu/>) under the BMRB accession number 52649.

## Declarations

**Competing interests** The authors declare no competing interests.

**Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

## References

- Barnett KC, Coronas-Serna JM, Zhou W, Ernandes MJ, Cao A, Kranzusch PJ, Kagan JC (2019) Phosphoinositide interactions position cGAS at the plasma membrane to ensure efficient distinction between self- and viral DNA. *Cell* 176:1432–1446.e11. <https://doi.org/10.1016/j.cell.2019.01.049>
- Bracken C, Palmer AG, Cavanagh J (1997)  $(\text{H})\text{N}(\text{COCA})\text{NH}$  and  $\text{HN}(\text{COCA})\text{NH}$  experiments for  $^1\text{H}$ - $^{15}\text{N}$  backbone assignments in  $^{13}\text{C}/^{15}\text{N}$ -labeled proteins. *J Biomol NMR* 9:94–100. <https://doi.org/10.1023/A:1018679819693/METRICS>
- Dass R, Mulder FAA, Nielsen JT (2020) ODINPred: comprehensive prediction of protein order and disorder. *Sci Rep* 2020 10:1(10):1–16. <https://doi.org/10.1038/s41598-020-71716-1>
- Decout A, Katz JD, Venkatraman S, Ablasser A (2021) The cGAS-STING pathway as a therapeutic target in inflammatory diseases. *Nat Rev Immunol* 21:548–569. <https://doi.org/10.1038/S41577-021-00524-Z>

- Gan Y, Li X, Han S, Liang Q, Ma X, Rong P, Wang W, Li W (2022) The cGAS/STING pathway: a novel target for cancer therapy. *Front Immunol* 12:795401. <https://doi.org/10.3389/FIMMU.2021.795401>
- Gao P, Ascano M, Wu Y, Barchet W, Gaffney BL, Zillinger T, Serganov AA, Liu Y, Jones RA, Hartmann G, Tuschl T, Patel DJ (2013) Cyclic [G(2',5')pA(3',5')p] is the metazoan second messenger produced by DNA-activated cyclic GMP-AMP synthase. *Cell* 153:1094–1107. <https://doi.org/10.1016/J.CELL.2013.04.046>
- Gentili M, Lahaye X, Nadalin F, Nader GFP, Puig Lombardi E, Herve S, De Silva NS, Rookhuizen DC, Zueva E, Goudot C, Maurin M, Bochnakian A, Amigorena S, Piel M, Fachinetti D, Londoño-Vallejo A, Manel N (2019) The N-terminal domain of cGAS determines preferential association with centromeric DNA and innate immune activation in the nucleus. *Cell Rep* 26:2377–2393e13. <https://doi.org/10.1016/j.celrep.2019.01.105>
- Jiang H, Xue X, Panda S, Kawale A, Hooy RM, Liang F, Sohn J, Sung P, Gekara NO (2019) Chromatin-bound cGAS is an inhibitor of DNA repair and hence accelerates genome destabilization and cell death. *EMBO J* 38(21). <https://doi.org/10.15252/EMBJ.2019102718>
- Michalski S, de Oliveira Mann CC, Stafford C, Witte G, Bartho J, Lammens K, Hornung V, Hopfner KP (2020) Structural basis for sequestration and autoinhibition of cGAS by chromatin. *Nature*. <https://doi.org/10.1038/s41586-020-2748-0>
- Nielsen JT, Mulder FAA (2018) Potenci: prediction of temperature, neighbor and ph-corrected chemical shifts for intrinsically disordered proteins. *J Biomol NMR* 70:141–165. <https://doi.org/10.1007/S10858-018-0166-5>
- Nielsen JT, Mulder FAA (2021) CheSPI: chemical shift secondary structure population inference. *J Biomol NMR* 75:273–291. <https://doi.org/10.1007/S10858-021-00374-W>
- Schmidt E, Güntert P (2012) A new algorithm for reliable and general NMR resonance assignment. *J Am Chem Soc* 134:12817–12829. <https://doi.org/10.1021/JA305091N>
- Skinner SP, Fogh RH, Boucher W, Ragan TJ, Mureddu LG, Vuister GW (2016) CcpNmr AnalysisAssign: a flexible platform for integrated NMR analysis. *J Biomol NMR* 66:111–124. <https://doi.org/10.1007/S10858-016-0060-Y>
- Song JX, Villagomes D, Zhao H, Zhu M (2022) cGAS in nucleus: the link between immune response and DNA damage repair. *Front Immunol* 13:1076784. <https://doi.org/10.3389/fimmu.2022.1076784>
- Stanek J, Koźmiński W (2010) Iterative algorithm of discrete fourier transform for processing randomly sampled NMR data sets. *J Biomol NMR* 47:65–77. <https://doi.org/10.1007/S10858-010-9411-2>
- Tamiola K, Acar B, Mulder FA (2010) Sequence-specific random coil chemical shifts of intrinsically disordered proteins. *J Am Chem Soc* 132:18000–18003. <https://doi.org/10.1021/ja105656t>

**Publisher's note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.