

Combined NMR and molecular dynamics conformational filter identifies unambiguously dynamic ensembles of Dengue protease NS2B/NS3pro

Tatiana Agback ^{1,7}, Dmitry Lesovoy^{2,3,7}, Xiao Han⁴, Alexander Lomzov ⁵, Renhua Sun ⁴, Tatyana Sandalova ⁴, Vladislav Yu. Orekhov^{3,6}, Adnane Achour ⁴✉ & Peter Agback ¹✉

The dengue protease NS2B/NS3pro has been reported to adopt either an ‘open’ or a ‘closed’ conformation. We have developed a conformational filter that combines NMR with MD simulations to identify conformational ensembles that dominate in solution. Experimental values derived from relaxation parameters for the backbone and methyl side chains were compared with the corresponding back-calculated relaxation parameters of different conformational ensembles obtained from free MD simulations. Our results demonstrate a high prevalence for the ‘closed’ conformational ensemble while the ‘open’ conformation is absent, indicating that the latter conformation is most probably due to crystal contacts. Conversely, conformational ensembles in which the positioning of the co-factor NS2B results in a ‘partially’ open conformation, previously described in both MD simulations and X-ray studies, were identified by our conformational filter. Altogether, we believe that our approach allows for unambiguous identification of true conformational ensembles, an essential step for reliable drug discovery.

¹Department of Molecular Sciences, Swedish University of Agricultural Sciences, PO Box 7015, SE-750 07 Uppsala, Sweden. ²Department of Structural Biology, Shemyakin-Ovchinnikov, Institute of Bioorganic Chemistry RAS, 117997 Moscow, Russia. ³Swedish NMR Centre, University of Gothenburg, Box 465, 40530 Gothenburg, Sweden. ⁴Science for Life Laboratory, Department of Medicine, Karolinska Institute, and Division of Infectious Diseases, Karolinska University Hospital, SE-171 76 Stockholm, Sweden. ⁵Laboratory of Structural Biology, Institute of Chemical Biology and Fundamental Medicine SB RAS, 630090 Novosibirsk, Russia. ⁶Department of Chemistry and Molecular Biology, University of Gothenburg, Box 465, 40530 Gothenburg, Sweden. ⁷These authors contributed equally: Tatiana Agback, Dmitry Lesovoy. ✉email: adnane.achour@ki.se; Peter.agback@slu.se

Infection by any of the four Dengue virus serotypes DENV1-4 can lead to Dengue fever, as well as to the significantly more severe Dengue haemorrhagic fever and/or Dengue shock syndrome. Appropriate maturation of DENV particles requires the multifunctional protein NS3 which comprises a serine protease domain (NS3pro) that is essential for polyprotein maturation and viral replication, making it an attractive drug target¹⁻⁴. Furthermore, *in vitro* studies revealed that the adequate function of the NS3pro domain requires the NS2B cofactor segment derived from the viral protein NS2 for full NS2B/NS3pro protease activity⁵.

The molecular bases underlying the function(s) of DENV-associated NS2B/NS3pro proteases, and a thorough understanding of the link(s) between overall conformation and function of this heterodimer remain unclear. A thorough understanding of unique or multiple conformations may be an essential step for the adequate development of inhibitory compounds^{1,6-8}. Several crystal structures of flaviviral serine proteases have been hitherto determined^{3,9}, including NS2B/NS3pro from all four DENV serotypes¹⁰⁻¹⁴. A comparison of these structures revealed that NS3pro always adopts a conserved fold comprising two β -barrels. Furthermore, the conformations of the catalytic residues His51 and Ser135, both localized within the active site, are similar in all four DENV and in several other Flaviviridae proteases, indicating that these structural features are conserved despite differences in sequences. Still, despite these structural similarities, it is well established that the most prevalent DENV-2 is significantly more infectious compared to the three other DENV serotypes¹⁵, indicating that differences in conformations and/or minor localized structural differences may have profound effects on the function of NS2B/NS3pro¹⁶.

In contrast to NS3pro, the overall conformation of NS2B varies between different crystal structures of this heterodimer. Therefore, the hypothetical existence of two distinct conformations for the C-terminus region of NS2B has been previously suggested¹⁷, where NS2B may adopt either a 'closed' conformation as found in a majority of inhibitor-bound complexes, or an 'open' conformation described in crystal structures of NS2B/NS3pro apo-forms^{11,17}. In the closed form, NS2B circumflexes more than 300° around the equatorial region of NS3pro, with the C-terminal β -hairpin of NS2B wrapped around the active site of NS3pro (Supplementary Fig. S1a). 'Open' NS2B conformations described in ligand-free flaviviral proteases can be divided in (i) disordered open and (ii) alternative open conformations. A stretch of only about 20 NS2B amino acids is visible in the electron density of proteases with 'disordered open' NS2B conformation (Supplementary Fig. S1b). In the 'alternative open' conformation the positioning of the first 16 NS2B residues is very similar compared to the 'closed' form, while the rest of NS2B folds away from NS3pro, retracting from the active site. Furthermore, the C-terminal stretch of residues 62-96 in NS2B forms a short α -helix followed by a β -strand composed of residues 70/73, resulting in a conformation that has been described as the 'fingerprint' for the inactive 'open' form of NS2B/NS3pro proteases¹¹ (Supplementary Fig. S1c). It should be noted that NS2B₆₂₋₉₆ is in close contact with symmetry-related molecules in all crystal structures of NS2B/NS3pro apo-forms, which may explain this alternative 'open' conformation^{11,17}. This can either result in high mobility and disorder of sections of NS2B (Supplementary Fig. S1d), or induce a possible artefactual conformation through tight contacts with surrounding molecules in the crystal (Supplementary Fig. S1e). We therefore argue here that the observed 'alternative open' conformation of the C-terminal part of NS2B is most probably induced by crystal contacts rather than describing adequate interactions with NS3pro.

The prevalence of the different possible conformations in solution remains thus elusive. Despite X-ray studies of the apo

form DENV-2 NS2B/NS3pro indicating that the C-terminus of NS2B was disordered¹¹, other NMR^{18,19} and paramagnetic labelling studies^{20,21} demonstrated that NS2B adopts predominantly a 'closed' conformation in solution. We hypothesized here that a higher localized flexibility corresponding to the 'open' conformation could be combined with the possibility for intermediate states. If valid, the 'open/closed' hypothesis would require conformational changes to occur upon *e.g.* activation by substrate or binding to an inhibitor, a concept initially supported by the crystal structure of NS2B/NS3pro in complex with an allosteric inhibitor (PDB code 4M9T)²², by molecular dynamics (MD) simulations of NS2B/NS3pro in 'open' conformation whence binding an allosteric inhibitor²³⁻²⁸ and through mutational studies forcing NS2B to assume an 'open' inactive-state conformation²⁹. This hypothesis would also imply the simultaneous presence of 'open' and 'closed' conformations, with a possible exchange time scale of milliseconds measured in ref. ³⁰. Finally, NS2B may also adopt an 'intermediate' conformation as described in the crystal structure of DENV-4 NS2B/NS3pro (PDB code 7VMV)³¹.

The combination of NMR spin relaxation spectroscopy with molecular dynamics (MD) simulations represents in our opinion methods of choice to assess the dynamics of biomolecules. Although MD simulations provided information on the motions of all atoms in NS2B/NS3pro in complex with inhibitors³²⁻³⁴, it should be noted that the results depend on the applied force field and the computational protocols³⁵. It is well established that the validation of such MD simulations by experimental results is critically important³⁶⁻⁴¹. Based on previous results^{40,42}, we have developed a protocol for conformation filtering in which MD simulation results were compared with NMR relaxation data. Since the three-dimensional structure of the apo form of DENV-2 NS2B/NS3pro complex in solution remained missing, we performed a detailed NMR investigation starting with the apo form of the Ser135Ala (NS2B/NS3proS135A) mutated protein variant. This mutation abolishes protease activity with minimal interference on the overall three-dimensional structure⁴³. We have previously reported a near complete description (>95%) of backbone ¹H^N, ¹⁵N, ¹³C ^{α} , ¹³C^O, ¹H ^{α} and sidechain ¹³C ^{β} chemical shift assignments for DENV-2 NS2B/NS3proS135A⁴⁴. We also assigned methyl resonances for the side chains of valine, leucine, and isoleucine residues.

Our results demonstrate that the choice of adequate ensembles of potential conformations for large proteins, combined with the use of force fields suitable for the task and stringent NMR data allows us to unambiguously probe the existence of different NS2B/NS3pro conformations. We assessed the binding pocket flexibility of NS3pro and the intrinsic flexibility of NS2B, and our results show the existence of mainly closed conformations, with relatively small and localized conformational shifts in limited sections of NS2B. Our results, based only on unlinked full-length NS2B/NS3pro protease heterodimers, demonstrate that the open inactive form of NS2B/NS3pro is not present in the solution, and that this alternative conformations rather appears to be mainly due to crystal packing. We believe that our approach can provide the scientific community with a more reliable template for the future development of inhibitory compounds, and that our protocol can be used for the unambiguous identification of different conformations for large proteins in solution.

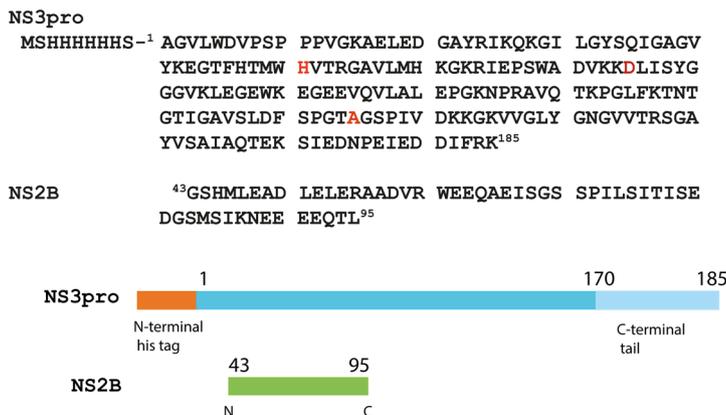
Results

Development of a protocol for NMR-restrained MD simulations. To optimally explore the conformational space and generate an ensemble of conformations, series of 100 simulated annealing (SA) steps with NOE-based restraints were performed

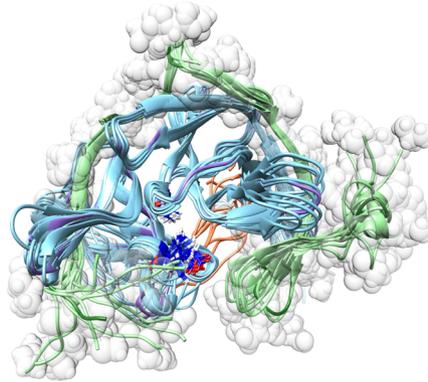
using as a starting point the three-dimensional structure of the ‘closed conformation’ of the DENV-2 NS2B/NS3proS135A heterodimer, which was created by homology modelling using SwissModel (<https://swissmodel.expasy.org/>)⁴⁵, and the crystal structure of DENV4 NS2B/NS3pro (PDB code 5YVU) as a template. Missing residues were added including the NS3pro N-terminal His-tag, and differing DENV-4 residues were replaced to the corresponding amino acids in DENV-2 (Fig. 1a) using UCSF Chimera 1.15⁴⁶. The MD simulation protocol for initial system preparation and equilibration is described in Supplementary Note S1. Our SA protocols included heating for 0.4 ns

from 300 to 500 K, followed by a cooling step to 300 K for 0.1 ns. Altogether, 366 NOE distances and 334 torsion angle restraints were applied with force constants 20 kcal/mol/Å² and 2 kcal/mol/rad², respectively (Supplementary Table S1). After SA analysis, three ensembles of NS2B/NS3pro conformations were selected, including one structure with a well-formed β-hairpin in NS2B, and two structures in which the positioning of the NS3 N-terminal his-tag differed (Supplementary Fig. S2). They were all subjected to 1μs MD simulations with similar NMR NOE distance and angle restraints. The analysis of the ten molecular models obtained by cluster analysis following the NMR-

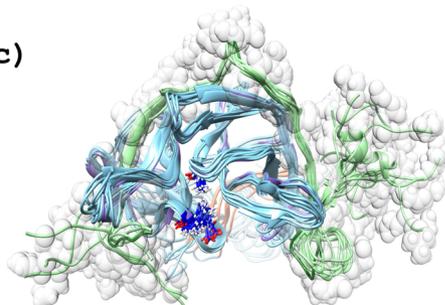
(a)



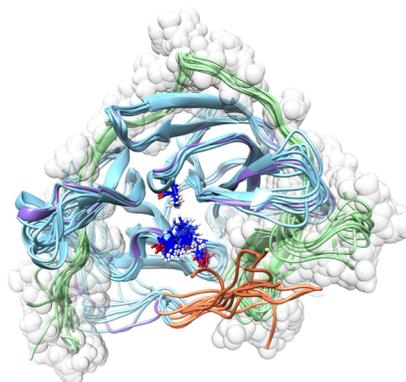
(b)



(c)



(d)



(e)

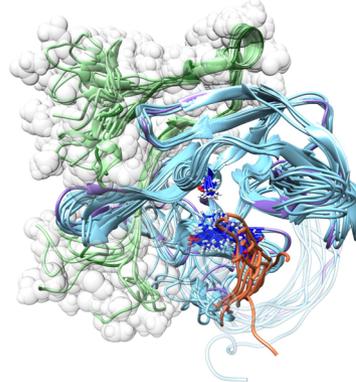


Fig. 1 The four main conformation ensembles obtained using free restraints MD simulations for DENV2 NS2B/NS3proS135A. (a) The sequences of the NS3proS135A and NS2B domains are displayed with corresponding numbering and schematic colours used in the structural models presented in (b)–(e). The catalytic triad comprising His51, Asp75 and the mutated S135A residues are indicated in dark blue. **b–e** Ribbon representations of the domain structures of four ensembles comprising 10 obtained molecular models of the DENV-2 NS2B/NS3proS135A heterodimer are presented, including: (b)–(e) ensembles I, II, III and IV, respectively. The globular NS3proS135A domain, comprising residues 1–170, is coloured in blue. N-terminal his tag is coloured in coral. The disordered NS3proS135A C-terminal tail is coloured in light blue. The starting structures used for free MD simulations presented in (b)–(e) were obtained as described in the Material and Method section, and are displayed in indigo and dark green for the NS3proS135A and NS2B domains, respectively. The surfaces corresponding to the Van der Waals radius of each heteroatom in the co-factor NS2B are displayed transparent white-grey.

restrained 1 μ s MD trajectory of the three conformation ensembles, I*, II* and III* are presented in Supplementary Note S1, Supplementary Table S2 and S3 and Supplementary Fig. S2. Importantly, based on the observed NOE in the NOESY spectrum of DENV-2 NS2B/NS3proS135A, seven of the nineteen identified intermolecular distances between NS3pro and NS2B in the range 109–117 and 68–88 aa, respectively, were below 5 Å. These distances are in agreement with the ‘closed’ but not the ‘open’ conformation. Indeed, the corresponding distances in the ‘open’ conformation calculated from the crystal structure of DENV-2 NS2B/NS3pro (PDB code 2FOM) exceed 7.7 Å. Applying SA protocol with NOE-based restraints converted the ‘open’ conformation to a partially ‘closed’ conformation ensemble. Instead it was used here directly as reported in 2FOM structure to the free MD simulation described below⁴⁷.

Identification of separate structural ensembles based on free MD calculations. Three structures obtained at the end of 1 μ s MD trajectories with NMR restraints from each of the three identified conformation ensembles I*, II* and III* were thereafter selected as starting structures for free 1 μ s MD simulations, creating novel conformations ensembles I, II and III. Ensemble I and III have a well-structured β -hairpin formed by residues 76–86 whereas for ensemble II this region is more flexible. In ensemble III the N-terminal his-tag at one end of the NS3pro domain points towards the C-terminal part of the NS2B co-factor.

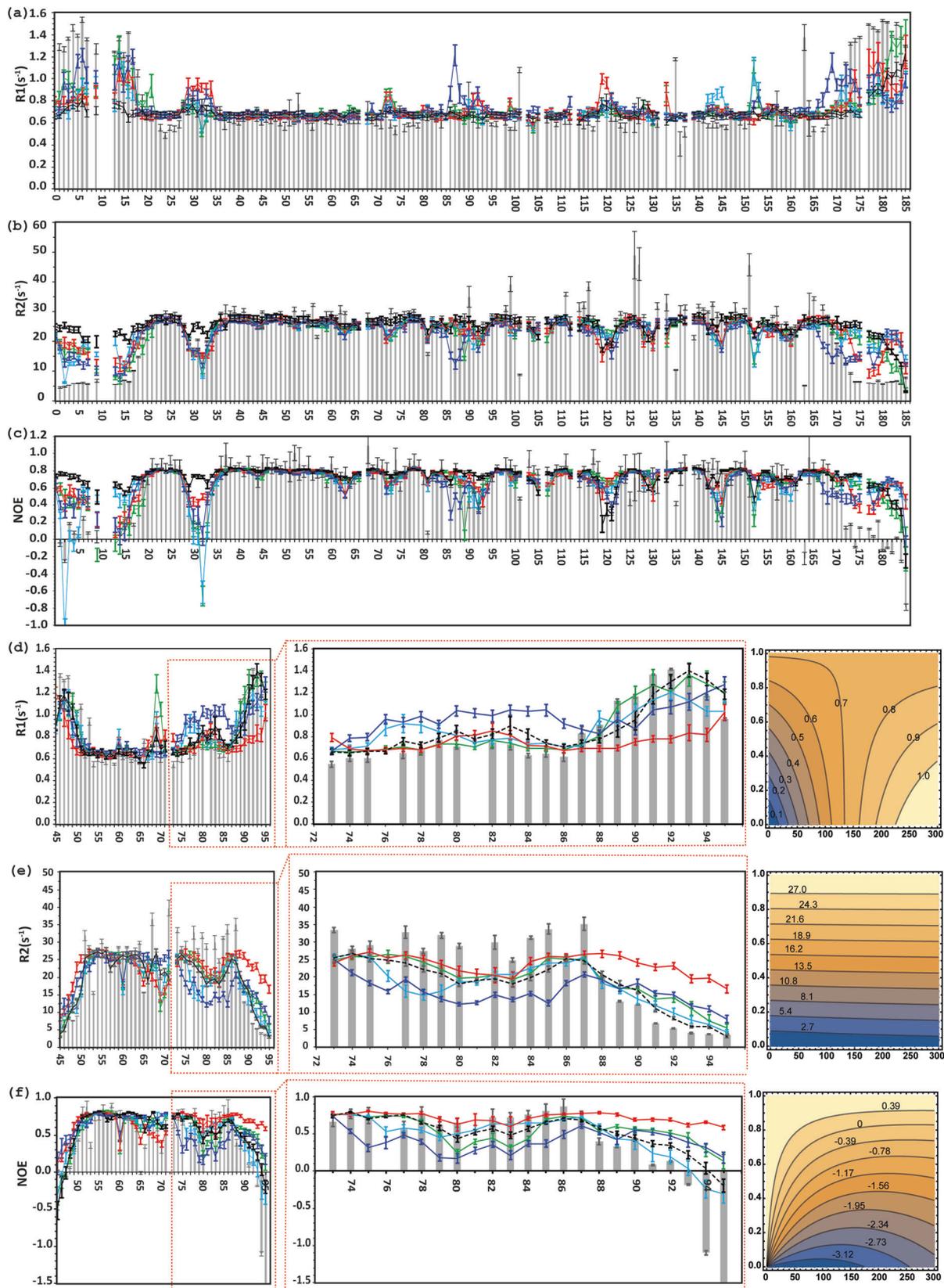
Although not identified through the protocol described above and in the supplementary material section, a fourth MD trajectory (called from now on IV) of an ‘open’ conformation was also obtained by starting from the crystal structure of the ‘open’ DENV-2 NS2B/NS3proS135A (PDB code 2FOM)¹¹. RMSD of backbone heavy atoms for dynamically stable residues 20–168 of NS3proS135A and 51–72 NS2B reached the plateau after up to 50 ns of free simulation for all ensembles, whereas all residues became stable after 500 ns (Supplementary Fig. S3). Thus, the last 500 ns were taken for further analysis. Altogether, 10 final structures were obtained by cluster analysis from free MD simulation trajectories for each of the four conformational ensembles I–IV (Fig. 1). Thereafter these MD simulated trajectories were used for back-calculation of the relaxation parameters for NS2B/NS3proS135A.

Review of relaxation parameters used for side chain methyl evaluation dynamics. To select the appropriate relaxation parameters to be evaluated and experimentally measured is crucial to this study, due to the time consuming computation processes, combined with extensive NMR experimental time. Side-chain methyl dynamics in MD simulations were verified using conventional R_1 and NOE, as well as dipole/dipolar (denoted as Γ_1 ⁴⁸ and Γ_2) and CSA/dipolar cross-correlation contributions to R_1 and R_2 (denoted by H1 and H2), respectively. We initially performed simulations of all the six relaxation parameter sets for methyl groups including R_1 , NOE, Γ_1 , Γ_2 , H1 and H2 (Supplementary Figs. S4, S5), using free MD trajectories for each different conformational ensemble (Fig. 1). Additionally, using the ‘closed’ conformation as a starting structure, identical to conformational ensemble I, the MD simulations were repeated with NMR restraints (called structural ensemble V). This additional step was performed to assess the influence of additional restraints in MD simulations on the obtained theoretical relaxation parameters as a test of force field perturbation. A value of $2.08e^{-08}$ s was used for the overall correlation time τ_c in all calculations, based on our isotropic tumbling model analysis (Supplementary Note S2). Importantly, the transverse relaxation rate R_2 for methyl groups was not included because these could not be reliably measured by

NMR due to dipole/dipolar cross-correlation effects^{42,49,50}. Careful perusing of our results presented in (Supplementary Figs. S4, S5), of the R_1 , NOE, Γ_1 , Γ_2 , H1 and H2 relaxation parameters for the different trajectories from the free MD calculations allowed us to conclude that R_1 and Γ_2 are the most sensitive values to evaluate MD trajectories profiles. These two parameters represent therefore the best sensors for conformational balance as well as potential differences among MD trajectories (Supplementary Figs. S4, S5). Additional criteria that were used within the frame of this study included the simulated curves (Supplementary Fig. S4), where the relaxation parameters R_1 , NOE, Γ_1 , Γ_2 , H1 and H2 were correlated with the internal motion of the methyl group τ_e and its amplitude S2 (Supplementary Fig. S6). According to the results from these simulations, the R_1 and Γ_2 relaxation parameters complement each other very well, showing higher sensitivities to τ_e and S2, respectively. In contrast, all the other relaxation parameters from methyl groups displayed less stringent features. Indeed, the NOE parameter was not sensitive to conformation and sequence diversities for both NS3proS135A and NS2B, and could fit to multiple values of τ_e and S2 (Supplementary Fig. S4b). Variations for H1, H2 and Γ_1 are limited in narrow ranges 0.0 to -0.5 s⁻¹, 0.0 to -1.2 s⁻¹ and 0.0 to -1.5 s⁻¹, respectively. Moreover, H1 and H2 back-calculated values strongly depend on the CSA parameter for every amino acid residue, which is not always readily available. Hence the R_1 and Γ_2 relaxation parameters were chosen for further comparison with the obtained corresponding experimental data.

Validation of structural ensembles based on backbone relaxation data. Using the four simulated trajectories I–IV with identical lengths obtained from free MD simulations (Fig. 1), the back correlation functions and dynamic parameters R_1 , R_2 and NOE of all the ¹H-¹⁵N amide backbone were calculated according to the developed protocols (See material and method section). Back-calculated theoretical versus experimental ¹⁵N R_1 , R_2 and NOE parameters for the backbone of NS2B/NS3proS135A at 600 MHz are presented for all five trajectories in Fig. 2. Furthermore, the theoretical profiles of relaxation rates simulated from different trajectories as functions of the internal motions τ_e and amplitude S² are also presented in Fig. 2.

The first conclusion that can be derived from the results presented in Fig. 2 is that the dynamic parameters ¹⁵N R_1 , R_2 and NOE (both back-calculated and experimental) for NS3proS135A residues 20–28 and 35–60 (Fig. 2a–c) and for the NS2B residues 50–64 (Fig. 2d–f) are not significantly different for all five trajectories. Thus, there are no systematic shifts between theoretical and experimental data, a conclusion that supports the validity of our back-calculation protocol presented in this study. Additionally, the S2 and τ_e parameters as qualitatively estimated from the R_1 , R_2 and NOE parameters (Fig. 2), are larger than 0.8 and 150 ps, respectively, which is in agreement with the corresponding parameters obtained with the Lipari and Szabo model^{51,52} (Supplementary Figs. S7d–h). Altogether these results clearly indicate a continuity of association between the N-terminal segment of NS2B (until residue Ser72) and the NS3 domain. The flexible nature of the backbone in some regions of NS3proS135A and NS2B is also revealed by the relaxation analysis, as identified by lower values for the heteronuclear NOE and R_2 (Fig. 2b, c, e, f), respectively. As expected, reduced NOE and R_2 values were observed for the N- and C-terminal residues 1–18 in NS3proS135A, 45–53 in NS2B, as well as for residues 170–185 in NS3proS135A and 86–95 in NS2B. It should also be noted that, clearly lower values for heteronuclear NOE were detected within the loop region linking the β 1- and β 2-strands spanning



from residues Lys28 to Ser34 in NS3proS135A, indicate its flexibility (Fig. 2c). The fast motions observed within this region were also corroborated by a substantial decrease in R_2 values (Fig. 2b) to ca $15 s^{-1}$ and an increase in R_1 values to ca $1.5 s^{-1}$ (Fig. 2a), indicating that these structural components exhibit extensive fast motion of random-coil on sub-nanosecond

timescales. Importantly this observed ‘drop’ can be reproduced by back-calculation in all four MD trajectories free of restraints.

However, the back-calculated relaxation parameters R_1 , R_2 and NOE for the molecular models corresponding to structural ensembles I-IV display clear differences compared to the fitting of the experimental data for NS2B starting from residue 72 until the

Fig. 2 DENV-2 NS2B/NS3proS135A amide backbone, $^{15}\text{N}(\text{H})$, dynamic parameters of the NS3proS135A and co factor NS2B obtained on 600 MHz spectrometers. The relaxation parameters for the NS3proS135A (a–c) and NS2B (d–f) are presented according to the following: Longitudinal relaxation rate $R_1(\text{s}^{-1})$ for NS3proS135A and NS2B are presented in (a) and (d), respectively. The transverse relaxation time $R_2(\text{s}^{-1})$ for NS3proS135A (and NS2B) are presented in (b) and (e). Heteronuclear ^1H – ^{15}N NOE values for NS3proS135A and NS2B are shown in (c) and (f). The experimentally obtained $R_1(\text{s}^{-1})$, $R_2(\text{s}^{-1})$ and NOE values are presented by grey solid brackets. The theoretically predicted dynamic parameters R_1 , R_2 and NOE, obtained through five trajectories, are displayed by solid lines and coloured in green, light blue, red, dark blue and dashed black for ensembles I, II, III, IV and V, respectively. NS2B regions corresponding to residues 72–95 are extended in boxes shown with red point lines within (d–f). The theoretical profiles of relaxation rates simulated as a function of overall correlation tumbling τ_c , internal motion τ_e and amplitude S^2 are presented in the right panels of NS2B. The error bars of the experimental data are one σ from the curve fitting and for the predicted parameters from the bootstrapping analysis.

end of the C-terminal region (Fig. 2d–f). The largest inconsistencies for experimental R_1 , R_2 and NOE were observed for the MD simulation trajectory based on the crystal structure of the ‘open’ conformation. Indeed, discrepancies between residues 72 and 88 in R_1 reach here up to 0.4 s^{-1} , and up to 10 – 15 s^{-1} in R_2 , with a reduction of the predicted NOE values. These results imply a higher mobility of the C-terminal part of NS2B compared to the overall structure of the heterodimer, and therefore demonstrate the absence of an ‘open’ conformation in solution, or at best in an extremely low amount. It should also be noted that back-calculated R_2 values for all trajectories were ca 5 s^{-1} lower compared to the corresponding experimental values for the stretch of residues 77–87 in NS2B. One potential reason underlying this result could be that back-calculated R_2 values reproduce a pure dipole-dipole relaxation term, and that the contribution of the slow exchange R_{ex} term is omitted.

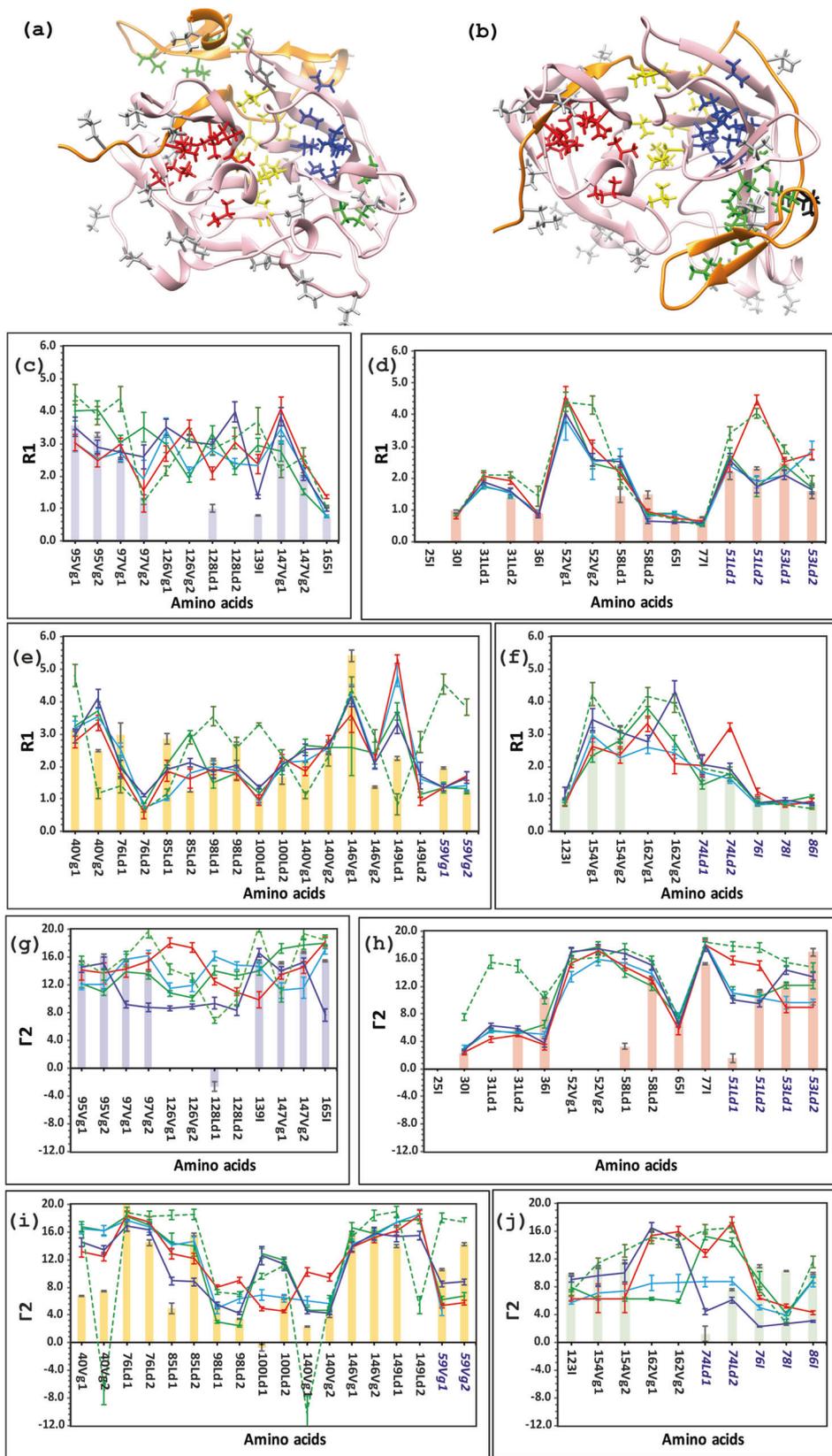
Validation of structural ensembles based on side chain methyl dynamics. The methyl longitudinal relaxation rate R_1 and the cross-cross correlated relaxation rate Γ_2 , both obtained at 800 MHz, are presented by grey solid brackets in Supplementary Fig. S8. The values for the R_1 and Γ_2 parameters span from ca 4.0 to $0.0 \text{ (s}^{-1}\text{)}$, and from ca 15.0 to $0.1 \text{ (s}^{-1}\text{)}$, respectively, in NS3proS135A (Supplementary Fig. S8a, b), covering a large range of the internal correlation time τ_c and order parameters S^2 (Supplementary Figs. S4a, S4e). Based on these results, the heterogeneity in behaviour of both the R_1 and Γ_2 profiles for NS3proS135A is evident. Nevertheless, all four trajectories that correspond to each of the conformation ensembles I–IV (Fig. 1b–d) are in agreement with the experimental data within the range of experimental uncertainties). We consider this result as remarkable in itself but also as unexpected, as it demonstrates that the position/conformation and dynamics of the NS2B co-factor have a minimal influence on the fast dynamics of the NS3proS135A domain.

The situation is different for NS2B where the heterogeneity of the Γ_2 parameters is notable (Supplementary Fig. S8c, d). Indeed, Γ_2 values for residues Ile67, Ile73 and Leu(CD1)74 are almost at zero as long as Γ_2 keeps on ca 10 s^{-1} for all other methyls of NS2B. This indicates that for Ile67, Ile73 and Leu(CD1)74 methyl the S^2 order parameters are dropping down according to estimation (Supplementary Fig. S4e). A similar trend was observed for the R_1 parameter for NS2B, where R_1 values for Ile67 and Ile73 were reduced to ca 1 s^{-1} , indicating that the internal correlation time τ_c values are below $20 \times 10^{-12} \text{ s}$ (Supplementary Fig. S4a). Furthermore, the back-calculated data for methyl side chains varied clearly in the trajectory of ensemble V (‘closed’ conformation model obtained with NMR restraints during MD calculations), compared to the corresponding experimental data for both NS3proS135A and NS2B (Supplementary Fig. S8). We consider this result important from a methodological point of view, as it indicates that free MD dynamics reproduces seemingly more correctly the dynamics of methyl side chains for Val, Leu and Ile residues compared to constrained MD calculations.

In order to identify the methyl groups that are mostly affected by interactions formed between NS2B and NS3proS135A, we combined the methyls for Val, Leu and Ile residues into four different hydrophobic clusters denominated Cl(1)–Cl(4) (Fig. 3a, b). The relaxation parameters R_1 and Γ_2 of the methyls in cluster Cl(1) are presented in Fig. 3c, g, respectively. In this cluster differences between experimental and back-calculated R_1 , Γ_2 data for the open (IV) but not the closed trajectories (I–III) are observed. Nevertheless, any conclusive choice between the trajectories is ambiguous due to the lack of a few experimental relaxation R_1 , Γ_2 data for methyls Val126 and Leu128. The relaxation R_1 , Γ_2 of methyls belonging to cluster Cl(2) are presented in Fig. 3d, h. For the four trajectories I–IV, back-calculated values of R_1 and Γ_2 fit well with the corresponding experimental relaxation R_1 and Γ_2 data. This is expected for cluster Cl(2) because it is located in the first β -barrel close to the NS2B and NS3proS135A conservative site of interaction. The relaxation R_1 , Γ_2 of methyls belonging to cluster Cl(3) are shown in Fig. 3e, i. It is located between the first and the second β -barrels, close to the active site, and thus this cluster could be very important regarding the interactions of NS2B/NS3proS135A with ligand mimicking substrates. The back-calculated R_1 , Γ_2 of methyls of four trajectories I–IV are very similar with the experimental relaxation R_1 , Γ_2 data. The most informative results were provided by cluster Cl(4) coloured in green in Fig. 3a, b. The relaxation R_1 , Γ_2 of methyls belonging to cluster Cl(4) are shown in Fig. 3f, j respectively. Despite the lack of experimental data for methyl Val162 and no difference in the R_1 relaxation parameter, the unambiguous differentiation between the four trajectories is observed for the Γ_2 relaxation parameters of Val154 in NS3proS135A, and Leu74, Ile76, Ile78 and Ile86 in NS2B. The fully ‘open’ conformation corresponding to trajectory IV displays the worse fit to the data. Indeed, this cluster is formed in Fig. 3b whereas this cluster is divided in the open conformation shown in Fig. 3a.

Discussion

Structure elucidation of proteins by NMR has traditionally focused on finding the structure with the best fit to the sets of distance and angle restraints extracted from NMR-related parameters. In general, the best fits for structures solved using NMR were reached for folded, stable proteins, resulting in good correlation with those obtained by X-ray crystallography. In contrast, more dynamic proteins have always been a challenge for structural studies. Recently, a highly dynamic nature for the NS2B/NS3pro fold in flaviviral proteases has been proposed⁵³, based on the analysis of variations in an ensemble of X-ray structures. Moreover, the recently determined crystal structure of DENV-4 NS2B/NS3pro, in which the NS2B cofactor adopted a clearly ‘intermediate’ conformation (PDB code 7VMV)³¹, opened the possibility for a revision of the commonly accepted simplistic two-step concept in which Dengue type proteases adopt only two conformations, either ‘open’ or ‘closed’. Keeping this in mind, the main goal of this study was to assess whether the inter-domain structures of the highly dynamic apo-form of the DENV-2 NS2B/NS3proS135A complex in



solution as obtained by NMR, differ from the results provided by crystal structure analyses for *e.g.* DENV-4.

To assess the flexibility of NS3proS135A and NS2B, we employed a novel tool by combining NMR spin relaxation spectroscopy and molecular dynamics (MD) computer simulations of dynamic parameters, an approach that was previously

developed and successfully applied to dynamic studies of proteins^{40,42}. We evaluated if the back-calculated dynamic parameters of different ensembles of the structures obtained in MD simulations, of equal trajectory length but starting from different starting structures, could be ranked according to the best fit to the measured dynamic parameters. Although the choice of starting

Fig. 3 DENV-2 NS2B/NS3proS135A methyl dynamic parameters of the R_1 and Γ_2 obtained on 800 MHz spectrometers. Experimental values of R_1 (s^{-1}) (c-f) and Γ_2 (s^{-1}) (g-j) for different methyl clusters in the NS2B/NS3proS135A are presented in blue (c, g), red (d, h), yellow (e, i) and green (f, j) solid brackets according to colours on panels (a, b). The theoretically predicted dynamical parameters, R_1 and Γ_2 , obtained through five trajectories, adopted different conformations ensembles shown by solid lines: green (I), light blue (II), red (III), dark blue (IV) and green dashed line (V). (a), (b) Annotation of the clusters of methyl groups in the DENV-2 mutant S135A. A model of (a) 'open' and (b) 'closed' forms of the DENV-2 S135A mutant. Four large hydrophobic clusters of methyl Ile, Val, Leu are coloured in blue, red, yellow and green for clusters Cl(1), Cl(2), Cl(3) and Cl(4), respectively. The Methyl of Ile, Val and Leu residues that are located outside of the hydrophobic cores of DENV-2 S135A mutant are coloured in light grey. The error bars of the experimental data are one σ from the curve fitting and for the predicted parameters from the bootstrapping analysis.

structures is arbitrary, it should however cover the conformational space of interest. In this study we created a family of starting structures based on all-atom MD simulations with NMR restraints of the proposed different structures of NS2B/NS3proS135A, i.e. open/semi-open/closed conformations (Fig. 1). The back-calculated relaxation parameters for every ensemble trajectory were compared with those experimentally obtained. Importantly, and advantageously, the choice of relaxation parameters of back bone and side chains of proteins to be measured can be evaluated and restricted. Only those which allow detection of the largest deviations of theoretical back-calculated parameters between different trajectories were chosen. After carefully pursuing the theoretical values for all the ensembles of NS2B/NS3proS135A (Supplementary Figs. S4–S5), the R_1 , R_2 , hetero NOE of amide bond and R_1 , Γ_2 of methyl side chain relaxation parameters, and in particular of the Val, Leu and Ile amino acids, were chosen as sufficient for the task (Figs. 2, 3, S8).

There is still an ongoing discussion about how the applied force field and the used computational protocols can influence the results of back-calculated relaxation parameters^{36–41}. We consider this discussion as outside the scope of this publication. Nevertheless, the validity of our back-calculation protocol used in this study was evaluated by comparing the theoretically obtained values against the corresponding experimental relaxation parameters located in the first β -barrel of NS3proS135A, between residues 20–28 and 35–60 (Fig. 2a–c). This approach has recently been reported⁵⁴. All trajectories in this region were expected to show similar results. Indeed, no systematic shifts between theoretical and experimental data were observed. This result also demonstrated the continuity of association between the N-terminal segment of NS2B until Ser72 and NS3proS135A for all investigated ensembles (Fig. 1). Additionally, it is evident that back-calculated parameters follow the heterogeneity values of experimental parameters outside the hot-spot of interactions formed between NS3 and NS2B.

Another issue that has been discussed is that experimentally obtained spin relaxation parameters for both backbone and methyl side chains represent an averaging of values through the superposition of different conformation ensembles. It is in our opinion evident that these parameters are dominated by the ensembles with higher populations. Nevertheless, one can argue that comparing with studies using luminescence assays⁵⁵ or ^{19}F NMR spectroscopy⁵⁶ to monitor conformational transitions, NMR obtained relaxation parameters reflect the local variation for every nuclei of sidechain and backbone⁴² directly related to structural changes due to allosteric interactions in proteins or conformational shifts between ensembles. In this study the back-calculated relaxation parameters for three conformation ensembles such as 'closed', 'partially open' or 'partially open with tag' fit to the experimental data (Figs. 2, 3). This means that all three ensembles could be present in an equilibrium. Importantly, the superposition of structurally similar states of the C-termini of NS2B includes the 'partially open' conformation similar to the one observed for DENV-4 NS2B/NS3pro in crystal state. However, the contribution of the ensemble with the 'fully open'

conformation of NS2B observed in X-ray structures should be reduced due to the large violations of the back-calculated *versus* experimental relaxation parameters in the C-termini of NS2B. Indeed, the largest differences are observed for the backbone ^1H - ^{15}N hetero NOE of Glu80 and Ser83 in NS2B stretching ca 0 *versus* 0.6 for the 'fully open' conformation *versus* the experimental values. This result allows us to conclude that if this conformation cannot be fully ruled out as irrelevant only due to the low score of fit, still this possible population can be estimated to be below the level of experimental error, ca 5%. Unfortunately, the present state of MD computation does not allow us to observe the slow transition between higher populated states and the lower populated ones, and this should be considered as a challenge for future development.

The possibility to detect by NMR low populated conformations in the slow exchange limit in proteins is well documented^{30,31,57,58}. It should be noted that the 'fully open' conformation is claimed to be a non-active conformation even if it is present according to our results only in a very minor population, if any. This is the opposite of the multiple observations for other proteins in which the 'invisible' minor forms of proteins are 'preferred' as functionally active forms. Keeping this in mind, the results obtained within the present study allow us to propose that the 'fully open' conformation is most likely a crystallisation artefact.

Our approach has also allowed us to establish that the most conserved part of NS2B/NS3proS135A where all four ensembles, including the 'fully open' conformation, fit equally to the experimental relaxation parameters. This is in our opinion a very important and unexpected result due to the reason that NS3proS135A is not properly folded in the absence of NS2B⁵⁹. Nevertheless, the position and interaction in the hot spots between NS2B and NS3 between the four different ensembles do not predict allosteric changes that may be located far away from the contacts. This could be due to that only the apo form of DENV-2 NS3/NS2B was used in the present study. Work is in progress to investigate if the approach presented in this study could be successfully applicable to monitor conformation shifts and changes in the dynamic of backbone and sidechains when the apo form interacts with allosteric ligands or peptide mimic substrates.

Our results indicate that crystal structures of the DENV-2 NS2B/NS3proS135A protease complex may have underestimated its high flexibility, as this protein can adopt multiple conformational states in the apo form in solution. Here we performed a detailed NMR study in solution of the apo form of the Ser135Ala mutated protein variant which abolishes functional activity with minimal interference on the overall three-dimensional structure. We employed a novel, to our best knowledge, newly developed tool, where back-calculated dynamic parameters obtained from free MD simulations of different trajectories of conformation ensembles were compared with experimentally obtained NMR relaxation data of the backbone and side chains of the protein. We expanded on the conformational changes occurring in DENV-2 NS2B/NS3proS135A to three conformational ensembles

obtained in long-timescale MD simulations equally satisfied with the NMR restraints. The three structures representing these ensembles showing differences in the C-termini of NS2B, called ‘closed’, ‘partially opened’ and ‘partially-open with-tag’, and one additional ‘fully open’ X-ray structure, were submitted to the follow up of free MD simulations of similar length of trajectory. R_1 , R_2 and hetero NOE of the amide bonds and R_1 and Γ_2 relaxation parameters of methyl groups were measured and theoretically predicted. First, we demonstrated the validity of our back-calculation protocol by comparing the theoretically obtained with the corresponding experimental relaxation parameters located in the first β -barrel of NS3proS135A, where no systematic shifts were observed and all trajectories in this region were expected to show similar results. We established an almost perfect correlation between all predicted *versus* experimental dynamical parameters for the three all-atom MD trajectories of conformations known as ‘closed’, ‘partially open’ and ‘partially open with tag’, indicating that they mainly contribute to conformation ensembles of NS2B in solution. Our results reveal that the main discrepancies in the C-termini of NS2B were observed for the trajectory of the ‘fully open’ conformational ensemble predicted in X-ray studies, indicating that if it exists in solution, it would correspond to a very low population. This allows us to argue that the altered conformation taken by NS2B in the crystal structure of DENV-2 NS2B/NS3proS135A protease complex observed in ‘fully open’ conformation is most probably due to crystal packing (Fig. S1). We believe that our approach to verify back-calculated relaxation parameters for chosen conformational ensembles provides a possibility for searching more realistic models that can be used for inhibitor screening and modelling, and will serve as a valuable tool in the future development of DENV protease inhibitors.

Methods

Expression constructs. The NS2B construct (comprising residues 47–95, which corresponds to amino acids 1394–1440 in the full-length DENV-2 polyprotein) and the NS3proS135A construct (residues 1–185 corresponding to amino acids 1476–1660 in DENV-2) were generated as described earlier^{19,44}. Briefly, the NS2B_{47–95} segment was subcloned into pET21b (Novagen) using the NdeI and BamHI sites. A His6-thrombin protease cleavage site was introduced at the N-terminus of NS2B. A His6 tag was also introduced at the N-terminus of the NS3proS135A domain, and thereafter subcloned into pET21b using the NdeI and XhoI sites. The S135A mutation was introduced using the QuikChange Lightning kit (Agilent, Santa Clara, CA, USA). All cloned constructs and introduced modifications were verified by DNA sequencing. All reagents were from Sigma (St. Louis, MO, USA) unless otherwise stated.

Protein expression and purification for NMR studies. The NS2B and NS3proS135A constructs were transformed into *E. coli* T7 express competent cells and expressed separately in different isotopic labelling combinations in $1/2\text{H}$, $^{15}\text{N}^{12}/^{13}\text{C}$ -labelled M9 medium. Chemicals for isotope labelling (ammonium chloride, ^{15}N (99%), D-glucose, ^{13}C (99%), deuterium oxide) were purchased from Cambridge Isotope Laboratories Inc. Protein expression was induced for 4–5 h at 37 °C by addition of β -D-1-thiogalactopyranoside (IPTG) to 1 mM final concentration, when the cell optical density at 600 nm (OD_{600}) reached 0.8. The cells were thereafter harvested by centrifugation at 6000 g.

A methyl protonated Ile δ 1- $^{13}\text{CH}_3$, Leu, Val- $^{13}\text{CH}_3/^{12}\text{CD}_3$, U- $^{15}\text{N},^{13}\text{C},^2\text{H}$ sample of NS3proS135A was obtained following previous protocols⁶⁰. The protein was expressed in 1 L D_2O M9 medium using 3 g/L of U- $^{13}\text{C},^2\text{H}$ -glucose as the main carbon

source and 1 g/L $^{15}\text{NH}_4\text{Cl}$ (CIL, Andover, MA, USA) as nitrogen source. One hour prior to induction, 70 mg/L of the precursor alpha-ketobutyric acid, sodium salt ($^{13}\text{C}_4,98\%$, 3,3- ^2H , 98%) and 120 mg/L alpha-ketoisovaleric acid, sodium salt (1,2,3,4- $^{13}\text{C}_4,99\%$, 3, 4, 4, 4- $^2\text{H}_97\%$) (CIL, Andover, MA, USA) were added to the growth medium. The growth was continued for 2 h at 37 °C and cells were thereafter harvested by centrifugation. NS2B and NS3proS135A were purified separately and refolded as previously described¹⁹. Briefly, cells were resuspended in 1x PBS buffer and lysed using ultra-sonicator followed by centrifugation at 40,000 g for 30 min. His-NS2B was collected and purified from a Ni^{2+} Sepharose 6 Fast Flow column. The target protein was eluted with 1 x PBS buffer containing 500 mM imidazole. His-NS2B and thrombin protease were added to the dialysis tubing (3.5 kD MWCO) to cleave off the His6 tag from NS2B. The protein was further purified with a size-exclusion chromatography using a HiLoad 16/60 Superdex 200 column, reaching >95% purity as checked by SDS-PAGE. His-NS3proS135A was expressed and purified as inclusion bodies, solubilization was performed overnight on a rolling board in a buffer comprising 8 M urea, 50 mM Tris pH 7.6 (RT), 20 mM imidazole, 0.5 M NaCl. NS2B and His-NS3proS135A were co-refolded by one-step dialysis overnight at 4 °C in a 2:1 molar ratio to maximize the formation of the active complex. The refolding buffer was 25 mM Tris pH 8.5 (pH set at 4 °C), 100 mM NaCl. The NS2B/His-NS3proS135A heterodimer was further purified on a HiLoad 16/60 Superdex 200 size exclusion column (Cytiva). The NS2B/His-NS3proS135A heterodimer was concentrated to 0.4–0.8 mM for data acquisition in NMR buffer containing 20 mM MES pH 6.5, 100 mM NaCl, 5 mM CaCl_2 , 1x cocktail, 0.02% NaN_3 and 10% D_2O . Deuterated MES (CIL, Andover, MA, USA) was used for all deuterated NMR samples. The His- part of the NS3proS135A will not be used anymore in the text.

NMR spectroscopy. NMR spectra of a methyl protonated Ile δ 1- $^{13}\text{CH}_3$, Leu, Val- $^{13}\text{CH}_3/^{12}\text{CD}_3$, U- $^{15}\text{N},^{13}\text{C},^2\text{H}$ sample of NS2B/NS3proS135A were recorded at 298 K on 600–900 MHz AVANCE III Bruker spectrometers equipped with 3 and 5 mm cryo-enhanced probes. Data were processed either by NMRPipe⁶¹ or TopSpin 4.06 (Bruker, Billerica, MA, USA), and analysed using CcpNmr2.4.2⁶², and Dynamics Center2.8 (Bruker, Billerica, MA, USA). Backbone and Methyl groups of Val, Leu and Ile resonance assignment for DENV-2 NS2B/NS3proS135A were obtained as previously described⁴⁴. Intra- and inter-molecular distance restraints were extracted from ^1H - ^1H Nuclear Overhauser Enhancement (NOE) data collected at 600 MHz with standard Bruker pulse sequence, 3D NOESY- ^{15}N -HSQC, with a mixing time of 100 ms. Intra- and inter-molecular distance restraints were calibrated using CcpNmr2.4.2⁶². Upper bound distances for NOE restraints were derived from NOE cross-peaks volumes using characteristic distances (sequential NH-NH NOEs in β -sheets or intra-residual NH-Methyl in alanine residues). Methyl-methyl distance restraints, derived from the 4D NUS(15%) ^{13}C , ^{13}C -methyl NOESY⁶³ experiments performed on a spectrometer at 900 MHz were processed by NMRPipe⁶¹ and the IST algorithm in the mddnmr software^{64,65} and were all set to 5.5 Å. Backbone dihedral angle restraints were predicted using the TALOS-N software⁶⁶ based on chemical shifts of ^1HN , ^{15}N , $^{13}\text{C}^\alpha$, $^{13}\text{C}^\beta$, $^{13}\text{C}'$ and $^1\text{H}^\alpha$ nuclei, previously derived for the DENV-2 NS2B/NS3proS135A heterodimer⁴⁴ (BioMagResBank accession code 51149). Predictions were converted to dihedral angle restraints with an error corresponding to twice the standard deviations given by TALOS-N. ^{15}N backbone spin relaxation measurements were performed using sensitivity-

enhanced TROSY-type pulse-sequences, with the addition of temperature compensation train of pulses before acquisition time in R_1 and R_2 relaxation measurement experiments⁶⁷. 3D pseudo spectra were recorded at a 600 and 700 MHz. For longitudinal relaxation rate (R_1) and transverse relaxation rate (R_2) measurements of the proton ^1H , spectral widths $\text{SW}(^1\text{H})$ of 16 ppm over 1024 complex points in the ^1H dimension, and 40 ppm over 128 complex points in the nitrogen ^{15}N , dimension with 24 transients (NS) were used. Relaxation delay (D1) with a duration of 1 s was used before temperature block. The R_1 relaxation value was determined from series of 11 relaxation delays including 10, 90, 192, 260, 380, 480, 690, 980, 1220 and 1444 ms with a repetition of delays corresponding to 10 ms. The R_2 relaxation value was measured using 12 Carr-Purcell-Meiboom-Gill (CPMG) delays of 8.48, 16.96, 25.44, 33.92, 42.4, 50.88, 59.36, 67.84, 76.32, 84.80 and 110.24 ms with a repetition of delays set at 8.48 ms. Both R_1 and R_2 experiments were repeated twice to estimate the experimental errors which were below 5%. Backbone (^1H) ^{15}N steady-state heteronuclear NOEs were measured using TROSY type experiments⁶⁷. 2D experiments consistent from acquisition of NOE-enhanced and an unsaturated spectra were collected using $\text{D1} = 3$ s, $\text{SW}(^1\text{H}) = 16$ ppm with 1024 complex points in the ^1H dimension and $\text{SW}(^{15}\text{N}) = 40$ ppm with 256 complex points, NS = 48. NOE values were obtained by dividing ^1H - ^{15}N peak intensities in a NOE-enhanced spectrum by the corresponding intensities in an unsaturated spectrum. An error of about 5% was set for all NOE experiments. All relaxation experiment data were evaluated using Bruker Dynamics Center2.8. The order parameter S^2 and the fast internal correlation time τ^E were obtained by fitting the relaxation parameters R_1 and R_2 , as well as the hetero-nuclear NOE values at two fields, using the Lipari-Szabo model-free approach with a NH bond length of 1.02 Å, and a ^{15}N chemical-shift anisotropy value (CSA) of -166 ppm^{51,52,68}.

Relaxation measurements for methyl $^{13}\text{C}^1\text{H}_3$ groups^{50,69} were performed at 800 MHz as a interleave pseudo-3D spectra. Experimental parameters in measurement of longitudinal relaxation rates R_1 for $^{13}\text{C}^1\text{H}_3$ were obtained as previously described⁴² by mono-exponential fitting of cross peak intensities in sets of 12 two-dimensional correlation spectra recorded with T1 relaxation delays ranging from 0.01, 0.04, 0.08, 0.13, 0.20, 0.29, 0.41, 0.57, 0.69, 0.99, 1.39 to 2.00 s. The number of transitions was set to 16. Dipolar CH, CH cross-correlation contribution to R_2 (named in this study as Γ_2) for $^{13}\text{C}^1\text{H}_3$ groups were measured as previously described^{42,69}, with a constant time period of 28.6 ms and 14 evolution delays Δ , ranging from 0.01, 0.6, 1.2, 1.8, 2.4, 3.0, 3.6, 4.2, 4.8, 5.6, 6.4, 7.2, 8.0 to 9.2 ms. The number of transitions was set to 32. For both R_1 and Γ_2 relaxation experiments on $^{13}\text{C}^1\text{H}_3$ groups, all parameters were set as previously described⁴²: the ^1H and ^{13}C carrier frequencies were set to water resonance, 4.7 and 16.5 ppm, respectively, $\text{SW}(^1\text{H})$, of 12 ppm over 1024 complex points in the ^1H dimension and 16 ppm over 80 complex points in the ^{13}C dimension, and the inter-scan delays were equal to 1 s. Processing of R_1 and Γ_2 spectra sets was performed in the TopSpin3.5pl7 software (Bruker) and analysed in the Mathematica software package (Wolfram Research Inc.) as previously described⁴².

Molecular dynamic calculations. The structure refinement and analysis of NS2B/NS3proS135A were performed using molecular dynamic (MD) simulations, based on the Amber20 software package⁷⁰ and parallel computing on central processors (CPU) and graphics accelerators (GPU) with the CUDA hardware and

software architecture, respectively. All MD simulations were performed using the ff14SBonlysc force field with improved side-chain calibration governing intermolecular interactions⁷¹. The charge of the protein residues was calculated using the pdb2pqr software v 3.1.0⁷². The length of NMR-based restrained MD calculations was 1 μs for every conformation. The length of free MD trajectories was chosen to be sufficient for collecting the data for the analyses presented in the next section and was more than 10 times the overall correlation tumbling time of molecule τ_c . Starting structures for free MD and simulation parameters can be found in Supplementary Note S3 and downloaded from the link in the data availability section. MD simulation trajectories were analysed by Hierarchical cluster analysis using the cpptraj software of Amber20⁷³.

Individual MD trajectory analyses with correlation function down-sampling. MD trajectory analyses with back-calculation of NMR spin-relaxation parameters for NH vectors were utilized as previously described⁴⁰, with the aid of the “Mathematica” software package [Wolfram Research] and the MD Analysis external library [mdanalysis.org]. The starting point of our analyses was the alignment of all protein frames onto the mean structure (for each corresponding MD trajectory), using heavy atoms from rigid backbone regions corresponding to the stretch of residues 47–87 in NS2B and 10–170 in NS3proS135A. This alignment step removes translational and rotational diffusion from MD trajectories. The auto-correlation functions $C(t)$ for the XH bond (where X can be ^{15}N , ^{13}C or ^1H) were calculated using normalized $\text{XH}(\tau)$ and $\text{XH}(\tau + t)$ vectors according to equation $2 C(t) = \langle P_2[\text{XH}(\tau)\text{XH}(\tau + t)] \rangle$ in⁴⁰ whereas for XH, XH' cross-correlation functions of two vectors $\text{XH}(\tau)$ and $\text{XH}'(\tau + t)$ were used. Moreover, the cross-correlation functions for the methyl groups were averaged over all CH vector pairs of the corresponding methyl group. Similarly, the cross-correlation functions were averaged over all HH' vector pairs. The length of correlation function $C(t)$ was calculated for the time intervals up to $7\tau_c$ was set to 145 ns (where τ_c is the overall rotational correlation of the molecule in solution) that was according to ref. ⁴⁰. For the subsequent multi-exponential approximation of the correlation functions $C(t_i)$, we used a weighted function $W(t_i)$ proportional to $|C(t_i) - C(t_{i+1})|/\sigma_i$, where σ_i is a bootstrapping standard deviation of $C(t_i)$ described below. This allowed both filtration and down-sampling of $C(t)$. More specifically, to obtain N points of the correlation function after down-sampling, the starting down-sampling point t_{dsstart} was derived from the following equation:

$$\begin{aligned} |C(t_{\text{dsstart}}) - C(t_{\text{dsstart}-1})| &= |C(t_{\text{dsstart}}) - C(t_{\text{final}})|/(N - \text{dsstart}) \\ &= |C(t_k) - C(t_{k+1})| \end{aligned}$$

for all k from d_{sstart} to N.

The values of the correlation function $C(t_k)$ in the center of the down-sampled k^{th} interval were calculated from the classical second-degree polynomial filter^{74–76}. We tested the number of $C(t_k)$ points N in ranges from 2 K till 32 K, which resulted in only insignificant differences compared to the back-calculated NMR relaxation parameters. Therefore, the N value was set to 2 K. The down-sampling for cross-correlation was performed in a similar way. Examples of auto- and cross-correlation function down-sampling are presented in Fig. 4.

Calculation of theoretical NMR spectral densities from correlation functions. The down-sampled auto-correlation function was fitted to a multi-exponential decay $C(t) = A_0 + \sum_{j=1}^m A_j e^{-t/\tau_j}$

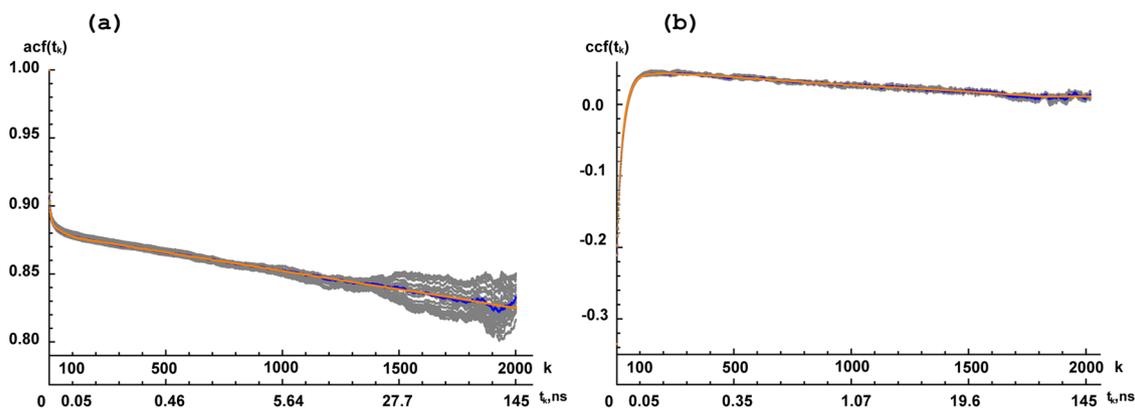


Fig. 4 Examples of auto- and cross-correlation functions down-sampling and approximation. Panel (a) displays the auto-correlation function, $acf(t_k)$, of 75 S NS2B NH vectors for trajectory I. Panel (b) presents the cross-correlation CHCH' function, $ccf(t_k)$, for 73ICD1 methyl group vectors for NS2B trajectory I. The abscissa axes represent the correlation function values, whereas the two ordinate axes represent the point number k after down-sampling and the time t as a function of k . The blue curves present data obtained from the calculated MD trajectories. The average of the results of four exponential approximations are represented by the yellow curves. The bootstrapping statistical analysis (subset of 32 curves, each using one-half fraction of MD trajectory) are shown in grey.

as described in equation 9 in⁴⁰. The exponent number m for each correlation function was tested from 2 to 7. An approximation was performed in the Mathematica program utilizing the Levenberg-Marquardt minimization approach as described in the Supplementary Note S4 section. Representative fits of different auto- and cross-correlations to multi-exponential models are presented in Fig. 4. Finally, the best-fit parameters A_0 , A_j , and τ_j provide spectral density functions according to equation 10 in⁴⁰:

$$J(\omega) = \frac{A_0 2\tau_c}{1 + (\omega\tau_c)^2} + \sum_{j=1}^m \frac{A_j 2\tau_j'}{1 + (\omega\tau_j')^2},$$

where $\tau_j' = \tau_c \tau_j / (\tau_c + \tau_j)$ τ_c is the experimental rotation correlation time.

Back-calculation of theoretical relaxation parameters from spectral density functions. For ¹⁵N uniformly-labelled proteins, classical NMR ¹⁵N relaxation rates R_1 , R_2 and heteronuclear NOE were provided as a function of spectral densities^{40,77,78}. For the methyl ¹³C¹H₃ groups in the (¹⁵N/¹³C/2H)-¹³CH₃-labelled protein, the longitudinal relaxation rate R_1 and the dipole-dipole cross-correlation contributions to R_2 (Γ_2) were calculated using the following equations previously described in^{50,69,79}:

$$R_1 = \frac{3}{20} \left(\frac{\mu_0 h \gamma_H \gamma_C}{8\pi^2 r_{CH}^3} \right)^2 [J(\omega_H - \omega_C) + 3J(\omega_C) + 6J(\omega_H + \omega_C)] + \frac{1}{20} \left(\frac{\mu_0 h \gamma_C \gamma_C}{8\pi^2 r_{CC}^3} \right)^2 [J(0) + 3J(\omega_C) + 6J(2\omega_C)] + \frac{\omega_C^2 \Delta\sigma^2}{15} J(\omega_C)$$

$$\Gamma_2 = \frac{1}{20} \left(\frac{\mu_0 h \gamma_H \gamma_C}{8\pi^2 r_{CH}^3} \right)^2 [4J_{CH,CH'}(0) + 3J_{CH,CH'}(\omega_C)] + \frac{1}{20} \left(\frac{\mu_0 h \gamma_H^2}{8\pi^2 r_{HH}^3} \right)^2 [3J_{HH',HH''}(\omega_H) + 3J(2\omega_H)],$$

where μ_0 is the vacuum permeability; h is Planck's constant; γ_H and γ_C are the gyromagnetic ratios of ¹H and ¹³C, respectively; $\Delta\sigma$ is the chemical shift anisotropy (CSA) of ¹³C = 25ppm and ¹⁵N CSA = -166ppm⁸⁰; $r_{NH} = 1.016\text{\AA}$ ⁸¹, for CH, CC' and HH distances in methyl group are $r_{CH} = 1.07\text{\AA}$, $r_{CC} = 1.533\text{\AA}$; and

$r_{HH} = 1.80\text{\AA}$ ⁸², respectively; ω_H and ω_C are Larmor frequencies of spins ¹H and ¹³C at 800 MHz, respectively; J is the auto-correlation spectral density; J' and $J_{HH',HH''}$ are cross-correlation spectral densities for pairs CH-CH' and HH'-HH'' according to^{50,69}, respectively. To exclude the R_1 systematic errors arising as a result of the methyl rotation torsion potential uncertainties discussed in detail in⁸³ the methyl MD back-calculated R_1 values were normalized *versus* the experimental values.

Statistics and reproducibility

Block bootstrapping statistical analysis of back-calculated relaxation parameters. The traditional analysis of a set of MD trajectories under the same conditions, known as classical bootstrapping, involves using random subsets of the MD trajectories. However, in our study, the primary focus of the manuscript is the development of a method for ranking individual MD trajectories that correspond to unique starting structures. These starting structures can be obtained through various methods including MD annealing results, NMR-derived structures, single X-ray structures, Cryo EM structures and even from AlphaFold prediction structures. Consequently, the dispersion of the back-calculated parameters need to be estimated within each individual trajectory corresponding to the single structure. This is in contrast to the tradition method of error estimation. To address this, we employed a moving block bootstrapping procedure with overlapping blocks^{84,85} to estimate the dispersion of the back-calculation parameters from individual MD trajectories. As mentioned above the equation $C(t) = \langle P_2[XH(\tau)XH(\tau+t)] \rangle$ was used for correlation function back-calculation with averaging over all τ values, whereas a bootstrapping used averaging over a half of τ values. The 32 bootstrapping blocks were used ranging from (0; $\tau_{\text{maximum}}/2$) till ($\tau_{\text{maximum}}/2$; τ_{maximum}) resulting in 32 correlation functions $C(t)$ presented in grey in Fig. 4. All 32 correlation functions undergo the same procedures as the main one: down-sampling preserving same t_k points, multi-exponential approximation, and the subsequent 32 relaxation parameter back-calculated. The standard deviation of the 32 relaxation parameter values divided by $\sqrt{2}$ were used as a discrepancy of the value in the particular single trajectory. The final errors (presented in figures N, N + 1, N + 2...) of back-calculated relaxation parameters were a geometric mean of bootstrapping discrepancy described above and

systematic errors estimated from the rNH, rCH and ^{15}N or ^{13}C CSA uncertainties.

Reporting summary. Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The starting structures that supports the findings in this study and the numerical sources for the graphs are available from: <https://doi.org/10.5878/5wea-fk76>. Other information and data are available upon request from the authors.

Code availability

For NMR TopSpin3.5pl7, TopSpin4.0.6 (Bruker.com) and NMRPipe (NMRbox.org) were used. For MD calculations Amber20 (ambermd.org). For analysis of MD: Mathematica (wolfram.com) with the MD Analysis external library (mdanalysis.org). Other information is available upon request from the authors.

Received: 16 June 2023; Accepted: 14 November 2023;

Published online: 24 November 2023

References

- Lim, S. Y. M., Chieng, J. Y. & Pan, Y. Recent insights on anti-dengue virus (DENV) medicinal plants: review on in vitro, in vivo and in silico discoveries. *All Life* **14**, 1–33 (2021).
- Luo, D., Vasudevan, S. G. & Lescar, J. The flavivirus NS2B-NS3 protease-helicase as a target for antiviral drug development. *Antivir. Res.* **118**, 148–158 (2015).
- Nitsche, C. Proteases from dengue, West Nile and Zika viruses as drug targets. *Biophys. Rev.* **11**, 157–165 (2019).
- Timiri, A. K., Sinha, B. N. & Jayaprakash, V. Progress and prospects on DENV protease inhibitors. *Eur. J. Medicinal Chem.* **117**, 125–143 (2016).
- Yusof, R., Clum, S., Wetzel, M., Murthy, H. M. K. & Padmanabhan, R. Purified NS2B/NS3 serine protease of dengue virus type 2 exhibits cofactor NS2B dependence for cleavage of substrates with dibasic amino acids in vitro. *J. Biol. Chem.* **275**, 9963–9969 (2000).
- Guzman, M. G. & Harris, E. Dengue. *Lancet* **385**, 453–465 (2015).
- Jacobs, M. G. & Young, P. R. Dengue: a continuing challenge for molecular biology. *Curr. Opin. Infect. Dis.* **11**, 319–324 (1998).
- van den Elsen, K., Quek, J. P. & Luo, D. Molecular Insights into the Flavivirus Replication Complex. *Viruses* **13**. <https://doi.org/10.3390/v13060956> (2021)
- Nitsche, C., Holloway, S., Schirmeister, T. & Klein, C. D. Biochemistry and medicinal chemistry of the dengue virus protease. *Chem. Rev.* **114**, 11348–11381 (2014).
- Chandramouli, S. et al. Serotype-specific structural differences in the protease-cofactor complexes of the dengue virus family. *J. Virol.* **84**, 3059–3067 (2010).
- Erbel, P. et al. Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. *Nat. Struct. Mol. Biol.* **13**, 372–373 (2006).
- Luo, D. H. et al. Crystal structure of the NS3 protease-helicase from Dengue virus. *Acta Crystallogr. A* **64**, C135–C135 (2008).
- Noble, C. G., Seh, C. C., Chao, A. T. & Shi, P. Y. Ligand-bound structures of the dengue virus protease reveal the active conformation. *J. Virol.* **86**, 438–446 (2012).
- Phoo, W. W. et al. Crystal structures of full length DENV4 NS2B-NS3 reveal the dynamic interaction between NS2B and NS3. *Antiviral Res.* **182**. <https://doi.org/10.1016/j.antiviral.2020.104900> (2020)
- Yung, C. F. et al. Dengue serotype-specific differences in clinical manifestation, laboratory parameters and risk of severe disease in adults, Singapore. *Am. J. Trop. Med Hyg.* **92**, 999–1005 (2015).
- Leung, D. et al. Activity of recombinant dengue 2 virus NS3 protease in the presence of a truncated NS2B co-factor, small peptide substrates, and inhibitors. *J. Biol. Chem.* **276**, 45762–45771 (2001).
- Aleshin, A. E., Shiryayev, S. A., Strongin, A. Y. & Liddington, R. C. Structural evidence for regulation and specificity of flaviviral proteases and evolution of the Flaviviridae fold. *Protein Sci.* **16**, 795–806 (2007).
- Kim, Y. M. et al. NMR analysis of a novel enzymatically active unlinked dengue NS2B-NS3 protease complex. *J. Biol. Chem.* **288**, 12891–12900 (2013).
- Woestenenk, E., Agback, P., Unnerstale, S., Henderson, I. & Agback, T. Co-refolding of a functional complex of Dengue NS3 protease and NS2S co-factor domain and backbone resonance assignment by solution NMR. *Protein Expr. Purif.* **140**, 16–27 (2017).
- de la Cruz, L., Chen, W. N., Graham, B. & Otting, G. Binding mode of the activity-modulating C-terminal segment of NS2B to NS3 in the dengue virus NS2B-NS3 protease. *FEBS J.* **281**, 1517–1533 (2014).
- Pilla, K. B., Leman, J. K., Otting, G. & Huber, T. Capturing conformational states in proteins using sparse paramagnetic NMR data. *PLoS ONE* **10**, 1–16 (2015).
- Yildiz, M., Ghosh, S., Bell, J. A., Sherman, W. & Hardy, J. A. Allosteric inhibition of the NS2B-NS3 protease from dengue virus. *Acc Chem. Biol.* **8**, 2744–2752 (2013).
- Dang, M., Lim, L. Z., Roy, A. & Song, J. X. Myricetin allosterically inhibits the dengue NS2B-NS3 protease by disrupting the active and locking the inactive conformations. *Acc Omega* **7**, 2798–2808 (2022).
- Jonniya, N. A. & Kar, P. Functional loop dynamics and characterization of the inactive state of the NS2B-NS3 dengue protease due to allosteric inhibitor binding. *J. Chem. Inf. Model* **62**, 3800–3813 (2022).
- Purohit, P., Sahoo, S., Panda, M., Sahoo, P. S. & Meher, B. R. Targeting the DENV NS2B-NS3 protease with active antiviral phytochemicals: structure-based virtual screening, molecular docking and molecular dynamics simulation studies. *J. Mol. Model* **28**. <https://doi.org/10.1007/s00894-022-05355-w> (2022)
- Gangopadhyay, A. & Saha, A. Exploring allosteric hits of the NS2B-NS3 protease of DENV2 by structure-guided screening. *Comput. Biol. Chem.* **104**, 107876 (2023).
- da Costa, R. A. et al. A computational approach applied to the study of potential allosteric inhibitors protease NS2B/NS3 from dengue virus. *Molecules* **27**. <https://doi.org/10.3390/molecules27134118> (2022)
- Purohit, P., Barik, D., Agasti, S., Panda, M. & Meher, B. R. Evaluation of the inhibitory potency of anti-dengue phytochemicals against DENV-2 NS2B-NS3 protease: virtual screening, ADMET profiling and molecular dynamics simulation investigations. *J. Biomol. Struct. Dyn.* <https://doi.org/10.1080/07391102.2023.2212798> (2023)
- Mohanty, A. K. & Kumar, M. S. Effect of mutation of NS2B cofactor residues on Dengue 2 NS2B-NS3 protease complex—an insight to viral replication. *J. Biomol. Struct. Dyn.* <https://doi.org/10.1080/07391102.2021.2008497> (2021)
- Lee, W. H. K., Liu, W., Fan, J. S. & Yang, D. W. Dengue virus protease activity modulated by dynamics of protease cofactor. *Biophys. J.* **120**, 2444–2453 (2021).
- Quek, J. P. et al. Dynamic interactions of post cleaved NS2B cofactor and NS3 protease identified by integrative structural approaches. *Viruses-Basel* **14**. <https://doi.org/10.3390/v14071440> (2022)
- Deng, J. et al. Discovery of novel small molecule inhibitors of dengue viral NS2B-NS3 protease using virtual screening and scaffold hopping. *J. Med. Chem.* **55**, 6278–6293 (2012).
- Ekonomiuk, D. et al. Flaviviral protease inhibitors identified by fragment-based library docking into a structure generated by molecular dynamics. *J. Medicinal Chem.* **52**, 4860–4868 (2009).
- Vilela, G. G. et al. Fragment-based design of alpha-cyanoacrylates and alpha-cyanoacrylamides targeting Dengue and Zika NS2B/NS3 proteases. *N. J. Chem.* **46**, 20322–20346 (2022).
- Adcock, S. A. & McCammon, J. A. Molecular dynamics: survey of methods for simulating the activity of proteins. *Chem. Rev.* **106**, 1589–1615 (2006).
- Ali, A. A. I., Hoffmann, F., Schafer, L. V. & Mulder, F. A. A. Probing methyl group dynamics in proteins by NMR cross-correlated dipolar relaxation and molecular dynamics simulations. *J. Chem. Theory Comput.* **18**, 7722–7732 (2022).
- Aliev, A. E. et al. Motional timescale predictions by molecular dynamics simulations: case study using proline and hydroxyproline sidechain dynamics. *Proteins* **82**, 195–215 (2014).
- Anderson, J. S., Hernandez, G. & LeMaster, D. M. C-13 NMR relaxation analysis of protein GB3 for the assessment of side chain dynamics predictions by current AMBER and CHARMM force fields. *J. Chem. Theory Comput.* **16**, 2896–2913 (2020).
- Anderson, J. S., Hernandez, G. & LeMaster, D. M. Molecular dynamics-assisted optimization of protein nmr relaxation analysis. *J. Chem. Theory Comput.* **18**, 2091–2104 (2022).
- Showalter, S. A. & Bruschweiler, R. Validation of molecular dynamics simulations of biomolecules using NMR spin relaxation as benchmarks: application to the AMBER99SB force field. *J. Chem. Theory Comput.* **3**, 961–975 (2007).
- Stenstroem, O. et al. How does it really move? Recent progress in the investigation of protein nanosecond dynamics by NMR and simulation. *Curr. Opin. Struct. Biol.* **77**. <https://doi.org/10.1016/j.sbi.2022.102459> (2022)

42. Lesovoy, D. M. et al. NMR relaxation parameters of methyl groups as a tool to map the interfaces of helix-helix interactions in membrane proteins. *J. Biomol. NMR* **69**, 165–179 (2017).
43. Agback, P., Woestenenk, E. & Agback, T. Probing contacts of inhibitor locked in transition states in the catalytic triad of DENV2 type serine protease and its mutants by ¹H, ¹⁹F and ¹⁵N NMR spectroscopy. *BMC Mol. Cell Biol.* **21**, 1–15 (2020).
44. Agback, P. et al. H-1, C-13 and N-15 resonance assignment of backbone and IVL-methyl side chain of the S135A mutant NS3pro/NS2B protein of Dengue II virus reveals unique secondary structure features in solution. *Biomol. NMR Assign.* **16**, 135–145 (2022).
45. Waterhouse, A. et al. SWISS-MODEL: homology modelling of protein structures and complexes. *Nucleic Acids Res.* **46**, W296–W303 (2018).
46. Pettersen, E. F. et al. UCSF chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
47. Bocharov, E. V. et al. Structural basis of the signal transduction via transmembrane domain of the human growth hormone receptor. *Biochim. Biophys Acta Gen. Subj.* **1862**, 1410–1420 (2018).
48. Idiyattullin, D., Daragan, V. A. & Mayo, K. H. A simple method to measure (CH₂)-C-13 heteronuclear dipolar cross-correlation spectral densities. *J. Magn. Reson.* **171**, 4–9 (2004).
49. Lesovoy, D. M., Dubinnyi, M. A., Nolde, S. B., Bocharov, E. V. & Arseniev, A. S. Accurate measurement of dipole/dipole transverse cross-correlated relaxation 2 in methylenes and primary amines of uniformly ¹⁵N-labeled proteins. *J. Biomol. NMR* **73**, 245–260 (2019).
50. Yang, D. W. Probing protein side chain dynamics via C-13 NMR relaxation. *Protein Pept. Lett.* **18**, 380–395 (2011).
51. Lipari, G. & Szabo, A. Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules. 2. Analysis of experimental results. *J. Am. Chem. Soc.* **104**, 4559–4570 (1982).
52. Lipari, G. & Szabo, A. Model-free approach to the interpretation of nuclear magnetic-resonance relaxation in macromolecules. 1. Theory and range of validity. *J. Am. Chem. Soc.* **104**, 4546–4559 (1982).
53. Behnam, M. A. M. & Klein, C. D. P. Conformational selection in the flaviviral NS2B-NS3 protease. *Biochimie* **174**, 117–125 (2020).
54. Kummerer, F., Orioli, S. & Lindorff-Larsen, K. Fitting force field parameters to NMR relaxation data. *J. Chem. Theory Comput.* <https://doi.org/10.1021/acs.jctc.3c00174> (2023)
55. Brecher, M. et al. A conformational switch high-throughput screening assay and allosteric inhibition of the flavivirus NS2B-NS3 protease. *PLoS Pathog.* **13**, <https://doi.org/10.1371/journal.ppat.1006411> (2017)
56. Zhu, L. et al. Conformational change study of dengue virus NS2B-NS3 protease using ¹⁹F NMR spectroscopy. *Biochem. Biophys. Res. Commun.* **461**, 677–680 (2015).
57. Hansen, D. F., Vallurupalli, P. & Kay, L. E. Using relaxation dispersion NMR spectroscopy to determine structures of excited, invisible protein states. *J. Biomol. NMR* **41**, 113–120 (2008).
58. Vallurupalli, P., Hansen, D. F., Stollar, E., Meirovitch, E. & Kay, L. E. Measurement of bond vector orientations in invisible excited states of proteins. *Proc. Natl Acad. Sci. USA* **104**, 18473–18477 (2007).
59. Gupta, G., Lim, L. & Song, J. NMR and MD studies reveal that the isolated dengue NS3 protease is an intrinsically disordered chymotrypsin fold which absolutely requests NS2B for correct folding and functional dynamics. *PLoS ONE* **10**, e0134823 (2015).
60. Tugarinov, V., Kanelis, V. & Kay, L. E. Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. *Nat. Protoc.* **1**, 749–754 (2006).
61. Delaglio, F. et al. Nmrpipe—a multidimensional spectral processing system based on unix pipes. *J. Biomol. NMR* **6**, 277–293 (1995).
62. Vranken, W. F. et al. The CCPN data model for NMR spectroscopy: development of a software pipeline. *Proteins* **59**, 687–696 (2005).
63. Zwahlen, C. et al. An NMR experiment for measuring methyl-methyl NOEs in C-13-labeled proteins with high resolution. *J. Am. Chem. Soc.* **120**, 7617–7625 (1998).
64. Kazimierzczuk, K. & Orekhov, V. Y. Accelerated NMR spectroscopy by using compressed sensing. *Angew. Chem. Int Ed.* **50**, 5556–5559 (2011).
65. Mayzel, M., Kazimierzczuk, K. & Orekhov, V. Y. The causality principle in the reconstruction of sparse NMR spectra. *Chem. Commun.* **50**, 8947–8950 (2014).
66. Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J. Biomol. NMR* **56**, 227–241 (2013).
67. Zhu, G., Xia, Y., Nicholson, L. K. & Sze, K. H. Protein dynamics measurements by TROSY-based NMR experiments. *J. Magn. Reson.* **143**, 423–426 (2000).
68. Hiyama, Y., Niu, C. H., Silverton, J. V., Bavoso, A. & Torchia, D. A. Determination of N-15 chemical-shift tensor via N-15-H-2 dipolar coupling in Boc-Glycylglycyl[¹⁵N]glycine benzyl ester. *J. Am. Chem. Soc.* **110**, 2378–2383 (1988).
69. Zhang, X., Sui, X. G. & Yang, D. W. Probing methyl dynamics from C-13 autocorrelated and cross-correlated relaxation. *J. Am. Chem. Soc.* **128**, 5073–5081 (2006).
70. AMBER2020 (University of California, 2020).
71. Maier, J. A. et al. ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* **11**, 3696–3713 (2015).
72. Jurrus, E. et al. Improvements to the APBS biomolecular solvation software suite. *Protein Sci.* **27**, 112–128 (2018).
73. Roe, D. R. & Cheatham, T. E. PTRAJ and CPPTRAJ: software for processing and analysis of molecular dynamics trajectory data. *J. Chem. Theory Comput.* **9**, 3084–3095 (2013).
74. Dubinnyi, M. A., Lesovoy, D. M., Dubovskii, P. V., Chupin, V. V. & Arseniev, A. S. Modeling of ³¹P-NMR spectra of magnetically oriented phospholipid liposomes: a new analytical solution. *Solid State Nucl. Magn. Reson.* **29**, 305–311 (2006).
75. Valiitti, S., Ovaska, S. J. & Vainio, O. Polynomial predictive filtering in control instrumentation: a review. *IEEE Trans. Ind. Electron.* **46**, 876–888 (1999).
76. Ziegler, H. Properties of digital smoothing polynomial (Dispo) filters. *Appl. Spectrosc.* **35**, 88–92 (1981).
77. Cavanagh, J., Fairbrother, W., Palmer III, A., Rance, M. & Skelton, N. *Protein NMR Spectroscopy* (Elsevier Academic Press, 2007).
78. Korzhnev, D. M., Billeter, M., Arseniev, A. S. & Orekhov, V. Y. NMR studies of Brownian tumbling and internal motions in proteins. *Prog. Nucl. Mag. Res. Sp.* **38**, 197–266 (2001).
79. Kumar, A., Grace, R. C. R. & Madhu, P. K. Cross-correlations Nmr. *Prog. Nucl. Mag. Res. Sp.* **37**, 191–319 (2000).
80. Tugarinov, V., Scheurer, C., Bruschweiler, R. & Kay, L. E. Estimates of methyl ¹³C and ¹H CSA values (Deltastigma) in proteins from cross-correlated spin relaxation. *J. Biomol. NMR* **30**, 397–406 (2004).
81. Case, D. A. Calculations of NMR dipolar coupling strengths in model peptides. *J. Biomol. NMR* **15**, 95–102 (1999).
82. Vermeeren, P., van Zeist, W. J., Hamlin, T. A., Fonseca Guerra, C. & Bickelhaupt, F. M. Not carbon s-p hybridization, but coordination number determines C-H and C-C bond length. *Chemistry* **27**, 7074–7079 (2021).
83. Hoffmann, F., Mulder, F. A. A. & Schafer, L. V. Dynamics in protein simulations with AMBER force fields. *J. Phys. Chem. B* **122**, 5038–5048 (2018). Accurate Methyl Group.
84. Efron, B. 1977 Rietz Lecture—bootstrap methods - another look at the Jackknife. *Ann. Stat.* **7**, 1–26 (1979).
85. Politis, D. N. & White, H. Automatic block-length selection for the dependent bootstrap. *Econom. Rev.* **23**, 53–70 (2004).

Acknowledgements

The authors thank the Swedish NMR centre for access to the instruments and support. This work was supported by the Swedish Foundation for Strategic Research grant ITM17-0218 to T.A. and P.A., grant RSF 23-44-10021 to D.M.L., project 121112900217-3 to A.A.L., Swedish Cancer Society grant 21 1605 Pj01H to A.A., and the Swedish Research Council grants 2021-05061 to A.A. and 2019-03561 to V.O. This study used NMRbox: National Center for Biomolecular NMR Data Processing and Analysis, a Biomedical Technology Research Resource (BTRR), which is supported by NIH grant P41GM111135 (NIGMS).

Author contributions

X.H. and R.S. have contributed to the production of all necessary protein variants, their purification, and developing a construct of labelled DENV II proteins. T.A. wrote the original manuscript draft. P.A. and A.A. contributed to the writing, reviewing, and final editing of the manuscript. T.A. and P.A. performed the NMR studies on DENV II. V.O., D.L. contributed with NMR measurements and methodology, developed and performed statistics analysis and back-calculation of relaxation parameters from MD trajectories making a concept of conformational filter. T.A. performed assignments using the ccpn program. A.L. performed MD calculations. T.S. did the analysis of earlier crystallographic data. P.A., T.A., T.S., A.A., and V.O. conceptualized the project, supervised different parts of the project and acquired the necessary funding acquisition.

Funding

Open access funding provided by Swedish University of Agricultural Sciences.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s42003-023-05584-6>.

Correspondence and requests for materials should be addressed to Adnane Achour or Peter Agback.

Peer review information *Communications Biology* thanks Thales Kronenberger and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Primary Handling Editor: Gene Chong. A peer review file is available.

Reprints and permission information is available at <http://www.nature.com/reprints>

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



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 license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license 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 license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023