Australian National University

Applications of paramagnetic NMR spectroscopy in drug discovery

A thesis submitted for the degree of Doctor of Philosophy of The Australian National University

By

Wan-Na Chen

Research School of Chemistry

The Australian National University

February 2016

Declaration of Originality

I declare that this thesis is composed of my original work and contains no material previously published or written by another person. The content of this thesis is the result of work that I have carried out since the commencement of my research degree candidature till now (Dec. 2012-Feb. 2016) and has not been previously submitted for another degree or diploma in any university or tertiary institution. Wherever possible, established methodologies have been acknowledged by citation of the original publication from which they derive.

Wan-Na Chen

February 2016

Acknowledgements

First of all, I would like to deliver my sincere appreciation to my supervisor/mentor Prof. Gottfried Otting. He has been always there for sharing fantastic ideas and bright viewpoints. I thank Gottfried for leading me into the field of paramagnetic NMR spectroscopy and also for the trust, support and inspiration. It has been a great experience working with you.

Secondly, I would like to thank for the generous support from my former and current colleagues in the Otting group, in particular Dr Laura de la Cruz, a forerunner in the NMR study of the dengue virus protease. I thank Dr Dimitry Shishmarev, Dr Karin Loscha, Dr Ruhu Qi and Dr Choy-Theng Loh for supervision; Dr Thomas Carruthers and Mr Elwy Abdelkader for being my English teachers and helpful discussions about research; and finally, other members in the biology lab for creating such a friendly work environment.

I would also like to acknowledge many wonderful collaborators, including Prof. Thomas Huber for computational support, Dr Bim Graham for providing unnatural amino acids and lanthanide tags, Dr Christoph Nitsche and Prof. Christian Klein for providing inhibitors against the dengue virus protease, Ms Kekini Kuppan and Dr Michael Lee for earlier work with DnaB and Dr Kristaps Jaudzems for bringing along the sortase A inhibitor. The support from professional and general staff in the Research School of Chemistry is also gratefully acknowledged.

Last but not least, I thank the company from friends; in particular, Dr Wan-Biao Hu, Dr Ping Lan and Dr Choy-Theng Loh helped me in all aspects of life and work. I apologize to my family and friends in China for being away for so long, especially to my mother.

I would not have come through all the hard time and finally completed my PhD studies without your strong support. Thank you all, again.

List of Publications

1. de la Cruz, L.; **Chen, W. N.**; Graham, B. and Otting, G. Binding mode of the activity‐modulating C‐terminal segment of NS2B to NS3 in the dengue virus NS2B–NS3 protease. *FEBS J.*, 2014, 281: 1517-1533.

I joined towards the end of the on-going project and contributed to protein expression, NMR experiments, data analysis and manuscript preparation. A preliminary account of this work was also part of the Ph.D. thesis of Dr L. de la Cruz submitted in 2013.

2. **Chen, W. N.**; Loscha, K. V.; Nitsche, C.; Graham, B. and Otting, G. The dengue virus NS2B–NS3 protease retains the closed conformation in the complex with BPTI. *FEBS Lett.*, 2014, 588: 2206-2211.

In this work I was the main investigator. I conducted all protein expression, performed NMR experiments, analysed all data, and wrote the initial draft of the manuscript.

3. **Chen, W. N.**; Kuppan, K. V.; Lee, M. D.; Jaudzems, K.; Huber, T. and Otting, G. *O-tert*-butyltyrosine, an NMR tag for high-molecular-weight systems and measurements of submicromolar ligand binding affinities. *J. Am. Chem. Soc.*, 2015, 137: 4581-4586.

In this work I was the main investigator. The project was based on initial work and results obtained by the second and third authors. I produced the key findings and extended the applications to explore the value of this methodology. A ^{19}F NMR spectrum from the supporting information is also found in the Masters thesis of Ms. K. V. Kuppan submitted in 2014.

4. **Chen, W. N.**; Nitsche, C.; Pilla, K. B.; Graham, B.; Huber, T.; Klein, C. D. and Otting, G. Sensitive NMR approach for determining the binding mode of tightly binding ligand molecules to protein targets. Manuscript submitted for publication.

In this work I was the main investigator. Using inhibitors synthesized by Dr Nitsche (second author, from Prof. C. D. Klein's laboratory in Heidelberg, Germany), I conducted the preliminary work, planned the research, performed NMR experiments, analyzed data and prepared the manuscript. Dr Kala Bharath Pilla (third author, from Prof. T. Huber's group) performed most of the computational work.

Abstract

NMR spectroscopy of proteins with paramagnetic metal ions, first performed with metalloproteins, is a unique technique to obtain long-range distance information for three-dimensional structure determinations. This thesis focuses on developing applications of paramagnetic NMR spectroscopy, particularly pseudocontact shifts, in drug discovery. The two-component dengue virus NS2B-NS3 protease (NS2B-NS3pro) from serotype 2 is a well-established drug target, but drug development has been hampered for many years by lack of structural information. In earlier work, pseudocontact shifts (PCSs) induced by lanthanide binding tags at multiple sites had successfully been used to determine the closed conformation of NS2B in the presence of a small inhibitor molecule. Subsequently, PCSs were used to prove that an unlinked construct of NS2B-NS3pro exists predominately in the closed conformation in solution, showing that the open conformation observed previously is an artefact generated by a covalent link between NS2B and NS3 (Paper 1). Next, PCSs generated for NS2B, NS3pro and bovine pancreatic trypsin inhibitor (BPTI) were used to show that the C-terminal segment of NS2B remains in the closed conformation in the presence of BPTI, correcting a crystallographic artefact (Paper 2). The work described in Papers 1 and 2 confirmed that the closed conformation of dengue virus NS2B-NS3pro is the best model for structure-guided drug design. As the sensitivity of NMR spectra of dynamic proteins, such as the dengue virus protease, is compromised by excessive line broadening, alternative NMR tags were sought. *O*-*tert*-butyltyrosine incorporated in proteins proved to be an outstanding NMR probe for conformational studies of high-molecular-weight systems and measurement of submicromolar ligand binding affinities in one-dimensional ^{1}H-NMR spectra without any isotope labelling (Paper 3). A *tert*-butyl probe was also introduced into a tightly binding lead compound against the dengue virus protease. Measurement of ligand PCSs from intense intramolecular NOESY cross-peaks with the *tert*-butyl group allowed positioning of the ligand on the protein with respect to the paramagnetic centre, while strong intermolecular NOEs validated the structural model of the complex established with the use of PCSs (Paper 4). In summary, the paramagnetic NMR approach, demonstrated on the dengue virus NS2B-NS3 protease, presents a broadly applicable and elegant way for structure-guided drug design at atomic resolution.

Table of Contents

Introduction

1.1 Biological NMR spectroscopy

Nuclear magnetic resonance (NMR) spectroscopy can determine three-dimensional (3D) structures of biological macromolecules in solution at atomic resolution and thus presents a structure determination method complementary to crystallography (Wüthrich 1990). Nowadays, high-resolution 3D structures of macromolecular systems are most often obtained by X-ray diffraction or cryo-electron microscopy (EM), whereas NMR spectroscopists tend to focus more on systems, which are not amenable to atomic-resolution studies by other methods or subject to artefacts (Sekhar and Kay 2013).

Single-crystal structures present frozen snapshots of protein states that sometimes do not reflect all physiologically relevant conformations. In addition, dense packing in a crystal lattice can produce crystallization artefacts (Carugo and Argos 2008; Chen et al. 2014). In contrast, the NMR method is uniquely suited for studies of protein dynamics (Skora et al. 2013; Saio et al. 2015), low-affinity interactions (Okude et al. 2015), denaturation by heat or pressure (Rochdi et al. 2000; Nucci et al. 2014; Maeno et al. 2015) , pH effects (Zong et al. 2004), effects from paramagnetic centres in metalloproteins (Ubbink et al. 2002; Bertini et al. 2005; Carruthers et al. 2014) and other properties that are best studied under near-physiological conditions.

In particular, NMR spectroscopy is widely used in drug discovery approaches (Diercks et al. 2001; Pellecchia et al. 2002), e.g. to screen compound libraries against protein targets (Kim and Wyss 2015; Pfaff et al. 2015). Even in the case of weak binding, chemical shift changes of either the protein targets or inhibitors can provide evidence for interactions (Viegas et al. 2011). In this thesis, Paper 3 describes examples, where the chemical shift change of a single intense NMR resonance in a protein serves as evidence of ligand binding either in slow exchange or in the fast exchange regime, establishing a rapid method to detect ligand binding in proteins in presence or absence of a metal ion and facilitating binding affinity measurements for large molecular targets (>30 kDa) (Chen et al. 2015). Conventionally, analysis of

ligand binding modes relies on analysis of the chemical shift perturbations (Shuker et al. 1996; Medek et al. 2000; Meyer and Peters 2003; Aguirre et al. 2014) and nuclear Overhauser effects (NOEs) (Breeze 2000; Eichmüller et al. 2001) observed in protein NMR spectra to yield detailed information about the binding sites. However, these methods rely on a large number of sequence-specific resonance assignments for the target proteins. Furthermore, line broadening effect and signal overlapping often make the analysis of intermolecular NOESY cross-peaks and chemical shift perturbations difficult on 2D spectra. A convenient way of detecting intermolecular NOEs and achieving site-specific resonance assignments in a dynamic protein target is described in Paper 4, which does not depend on a large number of resonance assignments.

High molecular weights present a barrier for studies by NMR spectroscopy (Wagner 1993; Riek et al. 2002). This problem has been reduced significantly by the development of high-field magnets (Markiewicz et al. 1996; Lin et al. 2000; Markiewicz et al. 2006), rapid data collection and processing methods for multidimensional NMR experiments, including automated projection spectroscopy (Hiller et al. 2005; Hiller and Wider 2011) and sparse sampling methods (Rovnyak et al. 2004; Hyberts et al. 2012; Mobli and Hoch 2014), dynamic nuclear polarisation (DNP) enhancement (Ardenkjær-Larsen et al. 2003; Maly et al. 2008; Idehara et al. 2015), magic angle spinning (MAS) (Medek et al. 1995; Griffin 1998; Comellas and Rienstra 2013) and stable isotope labelling (Yamazaki et al. 1998; Ozawa et al. 2004; Kainosho et al. 2006; Tugarinov et al. 2006; Ruschak and Kay 2012). Pushing the molecular weight limits has always been a motivating goal for NMR spectroscopists. For example, it has been shown that selectively protonated methyl groups in perdeuterated proteins combined with methyl-TROSY allow studies of supra-molecular systems up to 1 MDa (Sprangers and Kay 2007; Kay 2011). Paper 3 presents a highly sensitive NMR probe that is based on an unnatural amino acid and can be detected in high molecular weight systems (>300 kDa) without any isotope labelling (Chen et al. 2015).

Over more than 20 years and importantly for this thesis, NMR studies on metalloproteins containing paramagnetic metal ions have stimulated the development of paramagnetic NMR spectroscopy (Bertini et al. 1993; Banci et al. 1994; Banci et al.

1996; Banci et al. 1997). Specific applications and software have been developed to make best use of the structural information provided by paramagnetic centres.

1.2 Paramagnetic NMR spectroscopy

1.2.1 Paramagnetic effects

The magnetic moment of unpaired electrons is almost three orders of magnitude larger than that of protons. In a static magnetic field, magnetic dipolar fields generated by unpaired electrons are modulated by molecular re-orientation and thus perturb nuclear spins. Paramagnetic NMR spectroscopy takes advantage of the paramagnetism originating from unpaired electrons to study molecular structures and dynamics with respect to the paramagnetic centres.

Paramagnetism causes pronounced long-range paramagnetic effects, which are manifested in NMR spectra as changes in chemical shifts, line widths and coupling constants. Pseudocontact shifts (PCS), paramagnetic relaxation enhancements (PRE), cross-correlation between Curie spin and ^{1}H-^{15}N dipole-dipole relaxation (CCR) and residual dipolar couplings (RDC) are four paramagnetic effects studied widely (Figure 1). These paramagnetic effects can be described by equations **1**-**7**:

i) PCS (Bertini et al. 2002):

$$\delta^{\mathrm{PCS}} = \frac{1}{12\pi r^3}\left[\Delta\chi_{\mathrm{ax}}(3\cos^2\theta - 1) + \frac{3}{2}\Delta\chi_{\mathrm{rh}}\sin^2\theta\cos 2\varphi\right] \quad (1)$$

ii) PRE (Vega and Fiat 1976):

$$\lambda^{\mathrm{PRE}} = \frac{k}{r^6}\left(4\tau_{\mathrm{r}} + \frac{3\tau_{\mathrm{r}}}{1+\omega_{\mathrm{H}}^2\tau_{\mathrm{r}}^2}\right) \quad (2)$$

$$k = \frac{1}{5}\left(\frac{\mu_0}{4\pi}\right)^2 B_0^2\gamma_{\mathrm{H}}^2(g_J\mu_{\mathrm{B}})^4 J^2(J+1)^2/(3k_{\mathrm{B}}T)^2 \quad (3)$$

iii) CCR (Ghose and Prestegard 1997):

$$\eta^{\mathrm{CCR}} = \kappa\frac{3\cos^2\vartheta - 1}{r^3}\left(4\tau_{\mathrm{r}} + \frac{3\tau_{\mathrm{r}}}{1+\omega_{\mathrm{H}}^2\tau_{\mathrm{r}}^2}\right) \quad (4)$$

$$\kappa = \frac{1}{30}\left(\frac{\mu_0}{4\pi}\right)^2 B_0\gamma_H^2\gamma_N\hbar(g_J\mu_B)^2\, J(J+1)/(r_{NH}^3 k_B T) \tag{5}$$

iv) RDC (Tolman et al. 1995; Barbieri et al. 2002):

$$\Delta\nu^{RDC} = K\left[\Delta\chi_{ax}(3\cos^2\Theta - 1) + \frac{3}{2}\Delta\chi_{rh}\sin^2\Theta\cos 2\Phi\right] \tag{6}$$

$$K = -\frac{1}{120\pi^2} B_0^2\gamma_H\gamma_N\hbar S^2/(r_{NH}^3 k_B T) \tag{7}$$

where $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ are the axial and rhombic components of the anisotropic component of the magnetic susceptibility (χ) tensor; r, θ and φ are the polar coordinates of the nuclear spin with respect to the metal centre; τ_r is the rotational correlation time; ω_H is the proton Larmor frequency; μ_0 is the vacuum permeability; B_0 is the magnetic field; γ_H and γ_N are the proton and nitrogen magnetogyric ratios; g_J is the g-factor; μ_B is the Bohr magneton; J is the total spin moment; k_B is the Boltzmann constant; T is the temperature; $\hbar$ is Planck's constant divided by 2π; Θ and Φ are the polar angles of an internuclear vector with respect to the principal axes of the χ tensor; S is the order parameter of the molecular alignment (for backbone amide protons, S^2 is often set to 0.9 (Pintacuda et al. 2004)); r_{NH} is the distance between N and H nuclei.

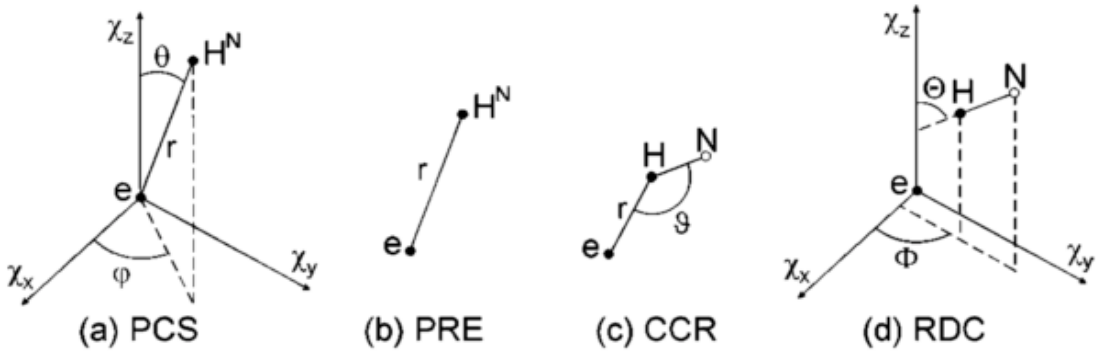


Figure 1. Geometric information contained in the four paramagnetic effects. (a) $^1H^N$ pseudocontact shift (PCS), (b) paramagnetic relaxation enhancement (PRE), (c) cross-correlation between Curie and ^1H-^{15}N dipole-dipole relaxation (CCR), (d) residual dipolar coupling (RDC) between a ^1H and a ^{15}N spin. Reproduced from Pintacuda et al. (2004).

As the chemical shifts of a protein are sensitive to the presence of a bound metal ion

regardless of whether it is diamagnetic or paramagnetic, paramagnetic effects are necessarily measured as differences in NMR spectra recorded for the target molecule in its paramagnetic and diamagnetic states (Otting 2010).

1.2.2 Unique advantages of PCSs

The great attraction of paramagnetic effects lies in their capability of providing long-range structural restraints. Unlike PRE, CCR and RDC (which increase with field strength), the PCS effect is field independent, if measured in ppm (Bertini et al. 2002; Otting 2008). In contrast to RDCs, PCSs provide information about the distances r of nuclear spins from the metal centres (Otting 2008). The r^{-6} distance dependence of PREs is much stronger than the r^{-3} distance dependence of PCSs (Equations 1 and 2). Therefore, PREs decay much more rapidly than PCSs with increasing distance of nuclear spins from the metal centre. Even though nuclear spins near the paramagnetic centre are easily broadened beyond detection, nuclear spins that are sufficiently far from the metal centre can deliver easily measurable PCSs with little PREs. PCSs are only little sensitive to small translational movements of the metal ions or small-amplitude motions of the nuclear spins, due to the relatively long distance from the metal centre. In addition, PCSs can be measured and analysed even for low-affinity protein-ligand complexes, even if they are averaged due to incomplete complex formation (Crowley and Ubbink 2003; John et al. 2006). Finally, PCSs can be measured for nuclear spins over 40 Å away from the paramagnetic metal centre (Biekofsky et al. 1999; Allegrozzi et al. 2000). All these properties make PCSs a unique and powerful tool.

The measurement of PCSs is simple, as they are measured as the chemical shift differences between paramagnetic and diamagnetic states. Following superimposition of the 2D NMR spectra of diamagnetic and paramagnetic samples, the diamagnetic peak and its corresponding shifted paramagnetic cross-peaks, originating from the same nuclear spin, can be connected by mostly straight lines. Most accurate results are obtained for nuclear spins with little chemical shift anisotropy (CSA), as the weak molecular alignment in the magnetic field arising from non-vanishing $\Delta\chi$ tensors combined with CSA can contribute small chemical shift changes in the paramagnetic

sample that do not stem from PCSs (John et al. 2005).

Experimental PCSs present extraordinarily powerful restraints for 3D structure determination. The power of PCSs lies in the $\Delta\chi$ tensors fitted from the limited number of observed PCSs. $\Delta\chi$ tensors can be characterized by 8 parameters, which include 3 parameters describing the position of the metal ion (x, y and z coordinates), 3 parameters describing the orientation of the $\Delta\chi$ tensors with respect to the molecular coordinates (Euler angles α, β and γ) and 2 parameters for the axial and rhombic components of the anisotropy tensor ($\Delta\chi_{ax}$ and $\Delta\chi_{rh}$). PCS isosurfaces, which depict the coordinates for which Equation 1 predicts the same PCS values, can be plotted onto the 3D structure of a target protein to give an alternative representation of a $\Delta\chi$ tensor and its orientation with respect to the protein (Figure 2). The $\Delta\chi$ tensor and its associated PCS isosurfaces correspond to a coordinate system, which is tied to the protein molecule and its metal ion centre (Otting 2010). Theoretically, measurement of eight PCSs is sufficient to fit the $\Delta\chi$-tensor parameters. Once a $\Delta\chi$ tensor has been obtained, its coordinate frame can be used to interpret other PCSs, which were not included in the fit, for example to guide the structure modelling using other nuclear spins for which PCSs are available. Paper 2 in this thesis presents an example of this methodology. Notably, any small structural variation of the protein or mobility of the metal ion can lead to discrepancies between the experimental PCSs and the predicted PCSs, which needs to be taken into account in the analysis.

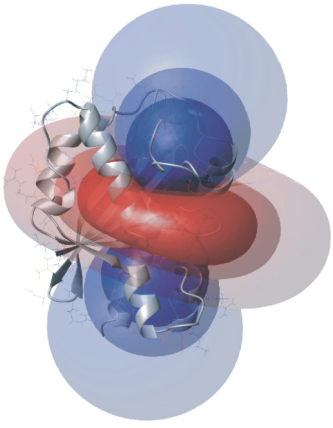


Figure 2. Isosurfaces depicting the PCSs induced by Dy^{3+} are plotted on a ribbon representation of a crystal structure. Blue and red surfaces identify the spatial locations of positive and negative PCSs of ±3, ±1.5 and ±0.5 ppm, respectively. The PCS isosurfaces constitute a representation of the $\Delta\chi$ tensors. Reproduced from Otting

(2010).

1.2.3 Lanthanides

Lanthanides comprise fifteen metallic chemical elements, all of which form trivalent cations (Ln^{3+}) and have unpaired electrons in f-orbitals (except lanthanum and lutetium). In the past, lanthanide ions have been used to introduce anisotropic magnetic susceptibilities in diamagnetic molecules partly because many of them exhibit very large PCS effects. The particular magnetic properties of Ln^{3+} ions make them powerful tools to induce paramagnetic effects in NMR spectra.

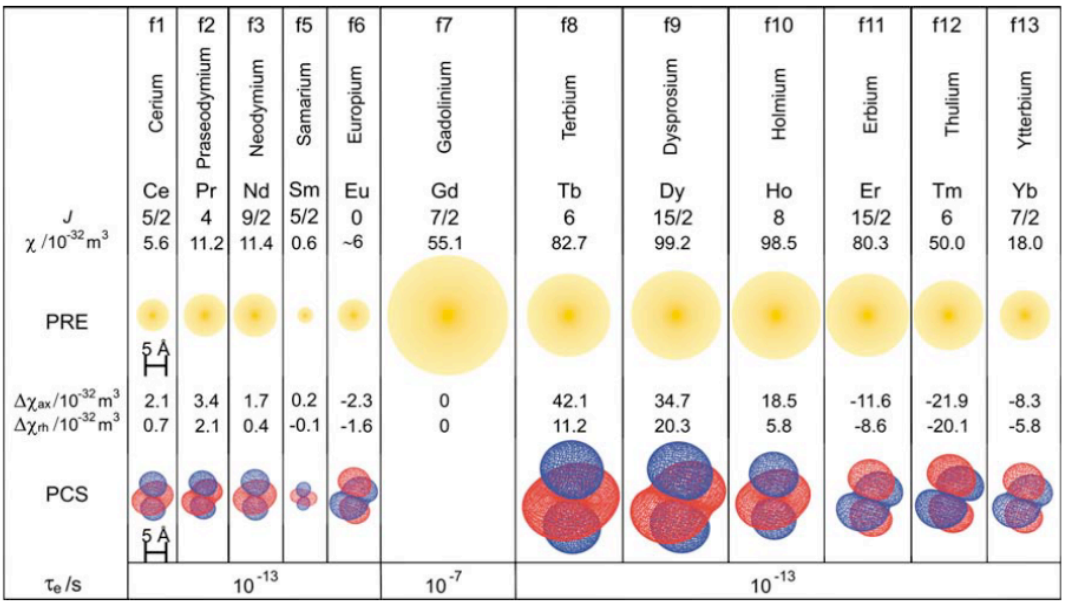


Figure 3. Paramagnetic properties of paramagnetic and non-radioactive lanthanide ions. Representative PCS isosurfaces are plotted at ±5 ppm (positive PCS in red and negative PCS in blue), using tensors reported by Bertini et al. (2001). The radii of the yellow spheres indicate the distances from the metal ion at which ^{1}H NMR signals of macromolecules with a rotational correlation time of 15 ns broaden by 80 Hz due to PRE at a magnetic field strength of 18.8 T. Typical electronic relaxation times representative for the same field strength are indicated underneath. Reproduced from Pintacuda et al. (2007).

The magnetic properties of lanthanide ions are well understood. Different lanthanide ions display very different magnitudes of the χ tensor and its anisotropy (Figure 3)

(Bleaney 1972; Bertini et al. 2001). Based on the magnitude of $\Delta\chi$ tensors, Dy^{3+}, Tb^{3+} and Tm^{3+} can be classed, somewhat arbitrarily, as highly paramagnetic, Er^{3+} and Yb^{3+} as moderately paramagnetic, and Eu^{3+}, Ce^{3+} and Sm^{3+} as little paramagnetic (Otting 2008). The magnitudes of the $\Delta\chi$ tensors of Dy^{3+} and Tb^{3+} are similar, but Dy^{3+} tends to cause more severe PRE effects than Tb^{3+}. Usually, the signs of the $\Delta\chi$ tensors for Tb^{3+} and Tm^{3+} are opposite. Therefore, Tb^{3+} is often used in combination with Tm^{3+} (de la Cruz et al. 2011; Graham et al. 2011). As a diamagnetic reference, other trivalent diamagnetic metal ions, such as Sc^{3+}, La^{3+}, Y^{3+} and Lu^{3+}, can be used. Due to the similar ionic radius of Y^{3+} compared with Tb^{3+} and Tm^{3+}, Tb^{3+}-Y^{3+}-Tm^{3+} is the diamagnetic and paramagnetic ion combination used most frequently in the paramagnetic studies of this thesis to generate PCSs.

1.2.4 Introduction of lanthanides onto proteins

Lanthanides are not readily available biologically and have not yet been found to occur naturally in, e.g., proteins. “Natural” lanthanide binding sites sometimes can be generated by the displacement of different metal ions (for example Ca^{2+} or Fe^{3+}) from the metal binding site of a metalloprotein (Figure 4C). In this case, the binding of a lanthanide ion can be achieved simply by titrating the corresponding lanthanide chloride or lanthanide sulfate into a solution of the metal-free target protein.

A famous example of metal replacement for paramagnetic NMR studies is the Ca_3Ln complex of vertebrate calmodulin (CaM), where the mutation N60D at site II of the four natural calcium-binding sites achieves selective and site-specific lanthanide binding even in the presence of excess calcium (Bertini et al. 2003). The introduction of paramagnetic lanthanides stimulated extensive paramagnetic NMR studies on the Ca_3Ln-CaM complex and various complexes with different peptides (Bertini et al. 2003; Bertini et al. 2004; Bertini et al. 2007; Russo et al. 2013). However, in general the lanthanide ions have to compete with the naturally occurring metal, which often results in low specificity of lanthanide labelling. In addition, not all proteins have natural binding sites that can accommodate lanthanide ions. Therefore, different methods for lanthanide tag attachments to proteins were developed (Figure 4) (Liu et al. 2014).

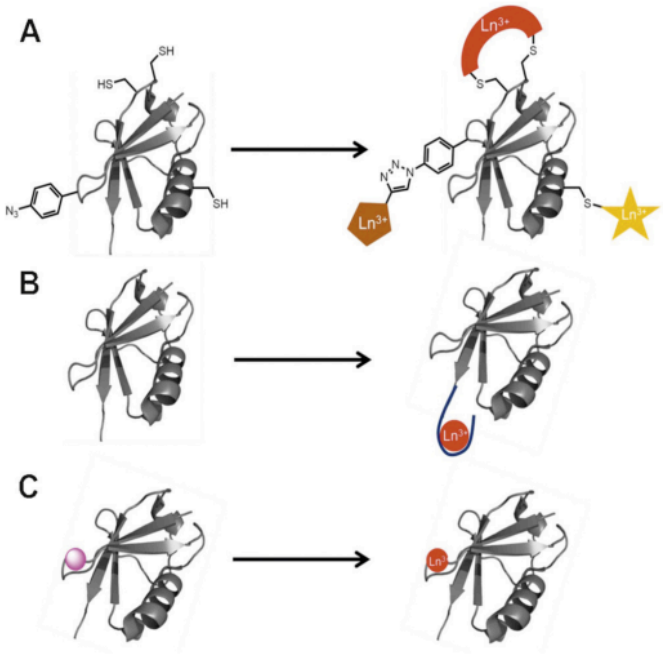


Figure 4. Approaches for introducing lanthanide ions on target proteins. (A) Synthetic tags can be linked to a cysteine residue or an unnatural amino acid. (B) A lanthanide binding peptide can be engineered genetically into a protein. (C) The metal ion of a metalloprotein can be replaced by a lanthanoid. Reproduced from Liu et al. (2014).

1.2.4.1 Lanthanide binding peptides

Early on, an "EF-hand" motif derived from calcium binding proteins consisting of 12 amino acid residues fused to the N-terminus of a membrane protein was shown to enable specific lanthanide binding (Figure 4B) (Ma and Opella 2000). Later, lanthanide binding peptides (LBPs), which are short peptides modified to improve the affinity for lanthanide ions (Franz et al. 2003; Nitz et al. 2003; Nitz et al. 2004), were appended to the N-terminus of ubiquitin to obtain better alignment tensors for RDC studies (Wöhnert et al. 2003). Obviously, fusion of LBPs at the N- or C-terminus limits the number of possible attachment sites to two.

It is important to note that movement of the metal ion with respect to the protein will lead to an averaged tensor. Therefore, tag mobility is problematic, especially when PCSs as well as PREs are interpreted (Ma and Opella 2000; Wöhnert et al. 2003). An improved LBP was later anchored to the target protein via a disulfide bridge and an

N-terminal fusion; it was shown that this two-point anchoring yielded two- to three-fold larger anisotropic paramagnetic effects due to better defined localisation of the LBPs and their bound metal ions (Saio et al. 2009; Saio et al. 2010). A double-armed LBP was also successfully engineered into the turn of two β strands to immobilize lanthanides (Barthelmes et al. 2010; Kobashigawa et al. 2012). However, LBPs typically contain between 12 and 20 amino acid residues. Insertion of LBPs can thus easily perturb the protein structure in their vicinity.

1.2.4.2 Chemical lanthanide tags

A more general approach for the introduction of lanthanide ions involves chemically synthesised metal chelators. Key aims in lanthanide tag design are to improve the affinity of lanthanide ions and also reduce tag mobility. The more rigid the metal ion, the larger anisotropic paramagnetic effects can be induced in the protein. The crystal structures of the protein targets, which are frequently known prior to NMR study, make it possible to select and engineer specific sites for chemical lanthanide tags, which can be ligated to a single cysteine residue or unnatural amino acid.

EDTA, DTPA, TAHA and DOTA have been shown to produce suitable Ln^{3+} chelating ligands (Cotton 2013). A number of synthetic lanthanide tags have been developed to enable site-specific introduction of lanthanide ions (Figure 5) (Su and Otting 2010; Liu et al. 2014). Similar to lanthanide binding peptides, chemical lanthanide tags can be introduced via single and double-arm tethers. In the case of synthetic lanthanide tags, however, single-arm attachment can be more attractive by offering more options for attachment and the possibility to use an unnatural amino acid for attachment to avoid mutations of naturally occurring cysteines. Lanthanide tags attached via a single tether produce effective $\Delta\chi$ tensors, which can describe PCSs and RDCs of protein nuclear spins in the presence of considerable tag mobility (Shishmarev and Otting 2013). In the case of the C1 tag (**10**; Figure 5), three chiral amide pendants are sufficient to provide an 8-coordinate lanthanide binding site and maintain a single enantiomeric species to produce a single set of peaks in the NMR spectrum of the protein (Graham et al. 2011). The C1 tag and its opposite enantiomer, named C2 tag, have successfully been used in different protein conformation studies

(de la Cruz et al. 2011; Yagi et al. 2013; Chen et al. 2014; de la Cruz et al. 2014). Due to the sizeable PCSs induced by the C2 tag (de la Cruz et al. 2011), it was the tag of choice for the dengue virus protease studies in this thesis.

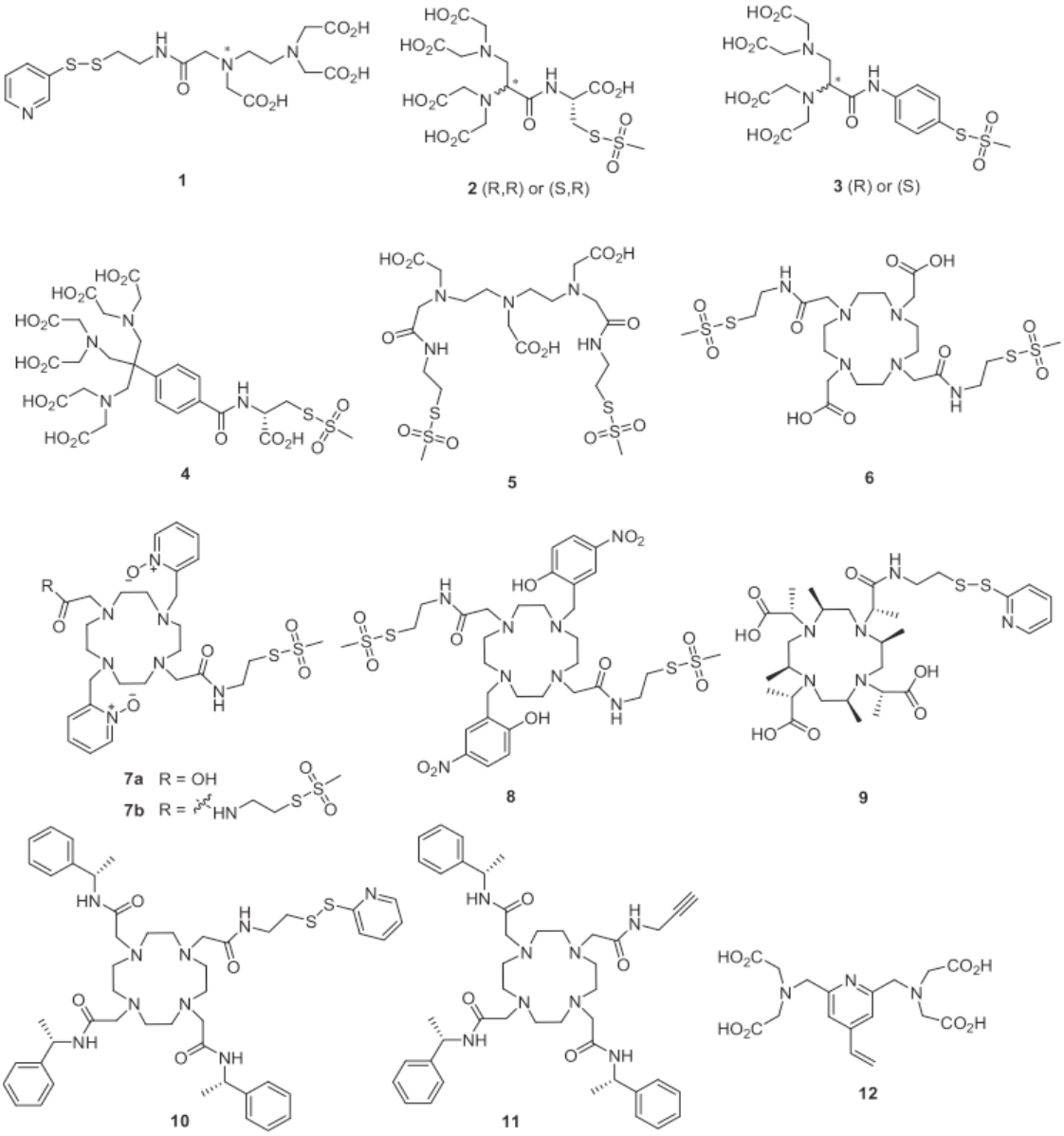


Figure 5. Chemical structures of lanthanide tags. (**1**-**3**) EDTA-based probes: *S*-(2-pyridylthio)-cysteaminyl-EDTA, EDTA-CA-MTS and *ent*-6; (**4**) TAHA based probe: Cys-Ph-TAHA; (**5**) DTPA-based probe: CLaNP1; (**6**-**11**) DOTA-based probes: CLaNP3, CLaNP7, DOTAM8, C1 and C4; (**12**) pyridine-core probe: 4VPyMTA. Reproduced from Liu et al. (2014).

1.2.5 Assignment of paramagnetic NMR spectra

The measurement of PCSs requires the resonance assignment of the paramagnetic NMR spectra. In practice, assigning paramagnetic NMR spectra can be hard because NMR signals near the paramagnetic ion centre are significantly broadened or even missing due to dipole moment of the metal ion. Although it is not necessary to measure PCSs for every single nuclear spin in the system of interest, the accuracy of $\Delta\chi$ tensors determination increases greatly, if more than eight experimental PCSs are available. Therefore, it is essential to get as many initial experimental PCS data as possible. A robust $\Delta\chi$-tensor fit should be insensitive with regard to the deletion or addition of some PCS data.

Even though unambiguous PCS assignments could be obtained using conventional 3D NMR spectra (Russo et al. 2013), such experiments have limited sensitivity. An alternative and rapid assignment strategy is to use ^{15}N-HSQC spectra to transfer the assignment of cross-peaks from the diamagnetic to the paramagnetic state of the target protein (Figure 6) (Pintacuda et al. 2004). This method requires no ^{13}C labelling and no NMR data other than a limited number of assigned ^{15}N-HSQC cross-peaks. Nuclear spins, which are close in space compared to their distances from the paramagnetic centre such as the amide protons and amide nitrogens of backbone NH groups, exhibit similar PCSs. ^{15}N-HSQC cross-peaks are thus displaced by similar PCSs in both the ^{1}H and ^{15}N dimension, resulting in a diagonal shift from their diamagnetic position (Allegrozzi et al. 2000; Bertini et al. 2001). For resolved cross-peaks, the pairs of diamagnetic and paramagnetic peaks can thus be identified unambiguously by simple superimposition of the ^{15}N-HSQC spectra. For NH groups in the vicinity of the paramagnetic centre, PCSs can be large and the shifts may no longer align along a diagonal line (John et al. 2007). In this situation it is helpful to superimpose three ^{15}N-HSQC spectra, recorded with two different paramagnetic ions and one diamagnetic ion. If the $\Delta\chi$ tensors of the two paramagnetic ions are proportional to each other, it will be possible to draw a parallel or near-parallel line through the three cross-peaks to correlate the paramagnetic and the diamagnetic peaks (Su et al. 2009b; Man et al. 2010). Combinations of paramagnetic lanthanide ions with opposite sign of $\Delta\chi$ tensor are favoured, since this will position the two paramagnetic peaks on either side of the diamagnetic peak. In connecting the three cross-peaks, one should also consider the relative magnitude of the $\Delta\chi$ tensors as an

aid in the assignment. For example, we frequently use tags loaded with Tb^{3+}, Tm^{3+} and Y^{3+}, where the $\Delta\chi$ tensors of Tb^{3+} tags tend to be larger and of opposite sign than the $\Delta\chi$ tensors of Tm^{3+} tags.

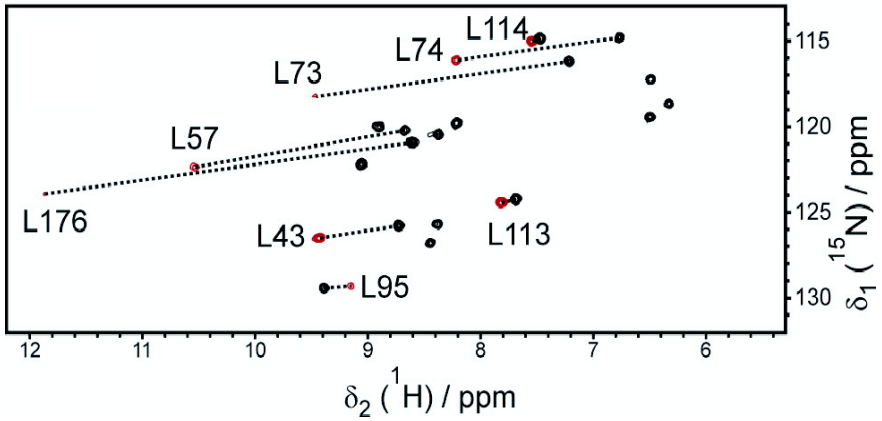


Figure 6. PCSs measured for a protein labelled with ^{15}N-Leu. The diamagnetic and paramagnetic spectra recorded with La^{3+} and Tb^{3+} were plotted in black and red, respectively. Reproduced from Pintacuda et al. (2007).

1.3 *E. coli* cell-free protein synthesis and genetically encoded unnatural amino acids for NMR spectroscopy

1.3.1 *E. coli* cell-free protein synthesis

Protein synthesis can be achieved in three ways, including chemical synthesis, *in vivo* expression and cell-free protein synthesis. Cell-free protein synthesis systems can synthesize proteins with high speed (Pavlov and Ehrenberg 1996) and accuracy (Kurland 1982). In cell-free protein synthesis systems, protein is produced using crude lysate from cells to provide the translational machinery, accessory enzymes, tRNA, and cofactors in combination with added RNA template, amino acids and energy supply (Nirenberg and Matthaei 1961).

Cell extracts used in cell-free synthesis systems can be derived both from prokaryotic (*Escherichia coli*) and eukaryotic (e.g. wheat germ, rabbit reticulocyte lysate, insect cells, HeLa cells, Xenopus eggs) origin. Since the mid-nineteenth century, cell-free systems have been developed extensively (Figure 7) to increase the protein yields by modification and improvement of the reaction conditions. Coupled

transcription/translation from DNA templates has been developed (DeVries and Zubay 1967; Kigawa and Yokoyama 1991) as well as dialysis methods (Spirin et al. 1988; Kigawa and Yokoyama 1991; Davis et al. 1996; Kim and Choi 1996; Kigawa et al. 1999), alternative energy regenerating systems (Kim et al. 1996; Kigawa et al. 1999) and optimized cell extract preparation protocols (Kigawa et al. 2004).

Components	Prokaryotic	Eukaryotic
Ribosomes	70S	80S
mRNA [b]	Incl. SD sequence	Incl. proper 5′- and 3′-UTRs
Elongator tRNAs, total	Prokaryotic set	Eukaryotic set
Initiator tRNA [c]	Prokaryotic $tRNA_f^{Met}$	Eukaryotic $tRNA_i^{Met}$
20 individual ARSes	Prokaryotic set	Eukaryotic set
Initiation factors	IFs: 1, 2, 3	eIFs: 1, 1A, 2, 2A, 2B, 3, 4A, 4B, 4F (4A+4E+4G), 5, 5A, 6
Elongation factors	EF-Tu, EF-Ts, EF-G	eEF1A, eEF1B, eEF2
Termination factors	RF1, RF2, RF3, RRF	eRF1, eRF3
Amino acids	All 20	All 20
NTPs [d]	ATP, GTP, or ATP, GTP, CTP and UTP	ATP, GTP, or ATP, GTP, CTP and UTP
NTP-regenerating sub-system [e]	PEP+PK, or CrP+CK, or AcP, or pyruvate+NAD+CoA	CrP+CK
Sulfhydryl compounds	ME, or DTT, or GSH	DTT, or GSH/GSSG mixture
Mg^{2+}	8 – 15 mM	1 – 4 mM
K^+ and/or NH_4^+ [f]	100 – 250 mM	ca. 100 mM

[a] Abbreviations: SD,polypurine Shine-Dalgarno sequence preceding the initiation codon; UTR, untranslated region of mRNA; ARS, aminoacyl-tRNA synthetase; IF and eIF, prokaryotic and eukaryotic initiation factors, respectively; EF and eEF, prokaryotic and eukaryotic elongation factors, respectively; RF and eRF, prokaryotic and eukaryotic termination factors, respectively; RRF, ribosome recycling factor; PEP, phosphoenolpyruvate; CrP, creatine phosphate; AcP, acetylphosphate; ME, mercaptoethanol; DTT, dithiotreitol; GSH and GSSG, reduced and oxidized glutathione, respectively.

[b] Instead of isolated mRNA, an mRNA-synthesizing sub-system consisting of a plasmid or a DNA fragment (gene) and a DNA-dependent RNA polymerase is present in coupled and combined transcription-translation systems. In this case all four NTPs (ATP, GTP, CTP and UTP) must be present.

[c] Prokaryotic (bacterial) systems must contain also Met-$tRNA_f$-formylation sub-system including Met-$tRNA_f$ transformylase and formyltetrahydrofolate or its congener (e.g., methenyltetrahydrofolate or folinic acid).

[d] ATP and GTP are sufficient for translation systems suppled with mRNA, whereas all coupled and combined transcription/translation systems require all four NTPs.

[e] In addition to the listed in the table, an ATP/GTP-regenerating sub-system can be also base on other glycolytic intermediates, including glucose-6-phosphate.

[f] In addition to the above-listed components, protease inhibitors, RNAase inhibitors (e.g., RNAsin) and antiseptics (e.g., NaN_3) are usually included into the cell-free translation mixture. In the case of the synthesis of some disulfide-bonded proteins a correct redox potential must be maintained with a mixture of GSH/GSSG. The incubation mixture must be buffered at pH 7.4 – 8. Incubation temperature is 26°C to 37°C.

Figure 7. Principal composition of cell-free translation systems. Reproduced from Kigawa et al. (1999).

While different cell-free systems have specific advantages, the *E. coli*-based system is most commonly used because *E. coli* grow quickly, are well characterised, allow cost-effective extract preparation, and generally produce high recombinant protein yields. Milligram quantities of proteins with stable isotope labelling and/or unnatural amino acids (UAAs) can be produced by *E. coli* based cell-free protein synthesis, which is sufficient for NMR studies. In particular, *E. coli*-based cell-free protein synthesis offers inexpensive access to selectively isotope labelled protein. Depending on the

requirements of NMR experiments, single or combinatorial isotope labelling can be obtained by replacing unlabelled amino acids in the amino acid mixture with isotope-labelled ones (Kigawa et al. 1995; Ozawa et al. 2004; Etezady-Esfarjani et al. 2007; Apponyi et al. 2008). Isotope scrambling can be efficiently suppressed by inhibiting 5'-phosphate (PLP) enzymes (Su et al. 2011). For example, in this thesis, partially deuterated NMR samples were easily made in an overnight cell-free expression, which highlights the time saving achievable by cell-free synthesis (Paper 4).

1.3.2 Genetically encoded unnatural amino acids

Genetic incorporation of UAAs into proteins also has become popular with the arrival of orthogonal aminoacyl-tRNA synthetase (aaRS)-amber suppressor tRNA ($tRNA_{CUA}$) pairs (Figure 8) (Furter 1998; Wang and Schultz 2001; Neumann et al. 2008). The most extensively used UAA incorporation system is based on suppressor tRNA recognising the amber stop codon TAG in *E. coli* by an orthogonal $tRNA_{CUA}^{Tyr}$/TyrRS pair derived from *Methanocaldococcus jannaschii* (Liu and Schultz 2010; Hoesl and Budisa 2012).

In living cells, the incorporation of an UAA is usually achieved in response to a TAG stop codon in competition with termination of translation by the endogenous release factor 1 (RF1), which can result in low productivity of full-length protein particularly for multiple incorporations and for in-cell approches (Liu et al. 2007; Lin and Wang 2008). Ueda and co-workers reported that, in a PURE cell-free system reconstituted from purified recombinant protein factors (IF1, IF2, IF3, EF-G, EF-Tu, EF-Ts, RF1, RF3, RRF, 20 aaRS, MTF and T7 RNA polymerase), omission of RF1 allowed efficient incorporation of a UAA (Shimizu et al. 2001). Complete removal of RF1 from the S30 extract derived from *E. coli* cells successfully suppresses the production of truncated products; moreover, by adding purified orthogonal amber suppressor tRNA/aaRS pairs into the standard cell-free system, one UAA can be incorporated into multiple sites with high yields (Loscha et al. 2012).

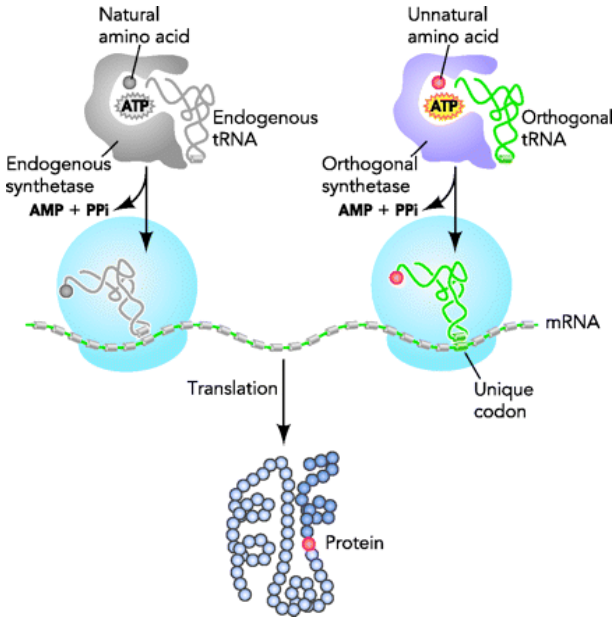


Figure 8. Site-specific incorporation of an unnatural amino acid into proteins. Reproduced from Lin et al. (2008).

In an alternative approach, many researchers attempted to create RF1-free *E. coli* strains to facilitate incorporation of UAAs. In the *rEc.E13.*Δ*prf*A strain, the TAG codons of seven essential genes were reassigned to allow their expression in the absence of RF1, while other house-hold proteins, requiring UAG stop signals, were synthesized with unnecessary read-through peptide tails terminated by in-frame UAA or UGA codons (Mukai et al. 2010). In the JX33 strain, RF2 was engineered to recognise the TAG stop codon, but this renews the competition of a release factor with suppressor tRNA (Johnson et al. 2011). Ultimately, the genomically recoded *E. coli* strain C321.ΔA was successfully constructed, in which all TAG stop codons were replaced and RF1 was deleted to abolish UAG-mediated translation termination (Lajoie et al. 2013). These RF1-depleted strains have successfully been used for the incorporation of UAAs, achieving enhanced suppression in response to UAG or UAGN codons both *in vivo* (Chatterjee et al. 2014; Schmidt et al. 2014; Lim et al. 2015; Pirman et al. 2015; Wilkins et al. 2015) and in cell-free synthesis setups (Hong et al. 2014; Chemla et al. 2015; Li et al. 2015; Sullivan et al. 2015). For NMR applications, UAAs containing reactive groups (Loh et al. 2013; Abdelkader et al. 2016) or acting as specific NMR probes, e.g. by introducing ^{19}F (Jackson et al. 2007; Shi et al. 2011; Li et al. 2013; Yang et al. 2015), ^{13}C/^{15}N (Jones et al. 2010) or *tert*-

butyl group labels (Chen et al. 2015) have successfully been incorporated into many proteins.

The ongoing development of both cell-free protein synthesis and genetic incorporation of UAAs will further broaden their applications in NMR spectroscopy. In particular, the unnatural amino acid *O-tert*-butyl tyrosine, incorporated both by *in vivo* and *in vitro* techniques, proved to deliver a highly sensitive NMR probe for high-molecular-weight systems and measurement of sub-micromolar binding affinities, which is described in Paper 3 (Chen et al. 2015).

1.4 Software supporting the experimental approaches

Computational methods play an important role in paramagnetic NMR studies, for example to help with the selection and design of tagging sites, as well as the analysis of experimental data.

1.4.1 Mutation guided by PyMOL

Point mutation or site-directed mutagenesis is widely used to support resonance assignments by NMR spectroscopy and to produce a single site for tagging with a lanthanide complex or for UAA incorporation. In order to increase the success rate of lanthanide tagging, it is of great importance to predict the outcome of a mutation based on the protein structure. For many protein targets, crystal structures, NMR structures or computational models are available. Crystal structures tend to be the best choice because they provide the most accurate atomic positions.

PyMOL is a molecular visualization software, built on an open-source platform. The mutagenesis wizard function in PyMOL has been developed to display the effect of a mutation prior to experiment. By inspecting the rotamers of specific amino acid substitutions in PyMOL, one can gain an idea about possible van der Waals violations and the possible χ_1 angles compatible with the site-directed mutagenesis or unnatural amino acid incorporation.

1.4.2 $\Delta\chi$-tensors fitting and PCS prediction by Numbat

For the situation where the 3D structure of the protein is known, several integrated software tools have been developed to determine the $\Delta\chi$ tensors from PCSs, including Fantasia (Banci et al. 1996), Fantasian (Banci et al. 1997), Platypus (Pintacuda et al. 2004), Echidna (Schmitz et al. 2006), Xplor-NIH (Schwieters et al. 2003; Schwieters et al. 2006) with PARA restraint (Banci et al. 2004), GROMACS (Van Der Spoel et al. 2005) with an implementation of orientation restraints (Hess and Scheek 2003) and DYANA (Güntert et al. 1997) with PSEUDYANA module (Banci et al. 1998). All these programs are very powerful and capable of doing much more than fitting $\Delta\chi$ tensors to protein structures, making them relatively cumbersome to use. The program Numbat (**N**ew **U**ser-friendly **M**ethod **B**uilt for **A**utomatic $\Delta\chi$-**T**ensor determination) is designed for the determination of $\Delta\chi$ tensors and metal ion position in an easy and user-friendly way (Schmitz et al. 2008).

The input files for Numbat are simply pdb files that comprise the atomic coordinates of the protein and text files (with the extension .npc) containing the PCS data. Two particularly helpful features of Numbat include PCS prediction and a combined fit using multiple PCS data sets obtained from different samples with different metal ions located at the same site. Fitting the metal position using PCS data sets from multiple different metals at the same tagging site improves the accuracy of the fitting. This approach helps to overcome limitations arising from small numbers of PCSs available in each data set. The PCS prediction can be done automatically for all atoms after fitting of the $\Delta\chi$ tensors. The quality of the fitting can easily be evaluated by the correlation between the experimental and calculated PCS values. The PCS predictions made for atoms, for which PCSs are not provided in the input, are useful to guide additional PCS assignments (de la Cruz et al. 2011) or to validate the fit by comparison with the experimental PCSs that are not included in the fit (described in Paper 2) (Chen et al. 2014).

1.4.3 $\Delta\chi$-tensors fitting and structure calculation by PCS-Rosetta and GPS-Rosetta

A breakthrough in 3D structure determinations of proteins was obtained by Shen et al.

(2008) who developed a protocol that combines the structure prediction program Rosetta with experimental $^{13}C^{\alpha}$, $^{13}C^{\beta}$, $^{13}C'$, ^{15}N, $^{1}H^{\alpha}$ and $^{1}H^{N}$ NMR chemical shifts to assist in the initial fragment selection step of Rosetta. The resulting program, CS-Rosetta, has been shown to successfully yield structures closely similar to X-ray or NMR structures determined primarily from nuclear Overhauser effects (NOEs) (Shen et al. 2008). In a later step, PCS-Rosetta was developed, which incorporates the long-range distance information inherent in PCSs in the fragment assembly step to fit $\Delta\chi$ tensors and thus yield models in agreement with experimental PCSs. This was shown to improve the yield of near-native structures over CS-Rosetta as measured by C^{α} rmsd values to the target structures (Schmitz et al. 2012). Recently, PCS-Rosetta was developed further into a version called GPS-Rosetta, which uses a similar algorithm as PCS-Rosetta but accepts amide proton PCSs generated by paramagnetic centres at multiple sites (Yagi et al. 2013). Remarkably, GPS-Rosetta can determine complete 3D structures of proteins from backbone amide PCSs only, which can readily be measured in ^{15}N-HSQC spectra (Pilla et al. 2015). GPS-Rosetta has also proved a powerful tool for the determination of protein-ligand complexes, which is described in Paper 4.

1.5 A two-component drug target: the dengue virus NS2B-NS3 protease

The four serotypes of dengue virus (DENV) constitute most significant mosquito-borne viral pathogens. Each of the four DENV serotypes produces acute, self-limited infections. The first infection with any of the serotypes typically produces a mild fever, rash or muscle/joint pain, and patients subsequently gain a lifelong immunity to that serotype. A second infection by another serotype, however, can cause deadly diseases due to dengue haemorrhagic fever and shock syndrome (DHF/DSS). It is estimated that there are 390 millions new dengue infections per year, of which roughly 96 million produce clinically apparent symptoms (Bhatt et al. 2013).

Recent reports of phase IIb and III trials with the candidate dengue vaccine CYD-TDV indicate that the vaccine is more serotype-specific than hoped, providing only limited protection against dengue 2 and no effective protection for children below the age of six (Sabchareon et al. 2012; Capeding et al. 2014; Villar et al. 2014). Drugs

against dengue are thus still of undiminished interest. As evidenced by the clinical availability of ten human immunodeficiency virus (HIV) protease inhibitors (PIs) (De Clercq 2009) and two hepatitis C virus (HCV) PIs (Kwong et al. 2008), viral proteases are proven drug targets, thus making the DENV protease one of the most attractive drug targets to overcome dengue.

1.5.1 Roles of NS3pro and the NS2B cofactor

In the polyprotein translated from the DENV genome (Figure 9), the N-terminal 180 amino acids of NS3 constitute a protease domain (NS3pro), which mediates the cleavage at several protein boundaries, including capsid protein/precursor of the viral membrane protein (prM), NS2A/NS2B, NS2B/NS3, NS3/NS4A, NS4A/NS4B and NS4B/NS5 junctions (Falgout et al. 1991; Cahour et al. 1992; Amberg and Rice 1999; Muñoz-Jordán et al. 2005). Deletion of NS2B abolishes the NS3 protease activity (Falgout et al. 1991) and further study revealed that the conserved hydrophilic domain of 40 amino acids from NS2B, with deletion of its hydrophobic region, is essential for the proteolytic activity of NS3pro (Falgout et al. 1993; Clum et al. 1997). Thus, the dengue virus NS3pro in complex with its NS2B cofactor was established as an attractive therapeutic target for the rational design of dengue drugs (Yusof et al. 2000; Leung et al. 2001).

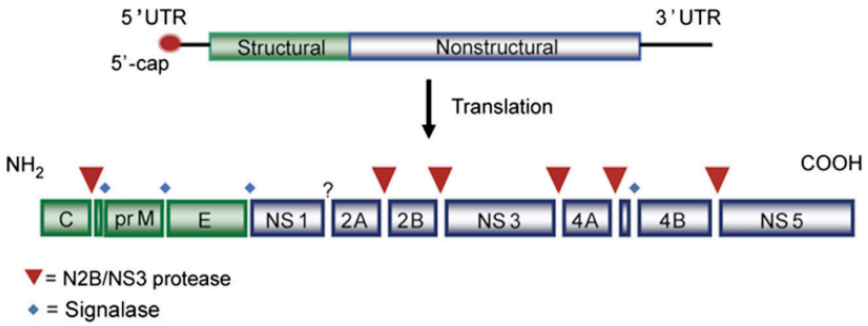


Figure 9. Schematic diagram of the DENV genome organization and polyprotein processing. The top depicts the viral genome including the 5'-cap, 5'UTR, the structural and non-structural protein coding regions, and the 3'UTR. The bottom shows the genome-translated polyprotein, which is processed by the viral NS2B/NS3 protease, host signalases and unknown proteases. C, prM and E denote the capsid protein, precursor of the viral membrane protein, and envelope protein on the viral surface, respectively. Reproduced from Lim et al. (2013).

The proteases from all four serotypes of DENV share high sequence identity as well as similar substrate specificity (Li et al. 2005). As infection by DENV2 has been reported to be more serious than the other three serotypes (Vaughn et al. 2000), DENV2 NS2B-NS3pro has become the target of particularly intensive efforts in drug development. The construct that used in this thesis consists of 47 residues from DENV2 NS2B, a cleavable linker (EVKKQR) and 185 residues from DENV2 NS3pro (de la Cruz et al. 2014).

1.5.2 Structural studies of dengue virus NS2B-NS3pro

Molecular structure determination plays an important role in the structure-based strategies for drug discovery (Kuntz 1992). To understand the role of the NS2B cofactor in DENV NS3pro activation from a structural point of view, many efforts have been made to determine the structure of the complex between NS2B and NS3pro.

Crystal structures of the highly homologous NS2B-NS3pro from West Nile virus (WNV), which shares 56% sequence identity with DENV NS2B-NS3pro, revealed two distinct conformations, open and closed (Figure 10), in the apo and substrate-bound states, respectively (Erbel et al. 2006; Aleshin et al. 2007). However, NMR studies could only confirm the closed conformation of the WNV protease irrespective of the presence or absence of low-molecular-weight inhibitors (Su et al. 2009a; Su et al. 2009c). Similarly, a high quality crystal structure of DENV2 NS2B-NS3pro showed an open conformation without inhibitor (Figure 10) (Erbel et al. 2006). The N-terminal part of the NS2B cofactor (residues 49-66), which is sufficient to stabilize NS3pro, forms a β strand and inserts into the N-terminal β barrel of NS3pro. Unfortunately, the electron density of the C-terminus of NS2B ($NS2B_c$), which is enzymatically essential, was discontinuous and located far away from the active site of NS3pro. Crystallization of DENV NS2B-NS3pro is difficult and the question of open and closed conformation is more easily addressed by NMR spectroscopy. Early NMR studies of a DENV NS2B-NS3pro construct with a covalent linker between NS2B and NS3pro indicated that NS2B was unfolded in the absence of inhibitor (Wu et al. 2007). A subsequent NMR analysis of this construct in the presence of a low-

molecular-weight inhibitor clearly demonstrated the closed conformation, which proved to be similar to the closed conformation of WNV protease (de la Cruz et al. 2011). Subsequently, a crystal structure of DENV3 NS2B-NS3pro in complex with a peptide inhibitor confirmed the closed conformation, showing that $NS2B_c$ forms a β hairpin and wraps around the core of NS3pro to line the substrate binding site (Figure 11) (Noble et al. 2012).

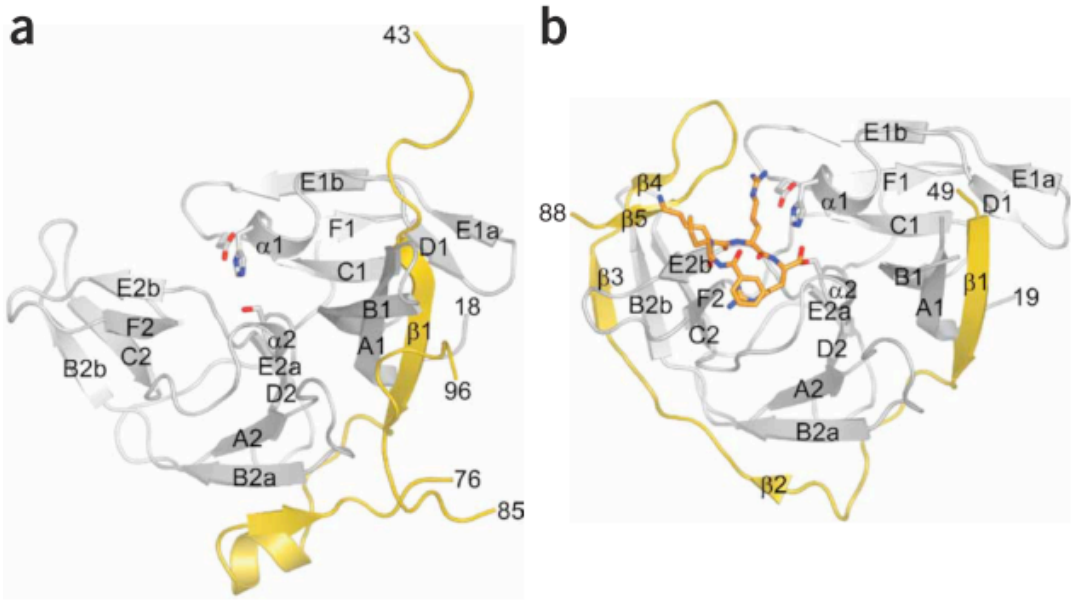


Figure 10. Structure of NS2B-NS3pro in the absence and presence of an inhibitor. (a) DENV NS2B-NS3pro. Gray, NS3pro; yellow, NS2B. PDB code: 2FOM. (b) WNV NS2B-NS3pro in complex with Bz-Nle-Lys-Arg-Arg-H (orange). PDB code: 2FP7. Reproduced from Erbel et al. (2006).

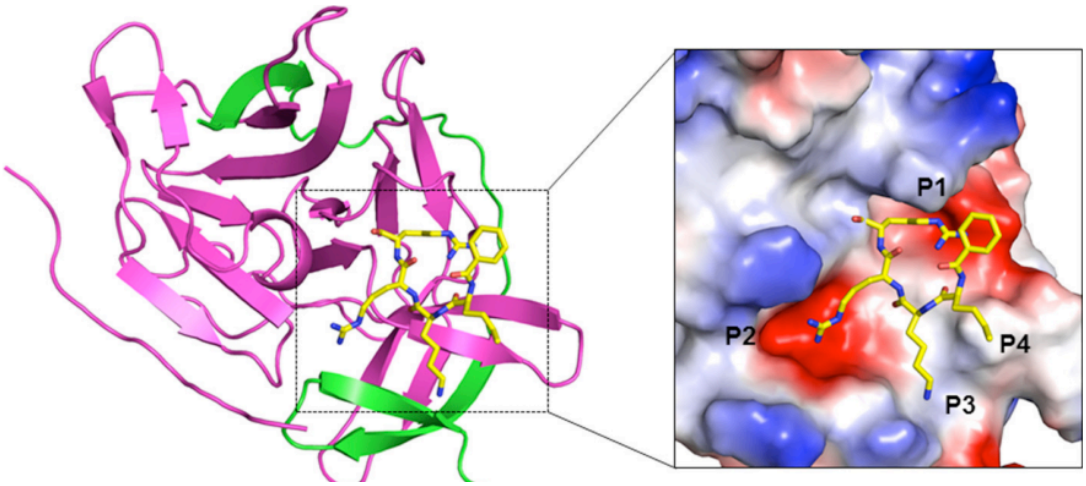


Figure 11. Structure of DENV3 NS2B-NS3pro, highlighting the S2 and S3 pockets, with the bound nKRR-H inhibitor shown as sticks (PDB code: 3U1I). The structure is shown in cartoon representation, with NS3 in pink and NS2B in green. Reproduced from Lim et al. (2013).

In contrast, more recent NMR studies indicated that the open conformation in the absence of any inhibitor may be an artefact of the Gly_4-Ser-Gly_4 linker between NS2B and NS3pro, as constructs without a covalent linker overwhelmingly populate the closed conformation in solution, even in the absence of inhibitor (Kim et al. 2013; de la Cruz et al. 2014). In addition, narrow resonances previously thought to represent the open state were reassigned as protein degradation signals (de la Cruz et al. 2014).

The two distinct confirmations of DENV NS2B-NS3pro evident by both crystallographic and/or NMR studies compromised early structure-guided approaches for drug discovery. In this thesis, the study on the unlinked construct of DENV2 confirmed that the closed conformation of DENV NS2B-NS3pro is the best model for structure-guided drug design (chapters 2 and 3).

1.5.3 Inhibitors targeting NS2B-NS3pro

After more than a decade of research, there still is no dengue virus protease inhibitor that has advanced into clinical trials. The big challenges of the inhibitor design are the shallow active site and the difficulty to replace the two basic P1 and P2 residues (Figure 11) (Lim et al. 2013).

Over the years, many research groups have focused on developing DENV inhibitors. Numerous non-peptidic inhibitors were identified via high-throughput library screening or *in silico* docking. However, none of the non-peptidic inhibitors reported showed inhibition constants in the nanomolar range. False positives and non-drug-like properties of the hits were major problems slowing down the progress (Lim et al. 2013; Nitsche et al. 2014). In contrast, using the substrate recognition sequence as starting point for the design of protease inhibitors, quite a number of peptidic inhibitors have been developed with low micromolar inhibition constants (Figure 12)

(Nitsche et al. 2014). The most potent peptidic inhibitor, a tetra-peptide derived from Bz-Nle-Lys-Arg-Arg (**6**, Figure 12) with an electrophilic boronic acid warhead, is unsuitable for oral dosing and also not DENV-specific (Yin et al. 2006). It is interesting that the sequence WYCW-NH_2 (**24**, Figure 12), which successfully removed Arg residues from the nonprime positions, exhibited a promisingly low K_i value against DENV2 (4.8 μM), indicating that there is no absolute requirement for basic side chains in the dengue virus inhibitors. The binding mode of this peptidic inhibitor, however, remains unknown (Prusis et al. 2013).

The key assumption of structure-based design is that good inhibitors must possess significant chemical as well as structural complementarity to their target receptors (Dean 1987). Due to the difficulties in co-crystallizing the dengue virus protease with any inhibitors, the binding modes of the dengue virus protease inhibitors relies mostly on computer docking algorithms. Even when molecular dynamic simulations were applied to explore conformational changes of the protein induced by specific interactions with inhibitors, the reliability of the resulting complex models remains a concern. Clearly, experimentally confirmed models of the complex between NS2B-NS3pro and inhibitors are needed to help with the development of lead compounds and thus guide lead optimisation. It is the central aim of this thesis to use information obtained by paramagnetic NMR spectroscopy as structural restraints for the modelling of protein-ligand complexes to provide a guide for rational drug design against dengue (Paper 4).

number	structure	inhibition potential (DENV-2)
6	Bz-nKRR-H	K_i = 5.8 μM
11	Bz-nKR(4-Me)F-H	K_i = 6.0 μM
12	Bz-nKRW-H	K_i = 7.5 μM
13	Bz-nKR(4-guanidinyl)F-H	K_i = 2.8 μM
14	Bz-KRR-H	K_i = 1.5 μM
15	4-Phenylphenylacetyl-KKR-H	K_i = 12.2 μM
16	2-Phenylacetyl-KRR-H	K_i = 6.7 μM
17	Itz-CTRRGLPVCGSEE GPTNCGRRS	K_i = 1.4 μM
18	Itz-CRRSWPVCTRRS GRRSEESGCVPLG	K_i = 3.0 μM
Tyrothricin	mixture of peptides	K_i = 3.0 μM
Protegrin-1	RGGRLCYCRRRFCVCVGR	K_i = 5.85 μM
Retrocyclin-1	GICRCICGRGICRCICGR	IC_{50} = 46.1 μM (20 °C) IC_{50} = 14.1 μM (40 °C)
Plectasin	linear, 40 amino acids	K_i = 5.03 μM
Protegrin-1-MAP30-Plectasin	fusion peptide, 37.7 kDa	IC_{50} = 0.5 μM
19	CGYKLC	K_i = 12.5 μM
20	CGKRKLC	K_i = 2.5 μM
21	CGKRKSC	K_i = 1.4 μM
22	CAGKRKSG	K_i = 2.2 μM
23	Ac-RRRRHWCW-NH_2	K_i = 0.3 μM
24	WYCW-NH_2	K_i = 4.8 μM

Figure 12. Selected peptidic inhibitors of the dengue virus protease. Reproduced from Nitsche et al. (2014).

Conclusion and Outlook

2.1 Additional lanthanide tagging sites on the dengue virus protease

2.1.1 Site B is not an ideal site for inhibitor binding studies

When studying the conformation of the C-terminal segment of NS2B ($NS2B_c$) in the complex between the dengue virus protease and BPTI, the bulky lanthanide tag at site B, which produces the largest PCSs for residues of $NS2B_c$ (Paper 1), was noted to be fairly close to ligands at the active site, which could lead to clashes between ligand and tag, in particular for bulky ligands like BPTI. Therefore, only site C was chosen to generate PCSs for the complex with BPTI. In hindsight, PCSs induced from a single site were sufficient to position $NS2B_c$ in that particular study (Paper 2). In the subsequent small-molecule binding study, large PREs resulted from tags at site B, which hampered the unambiguous assignment of PCSs in 1D ^{1}H NMR spectra (Paper 4). This indicated an unexpectedly short distance between the paramagnetic centre and the nuclei of interest, which may be attributed to tag flexibility. PREs are exquisitely sensitive to the presence of minor species, even with populations as low as 1%, if the distances between the paramagnetic centre and the protons of interest are much shorter in the minor than in the major species (Tang et al. 2007). At site B, the tag has the chance to move as close as 15 Å to the *tert*-butyl group, which would produce a large PRE effect and render the observation of PCSs in 1D ^{1}H NMR spectra more difficult (Figure 13).

Very recently, Koh-Stenta *et al*. (2015) reported an NMR study of the complex of the dengue virus protease with another covalently bonded inhibitor carrying a ester warhead (Koh-Stenta et al. 2015). They presented a docking model, which showed a 4-methoxyphenyl group interacting with residue 34 in the P2' site. Unfortunately, residue 34 is also site B. In other words, a bulky lanthanide tag ligated at site B could easily interfere with the binding of ligands in the P2' pocket.

Consequently, site B may need to be abandoned in future ligand binding studies of the dengue virus protease, and at least one or more alternative solvent-exposed sites suitable for lanthanide tag ligation should be identified.

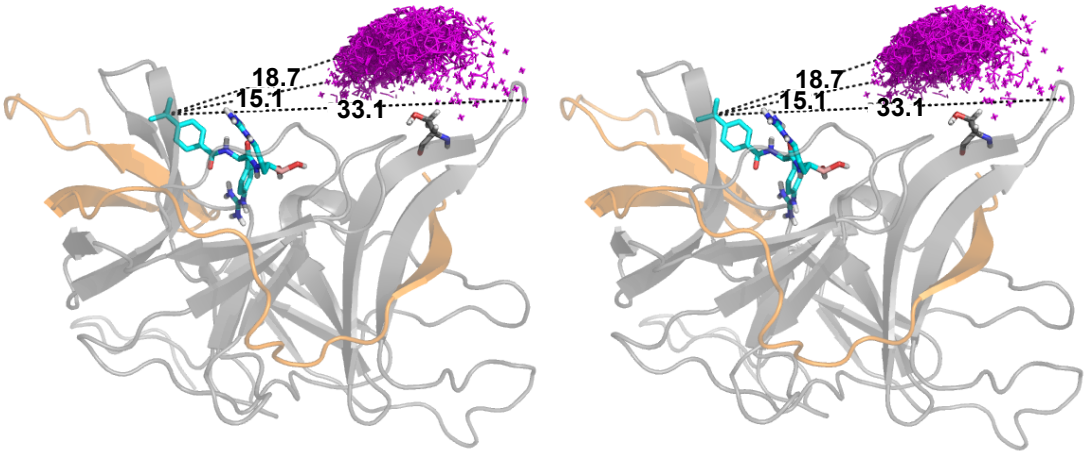


Figure 13. Stereo view of the metal position distribution at site B on the dengue virus protease complex. NS2B is coloured orange, while NS3pro is shown in grey. Ligand **1** and Ser34 are shown in sticks. Three examples of distances are drawn as black dotted lines. The positions of the metal were sampled from 100,000 rotamers. The most highly populated distances were between 18-20 Å, but the tag also has the chance to move towards the *tert*-butyl group of ligand **1** as close as 15 Å and as far as 33 Å. Minor populations of short distances easily dominate the line broadening effect for the *tert*-butyl group.

2.1.2 Possible new sites for lanthanide tag conjugation

In theory, the more tagging sites are available the better, since each of them would provide unique and helpful 3D restraints. Previously, three single cysteine mutations were established for the introduction of C1 and C2 lanthanide tags to study the conformation of NS2B: Ala57*Cys (site A), Ser34Cys (site B) and Ser68Cys (site C) (labelled as A-C in Figure 14) (de la Cruz et al. 2011). Based on the known ligand binding pockets, the following sites may possibly be suitable as well: Lys42, Thr45, Val109 and Leu115 (labelled as 1-4 in Figure 14). The rational for this proposal is based on a number of observations. Firstly, the conformation of $NS2B_c$ needs to be monitored in any ligand binding study, making NS2B unsuitable for tagging. In

addition, the mutation sites should be sufficiently far from the catalytic triad to avoid interference with the binding pocket due to the incorporation of a bulky lanthanide tag.

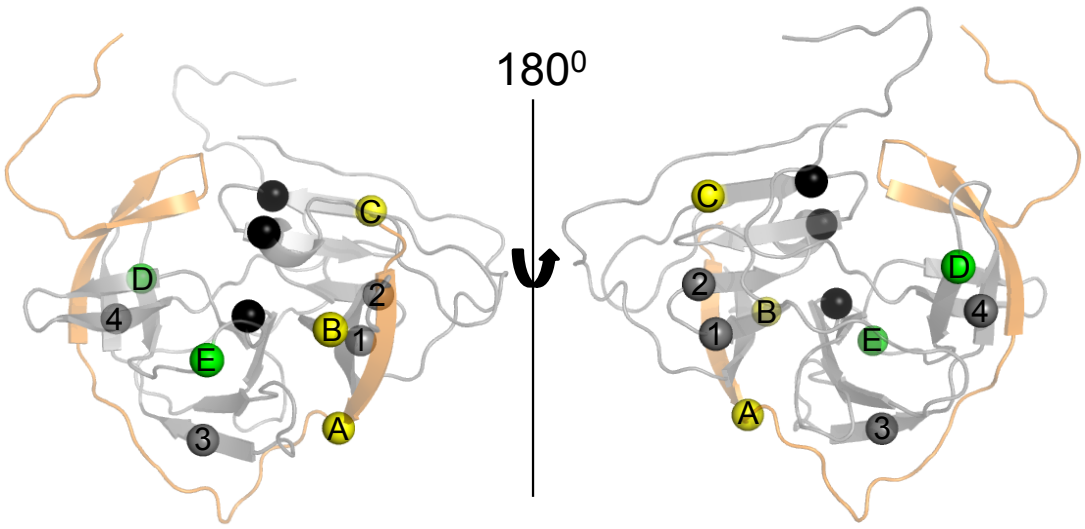


Figure 14. Proposed tagging sites mapped on the dengue virus protease. NS2B is coloured orange, while NS3pro is shown in grey. The new proposed tagging sites are numbered 1-4. They correspond to the sites of Lys42, Thr45, Val109 and Leu115. The catalytic triad (consisting His51, Asp75 and Ser135) is identified by black spheres. The three established sites A (Ala57*), B (Ser34) and C (Ser68) for single cysteine mutations are shown as yellow spheres. Two other sites (D: Thr122 and E: Ser189), which proved unsuitable for conformational studies of NS2B (de la Cruz, 2013) are identified by green spheres.

Lys42 is located in the same β-strand as Ser34 (site B), but it is on a different face of the protein and opposite to the active site. It must be noted, however, that charged residues may not the best choice for mutation, as charge-charge interactions can be important for protein stability. Thr45, the first residue of the next β-strand, is located close to Lys42 and may be equally suitable as a tagging site. A cysteine residue replacing Val109 has a good chance of directing a lanthanide tag to point in a direction parallel to the binding pocket, thus avoiding clashes with NS2B. Lys42, Thr45 and Val109 are all at similar distances to $NS2B_c$ as the sites B and C, suggesting that all of these sites might be suitable for assessing the conformation of $NS2B_c$ even when inhibitors bind to the active site of NS3pro.

In the Ph.D. thesis of Dr L. de la Cruz, Thr122 and Ser189 (labelled as C and E in Figure 14) were shown to be unsuitable tagging sites for conformational studies of $NS2B_c$ (de la Cruz, 2013). When re-evaluated, Ser189 is still considered an unsuitable tagging site, since it is too close to the binding pocket. Leu115 (labelled 4 in Figure 14) is near Thr122, is opposite to the binding site and highly solvent exposed. Lanthanide tags at Leu115 may induce sizeable PCSs for ligands binding at the active site. Like Thr122, however, Leu115 is very close to $NS2B_c$. Nevertheless, using a more rigid tag and a less strongly paramagnetic metal ion like Ce^{3+}, PREs on $NS2B_c$ can provide distance information to help assess its conformation. Therefore, Leu115 might serve as a useful supplemental tagging site for ligand binding studies using both PREs and PCSs.

2.2 The potential of *tert*-butyl substitutions

2.2.1 Multiple *tert*-butyl substitutions on the same small-molecule ligand

The usefulness of *tert*-butyl groups for the structural analysis of proteins and protein-ligand complexes by NMR spectroscopy is a key result of this thesis (Chapters 4 and 5).

For small-molecule ligands, the introduction of a *tert*-butyl group would significantly increase the overall size. However, free rotations of a *tert*-butyl group could reduce the impact from the change in size as long as it is solvent-exposed (*i.e.* points out from the binding pocket or away from the protein surface). A ligand with *tert*-butyl group should be capable of binding in generally the same location as the ligand without substitution. Clearly, the chemical properties of this group may still affect the protein-ligand affinity. For example, the dengue virus protease ligand analogue with a *tert*-butyl group exhibited a slightly decreased binding affinity compared with the original ligand. The effect on affinity would depend on the position of the substitution on the ligand and the chemical and structural properties of the binding pocket. Therefore, the site for introduction of a *tert*-butyl group on the ligand must be carefully designed and its effect on inhibition and/or binding affinity should be evaluated prior to structural studies. Furthermore, similar to fluorine substitutions, the

tert-butyl substitutions could be varied chemically (Figure 15), which would result in different chemical shifts. For example, the 1D ^{1}H NMR signal of a *O-tert*-butyl group typically appears near 1.2 ppm, while a *tert*-butyl-benzyl group exhibits a small downfield shift by about 0.05 ppm. Boc groups display signals even further downfield due to the electron withdrawing effect of the ester group, i.e. at 1.4 ppm. These features can be helpful for assignment, when multiple *tert*-butyl substitutions are introduced are different sites on the same ligand. Indeed, ligands exist that feature a *tert*-butyl group and a Boc group at the same time, e.g. asunaprevir (Scola et al. 2014).

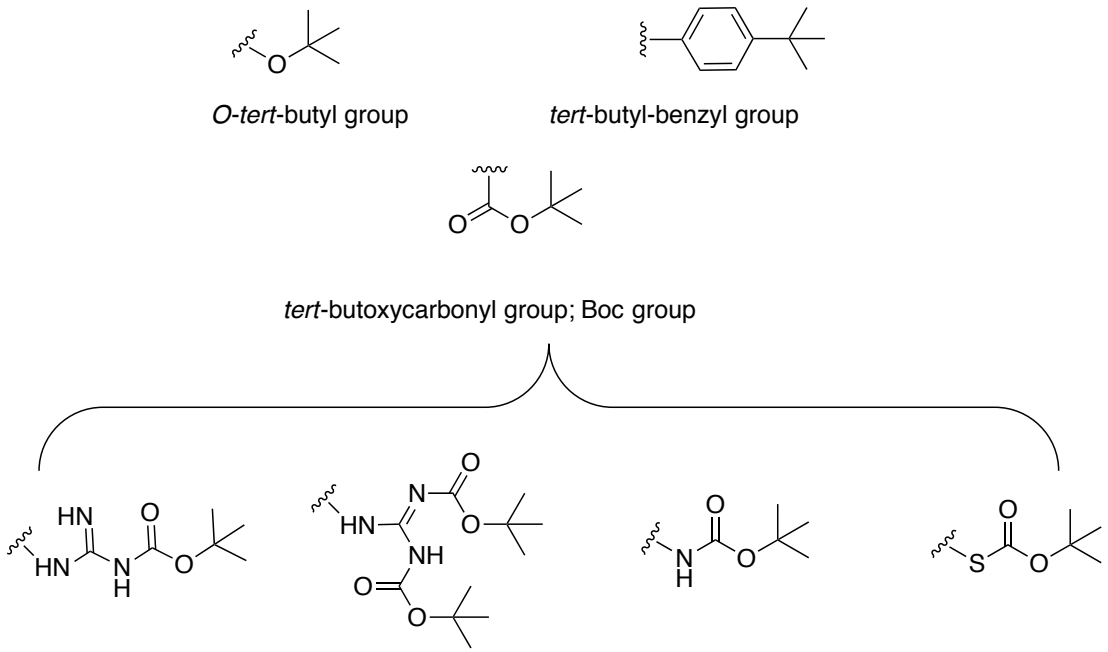


Figure 15. Chemical structures of various *tert*-butyl substitutions.

2.2.2 *Tert*-butyl substitutions on natural amino acids

A number of natural amino acids bearing *tert*-butyl substitutions have been developed and commercialised (Figure 15). *O-tert*-butyl substitutions could be introduced to modify Tyr, Ser and Thr. Boc substitutions could be performed on Cys, His, Lys, Arg, Asp and Glu. Different *tert*-butyl substitutions are possible. Chemical peptide synthesis uses *tert*-butyl containing unnatural amino acids, making it possible to simply retain *tert*-butyl groups during removal of the N-Boc protecting group (Gibson et al. 1994). Notably, the *tert*-butyl group is used as a common protection group because it can easily be removed; nonetheless, deprotection under mild NMR buffer

conditions (usually pH 6-8) is not easy. There is an increased interest in peptides in pharmaceutical research and development (R&D) and approximately 140 peptide therapeutics are currently being evaluated in clinical trials (Fosgerau and Hoffmann 2014). Therefore, peptides with *tert*-butyl substitutions may be straightforward to obtain and could produce useful structural information in interaction studies with the target proteins. *Tert*-butyl groups thus present a bright future for NMR spectroscopy in lead compound development.

2.3 Concluding remark

This thesis focuses on the development of paramagnetic NMR spectroscopy for structural analysis of proteins (Paper 1), protein-protein complexes (Paper 2) and protein-ligand complexes (Papers 3 and 4). Special tools used in these studies are pseudocontact shifts, lanthanide tags, unnatural amino acids, cell-free protein synthesis, as well as the programs Numbat and Rosetta. In particular, the combination of the highly NMR-sensitive *tert*-butyl label with PCSs broadens the scope of paramagnetic NMR spectroscopy in drug discovery to provide structural information for rational ligand design in situations, when atomic-resolution structural data of the protein-ligand complex cannot easily be obtained by other techniques.

References

Abdelkader EH, Yao X, Feintuch A, Adams LA, Aurelio L, Graham B, Goldfarb D and Otting G (2016) Pulse EPR-enabled interpretation of scarce pseudocontact shifts induced by lanthanide binding tags. J Biomol NMR 64:39–51.

Aguirre C, ten Brink T, Cala O, Guichou J-F and Krimm I (2014) Protein-ligand structure guided by backbone and side-chain proton chemical shift perturbations. J Biomol NMR 60:147–156.

Aleshin AE, Shiryaev SA, Strongin AY and Liddington RC (2007) Structural evidence for regulation and specificity of flaviviral proteases and evolution of the Flaviviridae fold. Protein Sci 16:795–806.

Allegrozzi M, Bertini I, Janik MBL, Lee Y-M, Liu G and Luchinat C (2000) Lanthanide-induced pseudocontact shifts for solution structure refinements of macromolecules in shells up to 40 Å from the metal ion. J Am Chem Soc 122:4154–4161.

Amberg SM and Rice CM (1999) Mutagenesis of the NS2B-NS3-mediated cleavage site in the flavivirus capsid protein demonstrates a requirement for coordinated processing. J Virol 73:8083–8094.

Apponyi MA, Ozawa K, Dixon NE and Otting G (2008) Cell-free protein synthesis for analysis by NMR spectroscopy. Methods Mol Biol 426:257–268.

Ardenkjær-Larsen JH, Fridlund B, Gram A, Hansson G, Hansson L, Lerche MH, Servin R, Thaning M and Golman K (2003) Increase in signal-to-noise ratio of > 10,000 times in liquid-state NMR. Proc Natl Acad Sci USA 100:10158–10163.

Banci L, Bertini I, Bren KL, Cremonini MA, Gray HB, Luchinat C and Turano P (1996) The use of pseudocontact shifts to refine solution structures of paramagnetic metalloproteins: Met80Ala cyano-cytochrome c as an example. J Biol Inorg Chem 1:117–126.

Banci L, Bertini I, Cavallaro G, Giachetti A, Luchinat C and Parigi G (2004) Paramagnetism-based restraints for Xplor-NIH. J Biomol NMR 28:249–261.

Banci L, Bertini I, Cremonini MA, Gori-Savellini G, Luchinat C, Wüthrich K and Güntert P (1998) PSEUDYANA for NMR structure calculation of paramagnetic metalloproteins using torsion angle molecular dynamics. J Biomol NMR 12:553–557.

Banci L, Bertini I, Eltis LD, Felli IC, Kastrau DHW, Luchinat C, Piccioli M, Pierattelli R and Smith M (1994) The three-dimensional structure in solution of the

paramagnetic high-potential iron-sulfur protein I from *Ectothiorhodospira halophila* through nuclear magnetic resonance. Eur J Biochem 225:715–725.

Banci L, Bertini I, Savellini GG, Romagnoli A, Turano P, Cremonini MA, Luchinat C and Gray HB (1997) Pseudocontact shifts as constraints for energy minimization and molecular dynamics calculations on solution structures of paramagnetic metalloproteins. Proteins-Struct Funct Bioinform 29:68–76.

Barbieri R, Bertini I, Cavallaro G, Lee Y-M, Luchinat C and Rosato A (2002) Paramagnetically induced residual dipolar couplings for solution structure determination of lanthanide binding proteins. J Am Chem Soc 124:5581–5587.

Barthelmes K, Reynolds AM, Peisach E, Jonker HRA, DeNunzio NJ, Allen KN, Imperiali B and Schwalbe H (2010) Engineering encodable lanthanide-binding tags into loop regions of proteins. J Am Chem Soc 133:808–819.

Bertini I, Del Bianco C, Gelis I, Katsaros N, Luchinat C, Parigi G, Peana M, Provenzani A and Zoroddu MA (2004) Experimentally exploring the conformational space sampled by domain reorientation in calmodulin. Proc Natl Acad Sci USA 101:6841–6846.

Bertini I, Gelis I, Katsaros N, Luchinat C and Provenzani A (2003) Tuning the affinity for lanthanides of calcium binding proteins. Biochemistry 42:8011–8021.

Bertini I, Gupta YK, Luchinat C, Parigi G, Peana M, Sgheri L and Yuan J (2007) Paramagnetism-based NMR restraints provide maximum allowed probabilities for the different conformations of partially independent protein domains. J Am Chem Soc 129:12786–12794.

Bertini I, Janik MBL, Lee YM, Luchinat C and Rosato A (2001) Magnetic susceptibility tensor anisotropies for a lanthanide ion series in a fixed protein matrix. J Am Chem Soc 123:4181–4188.

Bertini I, Luchinat C and Parigi G (2002) Magnetic susceptibility in paramagnetic NMR. Prog Nucl Magn Reson Spectrosc 40:249–273.

Bertini I, Luchinat C, Parigi G and Pierattelli R (2005) NMR spectroscopy of paramagnetic metalloproteins. ChemBioChem 6:1536–1549.

Bertini I, Turano P and Vila AJ (1993) Nuclear magnetic resonance of paramagnetic metalloproteins. Chem Rev 93:2833–2932.

Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, Drake JM, Brownstein JS, Hoen AG, Sankoh O, Myers MF, George DB, Jaenisch T, Wint GRW,

Simmons CP, Scott TW, Farrar JJ and Hay SI (2013) The global distribution and burden of dengue. Nature 496:504–507.

Biekofsky RR, Muskett FW, Schmidt JM, Martin SR, Browne JP, Bayley PM and Feeney J (1999) NMR approaches for monitoring domain orientations in calcium-binding proteins in solution using partial replacement of Ca^{2+} by Tb^{3+}. FEBS Lett 460:519–526.

Bleaney B (1972) Nuclear magnetic resonance shifts in solution due to lanthanide ions. J Magn Reson 8:91–100.

Breeze AL (2000) Isotope-filtered NMR methods for the study of biomolecular structure and interactions. Prog Nucl Magn Reson Spectrosc 36:323–372.

Cahour A, Falgout B and Lai CJ (1992) Cleavage of the dengue virus polyprotein at the NS3/NS4A and NS4B/NS5 junctions is mediated by viral protease NS2B-NS3, whereas NS4A/NS4B may be processed by a cellular protease. J Virol 66:1535–1542.

Capeding MR, Tran NH, Hadinegoro SRS, Ismail HIHM, Chotpitayasunondh T, Chua MN, Luong CQ, Rusmil K, Wirawan DN, Nallusamy R, Pitisuttithum P, Thisyakorn U, Yoon IK, van der Vliet D, Langevin E, Laot T, Hutagalung Y, Frago C, Boaz M, Wartel TA, Tornieporth NG, Saville M and Bouckenooghe A (2014) Clinical efficacy and safety of a novel tetravalent dengue vaccine in healthy children in Asia: a phase 3, randomised, observer-masked, placebo-controlled trial. Lancet 384:1358–1365.

Carruthers TJ, Carr PD, Loh CT, Jackson CJ and Otting G (2014) Iron (III) located in the dinuclear metallo-β-lactamase IMP-1 by pseudocontact shifts. Angew Chemie Int Ed 53:14269–14272.

Carugo O and Argos P (2008) Protein-protein crystal-packing contacts. Protein Sci 6:2261–2263.

Chatterjee A, Lajoie MJ, Xiao H, Church GM and Schultz PG (2014) A bacterial strain with a unique quadruplet codon specifying non-native amino acids. ChemBioChem 15:1782–1786.

Chemla Y, Ozer E, Schlesinger O, Noireaux V and Alfonta L (2015) Genetically expanded cell-free protein synthesis using endogenous pyrrolysyl orthogonal translation system. Biotechnol Bioeng 112:1663–1672.

Chen W-N, Kuppan KV, Lee MD, Jaudzems K, Huber T and Otting G (2015) O-*tert* -butyltyrosine, an NMR tag for high-molecular-weight systems and measurements of

submicromolar ligand binding affinities. J Am Chem Soc 137:4581–4586.

Chen W-N, Loscha K V, Nitsche C, Graham B and Otting G (2014) The dengue virus NS2B-NS3 protease retains the closed conformation in the complex with BPTI. FEBS Lett 588:2206–2211.

Clum S, Ebner KE and Padmanabhan R (1997) Cotranslational membrane insertion of the serine proteinase precursor NS2B-NS3(pro) of dengue virus type 2 is required for efficient in vitro processing and is mediated through the hydrophobic regions of NS2B. J Biol Chem 272:30715–30723.

Comellas G and Rienstra CM (2013) Protein structure determination by magic-angle spinning solid-state NMR, and insights into the formation, structure, and stability of amyloid fibrils. Annu Rev Biophys 42:515–536.

Cotton S (2013) Lanthanide and actinide chemistry. John Wiley & Sons

Crowley PB and Ubbink M (2003) Close encounters of the transient kind: protein interactions in the photosynthetic redox chain investigated by NMR spectroscopy. Acc Chem Res 36:723–730.

Davis J, Thompson D and Beckler GS (1996) Large scale dialysis reactions using *E. coli* S30 Extract Systems. Promega Notes 56:14–21.

De Clercq E (2009) Anti-HIV drugs: 25 compounds approved within 25 years after the discovery of HIV. Int J Antimicrob Agents 33:307–320.

de la Cruz L, Chen W-N, Graham B and Otting G (2014) 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.

de la Cruz L, Nguyen THD, Ozawa K, Shin J, Graham B, Huber T and Otting G (2011) Binding of low molecular weight inhibitors promotes large conformational changes in the dengue virus NS2B-NS3 protease: fold analysis by pseudocontact shifts. J Am Chem Soc 133:19205–19215.

Dean PM (1987) Molecular foundations of drug-receptor interaction. Cambridge Univ Press

DeVries JK and Zubay G (1967) DNA-directed peptide synthesis. II. The synthesis of the α-fragment of the enzyme β-galactosidase. Proc Natl Acad Sci USA 57:1010–1012.

Diercks T, Coles M and Kessler H (2001) Applications of NMR in drug discovery. 5:285–291.

Eichmüller C, Tollinger M, Kräutler B and Konrat R (2001) Mapping the ligand binding site at protein side-chains in protein-ligand complexes through NOE difference spectroscopy. J Biomol NMR 20:195–202.

Erbel P, Schiering N, D'Arcy A, Renatus M, Kroemer M, Lim SP, Yin Z, Keller TH, Vasudevan SG and Hommel U (2006) Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. Nat Struct Mol Biol 13:372–373.

Etezady-Esfarjani T, Hiller S, Villalba C and Wüthrich K (2007) Cell-free protein synthesis of perdeuterated proteins for NMR studies. J Biomol NMR 39:229–238.

Falgout B, Miller RH and Lai C (1993) Deletion analysis of dengue virus type 4 nonstructural protein NS2B: identification of a domain required for NS2B-NS3 protease activity. J Virol 67:2034–2042.

Falgout B, Pethel M, Zhang YM and Lai CJ (1991) Both nonstructural proteins NS2B and NS3 are required for the proteolytic processing of dengue virus nonstructural proteins. J Virol 65:2467–2475.

Fosgerau K and Hoffmann T (2014) Peptide therapeutics: current status and future directions. Drug Discov Today 20:122–128.

Franz KJ, Nitz M and Imperiali B (2003) Lanthanide-binding tags as versatile protein coexpression probes. Chembiochem 4:265–71.

Furter R (1998) Expansion of the genetic code: site-directed *p*-fluoro-phenylalanine incorporation in *Escherichia coli*. Protein Sci 7:419–426.

Ghose R and Prestegard JH (1997) Electron spin-nuclear spin cross-correlation effects on multiplet splittings in paramagnetic proteins. J Magn Reson 128:138–143.

Gibson FS, Bergmeier SC and Rapoport H (1994) Selective removal of an N-Boc protecting group in the presence of a *tert*-butyl ester and other acid-sensitive groups. J Org Chem 59:3216–3218.

Graham B, Loh CT, Swarbrick JD, Ung P, Shin J, Yagi H, Jia X, Chhabra S, Barlow N, Pintacuda G, Huber T and Otting G (2011) DOTA-amide lanthanide tag for reliable generation of pseudocontact shifts in protein NMR spectra. Bioconjugate Chem 22:2118–2125.

Griffin RG (1998) Dipolar recoupling in MAS spectra of biological solids. Nat Struct Biol 5 Suppl:508–512.

Güntert P, Mumenthaler C and Wüthrich K (1997) Torsion angle dynamics for NMR structure calculation with the new program DYANA. J Mol Biol 273:283–298.

Hess B and Scheek RM (2003) Orientation restraints in molecular dynamics simulations using time and ensemble averaging. J Magn Reson 164:19–27.

Hiller S, Fiorito F, Wüthrich K and Wider G (2005) Automated projection spectroscopy (APSY). Proc Natl Acad Sci USA 102:10876–10881.

Hiller S and Wider G (2011) Automated projection spectroscopy and its applications. Top Curr Chem 316:21–47.

Hoesl MG and Budisa N (2012) Recent advances in genetic code engineering in *Escherichia coli*. Curr Opin Biotechnol 23:751–757.

Hong SH, Ntai I, Haimovich AD, Kelleher NL, Isaacs FJ and Jewett MC (2014) Cell-free protein synthesis from a release factor 1 deficient *Escherichia coli* activates efficient and multiple site-specific nonstandard amino acid incorporation. ACS Synth Biol 3:398–409.

Hyberts SG, Arthanari H and Wagner G (2012) Applications of non-uniform sampling and processing. Top Curr Chem 316:125–148.

Idehara T, Khutoryan EM, Tatematsu Y, Yamaguchi Y, Kuleshov AN, Dumbrajs O, Matsuki Y and Fujiwara T (2015) High-speed frequency modulation of a 460-GHz gyrotron for enhancement of 700-MHz DNP-NMR spectroscopy. J Infrared Millim Te 36:819–829.

Jackson JC, Hammill JT and Mehl RA (2007) Site-specific incorporation of a ^{19}F-amino acid into proteins as an NMR probe for characterizing protein structure and reactivity. J Am Chem Soc 129:1160–1166.

John M, Headlam MJ, Dixon NE and Otting G (2007) Assignment of paramagnetic ^{15}N-HSQC spectra by heteronuclear exchange spectroscopy. J Biomol NMR 37:43–51.

John M, Park AY, Pintacuda G, Dixon NE and Otting G (2005) Weak alignment of paramagnetic proteins warrants correction for residual CSA effects in measurements of pseudocontact shifts. J Am Chem Soc 127:17190–17191.

John M, Pintacuda G, Park AY, Dixon NE and Otting G (2006) Structure determination of protein-ligand complexes by transferred paramagnetic shifts. J Am Chem Soc 128:12910–12916.

Johnson DBF, Xu J, Shen Z, Takimoto JK, Schultz MD, Schmitz RJ, Xiang Z, Ecker JR, Briggs SP and Wang L (2011) RF1 knockout allows ribosomal incorporation of unnatural amino acids at multiple sites. Nat Chem Biol 7:779–786.

Jones DH, Cellitti SE, Hao X, Zhang Q, Jahnz M, Summerer D, Schultz PG, Uno T and Geierstanger BH (2010) Site-specific labeling of proteins with NMR-active unnatural amino acids. J Biomol NMR 46:89–100.

Kainosho M, Torizawa T, Iwashita Y, Terauchi T, Ono AM and Güntert P (2006) Optimal isotope labelling for NMR protein structure determinations. Nature 440:52–57.

Kay LE (2011) Solution NMR spectroscopy of supra-molecular systems, why bother? A methyl-TROSY view. J Magn Reson 210:159–170.

Kigawa T, Muto Y and Yokoyama S (1995) Cell-free synthesis and amino acid-selective stable isotope labeling of proteins for NMR analysis. J Biomol NMR 6:129–134.

Kigawa T, Yabuki T, Matsuda N, Matsuda T, Nakajima R, Tanaka A and Yokoyama S (2004) Preparation of *Escherichia coli* cell extract for highly productive cell-free protein expression. J Struct Funct Genomics 5:63–68.

Kigawa T, Yabuki T, Yoshida Y, Tsutsui M, Ito Y, Shibata T and Yokoyama S (1999) Cell-free production and stable-isotope labeling of milligram quantities of proteins. FEBS Lett 442:15–19.

Kigawa T and Yokoyama S (1991) A continuous cell-free protein synthesis system for coupled transcription-translation. J Biochem 110:166–168.

Kim D-M and Choi C-Y (1996) A semicontinuous prokaryotic coupled transcription/translation system using a dialysis membrane. Biotechnol Prog 12:645–649.

Kim DM, Kigawa T, Choi CY and Yokoyama S (1996) A highly efficient cell-free protein synthesis system from *Escherichia coli*. Eur J Biochem 239:881–886.

Kim H-Y and Wyss DF (2015) NMR screening in fragment-based drug design: a practical guide. Chem Biol Methods Protoc 197–208.

Kim YM, Gayen S, Joy J, Huang Q, Shuyi A, Liang J, Wee K, Jin M, Ang Y, Lim HA, Alvin W, Li R, Noble CG, Tian L, Yip A, Wang Q, San C, Chia B, Hill J, Shi P, Keller TH, Kang C, Chen AS, Hung AW and Lee LT (2013) NMR analysis of a novel enzymatically active unlinked dengue NS2B-NS3 protease complex. J Biol Chem 288:12891–12900.

Kobashigawa Y, Saio T, Ushio M, Sekiguchi M, Yokochi M, Ogura K and Inagaki F (2012) Convenient method for resolving degeneracies due to symmetry of the

magnetic susceptibility tensor and its application to pseudocontact shift-based protein-protein complex structure determination. J Biomol NMR 53:53–63.

Koh-Stenta X, Joy J, Wang SF, Shovanlal G, Chen AS, Kang CB, Lee MA, Poulsen A, Vasudevan SG, Hill J and Nacro K (2015) Identification of covalent active site inhibitors of dengue virus protease. Drug Des Devel Ther 9:6389–6399.

Kuntz ID (1992) Structure-based strategies for drug design and discovery. Science 257:1078–1082.

Kurland CG (1982) Translational accuracy in vitro. Cell 28:201–202.

Kwong AD, McNair L, Jacobson I and George S (2008) Recent progress in the development of selected hepatitis C virus NS3·4A protease and NS5B polymerase inhibitors. Curr Opin Pharmacol 8:522–531.

Lajoie MJ, Rovner AJ, Goodman DB, Aerni H, Haimovich AD, Kuznetsov G, Mercer JA, Wang HH, Carr PA, Mosberg J, Rohland N, Schultz PG, Jacobson JM, Rinehart J, Church GM and Isaacs FJ (2013) Genomically recoded organisms expand biological functions. Science 342:357–360.

Leung D, Schroder K, White H, Fang NX, Stoermer MJ, Abbenante G, Martin JL, Young PR and Fairlie DP (2001) 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.

Li F, Shi P, Li J, Yang F, Wang T, Zhang W, Gao F, Ding W, Li D, Li J, Xiong Y, Sun J, Gong W, Tian C and Wang J (2013) A genetically encoded ^{19}F NMR probe for tyrosine phosphorylation. Angew Chemie Int Ed 3958–3962.

Li J, Lawton TJ, Kostecki JS, Nisthal A, Fang J, Mayo SL, Rosenzweig AC and Jewett MC (2015) Cell-free protein synthesis enables high yielding synthesis of an active multicopper oxidase. Biotechnol J 1–7.

Li J, Lim SP, Beer D, Patel V, Wen D, Tumanut C, Tully DC, Williams JA, Jiricek J, Priestle JP, Harris JL and Vasudevan SG (2005) Functional profiling of recombinant NS3 proteases from all four serotypes of dengue virus using tetrapeptide and octapeptide substrate libraries. J Biol Chem 280:28766–28774.

Lim SI, Hahn YS and Kwon I (2015) Site-specific albumination of a therapeutic protein with multi-subunit to prolong activity in vivo. J Control Rel 207:93–100.

Lim SP, Wang QY, Noble CG, Chen YL, Dong H, Zou B, Yokokawa F, Nilar S, Smith P, Beer D, Lescar J and Shi PY (2013) Ten years of dengue drug discovery:

Progress and prospects. Antiviral Res 100:500–519.

Lin MZ and Wang L (2008) Selective labeling of proteins with chemical probes in living cells. Physiology 23:131–141.

Lin Y, Ahn S, Murali N, Brey W, Bowers CR and Warren WS (2000) High-resolution, >1 GHz NMR in unstable magnetic fields. Phys Rev Lett 85:3732–3735.

Liu CC and Schultz PG (2010) Adding new chemistries to the genetic code. Annu Rev Biochem 79:413–444.

Liu W, Brock A, Chen S and Schultz PG (2007) Genetic incorporation of unnatural amino acids into proteins in mammalian cells. Nat Methods 4:239–244.

Liu WM, Overhand M and Ubbink M (2014) The application of paramagnetic lanthanoid ions in NMR spectroscopy on proteins. Coord Chem Rev 273:2–12.

Loh CT, Ozawa K, Tuck KL, Barlow N, Huber T, Otting G and Graham B (2013) Lanthanide tags for site-specific ligation to an unnatural amino acid and generation of pseudocontact shifts in proteins. Bioconjugate Chem 24:260–268.

Loscha K V, Herlt AJ, Qi R, Huber T, Ozawa K and Otting G (2012) Multiple-site labeling of proteins with unnatural amino acids. Angew Chemie Int Ed 51:2243–2246.

Ma C and Opella SJ (2000) Lanthanide ions bind specifically to an added “EF-hand” and orient a membrane protein in micelles for solution NMR spectroscopy. J Magn Reson 146:381–384.

Maeno A, Sindhikara D, Hirata F, Otten R, Dahlquist FW, Yokoyama S, Akasaka K, Mulder FAA and Kitahara R (2015) Cavity as a source of conformational fluctuation and high-energy state: high-pressure NMR study of a cavity-enlarged mutant of T4 lysozyme. Biophys J 108:133–145.

Maly T, Debelouchina GT, Bajaj VS, Hu K-N, Joo C-G, Mak–Jurkauskas ML, Sirigiri JR, van der Wel PCA, Herzfeld J, Temkin RJ and Griffin RG (2008) Dynamic nuclear polarization at high magnetic fields. J Chem Phys 128:052211.

Man B, Su X-C, Liang H, Simonsen S, Huber T, Messerle BA and Otting G (2010) 3-Mercapto-2,6-pyridinedicarboxylic acid: a small lanthanide-binding tag for protein studies by NMR spectroscopy. Chem - A Eur J 16:3827–32.

Markiewicz WD, Bonney EA, Dixon IR, Eyssa YM and Swenson CA (1996) Technology of 1 GHz NMR superconducting magnets. Proc Int Work Adv High Magn Fields 216:200–202.

Markiewicz WD, Miller JR, Schwartz J, Trociewitz UP and Weijers HW (2006) Perspective on a superconducting 30 T/1.3 GHz NMR spectrometer magnet. Appl Supercond IEEE Trans 16:1523–1526.

Medek A, Hajduk PJ, Mack J and Fesik SW (2000) The use of differential chemical shifts for determining the binding site location and orientation of protein-bound ligands. J Am Chem Soc 122:1241–1242.

Medek A, Harwood JS and Frydman L (1995) Multiple-quantum magic-angle spinning NMR: a new method for the study of quadrupolar nuclei in solids. J Am Chem Soc 117:12779–12787.

Meyer B and Peters T (2003) NMR spectroscopy techniques for screening and identifying ligand binding to protein receptors. Angew Chemie Int Ed 42:864–890.

Mobli M and Hoch JC (2014) Nonuniform sampling and non-Fourier signal processing methods in multidimensional NMR. Prog Nucl Magn Reson Spectrosc 83:21–41.

Mukai T, Hayashi A, Iraha F, Sato A, Ohtake K, Yokoyama S and Sakamoto K (2010) Codon reassignment in the *Escherichia coli* genetic code. Nucleic Acids Res 38:8188–8195.

Muñoz-Jordán JL, Laurent-Rolle M, Ashour J, Martínez-Sobrido L, Ashok M, Lipkin WI and García-Sastre A (2005) Inhibition of α/β interferon signaling by the NS4B protein of flaviviruses inhibition of α/β interferon signaling by the NS4B protein of flaviviruses. J Virol 79:8004–8013.

Neumann H, Peak-Chew SY and Chin JW (2008) Genetically encoding N^{ϵ}-acetyllysine in recombinant proteins. Nat Chem Biol 4:232–234.

Nirenberg MW and Matthaei JHH (1961) The dependence of cell-free protein synthesis in E. coli upon naturally occurring or synthetic polyribonucleotides. Proc Natl Acad Sci USA 47:1588–1602.

Nitsche C, Holloway S, Schirmeister T and Klein CD (2014) Biochemistry and medicinal chemistry of the dengue virus protease. Chem Rev 114:11348–11381.

Nitz M, Franz KJ, Maglathlin RL and Imperiali B (2003) A powerful combinatorial screen to identify high-affinity terbium(III)-binding peptides. Chembiochem 4:272–276.

Nitz M, Sherawat M, Franz KJ, Peisach E, Allen KN and Imperiali B (2004) Structural origin of the high affinity of a chemically evolved lanthanide-binding

peptide. Angew Chemie Int Ed 43:3682–3685.

Noble CG, Seh CC, Chao AT and Shi PY (2012) Ligand-bound structures of the dengue virus protease reveal the active conformation. J Virol 86:438–46.

Nucci N, Fuglestad B, Athanasoula EA and Wand AJ (2014) Role of cavities and hydration in the pressure unfolding of T4 lysozyme. Proc Natl Acad Sci USA 111:13846–13851.

Okude J, Ueda T, Kofuku Y, Sato M, Nobuyama N, Kondo K, Shiraishi Y, Mizumura T, Onishi K, Natsume M and Maeda M (2015) Identification of a conformational equilibrium that determines the efficacy and functional selectivity of the μ-opioid receptor. Angew Chemie Int Ed 127:15997–16002.

Otting G (2010) Protein NMR using paramagnetic ions. Annu Rev Biophys 39:387–405.

Otting G (2008) Prospects for lanthanides in structural biology by NMR. J Biomol NMR 42:1–9.

Ozawa K, Headlam MJ, Schaeffer PM, Henderson BR, Dixon NE and Otting G (2004) Optimization of an *Escherichia coli* system for cell-free synthesis of selectively ^{15}N-labelled proteins for rapid analysis by NMR spectroscopy. Eur J Biochem 271:4084–4093.

Pavlov MY and Ehrenberg M (1996) Rate of translation of natural mRNAs in an optimized in vitro system. Arch Biochem Biophys 328:9–16.

Pellecchia M, Sem DS and Wüthrich K (2002) NMR in drug discovery. Nat Rev Drug Discov 1:211–219.

Pfaff SJ, Chimenti MS, Kelly MJS and Arkin MR (2015) Biophysical methods for identifying fragment-based inhibitors of protein-protein interactions. Methods Mol Biol 1278:587–613.

Pilla KB, Leman JK, Otting G and Huber T (2015) Capturing conformational states in proteins using sparse paramagnetic NMR data. PLoS ONE 10:e0127053.

Pintacuda G, Keniry MA, Huber T, Park AY, Dixon NE and Otting G (2004) Fast structure-based assignment of ^{15}N HSQC spectra of selectively ^{15}N-labeled paramagnetic proteins. J Am Chem Soc 126:2963–2970.

Pirman NL, Barber KW, Aerni HR, Ma NJ, Haimovich AD, Rogulina S, Isaacs FJ and Rinehart J (2015) A flexible codon in genomically recoded *Escherichia coli* permits programmable protein phosphorylation. Nat Commun 6:8130.

Prusis P, Junaid M, Petrovska R, Yahorava S, Yahorau A, Katzenmeier G, Lapins M and Wikberg JES (2013) Design and evaluation of substrate-based octapeptide and non substrate-based tetrapeptide inhibitors of dengue virus NS2B-NS3 proteases. Biochem Biophys Res Commun 434:767–772.

Riek R, Fiaux J, Bertelsen EB, Horwich AL and Wüthrich K (2002) Solution NMR techniques for large molecular and supramolecular structures. J Am Chem Soc 124:12144–12153.

Rochdi A, Foucat L and Renou J-P (2000) NMR and DSC studies during thermal denaturation of collagen. Food Chem 69:295–299.

Rovnyak D, Frueh DP, Sastry M, Sun ZYJ, Stern AS, Hoch JC and Wagner G (2004) Accelerated acquisition of high resolution triple-resonance spectra using non-uniform sampling and maximum entropy reconstruction. J Magn Reson 170:15–21.

Ruschak AM and Kay LE (2012) Proteasome allostery as a population shift between interchanging conformers. Proc Natl Acad Sci USA 109:E3454–E3462.

Russo L, Maestre-Martinez M, Wolff S, Becker S and Griesinger C (2013) Interdomain dynamics explored by paramagnetic NMR. J Am Chem Soc 135:17111–17120.

Sabchareon A, Wallace D, Sirivichayakul C, Limkittikul K, Chanthavanich P, Suvannadabba S, Jiwariyavej V, Dulyachai W, Pengsaa K, Wartel TA, Moureau A, Saville M, Bouckenooghe A, Viviani S, Tornieporth NG and Lang J (2012) Protective efficacy of the recombinant, live-attenuated, CYD tetravalent dengue vaccine in Thai schoolchildren: a randomised, controlled phase 2b trial. Lancet 380:1559–1567.

Saio T, Ogura K, Kumeta H, Kobashigawa Y, Shimizu K, Yokochi M, Kodama K, Yamaguchi H, Tsujishita H and Inagaki F (2015) Ligand-driven conformational changes of MurD visualized by paramagnetic NMR. Sci Rep 5:16685.

Saio T, Ogura K, Yokochi M, Kobashigawa Y and Inagaki F (2009) Two-point anchoring of a lanthanide-binding peptide to a target protein enhances the paramagnetic anisotropic effect. J Biomol NMR 44:157–166.

Saio T, Yokochi M, Kumeta H and Inagaki F (2010) PCS-based structure determination of protein–protein complexes. J Biomol NMR 46:271–280.

Schmidt MJ, Borbas J, Drescher M and Summerer D (2014) A genetically encoded spin label for electron paramagnetic resonance distance measurements. J Am Chem Soc 136:1238–1241.

Schmitz C, John M, Park AY, Dixon NE, Otting G, Pintacuda G and Huber T (2006) Efficient χ-tensor determination and NH assignment of paramagnetic proteins. J Biomol NMR 35:79–87.

Schmitz C, Stanton-Cook MJ, Su X-C, Otting G and Huber T (2008) Numbat: an interactive software tool for fitting Δχ-tensors to molecular coordinates using pseudocontact shifts. J Biomol NMR 41:179–189.

Schmitz C, Vernon R, Otting G, Baker D and Huber T (2012) Protein structure determination from pseudocontact shifts using ROSETTA. J Mol Biol 416:668–677.

Schwieters CD, Kuszewski JJ and Clore GM (2006) Using Xplor-NIH for NMR molecular structure determination. Prog Nucl Magn Reson Spectrosc 48:47–62.

Schwieters CD, Kuszewski JJ, Tjandra N and Clore GM (2003) The Xplor-NIH NMR molecular structure determination package. J Magn Reson 160:65–73.

Scola PM, Sun L-Q, Wang AX, Chen J, Sin N, Venables BL, Sit S-Y, Chen Y, Cocuzza A, Bilder DM, D'Andrea S V, Zheng B, Hewawasam P, Tu Y, Friborg J, Falk P, Hernandez D, Levine S, Chen C, Yu F, Sheaffer AK, Zhai G, Barry D, Knipe JO, Han Y-H, Schartman R, Donoso M, Mosure K, Sinz MW, Zvyaga T, Good AC, Rajamani R, Kish K, Tredup J, Klei HE, Gao Q, Mueller L, Colonno RJ, Grasela DM, Adams SP, Loy J, Levesque PC, Sun H, Shi H, Sun L, Warner W, Li D, Zhu J, Meanwell NA and McPhee F (2014) The discovery of asunaprevir (BMS-650032), an orally efficacious NS3 protease inhibitor for the treatment of hepatitis C virus infection. J Med Chem 57:1730–52.

Sekhar A and Kay LE (2013) NMR paves the way for atomic level descriptions of sparsely populated, transiently formed biomolecular conformers. Proc Natl Acad Sci USA 110:12867–12874.

Shen Y, Lange O, Delaglio F, Rossi P, Aramini JM, Liu G, Eletsky A, Wu Y, Singarapu KK, Lemak A, Ignatchenko A, Arrowsmith CH, Szyperski T, Montelione GT, Baker D and Bax A (2008) Consistent blind protein structure generation from NMR chemical shift data. Proc Natl Acad Sci USA 105:4685–4690.

Shi P, Wang H, Xi Z, Shi C, Xiong Y and Tian C (2011) Site-specific ^{19}F NMR chemical shift and side chain relaxation analysis of a membrane protein labeled with an unnatural amino acid. Protein Sci 20:224–228.

Shimizu Y, Inoue A, Tomari Y, Suzuki T, Yokogawa T, Nishikawa K and Ueda T (2001) Cell-free translation reconstituted with purified components. Nat Biotechnol

19:751–755.

Shishmarev D and Otting G (2013) How reliable are pseudocontact shifts induced in proteins and ligands by mobile paramagnetic metal tags? A modelling study. J Biomol NMR 56:203–216.

Shuker SB, Hajduk PJ, Meadows RP and Fesik SW (1996) Discovering high-affinity ligands for proteins: SAR by NMR. Science 274:1531–1534.

Skora L, Mestan J, Fabbro D, Jahnke W and Grzesiek S (2013) NMR reveals the allosteric opening and closing of Abelson tyrosine kinase by ATP-site and myristoyl pocket inhibitors. Proc Natl Acad Sci USA 110:E4437–E4445.

Spirin AS, Baranov VI, Ryabova LA, Ovodov SY and Alakhov YB (1988) A continuous cell-free translation system capable of producing polypeptides in high yield. Science 242:1162–1164.

Sprangers R and Kay LE (2007) Quantitative dynamics and binding studies of the 20S proteasome by NMR. Nature 445:618–622.

Su X-C, Loh C-T, Qi R and Otting G (2011) Suppression of isotope scrambling in cell-free protein synthesis by broadband inhibition of PLP enymes for selective ^{15}N-labelling and production of perdeuterated proteins in H_2O. J Biomol NMR 50:35–42.

Su X-C, Ozawa K, Qi R, Vasudevan SG, Lim SP and Otting G (2009a) NMR analysis of the dynamic exchange of the NS2B cofactor between open and closed conformations of the West Nile virus NS2B-NS3 protease. PLoS Negl Trop Dis 3:e561.

Su XC, Liang H, Loscha K V. and Otting G (2009b) $[Ln(DPA)_3]^{3-}$ is a convenient paramagnetic shift reagent for protein NMR studies. J Am Chem Soc 131:10352–10353.

Su XC, Ozawa K, Yagi H, Lim SP, Wen D, Ekonomiuk D, Huang D, Keller TH, Sonntag S, Caflisch A, Vasudevan SG and Otting G (2009c) NMR study of complexes between low molecular mass inhibitors and the West Nile virus NS2B-NS3 protease. FEBS J 276:4244–4255.

Su XC and Otting G (2010) Paramagnetic labelling of proteins and oligonucleotides for NMR. J Biomol NMR 46:101–112.

Sullivan CJ, Pendleton ED, Sasmor HH, Hicks WL, Farnum JB, Muto M, Amendt EM, Schoborg JA, Martin RW, Clark LG, Anderson MJ, Choudhury A, Fior R, Lo YH, Griffey RH, Chappell SA, Jewett MC, Mauro VP and Dresios J (2015) A cell-

free expression and purification process for rapid production of protein biologics. Biotechnol J 1–11.

Tang C, Schwieters CD and Clore GM (2007) Open-to-closed transition in apo maltose-binding protein observed by paramagnetic NMR. Nature 449:1078–1082.

Tolman JR, Flanagan JM, Kennedy MA and Prestegard JH (1995) Nuclear magnetic dipole interactions in field-oriented proteins: information for structure determination in solution. Proc Natl Acad Sci USA 92:9279–9283.

Tugarinov V, Kanelis V and Kay LE (2006) Isotope labeling strategies for the study of high-molecular-weight proteins by solution NMR spectroscopy. Nat Protoc 1:749–754.

Ubbink M, Worrall JAR, Canters GW, Groenen EJJ and Huber M (2002) Paramagnetic resonance of biological metal centers. Anu Rev 31:393–422.

Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE and Berendsen HJC (2005) GROMACS: fast, flexible, and free. J Comput Chem 26:1701–1718.

Vaughn DW, Green S, Kalayanarooj S, Innis BL, Nimmannitya S, Suntayakorn S, Endy TP, Raengsakulrach B, Rothman AL, Ennis FA and Nisalak A (2000) Dengue viremia titer, antibody response pattern, and virus serotype correlate with disease severity. J Infect Dis 181:2–9.

Vega AJ and Fiat D (1976) Nuclear relaxation processes of paramagnetic complexes. The slow-motion case. Mol Phys 31:347–355.

Viegas A, Manso J, Nobrega FL and Cabrita EJ (2011) Saturation-transfer difference (STD) NMR: a simple and fast method for ligand screening and characterization of protein binding. J Chem Educ 88:990–994.

Villar L, Dayan GH, Arredondo-García JL, Rivera DM, Cunha R, Deseda C, Reynales H, Costa MS, Morales-Ramírez JO, Carrasquilla G, Rey LC, Dietze R, Luz K, Rivas E, Montoya MCM, Supelano MC, Zambrano B, Langevin E, Boaz M, Tornieporth N, Saville M and Noriega F (2014) Efficacy of a tetravalent dengue vaccine in children in Latin America. N Engl J Med 372:113–123.

Wagner G (1993) Prospects for NMR of large proteins. J Biomol NMR 3:375–85.

Wang L and Schultz PG (2001) A general approach for the generation of orthogonal tRNAs. Chem Biol 8:883–890.

Wilkins BJ, Hahn LE, Heitmüller S, Frauendorf H, Valerius O, Braus GH and Neumann H (2015) Genetically encoding lysine modifications on histone H4. ACS

Chem Biol 10:939–944.

Wöhnert J, Franz KJ, Nitz M, Imperiali B and Schwalbe H (2003) Protein alignment by a coexpressed lanthanide-binding tag for the measurement of residual dipolar couplings. J Am Chem Soc 125:13338–13339.

Wu PSC, Ozawa K, Lim SP, Vasudevan SG, Dixon NE and Otting G (2007) Cell-free transcription/translation from PCR-amplified DNA for high-throughput NMR studies. Angew Chemie Int Ed 46:3356–3358.

Wüthrich K (1990) Protein structure determination in solution by NMR spectroscopy. J Biol Chem 265:22059–22062.

Yagi H, Pilla KB, Maleckis A, Graham B, Huber T and Otting G (2013) Three-dimensional protein fold determination from backbone amide pseudocontact shifts generated by lanthanide tags at multiple sites. Structure 21:883–890.

Yamazaki T, Otomo T, Oda N, Kyogoku Y, Uegaki K, Ito N, Ishino Y and Nakamura H (1998) Segmental isotope labeling for protein NMR using peptide splicing. J Am Chem Soc 120:5591–5592.

Yang F, Yu X, Liu C, Qu C-X, Zheng G, Liu H-D, Li F-H, Wang H-M, He D-F, Yi F, Song C, Tian C-L, Xiao K-H, Wang J-Y and Sun J-P (2015) Phospho-selective mechanisms of arrestin conformations and functions revealed by unnatural amino acid incorporation and ^{19}F-NMR. Nat Commun 6:8202.

Yin Z, Patel SJ, Wang W-L, Wang G, Chan W-L, Rao KRR, Alam J, Jeyaraj D a, Ngew X, Patel V, Beer D, Lim SP, Vasudevan SG and Keller TH (2006) Peptide inhibitors of dengue virus NS3 protease. Part 1: warhead. Bioorg Med Chem Lett 16:36–39.

Yusof R, Clum S, Wetzel M, Murthy HK and Padmanabhan RJ (2000) 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.

Zong X-H, Zhou P, Shao Z-Z, Chen S-M, Chen X, Hu B-W, Deng F and Yao W-H (2004) Effect of pH and copper (II) on the conformation transitions of silk fibroin based on EPR, NMR, and Raman spectroscopy. Biochemistry 43:11932–11941.

the FEBS Journal

EDITOR'S CHOICE

Binding mode of the activity-modulating C-terminal segment of NS2B to NS3 in the dengue virus NS2B–NS3 protease

Laura de la Cruz[1], Wan-Na Chen[1], Bim Graham[2] and Gottfried Otting[1]

1 Research School of Chemistry, Australian National University, Canberra, Australia
2 Medicinal Chemistry and Drug Action, Monash Institute of Pharmaceutical Sciences, Parkville, Australia

Keywords
dengue virus; NMR spectroscopy; NS2B–NS3 protease; paramagnetic tags; pseudocontact shifts

Correspondence
G. Otting, Research School of Chemistry, Australian National University, Canberra, ACT 0200, Australia
Fax: +61 2 6125 0750
Tel: +61 2 6125 6508
E-mail: gottfried.otting@anu.edu.au

(Received 27 November 2013, revised 21 January 2014, accepted 24 January 2014)

doi:10.1111/febs.12729

The two-component dengue virus NS2B–NS3 protease (NS2B–NS3pro) is an established drug target but inhibitor design is hampered by uncertainties about its 3D structure in solution. Crystal structures reported very different conformations for the functionally important C-terminal segment of the NS2B cofactor (NS2Bc), indicating open and closed conformations in the absence and presence of inhibitors, respectively. An earlier NMR study in solution indicated that a closed state is the preferred conformation in the absence of an artificial linker engineered between NS2B and NS3pro. To obtain direct structural information on the fold of unlinked NS2B–NS3pro in solution, we tagged NS3pro with paramagnetic tags and measured pseudocontact shifts by NMR to position NS2Bc relative to NS3pro. NS2Bc was found to bind to NS3pro in the same way as reported in a previously published model and crystal structure of the closed state. The structure is destabilized, however, by high ionic strength and basic pH, showing the importance of electrostatic forces to tie NS2Bc to NS3pro. Narrow NMR signals previously thought to represent the open state are associated with protein degradation. In conclusion, the closed conformation of the NS2B–NS3 protease is the best model for structure-guided drug design.

Introduction

Dengue virus is a health burden that sends almost 100 million people to clinics every year [1] with an estimated 40% of the world's population at risk, especially in the tropical and subtropical regions [2,3]. The infection involves translation of the flavivirus RNA genome into a poly-protein that must be cleaved into several individual components. The N-terminal part of the nonstructural protein 3 (NS3) encodes a serine protease, NS3pro, which is required for cleavage of the poly-protein. This makes NS3pro a prime drug target [4]. Its activity towards peptide substrates is boosted ~3300–7600-fold by the presence of a second viral protein, NS2B, which comprises ~40 residues [5]. The NS2B–NS3 proteases of dengue serotypes 1–4 are closely related to each other (Fig. S1) and to the NS2B–NS3 protease from the West Nile virus (WNV; ~40% sequence identity). This study focuses on the NS2B–NS3 protease of dengue virus serotype 2, which is the most studied serotype.

The wild-type poly-protein contains a recognition site for the NS2B–NS3 protease between NS2B and NS3 [6]. To stabilize the protein, the most widely used construct of the NS2B–NS3 protease contains a 47-residue segment of NS2B fused to the N-terminus of

Abbreviations
NS3, nonstructural protein 3; PCS, pseudocontact shift; PRE, paramagnetic relaxation enhancement; WNV, West Nile virus.

NS3pro via an artificial Gly_4–Ser–Gly_4 linker. This and similar constructs retain the activity of the protease [7] and have been used for crystal structure determinations [8,9], NMR spectroscopic studies [10,11] and ligand-binding studies [12–28]. In the following, we refer to this construct as DENpro (WNVpro for the corresponding construct from WNV). We use the name DENp for an alternative construct that is devoid of the covalent linkage between NS2B and NS3pro. The distinction is important, because the functionally important NS2B cofactor has been reported to assume very different conformations in DENpro and DENp (Fig. 1).

The crystal structures of DENpro and WNVpro in the absence of inhibitors revealed an open conformation [8,29], in which the N-terminal part of the NS2B cofactor inserts into the N-terminal β-barrel of the NS3 protease, but the C-terminal segment of NS2B (NS2Bc, corresponding to the segment Glu66-Leu95 in DENpro serotype 2) is located far from the substrate-binding site of NS3pro. We refer to the structures with NS2Bc far from the active site of NS3pro as the 'open state' (Fig. 1A). By contrast, structures determined in the presence of inhibitors displayed NS2Bc forming a β-hairpin that lines the substrate-binding site [8,29–32]. We refer to this conformation as the 'closed state' (Fig. 1B). In view of the importance of NS2Bc for full proteolytic activity of the proteases [5,33–37], the closed conformation is assumed to represent the enzymatically active structure. As expected for homologous proteins, the closed conformations reported for DENpro and WNVpro in the presence of inhibitors are very similar [8,30]. Remarkably, the crystal structures of the open conformations without inhibitors are also very similar [8,29].

In solution, NMR experiments unambiguously showed that DENpro can assume the closed conformation in the presence of inhibitor **1** (Fig. S2A) [11]. In the absence of an inhibitor, however, not all signals could be observed, indicating extreme broadening of the NMR resonances of DENpro due to conformational exchange. A similar situation was observed in the case of WNVpro: the protein assumes the closed

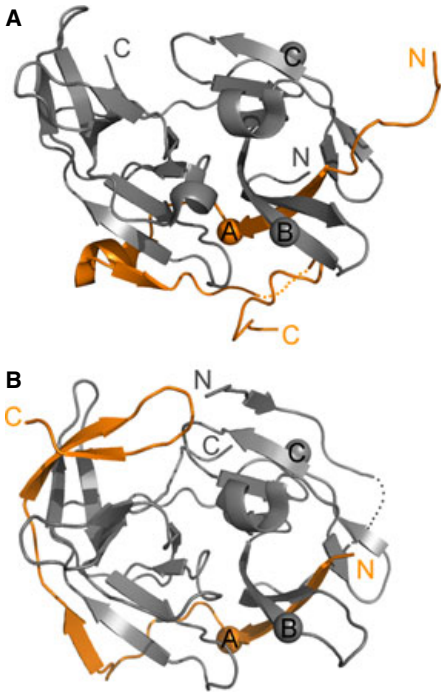


Fig. 1. Open and closed conformations observed in crystal structures of the NS2B–NS3 protease from dengue virus serotypes 2 and 3, respectively. In the orientation shown, the C- and N-terminal β-barrels of the protease are in the left and right halves, respectively. The parts from NS2B and NS3 are shown in orange and grey, respectively, and their N- and C-termini are labeled. Spheres identify the C^{α} atoms of the single cysteine residues for the attachment of lanthanide binding tags in the mutants A–C. Dotted lines indicate missing electron density in the crystal structure. (A) Open conformation (PDB ID 2FOM [8]). The C-terminal segment of NS2B, NS2Bc, is far from the active site. (B) Closed conformation (PDB ID 3U1I [30]). NS2Bc wraps around the active site.

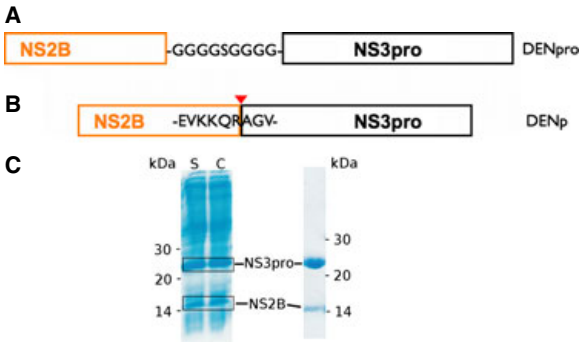


Fig. 2. Expression constructs of the dengue virus protease with and without covalent linker between NS2B and NS3. (A) Schematic representation of the protease construct with NS2B (orange) linked to NS3pro (black) via a flexible Gly_4–Ser–Gly_4 linker. We refer to this construct as DENpro. (B) Same as (A), except that the Gly_4–Ser–Gly_4 linker is replaced by the natural cleavage site between NS2B and NS3 in dengue virus serotype 2 (EVKKQR). We refer to this construct as DENp. The autoproteolytic cleavage site is identified by a red triangle. See Fig. S1 for the detailed amino acid sequence. (C) SDS/PAGE of DENp with Coomassie Brilliant Blue staining. (Left) Result from cell-free expression (S: supernatant, C: crude reaction mixture). (Right) Purified product from *in vivo* expression. In both cases, soluble DENp was obtained without covalent linkage between NS2B and NS3pro. The bands of NS2B and NS3pro are indicated.

conformation in the presence of inhibitors, but the disappearance of signals in the absence of inhibitors indicates conformational exchange. NMR experiments showed that, also without inhibitor, the closed conformation is more highly populated [38].

A recent study investigated a construct of DENp produced by coexpression of NS2B and NS3pro without a covalent link between both proteins [39]. Significant differences from earlier results obtained with DENpro were found. Importantly, no NMR signals were reported to be missing in the absence of inhibitors, allowing more complete resonance assignments and thus the measurement of secondary chemical shifts, ^{15}N-relaxation rates and paramagnetic relaxation enhancements (PREs) from a spin-label attached to NS2Bc. These data indicated that NS2B may exist in a conformation similar to the closed form of DENpro, but structural details were not obtained.

This study presents direct structural information for a related unlinked NS2B–NS3pro construct, using pseudocontact shifts induced by different paramagnetic lanthanide ions attached at different sites. The data fully support the structure of the closed state detected previously in the presence of inhibitors [11,30], but NS2Bc is found to undergo conformational exchange that becomes more pronounced at increased pH and ionic strength, without appreciably dissociating from NS3pro even under conditions of high salt. Evidence is presented that a set of very narrow signals at chemical shifts characteristic of random coil conformations previously attributed to open conformations [10,38] instead signifies degraded protein. The results show that the structure and dynamics of the dengue NS2B–NS3 protease are much more closely related to the corresponding protease from the West Nile virus than implied by previous analyses.

Results

Protein production

DENp was expressed as a single polypeptide chain to produce NS2B and NS3pro in an equimolar ratio. SDS/PAGE showed that the autoproteolytic cleavage site was quantitatively cleaved in the course of protein expression and purification, regardless of whether the protein was produced by cell-free synthesis or *in vivo* (Fig. 2). ESI-MS of DENp selectively labelled with ^{15}N-isoleucine showed two peaks of molecular mass 6670 and 20 753 Da (calculated: 6671 and 20 757 Da for NS2B and NS3pro, respectively). DENp was enzymatically active (Fig. S3).

DENp assumes the closed conformation

The ^{15}N-HSQC spectrum of DENp selectively labelled with ^{15}N-isoleucine (Fig. 3A) was strikingly similar to that of DENpro in the presence of inhibitor **1**

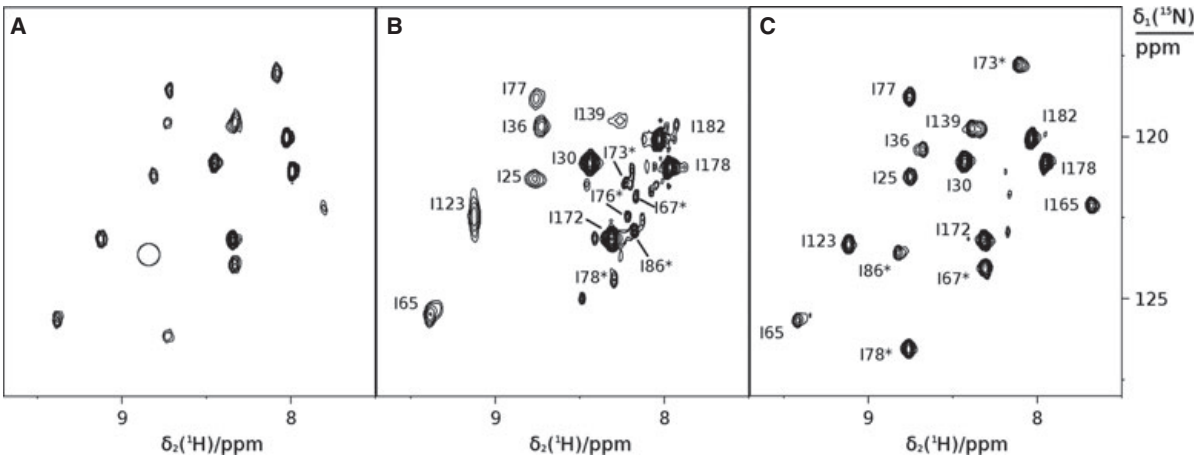


Fig. 3. Chemical shift conservation suggests that DENp predominantly assumes the closed conformation in the absence of an inhibitor. The spectra show the spectral region of the backbone amides in ^{15}N-HSQC spectra recorded of selectively ^{15}N-isoleucine-labelled samples of DENp and DENpro. All spectra were acquired using 0.2 mM protein solutions in 90% H_2O/10% D_2O with 20 mM MES, pH 6.5, 50 mM NaCl at 25 °C on a Bruker 800 MHz NMR spectrometer. Assignments of NS2B are identified by stars. (A) ^{15}N-HSQC spectrum of DENp. The circle indicates the position of the cross-peak of Ile86* which was observed with very weak intensity in a spectrum of a different sample. (B) As (A), except for DENpro. Weak narrow cross-peaks of NS2B appearing at random coil chemical shifts are from a minor species of variable intensity between different sample preparations. (C) As (B), except that a fivefold excess of inhibitor **1** was added to induce the closed conformation.

(Fig. 3C). The improved spectral quality of DENp is in line with earlier observations, in which a similar unlinked construct of the NS2B–NS3 protease was shown to produce better NMR spectra than the DENpro construct [39]. We previously established using pseudocontact shifts from multiple lanthanide binding sites [11], that the structure of DENpro with bound inhibitor **1** assumes the closed conformation, as observed in the crystal structure 3U1I [30]. The conservation of chemical shifts observed for the isoleucine residues in DENp without inhibitor (Fig. 3A) and DENpro with inhibitor **1** (Fig. 3C) thus indicates that DENp assumes the closed conformation too. Apart from small differences in chemical shifts expected to arise from the presence or absence of the inhibitor and the covalent linker peptide (e.g. Ile36, which is close to the active site and changes its chemical shift between DENp without inhibitor and DENpro with inhibitor **1**), there are pronounced differences in peak intensities. For example, the cross-peaks of Ile78* and Ile86* in NS2B and Ile165 in NS3pro were unexpectedly weak. Effects of this magnitude can only be explained by exchange broadening, which we attribute to conformational exchange of the β-hairpin of NS2Bc. Ile165 is buried in the C-terminal β-barrel of NS3pro which binds NS2Bc in the closed conformation. Sensitivity of the C-terminal β-barrel to the state of NS2Bc has been reported earlier [11].

Remarkably, greater spectral differences are observed between DENpro without inhibitor (Fig. 3B) and either DENpro with inhibitor **1** (Fig. 3C) or DENp without inhibitor (Fig. 3A). For example, the resonance of Ile123 is strongly affected in DENpro (Fig. 3B) and there are no signals of Ile165, Ile67* and Ile73* that would be expected for the closed conformation. Another notable feature of the spectra of DENpro (Fig. 3B,C) is a set of narrow signals for isoleucine residues of NS2B at ^{1}H chemical shifts characteristic of an unfolded polypeptide chain (8.1–8.3 ppm). In early NMR studies of DENpro, these signals were very intense and assigned to NS2B by systematic site-directed mutagenesis [10]. In the present spectra (Fig. 3B,C), however, these cross-peaks are much weaker than those of NS3pro. Having repeated these preparations and NMR measurements multiple times, we found the intensity of these NS2B signals to vary between different sample preparations and not to change when inhibitor **1** was added to the sample. Furthermore, previous experiments of selectively ^{15}N-Ser-labelled samples of DENpro without inhibitor displayed weak ^{15}N-HSQC cross-peaks of the serine residues of NS2B at the chemical shifts of the closed conformation, with little evidence of serine residues at random coil chemical shifts [11]. We conclude that the cross-peaks of isoleucine residues of NS2B at random coil chemical shifts stem from degradation of the NS2B–NS3pro complex, possibly resulting in free NS2B, which is known to be soluble, whereas NS3pro is insoluble. In agreement with this interpretation, the corresponding set of narrow peaks in selectively ^{15}N-isoleucine-labelled DENp grew noticeably in height after 4 h at 25 °C, in step with protein precipitation. This conclusion supersedes our earlier interpretation that the random coil signals of NS2B represent the open conformation of NS2B in the NS2B–NS3pro complex [11,38].

Discounting the signals from the degraded protease, there is no second set of isoleucine cross-peaks that would indicate the existence of the open conformation in either DENp or DENpro. Instead, the open conformation may be manifested in the apparent absence and weakness of some of the isoleucine cross-peaks in the ^{15}N-HSQC spectrum, which indicates excessive line broadening arising from conformational exchange. To rule out line broadening due to chemical exchange of the amide protons with the water, we performed a control experiment to assess the amide proton exchange rates. The experiment used selective excitation of the water resonance followed by a mixing time to allow magnetization transfer by chemical exchange followed by the ^{15}N-HSQC pulse sequence to record the amide cross-peaks arising from chemical exchange with water. The spectrum of selectively ^{15}N-Ile/^{15}N-Ser-labelled DENp was recorded. The scarcity of observable cross-peaks (Fig. 4C) confirmed that exchange with the water resonance was unimportant, even at pH 7.5, where exchange with water is more pronounced than at lower pH [40]. Weak cross-peak intensities in the ^{15}N-HSQC spectra of Fig. 3 are thus indicative of conformational rather than chemical exchange broadening.

Structural analysis of DENp by pseudocontact shifts

Although the overall conservation of chemical shifts between DENp and DENpro indicates structural conservation, we sought more direct evidence by measuring pseudocontact shifts (PCS) in DENp. To do this without mutations in the critical NS2Bc segment, we prepared three different single-cysteine mutants for labelling with paramagnetic C1- and C2-lanthanide tags. The mutation sites corresponded to the sites A, B and C (Ala57*Cys, Ser34Cys and Ser68Cys, respectively) used previously to establish the closed conformation of DENpro in the presence of the inhibitor **1** [11]. In the study of DENpro [11], on average, larger

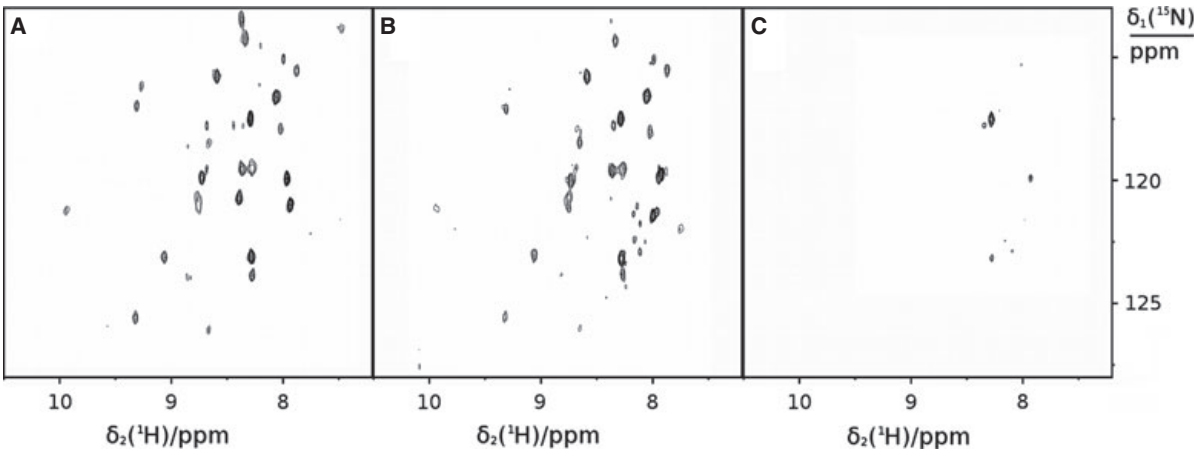


Fig. 4. Reduced peak intensities in the ^{15}N-HSQC spectra are caused by conformational, not chemical exchange. All spectra were recorded of selectively ^{15}N-Ser- and ^{15}N-Ile-labelled DENp, using the same sample condition and spectrometer as in Fig. 3. See Fig. S4 for resonance assignments. (A) ^{15}N-HSQC spectrum at pH 6.5. (B) As (A), except that the pH was 7.5. (C) ^{15}N-HSQC exchange spectrum at pH 7.5. The experiment was recorded with the CLEANEX-PM-FHSQC pulse sequence [62], in which selective water excitation by a 7.5 ms 180° pulse sequence was followed by 60 ms CLEANEX-PM mixing prior to the conventional HSQC pulse sequence. Cross-peaks in this spectrum arise from magnetization transfer from water to amide protons by proton exchange.

PCSs were obtained with the C2 tag than with the C1 tag. Therefore, we recorded data with the C1 tag only for mutant B, which delivered the largest number of PCSs for NS2B. For all three mutants, PCSs were measured with Tm^{3+} and Tb^{3+} tags. Figure 5 shows examples of PCSs measured for mutant B with the C2–Tb^{3+} tag in DENp without inhibitor and compares them with corresponding PCSs measured for DENpro with inhibitor **1**. Figure 6 shows an overview of all the PCSs that could be resolved for both DENp and DENpro with different tags. Clearly, the PCSs measured for DENp and DENpro were very similar throughout, including, in particular, resonances of NS2Bc. The PCSs thus provide very clear evidence that both NS3pro and NS2B predominantly assume the closed conformation in DENp. The model of the closed conformation of DENpro established by PCSs [11] is very similar to the crystal structure of the serotype 3 NS2B–NS3 protease determined in the presence of an inhibitor [30] with a backbone rmsd of 0.7 Å.

Conformational exchange of NS2B

DENp and DENpro display more intense ^{15}N-HSQC cross-peaks in the presence of the inhibitor **1**. Figure 7 shows a quantitative assessment of the relative peak heights of samples selectively labelled with ^{15}N-Ser and ^{15}N-Ile in the absence and presence of inhibitor. Clearly, signals from NS2B are more strongly affected than those of NS3, and DENpro suffers more from exchange broadening than DENp, including residues 123–165 of NS3pro. This observation reinforces the conclusion that the C-terminal β-barrel, which harbours those residues, senses the conformational exchange of NS2B. Notably, residual exchange broadening persists in DENp in the absence of inhibitor for NS2B and also for Ile139 in NS3.

Although it is common that binding of an inhibitor rigidifies a protein structure, it is remarkable that the low-molecular mass inhibitor used here achieves this for NS2B, because the inhibitor is too small to form direct contacts with NS2B. Based on mutation and MS data, the inhibitor hydrolyses upon binding, leaving behind only a *p*-guanidino-benzoyl moiety that is covalently bound to Ser135 [11]. We hypothesized that the positive charge of the guanidino group helps to tie NS2Bc to NS3, because the C-terminal β-hairpin of NS2B is highly negatively charged owing to the presence of Glu80 and Asp81 and four additional glutamates at the C-terminus of NS2B (Fig. S1). If electrostatic interactions are important, the association of NS2B to NS3 should be weakened by high salt concentrations. Furthermore, the association should be weakened at higher pH, when His51 in the active site loses its positive charge.

Figure 4 shows that the ^{15}N-HSQC cross-peaks of many residues became significantly weaker at pH 7.5 compared with pH 6.5 and that this loss in intensity was not due to fast proton exchange with the water. As expected, the effect was observed equally for

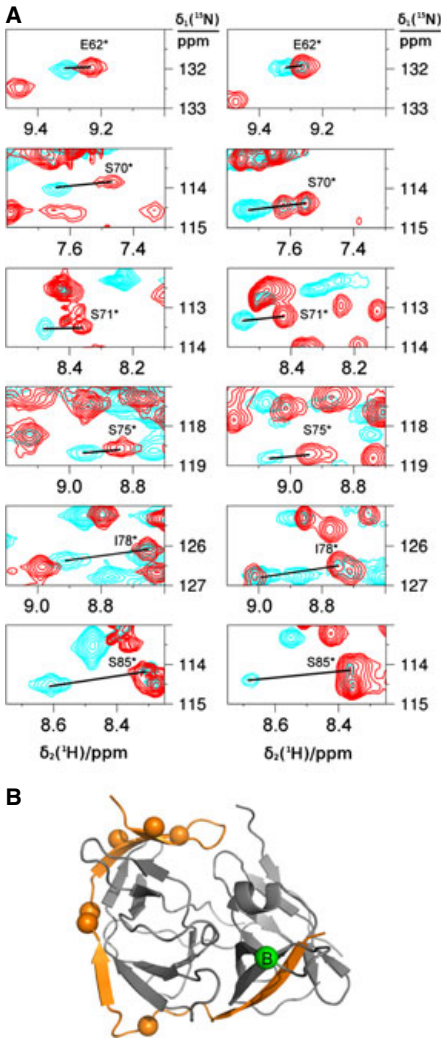


Fig. 5. Similar PCSs are observed for NS2B of DENp without inhibitor as for DENpro with inhibitor **1**. The spectra were recorded on a 600 MHz NMR spectrometer using the same conditions as in Fig. 3. (A) Superimpositions of selected spectral regions of ^{15}N-HSQC spectra of mutant B of DENp without inhibitor (left) and DENpro with inhibitor **1** (right). The samples were tagged with the C2 tag loaded with paramagnetic Tb^{3+} (cyan peaks) or diamagnetic Y^{3+} (red peaks). (B) Cartoon representation of the dengue virus protease in the closed conformation (PDB ID 3U1I [30]), showing NS2B in orange and NS3pro in grey. The locations of the residues analysed in (A) are highlighted by balls and the site of the C2-tag is indicated by a green sphere.

DENp and DENpro, whereas the signals retained their intensities in the presence of inhibitor **1** (Fig. 8). The presence of 300 mM NaCl at pH 6.5 weakened the signal of Ile78* in a selectively ^{15}N-Ile-labelled sample (Fig. S5). A uniformly ^{15}N-labelled sample confirmed this result, i.e. most cross-peaks of NS2Bc were significantly weaker or absent at 300 mM NaCl, whereas the cross-peaks from NS3pro were much less affected (Fig. 9B), except for some cross-peaks of amides in NS3pro that are sufficiently close to NS2B to sense its conformational exchange (Fig. 9C). These results indicate that the association of NS2Bc to NS3pro is indeed stabilized by electrostatic attraction.

In view of the open conformations found in crystal structures of DENpro and WNVpro [8,29], it is plausible that the pH- and salt-dependent conformational exchange of NS2B in DENp arises from temporary dissociation of NS2Bc from NS3pro, leading to the open conformation. As stated above, however, the PCSs measured for NS2B were in full agreement with the closed conformation. PCSs readily average to zero for flexible polypeptide chains, so that any disordered open conformation would not contribute to PCSs of NS2B and, therefore, the average PCSs of closed and open conformations would become smaller for increasing populations of open conformations. The similarity in PCSs observed in DENp without inhibitor and DENpro with inhibitor **1** (Fig. 6) indicates that, at pH 6.5 and low ionic strength, DENp does not populate open conformations to any significant extent.

In a fast equilibrium between open and closed conformations, a minor population of open conformation should be manifested in changes in chemical shifts, whereas a slow equilibrium would result in new cross-peaks for the open conformation. Because the protein chemical shifts are very highly conserved between pH 6.5 and pH 7.5 without any new peaks appearing at pH 7.5 (Fig. 4A,B), open conformations cannot be highly populated at the higher pH. Nonetheless, many signals broaden significantly at pH 7.5. A more complete picture was obtained by using uniformly ^{15}N-labelled DENp. Comparison of Figs 9 and 10 shows that high pH and high salt predominantly affect the residues of NS2Bc and nearby residues of NS3pro, suggesting that the electrostatic attraction of NS2Bc to NS3pro is weakened at increased pH. This result may be explained by a loss of charge of the side chain of His51. If the exchange broadening observed at high salt and high pH arises from small populations of open conformations, DENpro would be expected to be more sensitive than DENp. This is indeed the case (Fig. 8).

The chemical shifts of DENp changed slightly in the presence of high salt, raising the possibility that this effect could be attributed to a greater population of open conformations. It was difficult, however, to assess the amount of open population, because all NS2Bc resonances were too broad to be observable (Fig. 9A) and Ser70* and Ser71* were the last residues of NS2B for which amide cross-peaks could be detected in uniformly ^{15}N-labelled DENp at high salt (Fig. 9B). Their PCSs

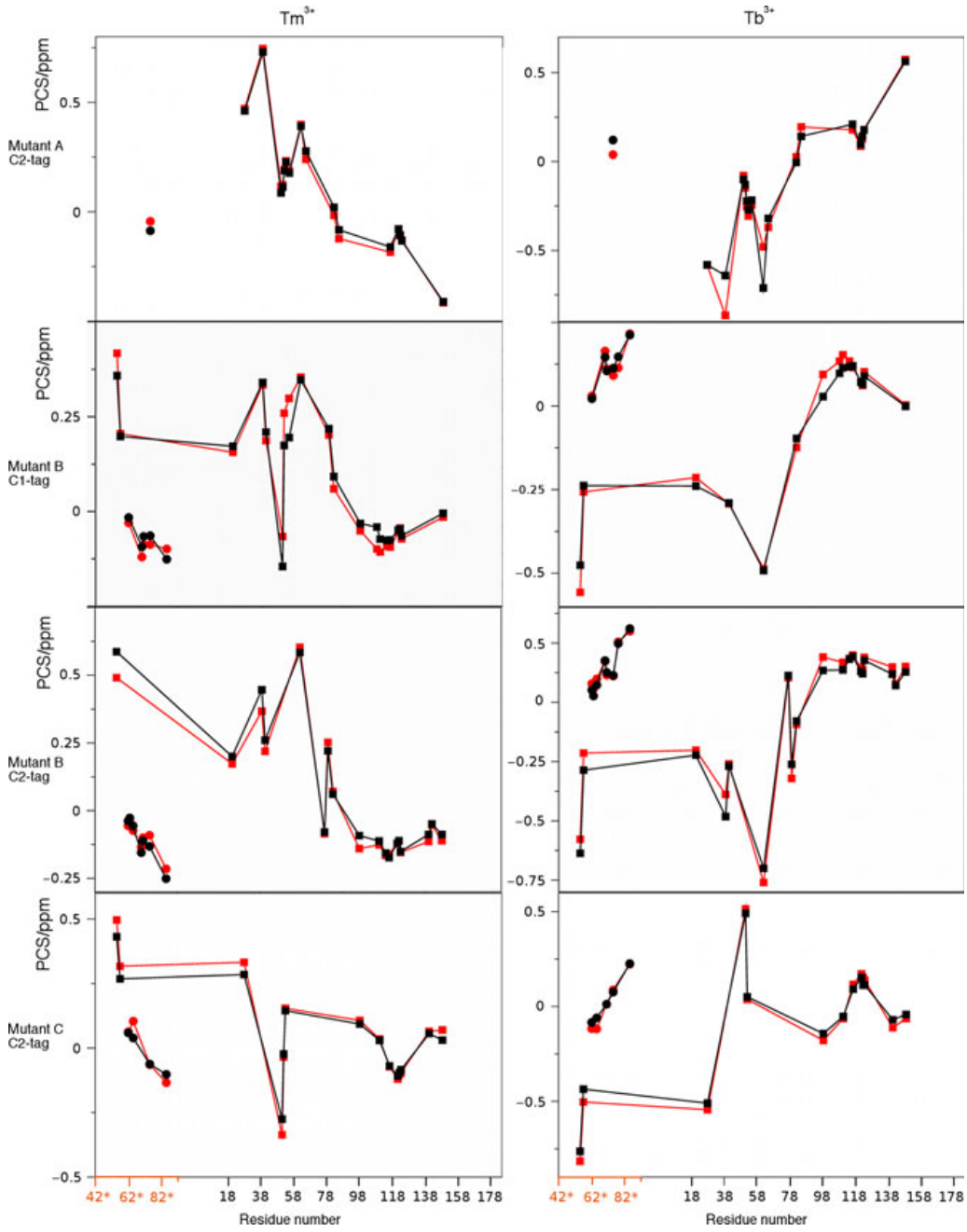


Fig. 6. PCSs demonstrate that DENp assumes a closed conformation very similar to that reported in the crystal structure of DENpro with bound inhibitor [30]. Figure plots of the PCSs measured with different tags and metal ions for mutants A, B and C of DENp in the absence of inhibitor (red symbols) and DENpro in the presence of inhibitor **1** (black symbols). Squares identify the PCSs of NS3pro and NS2B residues that form part of the N-terminal β-barrel of NS3pro; these residues are structurally similar between open and closed conformations (see Fig. 1). PCSs for NS2Bc are shown as circles. For clarity, data are shown only for residues for which PCSs could be measured both for DENp and DENpro. The close similarity between the PCSs indicates closely similar structures.

generated by the C2–Tb tag attached to a cysteine residue at site B (Fig. S6) changed slightly at high salt, but other PCSs in NS3pro also changed, reflecting a small change in the average location and orientation of the paramagnetic tag relative to the protein. This effect is not unexpected, because the lanthanide tags used were charged

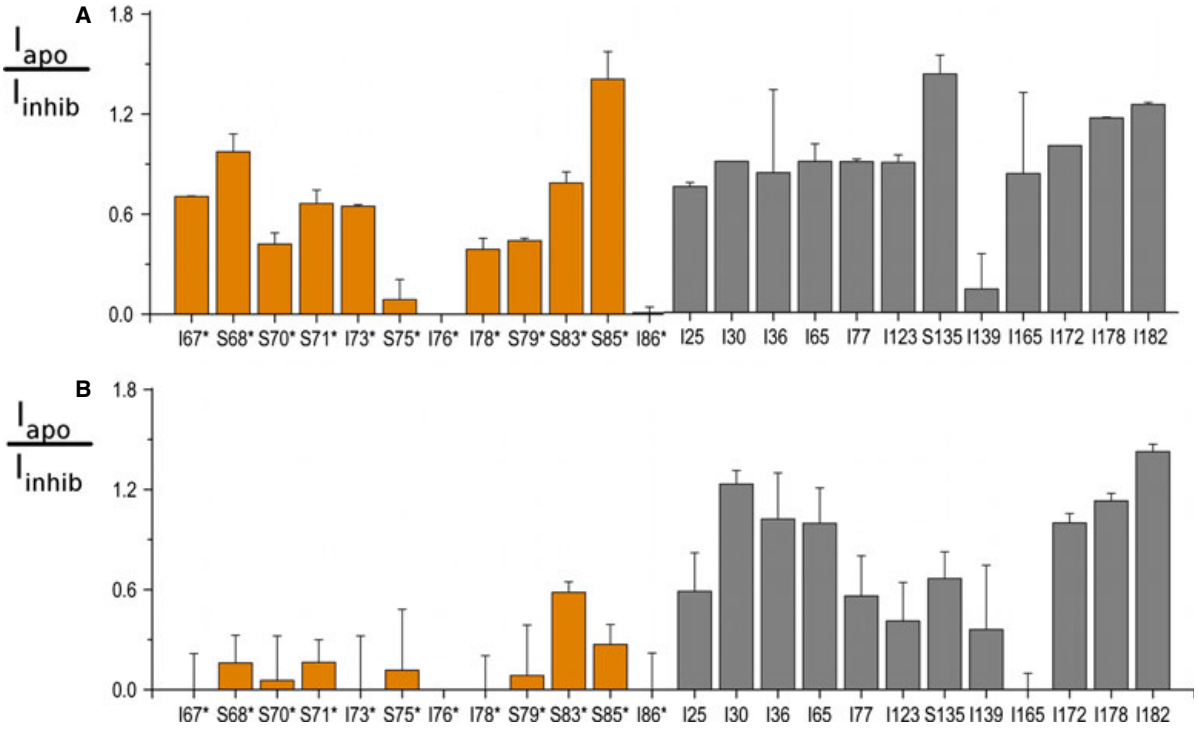


Fig. 7. Ratios of cross-peak heights in ^{15}N-HSQC spectra of DENp and DENpro without and with inhibitor **1** provide evidence for conformational exchange. All spectra were measured at pH 6.5 under the same sample conditions as in Fig. 3. In each spectrum, the cross-peak intensities were determined relative to the intense cross-peak of Ile172. Error ranges were estimated from the reproducibility of the peak heights in spectra recorded with two different samples, taking into account the level of white noise. Data for residues of NS2B and NS3 are shown in orange and grey, respectively. Residues, for which cross-peaks at or near the position characteristic of the closed conformation could not be observed in the absence of inhibitor are reported with a peak intensity ratio of zero. (A) Data for DENp. (B) Data for DENpro.

and thus sensitive to attractive and repulsive forces from the charged amino acid side chains on the protein surface. Using all observable PCSs to fit $\Delta\chi$ tensors yielded good correlations between back-calculated and experimental PCSs, including for residues 70* and 71* (Fig. S7, Tables S1 and S2). The PCS data thus indicate that Ser70* and Ser71* of NS2Bc retain their closed-state association with NS3pro even at high ionic strength. By contrast, the more C-terminal residues of NS2Bc are subject to substantial conformational exchange, which may or may not include dissociation from NS3pro. Clearly, however, the open conformation observed in the crystal structure [8] can be populated to a very small extent only (< 10%).

Discussion

Open and closed conformations

The question of open and closed conformations in the dengue virus and West Nile virus proteases has been vexed ever since their first observation in crystal structures [8]. Because NS2B greatly increases the activity of the proteases [5,41], it would be attractive to prevent the association of NS2Bc with the substrate-binding site by a drug molecule. Inhibitor designs for DENpro have used either conformation as template [14,17,19,20,23,26,42–52]. For both WNVpro and DENpro, the presence of low molecular mass inhibitors greatly improves the quality of the NMR spectra [11,38,44,53–55]. For WNVpro, however, the NMR experiments clearly showed that the closed conformation is predominant in solution regardless of the presence of an inhibitor [38]. The situation was thought to be different for DENpro, because very narrow signals indicative of NS2Bc in a random coil conformation were observed with high intensities in the absence of inhibitors [10]. In the course of this study, these signals proved to be little reproducible in intensity, with some sample preparations showing hardly any evidence for them at all. Furthermore, they were insensitive to the addition of inhibitor **1**. These observations clearly

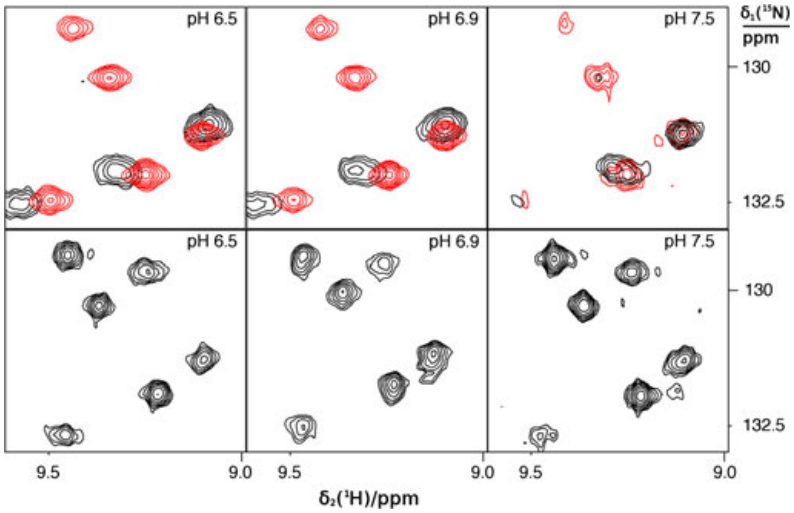


Fig. 8. The conformational equilibria in DENp and DENpro are pH dependent. (Upper) Superimposition of a selected spectral region from ^{15}N-HSQC spectra of DENp (red) and DENpro (black) recorded without inhibitor at different pH. Except for the pH, the sample conditions were the same as in Fig. 3. The decrease in peak intensities with increasing pH indicates the increasing population of alternative conformations with chemical shifts different from the closed state of the protease. (Lower) Corresponding region of DENpro in complex with inhibitor **1**, providing a reference for the closed conformation.

identify these signals as degradation products and eliminate them as evidence for a predominantly open conformation, in contrast to our previous interpretation [38]. Heterogeneity in DENpro sample preparations has been reported previously [56] but the evidence for open conformations in solution is not clear-cut.

Kim *et al.* [39] reported recently that an unlinked construct of the NS2B–NS3 protease yields better NMR spectra than the original linked DENpro construct. They also used NMR spectroscopy to establish that the unlinked protease most likely assumes the closed conformation. The key observations were: (a) that the ^{15}N-relaxation times of the NS2Bc resonances were overall comparable to those of NS3pro, showing that the mobility of NS2Bc is not greatly enhanced compared to that of NS3pro in agreement with the closed state; and (b) that PREs could be resolved for seven of the backbone amides in the globular part of the protease in a cysteine mutant with a nitroxide radical attached to residue 75 of NS2B. The PREs were also in agreement with the closed conformation and could not be explained by the open conformation of the crystal structure 2FOM [8]. In addition, approximate boundaries of secondary structure were derived from the chemical shifts of NS2B that are in agreement with the β-hairpin structure found for NS2Bc in the crystal structure 3U1I of the closed conformation [30].

The crystal structures of DENpro and WNVpro indicate that the association between NS2Bc and NS3pro is weak, because the closed conformation has only ever been observed in the presence of an inhibitor. Even with the strongly binding inhibitor aprotinin, a co-crystal structure of DENpro showed no electron density for NS2Bc, indicating that it is disordered [30]. Nonetheless, the crystal structures of WNVpro [29] and DENpro from serotype 1 [9] showed that the β-hairpin structure of NS2Bc forms readily also when NS2Bc is dissociated from NS3pro. It is thus conceivable that, in solution, the β-hairpin structure forms and associates nonspecifically with NS3pro. The NMR data of Kim *et al.* [39] could not discern this situation from the well-defined closed conformation observed with inhibitors [11,30], because distance restraints to nitroxide tags are difficult to quantify. In addition, the hydrophobicity of the tag could influence the interaction with NS3pro.

The PCSs observed in this study using multiple tags at multiple sites unambiguously establish that the dengue virus NS2B–NS3pro enzyme from serotype 2 assumes the closed conformation as observed in the crystal structure 3U1J [30], when the covalent linkage between NS2B and NS3pro is broken. At the same time, however, broad and missing NMR signals show that NS2Bc retains increased dynamics which increase with increasing salt and pH.

Conformational exchange

All backbone amide groups with the same relaxation times should display ^{15}N-HSQC cross-peaks of the same intensity, whereas increased relaxation rates reduce the cross-peak intensities by broadening the signals and because of loss of magnetization by fast relaxation during the spin-echo periods of the experiment. Because no relaxation mechanism can compete with the line broadening effects from chemical exchange, the weak signal intensities observed in the spectra of DENpro and DENp must arise from exchange broadening. Because hydrogen exchange with

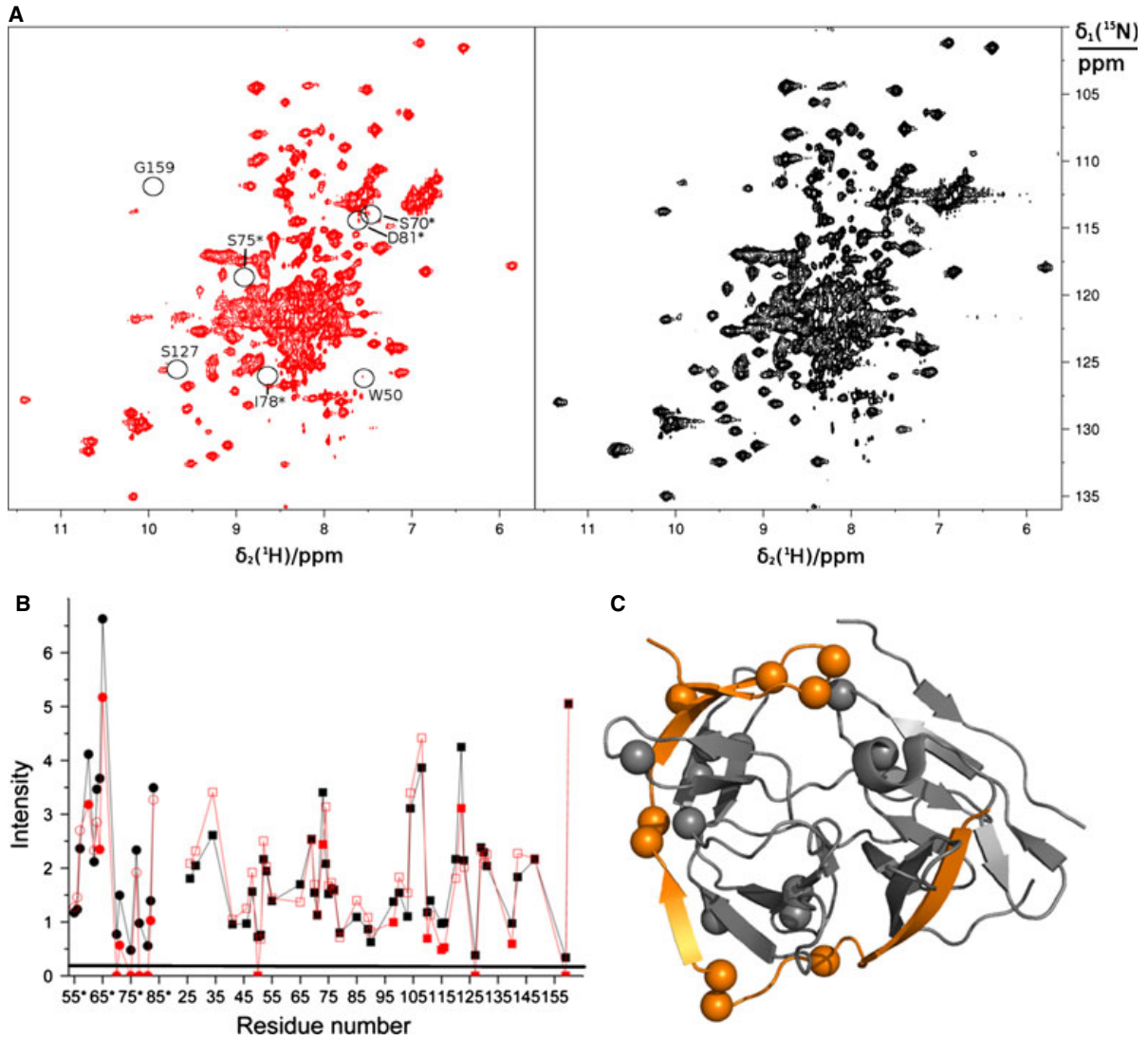


Fig. 9. High ionic strength weakens the association between NS2Bc and NS3pro. (A) ^{15}N-HSQC spectra of uniformly ^{15}N-labelled DENp at 300 mM NaCl (red spectrum) and 50 mM NaCl (black spectrum). All other sample conditions were the same as in Fig. 3. Circles identify peaks that broadened beyond detection after the addition of sodium chloride. (B) Cross-peak heights observed in the spectra of (A) plotted versus the amino acid sequence. Red and black symbols correspond to the data measured at high and low salt concentrations, respectively. A horizontal line indicates the level of white noise in the spectrum at high salt. Circles and squares refer to NS2B and NS3, respectively. The peaks in DENp displaying intensity reductions > 20% after the addition of salt are highlighted by filled red symbols. (C) Cartoon representation of the NS2B–NS3 protease in the closed conformation with NS2B and NS3 shown in orange and grey, respectively. Spheres identify the residues highlighted in (B) for greatly reduced peak intensities upon addition of salt.

water is inefficient at pH 7.5, conformational exchange remains the only plausible mechanism for the loss of cross-peak intensities.

It is tempting to identify the exchange broadening with an exchange between open and closed conformations. Our results show that the broadening becomes more pronounced by increasing the pH or the salt concentration, showing that the interaction between NS2Bc and NS3pro has a significant electrostatic component. Even in the presence of 300 mM NaCl, however, when most signals of NS2Bc become unobservable, the PCSs of the last backbone amides of NS2Bc that we were able to resolve (Ser70* and Ser71*), induced by a C2–Tb tag at site B, were not at all reduced, contrary to expectations when NS2Bc fully dissociates from NS3pro and assumes a random

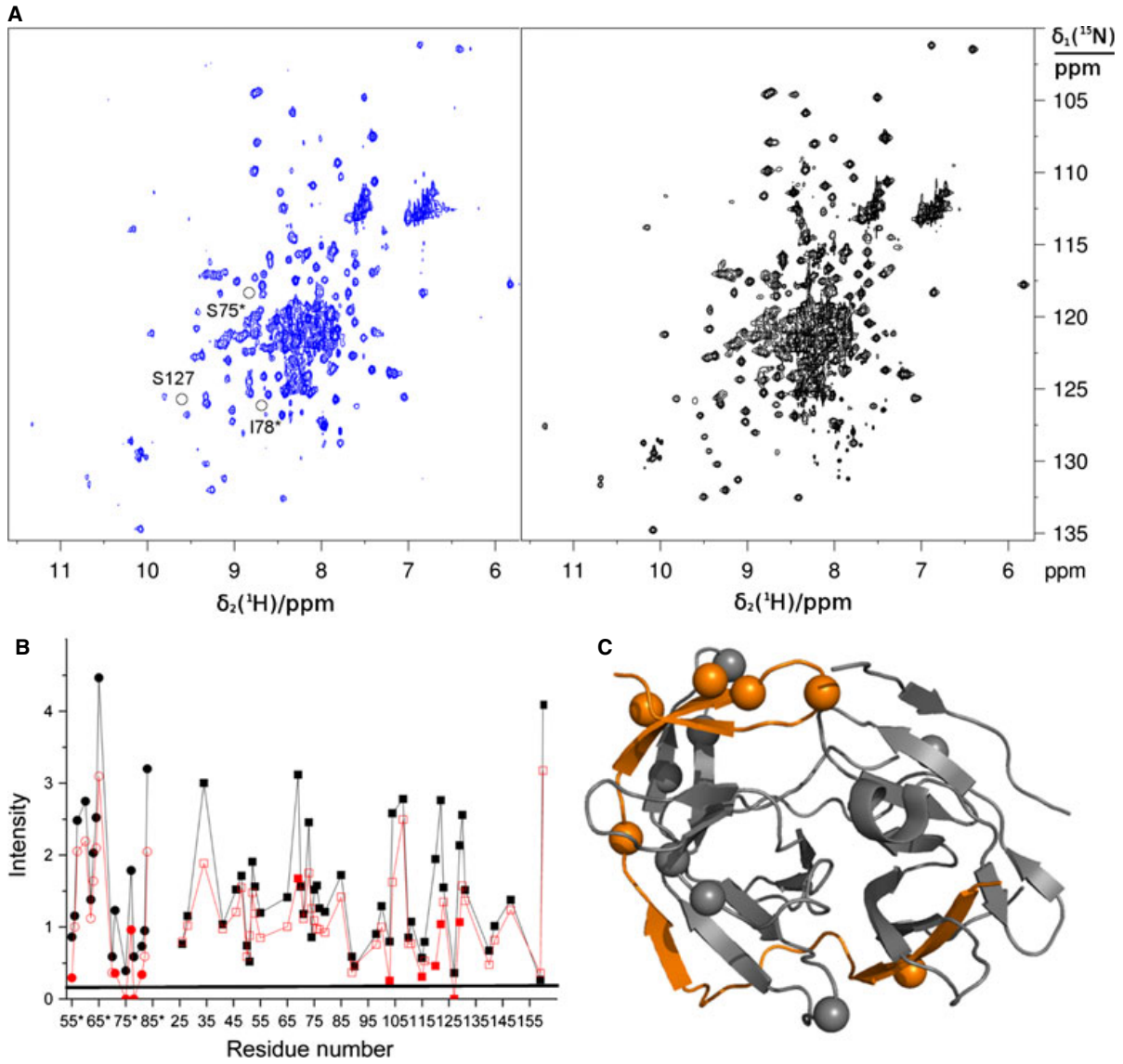


Fig. 10. High pH weakens the association between NS2Bc and NS3pro. The figure corresponds to Fig. 9, except that the peaks affected by pH rather than salt are analysed. (A) ^{15}N-HSQC spectra of uniformly ^{15}N-labelled DENp at pH 7.5 (blue spectrum) and pH 6.5 (black spectrum). All other sample conditions were the same as in Fig. 3. Circles identify peaks that broadened beyond detection at pH 7.5. (B) Cross-peak heights observed in the spectra of (A) plotted versus the amino acid sequence. Red and black symbols correspond to the data measured at high and low pH, respectively. A horizontal line indicates the level of white noise in the spectrum at high pH. Intensities are in arbitrary units and the spectra were scaled to give similar intensities for the peaks with highest intensities. The peaks in DENp displaying intensity reductions > 40% at high pH are highlighted by filled red symbols. (C) Cartoon representation of the NS2B–NS3 protease in the closed conformation with NS2B and NS3 shown in orange and grey, respectively. Spheres identify the residues highlighted in (B) for substantial reductions in peak intensities at pH 7.5.

position relative to the tag. This suggests that the exchange broadening arises from exchange between different conformations of the β-hairpin of NS2Bc while it is bound or from its selective dissociation from NS3pro. In either case, the fully open conformations reported in crystal structures are not the predominant form in solution even under conditions that promote conformational exchange of NS2Bc.

The ^{15}N-Ile-labelled DENpro construct shows no cross-peaks of NS2Bc in the closed conformation

(Fig. 3B), indicating that the Gly_4–Ser–Gly_4 linker enhances the conformational exchange of NS2Bc. Nonetheless, weak cross-peaks at chemical shifts characteristic of the closed conformation have been observed in ^{15}N-Ser-labelled samples [11], showing that the closed conformation is populated even in DENpro. Because there are no cross-peaks of NS2Bc in the spectrum of Fig. 3(B) that can be assigned to the intact NS2B–NS3 protease, it is hard to tell whether the closed or open state is more highly populated in DENpro. Certainly, the conformational state of DENpro is more sensitive to high pH, because cross-peaks from NS3pro disappear more readily at pH 7.5 than the corresponding signals in DENp (Fig. 8).

Pseudocontact shifts for structure analysis

This study illustrates the power of PCSs induced by lanthanide tags. Traditionally, most NMR studies collect long-range structure restraints by measuring PREs from site-specifically attached nitroxide spin-labels. PREs, however, can also arise from intermolecular interactions and artificial interactions can be promoted by the hydrophobic nature of conventional nitroxide spin-labels. By contrast, the PCS effect can be measured over greater distances, allowing attachment of the tag at a site far from the site of interest where it causes minimal perturbations, and nonspecific intermolecular interactions do not generate PCSs [57]. The PCSs proved particularly valuable in the case of DENp at high salt, where the signals of NS2Bc were either missing or strongly attenuated due to conformational exchange. This demonstrates how PCSs can provide structural information even under adverse conditions, where NMR signals can be observed only for isolated residues because of peak broadening, low protein concentrations or unstable preparations.

Dissociation of NS2Bc from NS3pro

In contrast to DENpro, which is stable against degradation at room temperature, our DENp construct precipitated rapidly during the NMR measurements. Precipitation was even faster at high ionic strengths, limiting the measurement time for each sample to ~4 h. A possible explanation for this effect is that the absence of the linker between NS2B and NS3pro facilitates the dissociation of NS2B, as NS3pro on its own is insoluble.

Could the open conformations observed in crystal structures be caused by the crystallization conditions used? It is well known that exchange with small populations of alternate conformations may be sufficient to broaden NMR signals beyond detection and that crystallization conditions can select unusual conformations. Notably, serotype 2 DENpro without inhibitor was crystallized at pH 8.5 and 48 °C [8], conditions which could promote dissociation of NS2Bc from NS3pro. The open conformation of WNVpro crystallized at pH 7.5 and 25 °C [29], but it was obtained of the His51Ala mutant, which might have weakened the binding of NS2Bc to NS3pro. The open conformation of serotype 1 DENpro was of the wild-type protein crystallized at pH 7.5 and 20 °C [9]. Based on our data, these conditions would not promote dissociation of NS2Bc from NS3pro, but the difference in amino acid sequence may confer different properties to this particular protein, favouring the association between the NS2Bc segments from different protein molecules as observed in the crystal structure.

Conclusion

Our results unequivocally demonstrate that the serotype 2 dengue virus NS2B–NS3 protease in solution predominantly assumes the closed conformation in the absence of inhibitors and that any open conformation is hardly populated. This has obvious and important implications for the choice of structural templates in the rational design of inhibitors.

Experimental procedures

Materials

Two different dengue 2 NS2B–NS3pro constructs were used. The first was the traditional construct used earlier for structural and functional work [7,8,11], comprising 255 residues with a $(Gly)_4$–Ser–$(Gly)_4$ linker between NS2B (47 residues) and NS3pro (185 residues). In addition, it contained a His_6-tag at the C-terminus and the T7 gene 10 N-terminal peptide MASMTG at the N-terminus followed by a two-residue cloning artefact (Leu–Glu), resulting in a 27 kDa protein. In this construct, NS2Bc comprises the segment from Glu66* to Leu95* (to discriminate between NS2B and NS3 residues, residue numbers of NS2B are marked by an asterisk throughout this text). We refer to this construct as DENpro. The second construct was the same as the first, except that the $(Gly)_4$–Ser–$(Gly)_4$ sequence in the linker between NS2B and NS3pro was replaced by the protease recognition sequence EVKKQR that precedes NS3 in the wild-type poly-protein (Fig. 2A,B). We refer to this construct as DENp.

DENpro and DENp samples were prepared either by cell-free protein synthesis from PCR-amplified DNA [10] or by high-cell density *in vivo* expression [58]. Cell-free protein

synthesis used the high-copy number plasmid pRSET-6d-DEN2 CF40GlyNS3pro as the template for PCR-amplification [8,59]. S30 cell extracts were prepared in-house from *Escherichia coli* strains Rosetta::λDE3/pRARE and BL21 Star::λDE3 [42], including concentration with poly(ethylene glycol) 8000. ^{15}N-labelled amino acids were purchased from Cambridge Isotope Laboratories (Andover, MA, USA) and ISOTEC (St. Louis, MO, USA). Synthetic oligonucleotides were purchased from IDT (Coralville, IA, USA) and Geneworks (Thebarton, SA, Australia). PCRs were performed using *Vent* DNA polymerase (New England Biolabs, Ipswich, MA, USA). PCR products were purified using the QiaQuick PCR purification kit (Qiagen, Hilden, Germany).

Single-cysteine mutants were prepared by *in vivo* expression experiments using the same gene constructs as cell-free expression but inserted into the pETMCSI T7 expression vector [60] and transformed into the *E. coli* strain Rosetta::λDE3/pRARE.

The C1- and C2-compounds for tagging with lanthanides were synthesized as described [61].

Cell-free protein synthesis

Production by cell-free synthesis allowed selective labelling with ^{15}N-labelled amino acids without significant isotope scrambling [62] and the rapid production of site-specific mutants [10]. Samples of DENp were synthesized in a cell-free *E. coli* coupled transcription–translation system using a previously described protocol [63,64]. ^{15}N-labelled isoleucine was used in the cell-free reaction mixture. Wild-type DENpro was synthesized using the pRSET-DEN2 CF40GlyNS3pro plasmid at a concentration of 16 μg·mL^{-1} of reaction mixture. Mutant DENpro was synthesized using reannealed DNA at a concentration of 20 μg·mL^{-1} of reaction mixture. Each synthesis was performed in two identical parallel reactions, each using a reaction volume of 2 mL in 20 mL of outer buffer and proceeding for 14–15 h at 30 °C. Typical protein yields were 1.6 mg per reaction. The S30 extract used for the expression of selectively ^{15}N-serine-labelled samples was reduced with sodium borohydride to suppress cross-labelling by transaminases as described previously [65].

Protein purification

Following synthesis, the protein samples were purified using IMAC Ni-NTA spin columns. The columns were equilibrated with binding buffer (50 mM HEPES, 300 mM NaCl, 10 mM imidazole at pH 7.5), loaded with the sample, washed five times with wash buffer (50 mM HEPES, 300 mM NaCl, 30 mM imidazole, pH 7.5) and eluted with elution buffer (50 mM HEPES, 300 mM NaCl, 300 mM imidazole, pH 7.5). The purified protein was dialysed extensively against NMR buffer (20 mM MES, pH 6.5, 50 mM NaCl or 20 mM Tris/HCl, pH 6.9 or pH 7.5, 50 mM NaCl). Samples at higher salt concentrations (300 mM NaCl) were prepared by dialysis against buffers with correspondingly increased salt concentration. Samples were concentrated using a Centricon-4 ultrafilter MWCO-10 kDa concentrator (Amicon, Billerica, MA, USA) to final volumes of 0.5 mL for samples at low salt concentration in 5 mm NMR tubes and 0.2 mL for samples at high salt concentration in 3 mm NMR tubes. The protein concentrations at low and high salt were 0.2 and 0.7 mM, respectively.

Tagging reaction

C1 and C2 tags loaded with Tb^{3+}, Tm^{3+} or Y^{3+} (Fig. S2) were attached to the single-cysteine mutants of DENp by adding the protein to a threefold excess of the respective metal complexes of C1 or C2 and incubating at room temperature for several hours. Excess low-molecular mass reactants and products were removed by washing with NMR buffer.

NMR spectroscopy

All NMR spectra were recorded at 25 °C on Bruker 600 and 800 MHz NMR spectrometers equipped with cryoprobes. NMR spectra were recorded in NMR buffer unless indicated otherwise. *p*-Nitrophenyl-*p*-guanidino benzoate (**1**; Fig. S2A) was used as the inhibitor [66]. The inhibitor was added in fivefold molar excess to the protein using a 100 mM stock solution in dimethylsulfoxide-d_6. A spectrum for measuring chemical exchange between water and amide protons was recorded using the CLEANEX-PM pulse sequence [67].

Pseudocontact shifts

PCSs were measured as the ^{1}H chemical shifts of backbone amides observed in ^{15}N-HSQC spectra of paramagnetic samples minus those observed for diamagnetic samples. The PCSs were generated by a lanthanide ion with nonisotropic magnetic susceptibility. For each nuclear spin, the PCS (measured in ppm) depends on its polar coordinates r, θ and φ with respect to the principal axes of the Δχ tensor:

$$\Delta\delta^{PCS} = \frac{1}{12\pi r^3}\left[\Delta\chi_{ax}(3\cos^2\theta - 1) + \frac{3}{2}\Delta\chi_{rh}\sin^2\theta\cos 2\varphi\right] \quad (1)$$

where $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ denote, respectively, the axial and rhombic components of the magnetic susceptibility tensor χ [68], and the Δχ tensor is defined as the χ tensor minus its isotropic component. Equation (1) shows that PCSs can be positive or negative, depending on the position of the nuclear spin with respect to the coordinate system defined by principal axes of the Δχ tensor. For mobile polypeptide chains, PCSs easily average to zero.

Acknowledgements

We thank Dr Siew-Pheng Lim at the Novartis Institute for Tropical Diseases (NITD) for the initial construct of the dengue virus protease. Funding by the Australian Research Council is gratefully acknowledged.

References

1 Bhatt S, Gething PW, Brady OJ, Messina JP, Farlow AW, Moyes CL, Drake JM, Brownstein JS, Hoen AG, Sankoh O *et al.* (2013) The global distribution and burden of dengue. *Nature* **496**, 504–507.
2 Kuno G (2007) Research on dengue and dengue-like illness in East Asia and the Western Pacific during the first half of the 20th century. *Rev Med Virol* **17**, 327–341.
3 Guzman MG, Halstead SB, Artsob H, Buchy P, Farrar J, Gubler DJ, Hunsperger E, Kroeger A, Margolis HS, Martínez E *et al.* (2010) Dengue: a continuing global threat. *Nat Rev Microbiol* **8**, S7–S16.
4 Lescar J, Luo D, Xu T, Sampath A, Lim SP, Canard B & Vasudevan SG (2008) Towards the design of antiviral inhibitors against flaviviruses: the case for the multifunctional NS3 protein from dengue virus as a target. *Antiviral Res* **80**, 94–101.
5 Yusof R, Clum S, Wetzel M, Murthy HMK & Padmanabhan R (2000) 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.
6 Preugschat F, Yao CW & Strauss JH (1990) *In vitro* processing of dengue virus type 2 nonstructural proteins NS2A, NS2B, and NS3. *J Virol* **64**, 4364–4374.
7 Leung D, Schroder K, White H, Fang NX, Stoermer MJ, Abbenante G, Martin JL, Young PR & Fairlie DP (2001) 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.
8 Erbel P, Schiering N, D'Arcy A, Renatus M, Kroemer M, Lim SP, Yin Z, Keller TH, Vasudevan SG & Hommel U (2006) Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. *Nat Struct Mol Biol* **13**, 372–373.
9 Chandramouli S, Joseph JS, Daudenarde S, Gatchalian J, Cornillez-Ty C & Kuhn P (2010) Serotype-specific structural differences in the protease–cofactor complexes of the dengue virus family. *J Virol* **84**, 3059–3067.
10 Wu PSC, Ozawa K, Lim SP, Vasudevan S, Dixon NE & Otting G (2007) Cell-free transcription/translation from PCR amplified DNA for high-throughput NMR studies. *Angew Chem Int Ed* **46**, 3356–3358.
11 de la Cruz L, Nguyen THD, Ozawa K, Shin J, Graham B, Huber T & Otting G (2011) Binding of low-molecular weight inhibitors promotes large conformational changes in the dengue virus NS2B–NS3 protease: fold analysis by pseudocontact shifts. *J Am Chem Soc* **133**, 19205–19215.
12 Li J, Lim SP, Beer D, Patel V, Wen D, Tumanut C, Tully DC, Williams JA, Jiricek J, Priestle JP *et al.* (2005) Functional profiling of recombinant NS3 proteases from all four serotypes of dengue virus using tetrapeptide and octapeptide substrate libraries. *J Biol Chem* **280**, 28766–28774.
13 Yin Z, Patel SJ, Wang WL, Wang G, Chan WL, Rao KRR, Alam J, Jeyaraj DA, Ngew X, Patel V *et al.* (2006) Peptide inhibitors of dengue virus NS3 protease. Part 1: warhead. *Bioorg Med Chem Lett* **16**, 36–39.
14 Yin Z, Patel SJ, Wang WL, Chan WL, Rao KRR, Wang G, Ngew X, Patel V, Beer D, Knox JE *et al.* (2006) Peptide inhibitors of dengue virus NS3 protease. Part 2: SAR study of tetrapeptide aldehyde inhibitors. *Bioorg Med Chem Lett* **16**, 40–43.
15 Bodenreider C, Beer D, Keller TH, Sonntag S, Wen D, Yap L, Yau YH, Geifman Shochat S, Huang D, Zhou T *et al.* (2009) A fluorescence quenching assay to discriminate between specific and non-specific inhibitors of dengue virus protease. *Anal Biochem* **395**, 195–204.
16 Tomlinson SM, Malmstrom RD, Russo A, Mueller N, Pang YP & Watowich SJ (2010) Structure-based discovery of dengue virus protease inhibitors. *Antivir Res* **82**, 110–114.
17 Tomlinson SM & Watowich SJ (2011) Anthracene-based inhibitors of dengue virus NS2B–NS3 protease. *Antivir Res* **89**, 127–135.
18 Steuer C, Gege C, Fischl W, Heinonen KH, Bartenschlager R & Klein CD (2011) Synthesis and biological evaluation of α-ketoamides as inhibitors of the dengue virus protease with antiviral activity in cell-culture. *Bioorg Med Chem* **19**, 4067–4074.
19 Knehans T, Schüller A, Doan DN, Nacro K, Hill J, Güntert P, Madhusudhan MS, Weil T & Vasudevan SG (2011) Structure-guided fragment-based *in silico* drug design of dengue protease inhibitors. *J Comput Aided Mol Des* **25**, 263–274.
20 Schüller A, Yin Z, Chia CSB, Doan DN, Kim HK, Shang L, Loh TP, Hill J & Vasudevan SG (2011) Tripeptide inhibitors of dengue and West Nile virus NS2B–NS3 protease. *Antivir Res* **92**, 96–101.
21 Yang CC, Hsieh YC, Lee SJ, Wu SH, Liao CL, Tsao CH, Chao YS, Chern JH, Wu CP & Yueh A (2011) Novel dengue virus-specific NS2B/NS3 protease inhibitor, BP2109, discovered by a high-throughput screening assay. *Antimicrob Agents Chemother* **55**, 229–238.
22 Nitsche C, Steuer C & Klein CD (2011) Arylcyanoacrylamides as inhibitors of the dengue and West Nile virus proteases. *Bioorg Med Chem* **19**, 7318–7337.

23 Doan DN, Li KQ, Basavannacharya C, Vasudevan SG & Madhusudhan MS (2012) Transplantation of a hydrogen bonding network from West Nile virus protease onto dengue-2 protease improves catalytic efficiency and sheds light on substrate specificity. *Protein Eng Des Sel* **25**, 843–850.
24 Rothan HA, Han HC, Ramasamy TS, Othman S, Rahman NA & Yusof R (2012) Inhibition of dengue NS2B–NS3 protease and viral replication in Vero cells by recombinant retrocyclin-1. *BMC Infect Dis* **12**, 314.
25 Tomlinson SM & Watowich SJ (2012) Use of parallel validation high-throughput screens to reduce false positives and identify novel dengue NS2B–NS3 protease inhibitors. *Antiviral Res* **93**, 245–252.
26 Xu S, Li H, Shao X, Fan C, Ericksen B, Liu J, Chi C & Wang C (2012) Critical effect of peptide cyclization on the potency of peptide inhibitors against dengue virus NS2B–NS3 protease. *J Med Chem* **55**, 6881–6887.
27 Nitsche C, Behnam MA, Steuer C & Klein CD (2012) Retro peptide-hybrids as selective inhibitors of the dengue virus NS2B–NS3 protease. *Antiviral Res* **94**, 72–79.
28 Pasfield LA, de la Cruz L, Ho J, Coote ML, Otting G & McLeod MD (2013) Synthesis of (±)-panduratin A and related natural products using the high pressure Diels-Alder reaction. *Asian J Org Chem* **2**, 60–63.
29 Aleshin AE, Shiryaev SA, Strongin AY & Liddington RC (2007) Structural evidence for regulation and specificity of flaviviral proteases and evolution of the Flaviviridae fold. *Protein Sci* **16**, 795–806.
30 Noble CG, Seh CC, Chao AT & Shi PY (2012) Ligand-bound structures of the dengue virus protease reveal the active conformation. *J Virol* **86**, 438–446.
31 Robin G, Chappell K, Stoermer MJ, Hu S, Young PR, Fairlie DP & Martin JL (2009) Structure of West Nile virus NS3 protease: ligand stabilization of the catalytic conformation. *J Mol Biol* **385**, 1568–1577.
32 Hammamy MZ, Haase C, Hammami M, Hilgenfeld R & Steinmetzer T (2013) Development and characterization of new peptidomimetic inhibitors of the West Nile virus NS2B–NS3 protease. *ChemMedChem* **8**, 231–241.
33 Falgout B, Miller RH & Lai C-J (1993) Deletion analysis of dengue virus type 4 nonstructural protein NS2B: identification of a domain required for NS2B–NS3 protease activity. *J Virol* **67**, 2034–2042.
34 Wu CF, Wang SH, Sun CM, Hu ST & Syu WJ (2003) Activation of dengue protease autocleavage at the NS2B–NS3 junction by recombinant NS3 and GST–NS2B fusion proteins. *J Virol Method* **114**, 45–54.
35 Niyomrattanakit P, Winoyanuwattikun P, Chanprapaph S, Angsuthanasombat C, Panyim S & Katzenmeier G (2004) Identification of residues in the dengue virus type 2 NS2B cofactor that are critical for NS3 protease activation. *J Virol* **78**, 13708–13716.
36 Phong WY, Moreland NJ, Lim SP, Wen D, Paradkar PN & Vasudevan SG (2011) Dengue protease activity: the structural integrity and interaction of NS2B with NS3 protease and its potential as a drug target. *Biosci Rep* **31**, 399–409.
37 Radichev I, Shiryaev SA, Aleshin AE, Ratnikov BI, Smith JW, Liddington RC & Strongin AY (2008) Structure-based mutagenesis identifies important novel determinants of the NS2B cofactor of the West Nile virus two-component NS2B–NS3 proteinase. *J Gen Virol* **89**, 636–641.
38 Su XC, Ozawa K, Qi R, Vasudevan SG, Lim SP & Otting G (2009) NMR analysis of the dynamic exchange of the NS2B cofactor between open and closed conformations of the West Nile virus NS2B–NS3 protease. *PLoS Negl Trop Dis* **3**, e561.
39 Kim YM, Gayen S, Kang C, Joy J, Huang Q, Chen AS, Wee JL, Ang MJ, Lim HA, Hung AW *et al.* (2013) NMR analysis of a novel enzymatically active unlinked dengue NS2B–NS3 protease complex. *J Biol Chem* **288**, 12891–12900.
40 Bai Y, Milne JS, Mayne L & Englander SW (1993) Primary structure effects on peptide group hydrogen exchange. *Proteins* **17**, 75–86.
41 Chappell KJ, Stoermer MJ, Fairlie DP & Young PR (2008) Mutagenesis of the West Nile virus NS2B cofactor domain reveals two regions essential for protease activity. *J Gen Virol* **89**, 1010–1014.
42 Ghanesh VK, Muller N, Judge K, Luan CH, Padmanabhan R & Murthy KHM (2005) Identification and characterization of nonsubstrate based inhibitors of the essential dengue and West Nile virus proteases. *Bioorg Med Chem* **13**, 257–264.
43 Frecer V & Miertus S (2010) Design, structure-based focusing and *in silico* screening of combinatorial library of peptidomimetic inhibitors of Dengue virus NS2B–NS3 protease. *J Comput Aided Mol Des* **24**, 195–212.
44 Tambunan USF & Alamudi S (2010) Designing cyclic peptide inhibitor of dengue virus NS3–NS2B protease by using molecular docking approach. *Bioinformation* **5**, 250–254.
45 Frimayanti N, Chee CF, Zain SM & Rahman NA (2011) Design of new competitive dengue Ns2b/Ns3 protease inhibitors – a computational approach. *Int J Mol Sci* **12**, 1089–1100.
46 Lai H, Prasad GS & Padmanabhan R (2013) Characterization of 8-hydroxyquinoline derivatives containing aminobenzothiazole as inhibitors of dengue virus type 2 protease *in vitro*. *Antiviral Res* **97**, 74–80.
47 Prusis P, Junaid M, Petrovska R, Yahorava S, Yahorau A, Katzenmeier G, Lapins M & Wikberg JE

(2013) Design and evaluation of substrate-based octapeptide and non substrate-based tetrapeptide inhibitors of dengue virus NS2B–NS3 proteases. *Biochem Biophys Res Commun* **434**, 767–772.
48 Wichapong K, Nueangaudom A, Pianwanit S, Sippl W & Kokpol S (2013) Identification of potential hit compounds for Dengue virus NS2B/NS3 protease inhibitors by combining virtual screening and binding free energy calculations. *Trop Biomed* **30**, 388–408.
49 Nitsche C, Schreier VN, Behnam MA, Kumar A, Bartenschlager R & Klein CD (2013) Thiazolidinone–peptide hybrids as dengue virus protease inhibitors with antiviral activity in cell culture. *J Med Chem* **56**, 8389–8403.
50 Pambudi S, Kawashita N, Phanthanawiboon S, Omokoko MD, Masrinoul P, Yamashita A, Limkittikul K, Yasunaga T, Takagi T, Ikuta K *et al.* (2013) A small compound targeting the interaction between nonstructural proteins 2B and 3 inhibits dengue virus replication. *Biochem Biophys Res Commun* **440**, 393–398.
51 Yildiz M, Ghosh S, Bell JA, Sherman W & Hardy JA (2013) Allosteric inhibition of the NS2B–NS3 protease from dengue virus. *ACS Chem Biol* **8**, 2744–2752.
52 Zhou GC, Weng Z, Shao X, Liu F, Nie X, Liu J, Wang D, Wang C & Guo K (2013) Discovery and SAR studies of methionine–proline anilides as dengue virus NS2B–NS3 protease inhibitors. *Bioorg Med Chem Lett* **23**, 6549–6554.
53 Su XC, Ozawa K, Yagi H, Lim SP, Wen D, Ekonomiuk D, Huang D, Keller TH, Sonntag S, Caflisch A *et al.* (2009) NMR study of complexes between low-molecular weight inhibitors and the West Nile virus NS2B–NS3 protease. *FEBS J* **276**, 4244–4255.
54 Ekonomiuk D, Su XC, Ozawa K, Bodenreider C, Lim SP, Yin Z, Keller TH, Beer D, Patel V, Otting G *et al.* (2009) Discovery of a non-peptidic inhibitor of West Nile virus NS2B–NS3 protease by high-throughput docking. *PLoS Negl Trop Dis* **3**, e356.
55 Kang C & Gayen S (2013) Exploring the binding of peptidic West Nile virus NS2B–NS3 protease inhibitors by NMR. *Antivir Res* **97**, 137–144.
56 Arakaki TL, Fang NX, Fairlie DP, Young PR & Martin JL (2002) Catalytically active Dengue virus NS3 protease forms aggregates that are separable by size exclusion chromatography. *Protein Expr Purif* **25**, 241–247.
57 Otting G (2008) Prospects for lanthanides in structural biology by NMR. *J Biomol NMR* **42**, 1–9.
58 Sivashanmugam A, Murray V, Cui C, Zhang Y, Wang J & Li Q (2009) Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. *Protein Sci* **18**, 936–948.
59 Schoepfer R (1993) The pRSET family of T7 promoter expression vectors for *Escherichia coli*. *Gene* **124**, 83–85.
60 Neylon C, Brown SE, Kralicek AV, Miles CS, Love CA & Dixon NE (2000) Interaction of the *Escherichia coli* replication terminator protein (Tus) with DNA: a model derived from DNA-binding studies of mutant proteins by surface plasmon resonance. *Biochemistry* **39**, 11989–11999.
61 Graham B, Loh CT, Swarbrick JD, Ung P, Shin J, Yagi H, Jia X, Chhabra S, Pintacuda G, Huber T *et al.* (2011) A DOTA-amide lanthanide tag for reliable generation of pseudocontact shifts in protein NMR spectra. *Bioconjugate Chem* **22**, 2118–2125.
62 Ozawa K, Wu PCS, Dixon NE & Otting G (2006) ^{15}N-labelled proteins by cell free protein synthesis: strategies for high-throughput NMR studies of proteins and protein–ligand complexes. *FEBS J* **273**, 4154–4159.
63 Apponyi M, Ozawa K, Dixon NE & Otting G (2008) Cell-free protein synthesis for analysis by NMR spectroscopy. *Methods Mol Biol* **426**, 257–268.
64 Ozawa K, Headlam MJ, Schaeffer PM, Henderson BR, Dixon NE & Otting G (2004) Optimization of an *Escherichia coli* system for cell-free synthesis of selectively ^{15}N-labelled proteins for rapid analysis by NMR spectroscopy. *Eur J Biochem* **271**, 4084–4093.
65 Su XC, Loh CT, Qi R & Otting G (2011) Suppression of isotope scrambling in cell-free protein synthesis by broadband inhibition of PLP enzymes for selective ^{15}N-labelling and production of perdeuterated proteins in H_2O. *J Biomol NMR* **50**, 35–42.
66 Ekonomiuk D, Su XC, Ozawa K, Bodenreider C, Lim SP, Otting G, Huang D & Caflisch A (2009) Flaviviral protease inhibitors identified by fragment-based library docking into a structure generated by molecular dynamics. *J Med Chem* **52**, 4860–4868.
67 Hwang TL, van Zijl PCM & Mori S (1998) Accurate quantitation of water–amide proton exchange rates using the phase-modulated CLEAN Chemical Exchange (CLEANEX-PM) approach with a Fast-HSQC (FHSQC) detection scheme. *J Biomol NMR* **11**, 221–226.
68 Bertini I, Luchinat C & Parigi G (2002) Magnetic susceptibility in paramagnetic NMR. *Prog NMR Spectrosc* **40**, 249–273.

Supporting information

Additional supporting information may be found in the online version of this article at the publisher's web site:

Fig. S1. Sequence alignment of the NS2B–NS3pro in the dengue virus serotypes 1–4.

Fig. S2. Chemical structures of inhibitor 1 and C1 lanthanide tag.

Fig. S3. DENp is enzymatically active.
Fig. S4. ^{15}N-HSQC spectrum of DENp selectively labelled with ^{15}N-serine and ^{15}N-isoleucine.
Fig. S5. ^{15}N-HSQC spectra of DENp selectively labelled with ^{15}N-isoleucine at different ionic strengths.
Fig. S6. PCSs observed for NS2B of DENp at 50 mM and 300 mM NaCl.
Fig. S7. Correlations between back-calculated and experimental PCSs of DENp at low and high salt.
Table S1. PCSs of backbone amide protons of ^{15}N-labelled DENp mutant B with C2 tag loaded with Tb^{3+} at 50 mM and 300 mM NaCl concentrations.
Table S2. $\Delta\chi$ tensor parameters of DENp mutant B with C2–Tb^{3+} tag.

the FEBS Journal

WILEY Blackwell

www.febsjournal.org

Binding mode of the activity-modulating C-terminal segment of NS2B to NS3 in the dengue virus NS2B–NS3 protease

Laura de la Cruz, Wan-Na Chen, Bim Graham and Gottfried Otting

DOI: 10.1111/febs.12729

Supporting Information

Binding mode of the activity-modulating C-terminal segment of NS2B to NS3 in the dengue virus NS2B-NS3 protease

Laura de la Cruz, Wan-Na Chen, Bim Graham, Gottfried Otting

```
                 51*         61*         71*         81*         91*
DENV1       AD LSLEKAAEVS WEEEAEHSGA SHNILVEVQD DGTMKIKDEE RDDTLTILLK
DENV2       AD LELERAADVR WEEQAEISGS SPILSITISE DGSMSIKNEE EEQTLTILIR
DENV3       AD LTVEKAADVT WEEEAEQTGV SHNLMITVDD DGTMRIKDDE TENILTVLLK
DENV4       AD LSLEKAANVQ WDEMADITGS SPIIEVKQDE DGSFSIRDVE ETNMITLLVK

            101*        111*        121*        1           11
DENV1       ATLLAVSGVY PLSIPATLFV WYFWQKKKQR SGVLWDTPSP PEVERAVLDD
DENV2       TGLLVISGLF PASIPITAAA WYLWEVKKQR AGVLWDVPSP PPVGKAELED
DENV3       TALLIVSGIF PYSIPATLLV WHTWQKQTQR SGVLWDVPSP PETQKAELEE
DENV4       LALITVSGLY PLAIPVTMTL WYMWQVKTQR SGALWDVPSP AAAQKATLTE

            21          31          41          51          61
DENV1       GIYRIMQRGL LGRSQVGVGV FQDGVFHTMW HVTRGAVLMY QGKRLEPSWA
DENV2       GAYRIKQKGI LGYSQIGAGV YKEGTFHTMW HVTRGAVLMH KGKRIEPSWA
DENV3       GVYRIKQQGI FGKTQVGVGV QKEGVFHTMW HVTRGAVLTH NGKRLEPNWA
DENV4       GVYRIMQRGL FGKTQVGVGI HMEGVFHTMW HVTRGSVICH ESGRLEPSWA

            71          81          91          101         111
DENV1       SVKKDLISYG GGWRFQGSWN TGEEVQVIAV EPGKNPKNVQ TAPGTFKTPE
DENV2       DVKKDLISYG GGWKLEGEWK EGEEVQVLAL EPGKNPRAVQ TKPGLFKTNT
DENV3       SVKKDLISYG GGWRLSAQWQ KGEEVQVIAV EPGKNPKNFQ TMPGIFQTTT
DENV4       DVRNDMISYG GGWRLGDKWD KEEDVQVLAI EPGKNPKHVQ TKPGLFKTLT

            121         131         141         151         161
DENV1       GEVGAIALDF KPGTSGSPIV NREGKIVGLY GNGVVTTSGT YVSAIAQAKA
DENV2       GTIGAVSLDF SPGTSGSPIV DKKGKVVGLY GNGVVTRSGA YVSAIAQTEK
DENV3       GEIGAIALDF KPGTSGSPII NREGKVVGLY GNGVVTKNGG YVSGIAQTNA
DENV4       GEIGAVTLDF KPGTSGSPII NRKGKVIGLY GNGVVTKSGD YVSAITQAER

            171
DENV1       SQEGPLPEIE DEVFRK
DENV2       SIED-NPEIE DDIFRK
DENV3       EPDGPTPELE EEMFKK
DENV4       IGEP-DYEVD EDIFRK
```

Figure S1. Sequence alignment of the NS2B-NS3pro in the dengue virus serotypes 1-4 (NCBI taxonomy IDs 11053, 11070, 408693 and 408688). Positively and negatively charged amino acids are shown in blue and red, respectively. Residues of NS2B are displayed in italics and identified by an asterisk. NS2B contains a hydrophilic domain acting as a cofactor (residues 49*-95*) followed by a hydrophobic transmembrane

domain and a C-terminal hydrophilic segment containing the autoproteolytic cleavage site (residues 125*-130*) leading to NS3. In previous studies of the dengue virus protease, the NS2B cofactor was linked to NS3pro via a covalent Gly_4-Ser-Gly_4-linker, a construct referred to as DENpro [1]. The construct presented here, referred to as DENp, uses the autoproteolytic cleavage site (residues 125*-130*) instead of the Gly_4-Ser-Gly_4-linker to connect NS2B and NS3pro. The resulting amino acid sequence of the DENp construct is shown in bold, underlined residues form the catalytic triad. Autoproteolysis of DENp during protein expression yields a non-covalent, active protease complex.

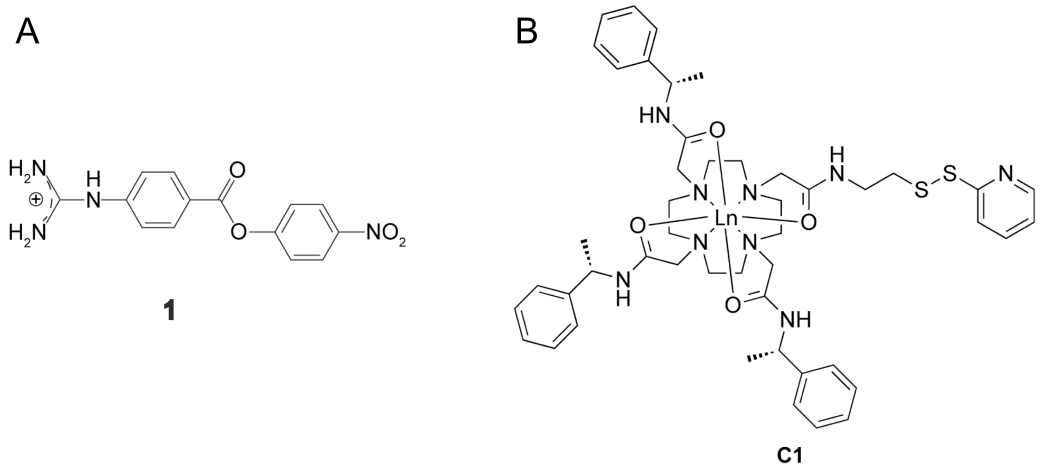


Figure S2. Chemical structures of inhibitor 1 and C1 lanthanide tag. (A) Inhibitor used in the present work. It is a generic inhibitor of serine proteases [2,3], which reacts with the protease with formation of an ester bond to Ser135, releasing yellow *p*-nitrophenolate [4]. (B) Structure of the C1 tag with bound lanthanide ion [5]. The tag is activated by a pyridin-2-yldisulfanyl group for spontaneous reaction with a cysteine thiol group to form a disulfide bond. The C2-tag is the opposite enantiomer of the C1-tag with ((**R**)-1-phenylethyl)acetamide pendants instead of ((**S**)-1-phenylethyl)acetamide pendants throughout.

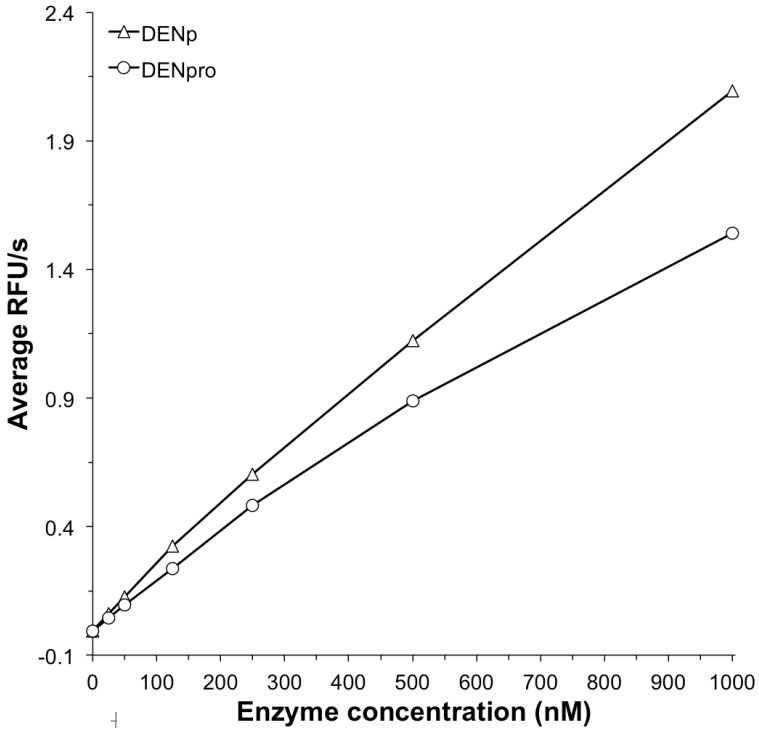


Figure S3. DENp is enzymatically active. The activities of DENp (triangles) and DENpro (circles) were compared using 10 μM Bz-norleucine-lysine-arginine-arginine-AMC [6] in 50 mM Tris·HCl (pH 7.5), 50 mM NaCl and 1 mM CHAPS in a final volume of 50 μL at 25 °C. The proteolytic reaction was monitored by the increase in fluorescence (relative fluorescence units per second, exciting at 380 nm and monitoring the emission at 450 nm) observed during 30 mins on a SpectraMax microplate reader. Each measurement was performed in three different wells and the data were averaged.

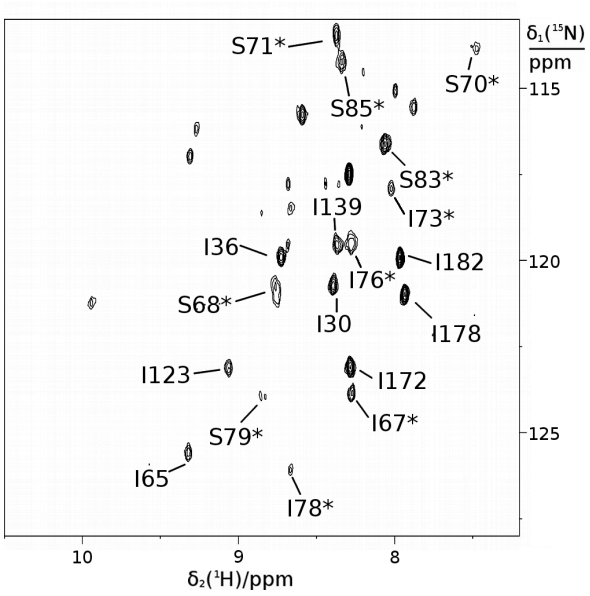


Figure S4. ^{15}N-HSQC spectrum of DENp selectively labelled with ^{15}N-serine and ^{15}N-isoleucine. Same as Figure 3A, except that the resonance assignments are shown.

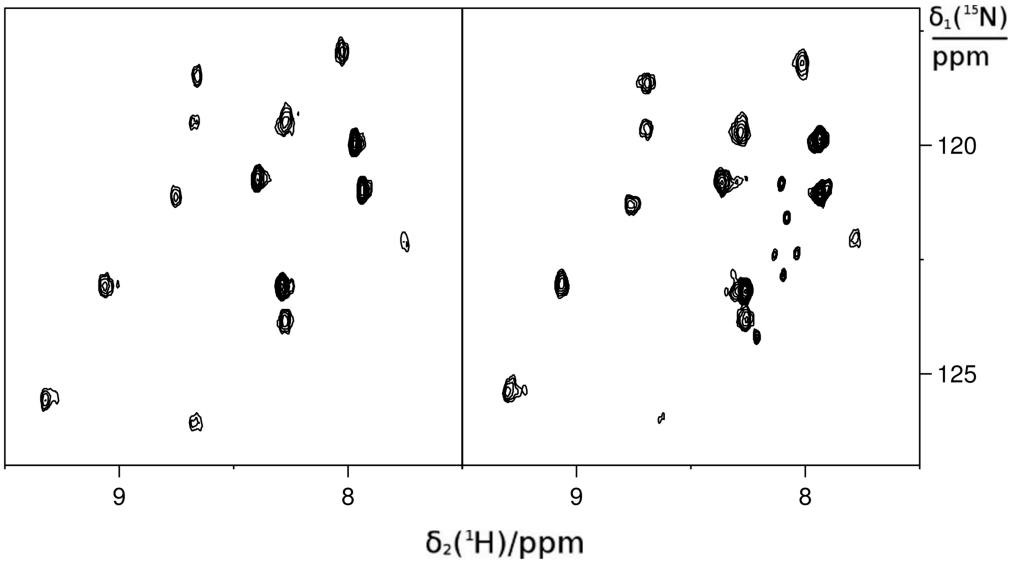


Figure S5. ^{15}N-HSQC spectra of DENp selectively labelled with ^{15}N-isoleucine at different ionic strengths. Both spectra were acquired using 0.2 mM protein solutions in 90% H_2O/10% D_2O with 20 mM MES, pH 6.5, at 25 ^{o}C on a Bruker 800 MHz NMR spectrometer. (A) ^{15}N-HSQC spectrum recorded in the presence of 50 mM NaCl. The spectrum is the same as the spectrum in Figure 3A. (B) Same as (A), except that the buffer contained 300 mM NaCl.

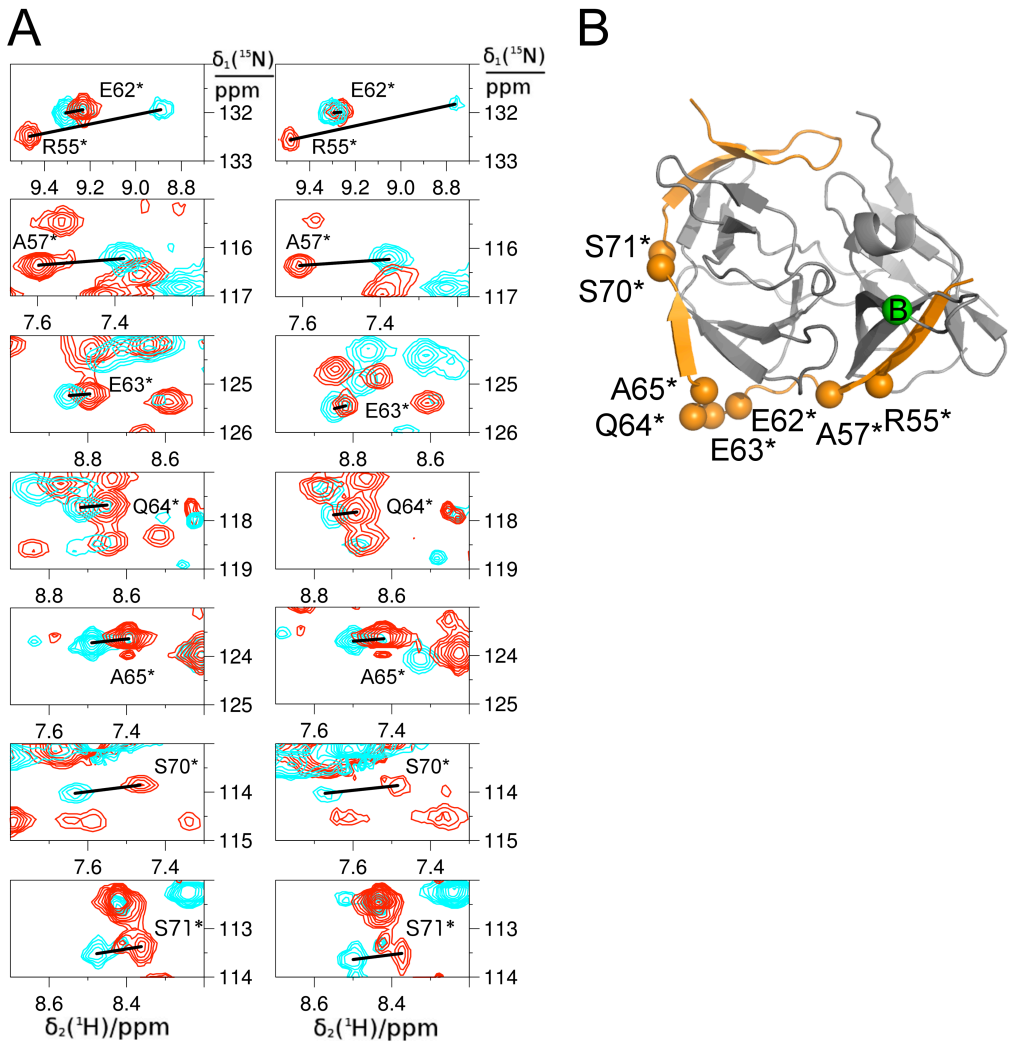


Figure S6. PCSs observed for NS2B of DENp at 50 mM and 300 mM NaCl. The spectra were recorded using the same conditions as in Figure 3. (A) Superimpositions of selected spectral regions of ^{15}N-HSQC spectra of mutant B of DENp with the C2 tag at 50 mM NaCl (left panels) and 300 mM NaCl (right panels). The samples were tagged with the C2-tag loaded with paramagnetic Tb^{3+} (cyan peaks) or diamagnetic Y^{3+} (red peaks). The resonance assignments of the selected cross-peaks are indicated. (B) Cartoon representation of the dengue virus protease in the closed conformation (PDB ID 3U1I [7]), showing NS2B in orange and NS3pro in grey. The locations of the residues analysed in (A) are highlighted by balls and the site of the C2-tag is indicated by a green sphere.

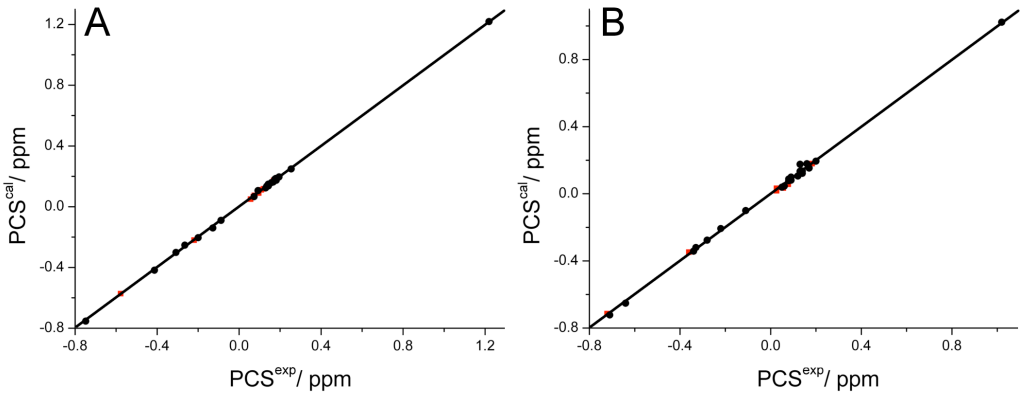


Figure S7. Correlations between back-calculated and experimental PCSs of DENp at low and high salt. The PCSs were measured for the backbone amide protons of a uniformly ^{15}N-labelled cysteine mutant of DENp with the cysteine positioned at site B (see Figure 1) and ligated with either a paramagnetic C2-Tb^{3+} or a diamagnetic C2-Y^{3+} tag. The PCSs of NS2B and NS3pro are shown as red squares and black circles, respectively. NS2B appears to be located in the same position at high and low salt, as this point fits equally well to the correlation line under both conditions. (A) Data at 50 mM NaCl. (B) Data at 300 mM NaCl.

Table S1. PCSs of backbone amide protons of ^{15}N-labelled DENp mutant B with C2-tag loaded with Tb^{3+} at 50 mM and 300 mM NaCl concentrations[a]

Residue[b]	PCS in low salt (ppm)	PCS in high salt (ppm)
Arg55*	-0.578	-0.722
Ala57*	-0.220	-0.361
Glu62*	0.075	0.026
Glu63*	0.056	0.026
Gln64*	0.070	0.056
Ala65*	0.096	0.078
Ser70*	0.169	0.183
Ser71*	0.119	0.140
Ser75*	0.108	-
Gly21	-0.200	-0.280
Gly39	-0.413	-0.640
Thr41	-0.265	-0.340
Thr48	-0.128	-0.220
His51	1.219	1.020
Gly62	-0.748	-0.710
Asp71	0.254	0.200
Ile77	0.092	0.060
Tyr79	-0.308	-0.330
Gly82	-0.088	-0.110
Leu98	0.175	0.090
Ala108	0.142	0.080
Gln110	0.166	0.140
Thr111	0.158	0.130
Gly114	0.179	0.130
Phe116	0.195	0.160
Gly121	0.141	0.140
Thr122	0.130	0.120
Ile123	0.181	0.170
Val140	0.135	0.090
Lys142	0.073	0.050
Gly148	0.143	0.080

[a] Other parameters were 20 mM MES, pH 6.5, 800 MHz ^{1}H NMR frequency. PCSs were calculated as the difference in chemical shifts measured with paramagnetic Tb^{3+} minus the chemical shifts with diamagnetic Y^{3+}.

[b] NS2B residues are identified by an asterisk.

Table S2. $\Delta\chi$ tensor parameters of DENp mutant B with C2-Tb^{3+} tag[a]

	$\Delta\chi_{ax}$	$\Delta\chi_{rh}$	x	y	z	α	β	γ
Low salt	-16.1	-1.5	29.2	34.0	-4.2	7.7	48.4	56.9
	(0.6)	(0.3)	(0.2)	(0.2)	(0.3)	(1.2)	(1.0)	(8.7)
High salt	-13.7	-1.4	30.4	34.2	-4.0	20.0	51.2	26.0
	(1.1)	(0.4)	(0.3)	(0.3)	(0.4)	(2.1)	(1.7)	(21.5)

[a] The tensor parameters of NS2B-NS3pro complex were calculated with the program Numbat [8], using the PCSs of Table S1 and the model of the closed conformation established previously [4]. This model deviates from the crystal structure of serotype 3 DENpro 3U1I [7] with a backbone RMSD of 0.7 Å. The axial and rhombic components of the $\Delta\chi$ tensors are given in 10^{-32} m^3 and the Euler angles in degrees, using the zyz convention and unique tensor representation [8]. Error estimates (in brackets) were determined from fits obtained by randomly omitting 10% of the PCS data.

References

1 Leung D, Schroder K, White H, Fang NX, Stoermer MJ, Abbenante G, Martin JL, Young PR & Fairlie DP (2001) 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.

2 Chase T & Shaw E (1967) p-Nitrophenyl-p'-guanidinobenzoate HCl: a new active site titrant for trypsin. *Biochem Biophys Res Commun* **29**, 508-514.

3 Scheiner CJ & Quigley JP (1982) Detection, identification and partial characterization of serine proteases and esterases in biological systems. *Analyt Biochem* **122**, 58-69.

4 de la Cruz L, Nguyen THD, Ozawa K, Shin J, Graham B, Huber T & Otting G (2011) Binding of low-molecular weight inhibitors promotes large conformational changes in the dengue virus NS2B-NS3 protease: fold analysis by pseudocontact shifts. *J Am Chem Soc* **133**, 19205-19215.

5 Graham B, Loh CT, Swarbrick JD, Ung P, Shin J, Yagi H, Jia X, Chhabra S, Pintacuda G, Huber T & Otting G (2011) A DOTA-amide lanthanide tag for reliable generation of pseudocontact shifts in protein NMR spectra. *Bioconj Chem* **22**, 2118-2125.

6 Löhr K, Knox JE, Phong WY, Ma NL, Yin Z, Sampath A, Patel SJ, Wang WL, Chan WL, Ranga Rao KR, Wang G, Vasudevan SG, Keller TH & Lim SP (2007) Yellow fever virus NS3 protease: peptide-inhibition studies. *J Gen Virol* **88**, 2223-2227.

7 Noble CG, Seh CC, Chao AT & Shi PY (2012) Ligand-bound structures of the dengue virus protease reveal the active conformation. *J Virol* **86**, 438-446.

8 Schmitz C, Stanton-Cook MJ, Su XC, Otting G & Huber T (2008) Numbat: an interactive software tool for fitting $\Delta\chi$ tensors to molecular coordinates using pseudocontact shifts. *J Biomol NMR* **41**, 179-189.

FEBS Letters 588 (2014) 2206–2211

journal homepage: www.FEBSLetters.org

The dengue virus NS2B–NS3 protease retains the closed conformation in the complex with BPTI

Wan-Na Chen [a], Karin V. Loscha [a], Christoph Nitsche [b], Bim Graham [c], Gottfried Otting [a,*]

[a] Australian National University, Research School of Chemistry, Canberra, ACT 0200, Australia
[b] Medicinal Chemistry, Institute of Pharmacy and Molecular Biotechnology, Heidelberg University, Im Neuenheimer Feld 364, 69120 Heidelberg, Germany
[c] Medicinal Chemistry and Drug Action, Monash Institute of Pharmaceutical Sciences, Parkville, VIC 3052, Australia

ARTICLE INFO

Article history:
Received 13 March 2014
Revised 16 April 2014
Accepted 7 May 2014
Available online 21 May 2014

Edited by Christian Griesinger

Keywords:
Bovine pancreatic trypsin inhibitor
Dengue virus protease
Lanthanide tag
NMR spectroscopy
Pseudocontact shift

ABSTRACT

The C-terminal β-hairpin of NS2B (NS2Bc) in the dengue virus NS2B–NS3 protease is required for full enzymatic activity. In crystal structures without inhibitor and in the complex with bovine pancreatic trypsin inhibitor (BPTI), NS2Bc is displaced from the active site. In contrast, nuclear magnetic resonance (NMR) studies in solution only ever showed NS2Bc in the enzymatically active closed conformation. Here we demonstrate by pseudocontact shifts from a lanthanide tag that NS2Bc remains in the closed conformation also in the complex with BPTI. Therefore, the closed conformation is the best template for drug discovery.

Structured summary of protein interactions:
NS2B, **BPTI** and **NS3** physically interact by nuclear magnetic resonance (View interaction)

1. Introduction

Dengue virus (DENV) is the most important mosquito-borne pathogen in terms of human suffering and cost, with a high rate of hospitalization and potentially deadly outcome [1]. There are four different closely related serotypes (DENV-1–DENV-4). To date, no approved vaccine or drug is available for any of them, but an established drug target is presented by the NS2B–NS3 protease (NS2B–NS3pro) formed from segments of the non-structural virus proteins NS2B and NS3 [2,3]. For a long time, however, development of an inhibitor of NS2B–NS3pro has been hampered by difficulties to ascertain the correct structure of the protein.

The first crystal structure of NS2B–NS3pro, determined for a construct from DENV-2, displayed an open conformation in which the C-terminal segment of NS2B (NS2Bc; residues 66*–95*; throughout this text, residue numbers of NS2B are identified by asterisks) was located far from the active site in an inactive conformation (Fig. 1) [4]. Subsequent NMR studies showed, however, that the protease assumes a closed conformation in solution, where NS2Bc lines the substrate binding site [5–7], in complete analogy to the structure of the related West Nile virus protease [4,8,9]. A crystal structure of NS2B–NS3pro from DENV-3 in complex with a peptide inhibitor confirmed the closed conformation [10], but the same protein in complex with the high-affinity (K_i = 26 nM [11]) trypsin inhibitor bovine pancreatic trypsin inhibitor (BPTI) lacked any electron density for NS2Bc, indicating once again an open conformation [10].

In the absence of firm information about the correct target structure in solution, computational ligand binding studies targeting the active site used either closed or open conformations of the NS2B–NS3 protease [12–28]. Unfortunately, the actual three-dimensional (3D) structure in solution cannot readily be determined by conventional NMR methods due to poor spectral resolution, line broadening by conformational exchange and limited protein stability. In the case of the NS2B–NS3pro complex with BPTI, its high molecular weight (35 kDa) further impedes NMR analysis.

Here we show that, in solution, the closed conformation prevails also in the presence of BPTI. This result was obtained by measuring pseudocontact shifts (PCS) of ^{15}N-HSQC cross-peaks of backbone amides in uniformly and selectively ^{15}N-labelled samples with a paramagnetic lanthanide tag attached to NS3pro. The PCSs provide a clear picture of the binding mode of BPTI at the active site and of the location of the β-hairpin of NS2Bc.

Abbreviations: DENV, dengue virus; WNV, West Nile virus; DENpro, dengue virus protease; PCS, pseudocontact shift; BPTI, bovine pancreatic trypsin inhibitor; DsbC, disulfide bond isomerase C

* Corresponding author. Fax: +61 (0)2 61250750.
E-mail address: gottfried.otting@anu.edu.au (G. Otting).

http://dx.doi.org/10.1016/j.febslet.2014.05.018
0014-5793/

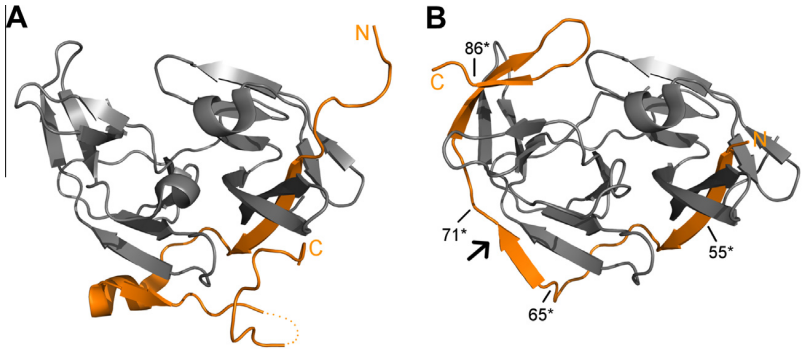


Fig. 1. Ribbon drawings of the dengue virus NS2B–NS3 protease as observed in crystal structures. NS3pro is shown in grey and NS2B in orange. The active site is located near the centre top in the orientation shown. The N-terminal segment of NS2B (NS2Bn, residues 49*–65*) includes a structurally conserved β-strand that inserts into the N-terminal β-barrel of NS3pro. The C-terminal segment of NS2B (NS2Bc, residues 66*–95*) differs greatly between the open (panel A, PDB ID: 2FOM) and closed (panel B, PDB ID: 3U1I [10]) conformations detected by X-ray crystallography [4,10]. The arrow identifies the last residue of NS2B, for which electron density was observed in the complex with BPTI (PDB ID: 3U1J [10]). The positions of selected residues are indicated.

2. Materials and methods

2.1. Expression plasmids

The DENV-2 NS2B–NS3 protease was produced without a covalent link between NS2B and NS3pro, using a previously described construct in which NS2B is separated from NS3pro by the natural protease recognition sequence EVKKQR [7]. The entire construct consists of the T7 gene 10N-terminal peptide MASMTG followed by a two-residue cloning artifact (Leu-Glu), 47 residues from NS2B, the cleavage sequence EVKKQR, 185 residues from NS3pro and a C-terminal His_6-tag. The gene was inserted into the pETMCSI expression vector [29] and the plasmid was transformed into the *Escherichia coli* strain Rosetta::λDE3/pRARE. An additional construct contained the single-cysteine mutation S68C in NS3pro for subsequent ligation with lanthanide tags [5]. The nucleotide sequence of BPTI [30] was also subcloned into the pETMCSI T7 expression vector and transformed into the *E. coli* strain TOP10 (Life Technologies, Carlsbad, CA, USA). The N-terminus was preceded by a His_6-tag followed by the tobacco etch virus (TEV) protease cleavage sequence ENLYFQG.

2.2. Protein sample preparation

Uniformly and selectively ^{15}N-labelled samples of the protease were made by high-cell density in vivo expression [31] and cell-free synthesis [32–34], respectively, and purified as described previously [7]. BPTI samples selectively labelled with ^{15}N-labelled Ala, Arg, Lys, Thr, Leu, Phe and Ile (Cambridge Isotope Laboratories, Andover, MA, USA; ISOTEC, St. Louis, MO, USA) were prepared by cell-free synthesis at 30 °C overnight following a published protocol that includes His_6-tagged disulfide bond isomerase C (DsbC) to catalyze correct disulfide bond formation [30]. The reaction volume was 1 mL in a dialysis tube of Spectra/Por 2 (MWCO: 12–14 kDa) suspended in 10 mL outer buffer. The protein was purified by diluting the supernatant of the cell-free reaction with 2 times buffer A (50 mM HEPES, pH 7.5, 300 mM NaCl), loading the mixture onto a Ni-NTA agarose spin column (Qiagen, Hilden, Germany), washing with buffer A plus 30 mM imidazole and eluting with buffer A plus 300 mM imidazole. BPTI was separated from DsbC by a spin column containing SP-650 M resin (TOSOH, Minato, Japan), using buffer B (50 mM HEPES, pH 8.0) with 50 mM NaCl for loading, buffer B with 100 mM NaCl for washing and buffer B with 1 M NaCl for elution. All purified proteins were exchanged into NMR buffer (20 mM MES, pH 6.5, 50 mM NaCl, 10% D_2O). Protein concentrations were determined by their UV absorbance at 280 nm.

Samples of the protease–BPTI complex were prepared either with ^{15}N-labelled protease and unlabelled BPTI or *vice versa*. Tagging of the S68C mutant of the protease was performed after formation of the 1:1 complex with BPTI by addition to a 3-fold excess of the C2 tag loaded with either Y^{3+} or Tb^{3+} [35,5]. Following 5 h incubation at room temperature, the complex was washed with NMR buffer using a centrifugal filter unit (Amicon Ultra with a MWCO of 3 kDa; Millipore, Billerica, USA).

2.3. NMR spectroscopy

All NMR spectra were recorded of 0.1–0.2 mM solutions of the protease-BPTI complex in NMR buffer at 25 °C, using a Bruker 800 MHz NMR spectrometer equipped with a TCI cryoprobe. PCSs were measured in ppm as the amide proton chemical shifts measured in the presence of a paramagnetic lanthanide minus the chemical shift observed in the presence of diamagnetic Y^{3+}.

2.4. Δχ-tensor fitting

The PCS of a nuclear spin, $\Delta\delta^{PCS}$, is measured in ppm as the difference in chemical shift between a sample with a paramagnetic ion (Tb^{3+} in the present work) and a sample with a diamagnetic ion. PCSs arise from an anisotropic magnetic susceptibility ($\Delta\chi$) tensor associated with the unpaired electrons of the paramagnetic ion. PCSs follow the equation [36]:

$$\Delta\delta^{PCS} = 1/(12\pi r^3)[\Delta\chi_{ax}(3\cos^2\theta - 1) + 1.5\Delta\chi_{rh}\sin^2\theta\cos 2\phi] \quad (1)$$

where $\Delta\delta^{PCS}$ is the PCS, r is the distance of the nuclear spin from the metal ion, $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ are the axial and rhombic components of the $\Delta\chi$ tensor, and θ and ϕ are the polar angles describing the position of the nuclear spin with respect to the principal axes of the $\Delta\chi$ tensor. PCSs from lanthanide tags can be measured for nuclear spins over 40 Å from the metal ion [37].

The PCSs measured for the complex of NS2B–NS3pro with BPTI were used to determine the position and orientation of the magnetic susceptibility anisotropy ($\Delta\chi$) tensors of the paramagnetic ions relative to the protein structure. The program Numbat [38] was used to fit the $\Delta\chi$ tensors to the crystal structure of the complex with BPTI (PDB ID: 3U1J [10]). As no electron density of NS2Bc was observed in the structure 3U1J, the coordinates of NS2Bc in the crystal structure with the peptide inhibitor Bz-nKRR-H (PDB ID: 3U1I [10]) were grafted onto the 3U1J structure for a complete model of the closed conformation.

3. Results

3.1. The complex of the protease with BPTI is stable

Producing the protease as a single polypeptide chain with a self-cleavage site between NS2B and NS3pro ensured a precise 1:1 ratio of both components in the final protease. Cell-free synthesis of BPTI produced correctly folded, selectively ^{15}N-labelled protein (Fig. S1), yielding well-resolved ^{15}N-HSQC cross-peaks that were straightforward to assign to individual residues by comparison with resonance assignments of the free protein (Fig. 2B).

In our hands, samples of the protease without covalent link between NS2B and NS3pro are stable for only a few hours at room temperature in the absence of inhibitor. Sample degradation results in additional narrow cross-peaks from NS2B at random coil chemical shifts which, previously, had erroneously been attributed to open conformations [39,40,5]. Notably, ^{15}N-HSQC spectra of the complex with BPTI were free of these additional cross-peaks and stable for at least a month, indicating that BPTI inhibits the degradation of the protease (Fig. S2).

3.2. The binding mode of BPTI in solution is the same as in the crystal

Fig. 2 illustrates the PCSs observed for the complex of the S68C mutant of the protease with BPTI, with the C2-Tb^{3+} tag ligated to Cys68. The same construct ligated with diamagnetic C2-Y^{3+} served as the diamagnetic reference. PCSs of NS2B–NS3pro were observed in the sample of ^{15}N-labelled protease with unlabelled BPTI (Fig. 2A), while PCSs of BPTI were observed in the sample of selectively ^{15}N-labelled BPTI with unlabelled protease (Fig. 2B). Using the previously established assignment of the protease [5], 21 PCSs could be resolved for NS3pro and residues 55*–65* of NS2Bn, and 8 PCSs for NS2Bc (Table S1). Owing to the better spectral resolution in the selectively ^{15}N-labelled BPTI sample, PCSs could be assigned and accurately measured for 23 residues of the BPTI construct.

To verify that BPTI binds to the protease in solution as reported by the crystal structure 3U1J, we used the PCSs from NS3pro and NS2Bn to fit an initial $\Delta\chi$ tensor to the crystal structure (Fit 1 in Table 1). The good quality of the fit is illustrated by the close correlation between experimental and back-calculated PCSs (Fig. 3A). Using the fitted $\Delta\chi$ tensor together with the crystal structure coordinates of the DENV-3 protease-BPTI complex (PDB ID: 3U1I [10]) to predict the PCSs of BPTI reproduced the expected trend (Fig. 3C). This shows that the inhibitor binds in solution as reported by the crystal structure (Fig. 3E). This conclusion is unaffected by the observation that the back-calculated PCSs of BPTI were consistently more positive than the experimental PCSs (Fig. 3C). Discrepancies of this magnitude are common for mobile tags like the C2 tag, as translational motions of the metal ion compromise the accuracy with which the PCS data can be approximated by a single effective $\Delta\chi$ tensor [41].

Table 1
$\Delta\chi$-tensor parameters of the protease–BPTI complex with the C2-Tb^{3+} tag attached to residue 68 of NS3pro[a].

	$\Delta\chi_{ax}$	$\Delta\chi_{rh}$	x	y	z	α	β	γ
Fit 1[b]	−12.3	−4.0	−16.098	−34.428	3.681	146.1	48.9	30.1
	(0.1)	(0.6)	(0.3)	(0.3)	(0.2)	(2.5)	(1.2)	(4.7)
Fit 2[c]	−12.9	−1.9	−16.800	−33.423	3.582	144.0	50.9	67.7
	(0.2)	(0.3)	(0.1)	(0.3)	(0.1)	(1.5)	(0.6)	(3.1)

[a] $\Delta\chi$-tensor parameters determined by the program Numbat using the crystal structure 3U1J. The axial and rhombic components of the $\Delta\chi$ tensors are given in 10^{-32} m^3 and the Euler angles in degrees, using the *zyz* convention and unique tensor representation [38]. Uncertainties (in brackets) were estimated by randomly omitting 10% of the PCS data in multiple tensor fits.

[b] Fit 1 used only PCSs from NS3pro and residues 55*–65* of NS2B. The root mean square deviation (rmsd) between the experimental and back-calculated PCSs was 0.014 ppm for the residues included in the fit.

[c] Fit 2 used only PCSs from NS3pro, residues 55*–65* of NS2B, and PCSs from those BPTI amides, which are within 30 Å from the paramagnetic metal centre determined by Fit 1. The rmsd between the experimental and back-calculated PCSs was 0.029 ppm for the residues included in the fit.

3.3. NS2Bc retains the closed conformation in the presence of BPTI

To obtain more accurate predictions for the PCSs of NS2Bc in the closed conformation, we accepted the correctness of the crystal structure 3U1I [10] with regard to the position of BPTI and determined a new $\Delta\chi$ tensor using amide PCSs of residues from NS3pro, NS2Bn and, in addition, BPTI, so long as they were within 30 Å of the metal ion position determined by Fit 1. The resulting fit (Fit 2 in Table 1) still produced good agreement between back-calculated and experimental PCSs (Fig. 3B) and accurately predicted the PCSs also for BPTI residues located further than 30 Å from the metal centre (Fig. S3B). Most important, however, the PCSs predicted for NS2Bc in the closed conformation of the crystal structure (3U1I [10]) very closely matched the experimental data (Fig. 3D). This indicates that NS2Bc is in the closed conformation as in the crystal

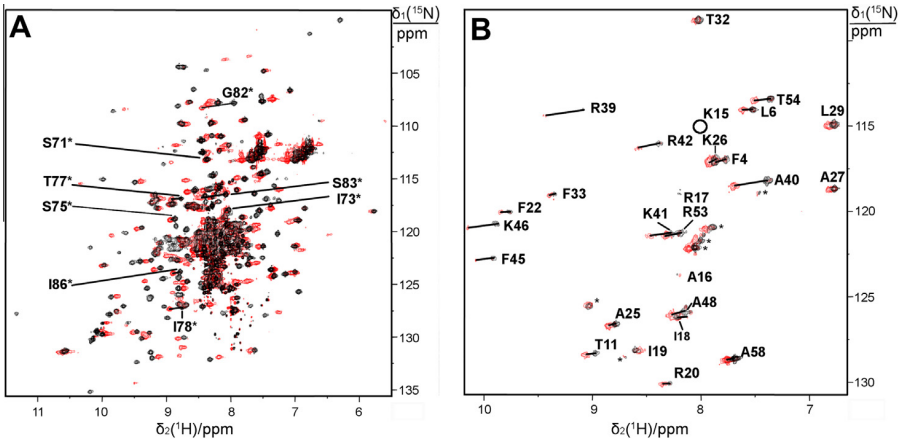


Fig. 2. PCSs observed for the complex of NS2B–NS3 protease with BPTI. (A) Superimposition of ^{15}N-HSQC spectra of uniformly ^{15}N-labelled protease S68C, tagged with either C2-Tb^{3+} (red spectrum) or C2-Y^{3+} (black spectrum), in complex with unlabelled BPTI. PCSs of NS2Bc residues are labelled. (B) Same as (A), except that the protease is at natural isotopic abundance and BPTI is selectively labelled with ^{15}N-Ala, Arg, Lys, Thr, Leu, Phe and Ile. PCSs of BPTI are identified by lines. Cross-peaks of residues 15 (circle) and 17 were observable only at the noise level. Asterisks identify unassigned resonances attributed to the modified N-terminus and the nearby C-terminus.

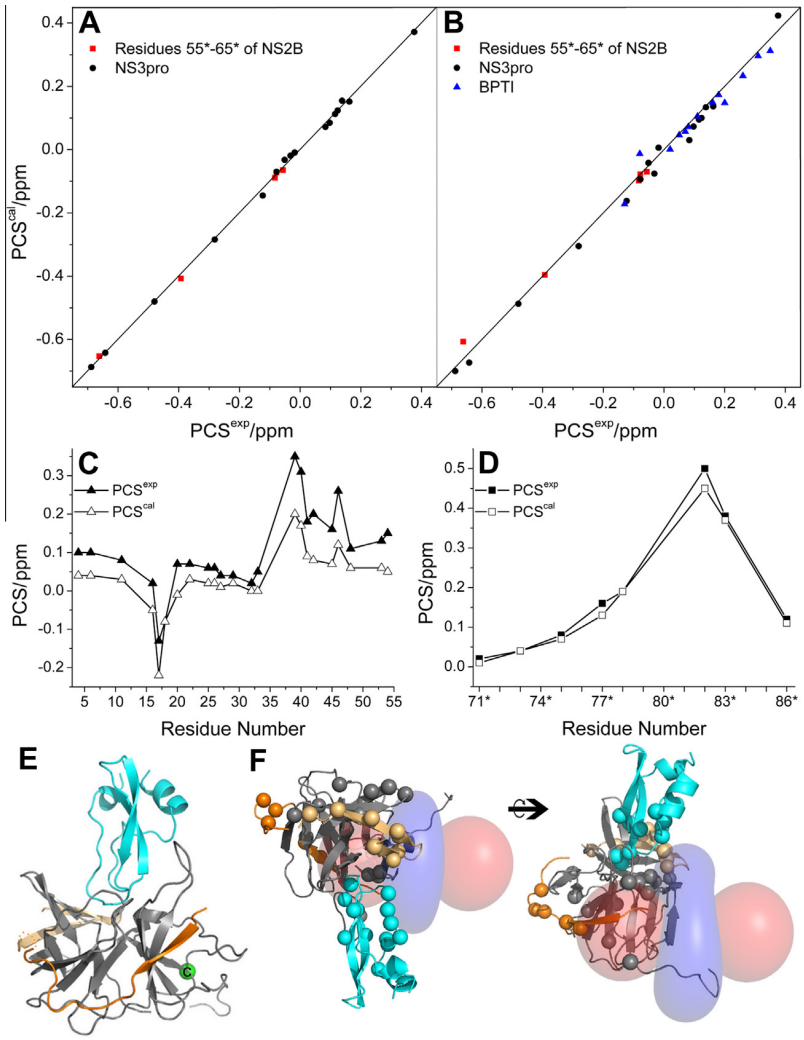


Fig. 3. Correlation between experimental and calculated PCSs for the complex of NS2B–NS3pro with BPTI. (A) Back-calculated versus experimental PCSs produced by Fit 1 of Table 1. Data of NS2B (residues 55*–65*) and NS3pro are shown as red squares and black circles, respectively. (B) Same as (A), except showing the data from Fit 2. PCSs from BPTI are shown as blue triangles. (C) PCSs of BPTI predicted by the $\Delta\chi$ tensor of Fit 1 versus the amino-acid sequence of BPTI. Experimental and calculated data are shown as black and white triangles, respectively. (D) PCSs of NS2Bc predicted by the $\Delta\chi$ tensor of Fit 2 versus the amino-acid sequence of NS2Bc. Experimental and calculated data are shown as black and white squares, respectively. (E) Cartoon representation of the structure 3U1J with NS2Bc copied from the structure 3U1I [10]. The N-terminal segment of NS2B is shown in orange, NS2Bc in light orange, NS3pro in grey and BPTI in cyan. The link between residues 69* and 70* is indicated by dots. The green sphere highlights the site of the cysteine residue carrying the lanthanide tag (residue 68). (F) PCS isosurfaces (±0.5 ppm) generated by the $\Delta\chi$ tensor of Fit 2 (Table 1) plotted on the structure of the protease–BPTI complex. Blue and red surfaces identify, respectively, positive and negative PCSs. The amides for which PCSs were measured and used in the fit are identified by spheres. Light orange balls illustrate amides of NS2Bc for which experimental PCSs were measured. The same data are displayed in two different orientations.

structure with the peptide inhibitor (3U1I [10]) and as determined by NMR in the absence of an inhibitor [7]. Using the PCSs to fit a $\Delta\chi$ tensor to NS2Bn and NS3pro in the open conformation of Fig. 1A (2FOM [4]) predicted completely wrong PCSs for NS2Bc, confirming that this conformation cannot be populated to any significant extent in the complex with BPTI (Fig. S4).

Our results suggest that the open conformation observed in the crystal structure of the protease–BPTI complex could be a crystallization artifact. Indeed, inspection of intermolecular contacts in the crystal lattice show that a symmetry-related molecule occupies much of the space where the β-hairpin of NS2Bc would be expected to be in the closed conformation (Fig. S5).

4. Discussion

The present work demonstrates that the C-terminal β-hairpin of NS2B retains the closed conformation in solution also when BPTI is bound to the dengue virus NS2B–NS3 protease. Previous solution NMR data showed that the protease assumes the closed conformation also in the absence of an inhibitor [6,7]. Without exception, the closed conformation thus is the overwhelmingly predominant conformation in solution, regardless of the presence or absence of inhibitors of high or low molecular mass. These results are in stark contrast to the crystal structure determinations of the NS2B–NS3 protease, which reported NS2Bc exclusively in open conformations except for a single case in which a small peptide inhibitor was bound to the protease [4,42,10]. Our results have important implications for rational drug design: while the crystallizations indicate that the association between NS2Bc and NS3pro may be weak, the closed conformation, as observed in the crystal structure with a small peptide inhibitor [10], is the most faithful representation of the protease structure to use in drug development efforts.

Notably, the NS2B–NS3 protease is subject to conformational changes, as indicated by non-uniform peak intensities in the

^{15}N-HSQC spectra (Fig. 2A). High pH (7.5 versus 6.5) or high concentrations of NaCl (300 mM) in the absence of any inhibitor have been observed to broaden the signals of NS2Bc after Ser71* beyond detection, which can be attributed to an equilibrium between open and closed conformations (with residues preceding Ser71* retaining their location in the closed state) or to conformational changes within the closed state [7]. The non-uniform peak intensities in the presence of BPTI indicate that the inhibitor does not eliminate all conformational exchange processes in the protease. As the PCSs of NS2Bc closely match the expectations for the closed conformation and we could not observe signals of an additional conformation, we estimate that at most 10% of NS2Bc in an open conformation could have escaped detection.

On a technical note, our data illustrate the power of pseudocontact shifts for the structure analysis of proteins and protein–ligand complexes in aqueous solution, even if the lanthanide tag is attached via a fairly flexible linker. An earlier analysis of the C1 tag (which is the opposite enantiomer of the C2 tag used here) attached to the NS2B–NS3 protease concluded that the mobility of the tag can compromise the prediction of accurate PCSs of a ligand, if a single effective $\Delta\chi$ tensor is fitted to all experimental PCS data of the protein [41]. This effect arises from the fact that, in principle, displacements of the metal requires the fitting of multiple $\Delta\chi$ tensors for accurate PCS back-calculations. Much improved PCS predictions are already expected, however, if a single effective $\Delta\chi$ tensor is fitted to a subset of PCSs measured for nuclear spins in the vicinity of the ligand. This expectation is borne out by the experimental data of the present work, as the PCSs predicted for NS2Bc matched the experimental results much better, when PCSs from nuclear spins of BPTI were included to fit the $\Delta\chi$ tensor while excluding those BPTI amides that were located far from the metal site. In this way, the approximation of the PCS data by a single effective $\Delta\chi$ tensor resulted in a good fit of PCSs for nuclear spins in the vicinity of NS2Bc (Fig. 3F), producing a meaningful prediction of PCSs for NS2Bc and comparison with experimental PCSs.

In conclusion, our results stress the stability of the closed conformation of the dengue virus NS2B–NS3 protease, at least for serotype 2 which is the serotype with the highest global growth rate over the last decade [43]. NMR studies of the other serotypes will be required to assess the effect of sequence differences on the prevalence of open conformations. Our results also highlight the way in which PCSs induced by fairly mobile lanthanide tags can be used to obtain useful structural information in solution for protein systems that are difficult to analyse in detail by conventional techniques.

Author contributions

W.N.C., K.L. and G.O. designed experiments. K.L., C.N. and B.G. generated reagents. W.N.C. performed experiments. W.N.C. and G.O. wrote and edited the manuscript.

Acknowledgements

We thank Prof. Takanori Kigawa for expression constructs of BPTI and DsbC. W.-N. C. thanks the government of the People's Republic of China for a CSC scholarship and C.N. thanks the Studienstiftung des deutschen Volkes for a fellowship. Financial support by the Australian Research Council is gratefully acknowledged.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at http://dx.doi.org/10.1016/j.febslet.2014.05.018.

References

[1] Bhatt, S., Gething, P.W., Brady, O.J., Messina, J.P., Farlow, A.W., Moyes, C.L., Drake, J.M., Brownstein, J.S., Hoen, A.G., Sankoh, O., Myers, M.F., George, D.B., Jaenisch, T., Wint, G.R., Simmons, C.P., Scott, T.W., Farrar, J.J. and Hay, S.I. (2013) The global distribution and burden of dengue. Nature 496, 504–507.
[2] Lescar, J., Luo, D., Xu, T., Sampath, A., Lim, S.P., Canard, B. and Vasudevan, S.G. (2008) Towards the design of antiviral inhibitors against flaviviruses: the case for the multifunctional NS3 protein from dengue virus as a target. Antiviral Res. 80, 94–101.
[3] Rothan, H.A., Han, H.C., Ramasamy, T.S., Othman, S., Rahman, N.A. and Yusof, R. (2012) Inhibition of dengue NS2B–NS3 protease and viral replication in Vero cells by recombinant retrocyclin-1. BMC Infect. Dis. 12, 314.
[4] Erbel, P., Schiering, N., D'Arcy, A., Renatus, M., Kroemer, M., Lim, S.P., Yin, Z., Keller, T.H., Vasudevan, S.G. and Hommel, U. (2006) Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. Nat. Struct. Mol. Biol. 13, 372–373.
[5] de la Cruz, L., Nguyen, T.H., Ozawa, K., Shin, J., Graham, B., Huber, T. and Otting, G. (2011) Binding of low molecular weight inhibitors promotes large conformational changes in the dengue virus NS2B–NS3 protease: fold analysis by pseudocontact shifts. J. Am. Chem. Soc. 133, 19205–19215.
[6] Kim, Y.M., Gayen, S., Kang, C., Joy, J., Huang, Q., Chen, A.S., Wee, J.L., Ang, M.J., Lim, H.A., Hung, A.W., Li, R., Noble, C.G., Lee, L.T., Yip, A., Wang, Q.Y., Chia, C.S., Hill, J., Shi, P.Y. and Keller, T.H. (2013) NMR analysis of a novel enzymatically active unlinked dengue NS2B–NS3 protease complex. J. Biol. Chem. 288, 12891–12900.
[7] de la Cruz, L., Chen, W.N., Graham, B. and Otting, G. (2014) 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.
[8] Aleshin, A.E., Shiryaev, S.A., Strongin, A.Y. and Liddington, R.C. (2007) Structural evidence for regulation and specificity of flaviviral proteases and evolution of the Flaviviridae fold. Protein Sci. 16, 795–806.
[9] Robin, G., Chappell, K., Stoermer, M.J., Hu, S., Young, P.R., Fairlie, D.P. and Martin, J.L. (2009) Structure of West Nile virus NS3 protease: ligand stabilization of the catalytic conformation. J. Mol. Biol. 385, 1568–1577.
[10] Noble, C.G., Seh, C.C., Chao, A.T. and Shi, P.Y. (2012) Ligand-bound structures of the dengue virus protease reveal the active conformation. J. Virol. 86, 438–446.
[11] Mueller, N.H., Yon, C., Ganesh, V.K. and Padmanabhan, R. (2007) Characterization of the West Nile virus protease substrate specificity and inhibitors. Int. J. Biochem. Cell Biol. 39, 606–614.
[12] Yin, Z., Patel, S.J., Wang, W.L., Chan, W.L., Rao, K.R.R., Wang, G., Ngew, X., Patel, V., Beer, D., Knox, J.E., Ma, N.L., Ehrhardt, C., Lim, S.P., Vasudevan, S.G. and Keller, T.H. (2006) Peptide inhibitors of dengue virus NS3 protease. Part 2: SAR study of tetrapeptide aldehyde inhibitors. Bioorg. Med. Chem. Lett. 16, 40–43.
[13] Tomlinson, S.M. and Watowich, S.J. (2011) Anthracene-based inhibitors of dengue virus NS2B–NS3 protease. Antiviral Res. 89, 127–135.
[14] Knehans, T., Schüller, A., Doan, D.N., Nacro, K., Hill, J., Güntert, P., Madhusudhan, M.S., Weil, T. and Vasudevan, S.G. (2011) Structure-guided fragment-based *in silico* drug design of dengue protease inhibitors. J. Comput. Aided Mol. Des. 25, 263–274.
[15] Schüller, A., Yin, Z., Chia, C.S.B., Doan, D.N., Kim, H.K., Shang, L., Loh, T.P., Hill, J. and Vasudevan, S.G. (2011) Tripeptide inhibitors of dengue and West Nile virus NS2B–NS3 protease. Antiviral Res. 92, 96–101.
[16] Doan, D.N., Li, K.Q., Basavannacharya, C., Vasudevan, S.G. and Madhusudhan, M.S. (2012) Transplantation of a hydrogen bonding network from West Nile virus protease onto dengue-2 protease improves catalytic efficiency and sheds light on substrate specificity. Protein Eng. Des. Sel. 25, 843–850.
[17] Xu, S., Li, H., Shao, X., Fan, C., Ericksen, B., Liu, J., Chi, C. and Wang, C. (2012) Critical effect of peptide cyclization on the potency of peptide inhibitors against dengue virus NS2B–NS3 protease. J. Med. Chem. 55, 6881–6887.
[18] Ghanesh, V.K., Muller, N., Judge, K., Luan, C.H., Padmanabhan, R. and Murthy, K.H.M. (2005) Identification and characterization of nonsubstrate based inhibitors of the essential dengue and West Nile virus proteases. Bioorg. Med. Chem. 13, 257–264.
[19] Frecer, V. and Miertus, S. (2010) Design, structure-based focusing and *in silico* screening of combinatorial library of peptidomimetic inhibitors of dengue virus NS2B–NS3 protease. J. Comput. Aided Mol. Des. 24, 195–212.
[20] Tambunan, U.S.F. and Alamudi, S. (2010) Designing cyclic peptide inhibitor of dengue virus NS3–NS2B protease by using molecular docking approach. Bioinformation 5, 250–254.
[21] Frimayanti, N., Chee, C.F., Zain, S.M. and Rahman, N.A. (2011) Design of new competitive dengue NS2B/NS3 protease inhibitors – a computational approach. Int. J. Mol. Sci. 12, 1089–1100.
[22] Lai, H., Prasad, G.S. and Padmanabhan, R. (2013) Characterization of 8-hydroxyquinoline derivatives containing aminobenzothiazole as inhibitors of dengue virus type 2 protease *in vitro*. Antiviral Res. 97, 74–80.
[23] Prusis, P., Junaid, M., Petrovska, R., Yahorava, S., Yahorau, A., Katzenmeier, G., Lapins, M. and Wikberg, J.E. (2013) Design and evaluation of substrate-based octapeptide and non substrate-based tetrapeptide inhibitors of dengue virus NS2B–NS3 proteases. Biochem. Biophys. Res. Commun. 434, 767–772.
[24] Wichapong, K., Nueangaudom, A., Pianwanit, S., Sippl, W. and Kokpol, S. (2013) Identification of potential hit compounds for dengue virus NS2B/NS3 protease inhibitors by combining virtual screening and binding free energy calculations. Trop. Biomed. 30, 388–408.

[25] Nitsche, C., Schreier, V.N., Behnam, M.A., Kumar, A., Bartenschlager, R. and Klein, C.D. (2013) Thiazolidinone-peptide hybrids as dengue virus protease inhibitors with antiviral activity in cell culture. J. Med. Chem. 56, 8389–8403.
[26] Pambudi, S., Kawashita, N., Phanthanawiboon, S., Omokoko, M.D., Masrinoul, P., Yamashita, A., Limkittikul, K., Yasunaga, T., Takagi, T., Ikuta, K. and Kurosu, T. (2013) A small compound targeting the interaction between nonstructural proteins 2B and 3 inhibits dengue virus replication. Biochem. Biophys. Res. Commun. 440, 393–398.
[27] Yildiz, M., Ghosh, S., Bell, J.A., Sherman, W. and Hardy, J.A. (2013) Allosteric inhibition of the NS2B–NS3 protease from dengue virus. ACS Chem. Biol. 8. 2744–2742.
[28] Zhou, G.C., Weng, Z., Shao, X., Liu, F., Nie, X., Liu, J., Wang, D., Wang, C. and Guo, K. (2013) Discovery and SAR studies of methionine–proline anilides as dengue virus NS2B–NS3 protease inhibitors. Bioorg. Med. Chem. Lett. 23, 6549–6554.
[29] Neylon, C., Brown, S.E., Kralicek, A.V., Miles, C.S., Love, C.A. and Dixon, N.E. (2000) Interaction of the *Escherichia coli* replication terminator protein (Tus) with DNA: a model derived from DNA-binding studies of mutant proteins by surface plasmon resonance. Biochemistry 39, 11989–11999.
[30] Matsuda, T., Watanabe, S. and Kigawa, T. (2013) Cell-free synthesis system suitable for disulfide-containing proteins. Biochem. Biophys. Res. Commun. 431, 296–301.
[31] Sivashanmugam, A., Murray, V., Cui, C.X., Zhang, Y.H., Wang, J.J. and Li, Q.Q. (2009) Practical protocols for production of very high yields of recombinant proteins using *Escherichia coli*. Protein Sci. 18, 936–948.
[32] Ozawa, K., Headlam, M.J., Schaeffer, P.M., Henderson, B.R., Dixon, N.E. and Otting, G. (2004) Optimization of an *Escherichia coli* system for cell-free synthesis of selectively ^{15}N-labelled proteins for rapid analysis by NMR spectroscopy. Eur. J. Biochem. 271, 4084–4093.
[33] Apponyi, M.A., Ozawa, K., Dixon, N.E. and Otting, G. (2008) Cell-free protein synthesis for analysis by NMR spectroscopy. Methods Mol. Biol. 426, 257–268.
[34] Jia, X., Ozawa, K., Loscha, K. and Otting, G. (2009) Glutarate and N-acetyl-L-glutamate buffers for cell-free synthesis of selectively ^{15}N-labelled proteins. J. Biomol. NMR 44, 59–67.
[35] Graham, B., Loh, C.T., Swarbrick, J.D., Ung, P., Shin, J., Yagi, H., Jia, X., Chhabra, S., Barlow, N., Pintacuda, G., Huber, T. and Otting, G. (2011) DOTA-amide lanthanide tag for reliable generation of pseudocontact shifts in protein NMR spectra. Bioconjug. Chem. 22, 2118–2125.
[36] Bertini, I., Luchinat, C. and Parigi, G. (2002) Magnetic susceptibility in paramagnetic NMR. Prog. NMR Spectrosc. 40, 249–273.
[37] Otting, G. (2008) Prospects for lanthanides in structural biology by NMR. J. Biomol. NMR 42, 1–9.
[38] Schmitz, C., Stanton-Cook, M.J., Su, X.C., Otting, G. and Huber, T. (2008) Numbat: an interactive software tool for fitting $\Delta\chi$ to molecular coordinates using pseudocontact shifts. J. Biomol. NMR 41, 179–189.
[39] Wu, P.S.C., Ozawa, K., Lim, S.P., Vasudevan, S.G., Dixon, N.E. and Otting, G. (2007) Cell-free transcription/translation from PCR-amplified DNA for high-throughput NMR studies. Angew. Chem. Int. Ed. 46, 3356–3358.
[40] Su, X.C., Ozawa, K., Qi, R., Vasudevan, S.G., Lim, S.P. and Otting, G. (2009) NMR analysis of the dynamic exchange of the NS2B cofactor between open and closed conformations of the West Nile virus NS2B–NS3 protease. PLoS Negl. Trop. Dis. 3, e561.
[41] Shishmarev, D. and Otting, G. (2013) How reliable are pseudocontact shifts induced in proteins and ligands by mobile paramagnetic metal tags? A modelling study. J. Biomol. NMR 56, 203–216.
[42] Chandramouli, S., Joseph, J.S., Daudenarde, S., Gatchalian, J., Cornillez-Ty, C. and Kuhn, P. (2010) Serotype-specific structural differences in the protease-cofactor complexes of the dengue virus family. J. Virol. 84, 3059–3067.
[43] Costa, R.L., Voloch, C.M. and Schrago, C.G. (2012) Comparative evolutionary epidemiology of dengue virus serotypes. Infect. Genet. Evol. 12, 309–314.

Supplementary Material

The dengue virus NS2B-NS3 protease retains the closed conformation in the complex with BPTI

Wan-Na Chen, Karin V. Loscha, Christoph Nitsche, Bim Graham, Gottfried Otting

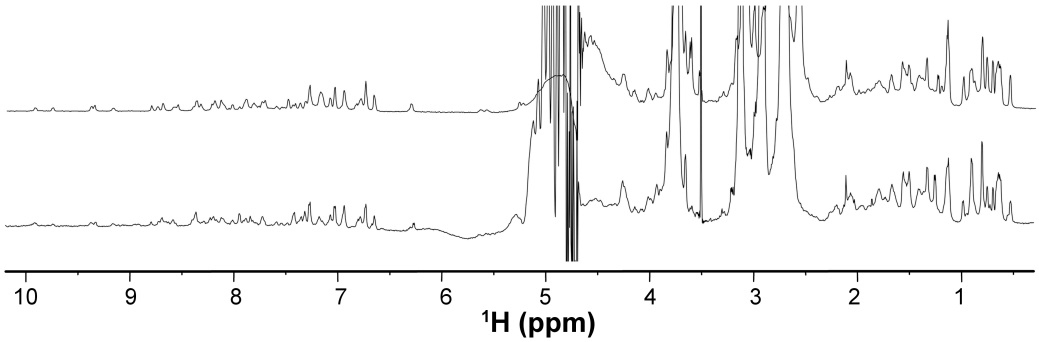


Fig. S1. BPTI produced by cell-free synthesis is correctly folded. The figure shows 1D ^{1}H NMR spectra of BPTI produced by cell-free synthesis on the top and of commercial BPTI (Sigma, derived from bovine lung) at the bottom. Peaks from solvent and buffer are truncated. The NMR buffer (90% H_2O/10% D_2O) contained 20 mM MES, pH 6.5, and 50 mM NaCl. The spectra were recorded of 0.2 mM protein solutions at 25 ^{o}C on a Bruker 800 MHz NMR spectrometer.

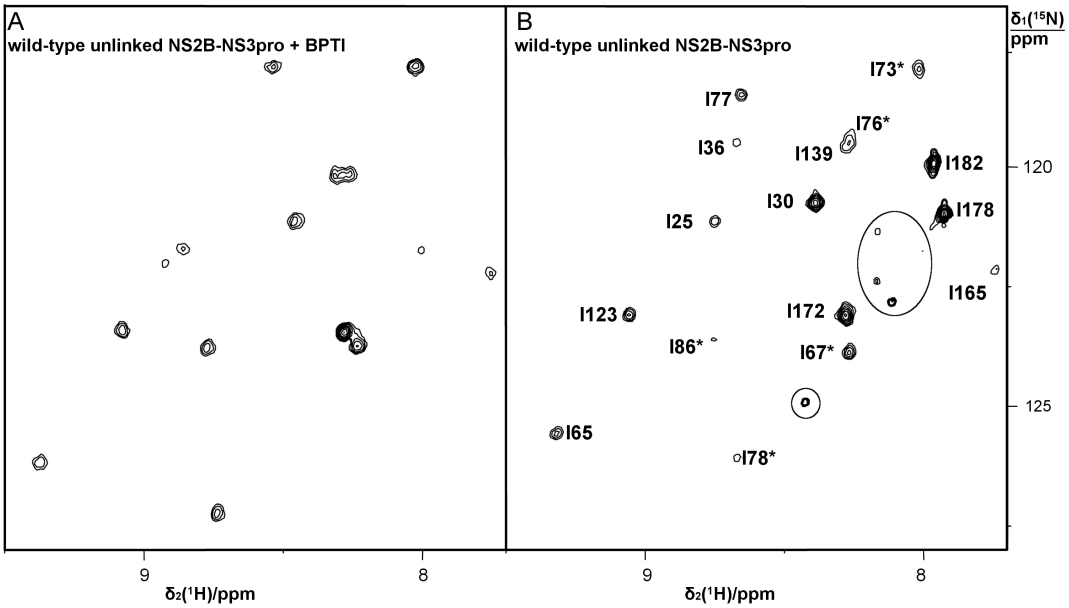


Fig. S2. BPTI protects the NS2B-NS3 protease against degradation. The figure shows ^{15}N-HSQC spectra of 0.1 mM solutions of selectively ^{15}N-Ile-labelled wild-type NS2B-NS3pro in the presence and absence of BPTI. Same sample conditions and temperature as in Fig. S1. (A) Sample purified in the presence of BPTI. The spectrum was recorded following addition of BPTI to a nominal ratio of about 1.1:1 (BPTI to protease). (B) Spectrum recorded of NS2B-NS3pro in the absence of inhibitor. The resonance assignment is taken from ref. [1]. The circles highlight weak narrow cross-peaks absent from the spectrum in (A), indicating sample degradation. The cross-peaks of I178 and I182 from the unstructured C-terminal segment of NS3pro were very weak in the spectrum in (A) for unknown reasons. The same experiment performed with the mutant S68C displayed these cross-peaks with usual intensities (data not shown).

Table S1. PCSs of backbone amide protons in the complex between NS2B-NS3pro (S68C) tagged with C2-Tb^{3+} and BPTI[a]

Residue	PCS^{exp} (ppm)	Fit 1[b]		Fit 2[c]	
		PCS^{cal} (ppm)	Distance from metal (Å)[d]	PCS^{cal} (ppm)	Distance from metal (Å)[e]
NS2B					
Arg55*	-0.66	-0.65	21.6	-0.61	22.4
Ala57*	-0.39	-0.41	24.6	-0.40	25.2
Glu62*	-0.08	-0.09	34.5	-0.10	34.9
Glu63*	-0.08	-0.07	38.4	-0.08	38.8
Ala65*	-0.06	-0.06	38.6	-0.07	38.8
Ser71*	0.02	**0.01**	40.4	**0.01**	40.1
Ile73*	0.04	**0.04**	39.6	**0.04**	39.1
Ser75*	0.08	**0.07**	37.4	**0.07**	36.8
Thr77*	0.16	**0.12**	31.7	**0.13**	30.9
Ile 78*	0.19	**0.19**	28.5	**0.19**	27.7
Gly82*	0.50	**0.39**	21.8	**0.45**	20.9
Ser83*	0.38	**0.30**	23.0	**0.37**	22.2
Ile86*	0.12	**0.09**	33.0	**0.11**	32.3
NS3pro					
Lys28	-0.48	-0.48	16.2	-0.49	16.7
Tyr41	-0.69	-0.69	19.1	-0.70	19.7
His51	0.38	0.37	16.1	0.42	15.7
Val52	-0.02	-0.01	15.5	0.01	15.3
Thr53	-0.28	-0.28	13.1	-0.30	13.0
Gly62	-0.64	-0.64	15.7	-0.67	16.8
Trp89	0.08	0.07	28.7	0.03	28.5
Leu98	-0.12	-0.15	27.7	-0.16	27.9
Thr111	-0.05	-0.03	35.5	-0.04	35.5
Phe116	0.10	0.09	32.3	0.07	31.8
Thr120	0.16	0.15	31.3	0.14	30.4
Gly121	0.14	0.16	30.8	0.14	30.0
Thr122	0.15	0.12	32.1	0.10	31.5
Ile123	0.12	0.11	31.3	0.10	30.7

Val140	-0.08	-0.07	29.1	-0.09	29.2
Gly148	-0.03	-0.02	24.8	-0.08	24.8
BPTI					
Phe4	0.10	**0.04**	35.7	**0.08**	35.1
Leu6	0.10	**0.04**	37.0	**0.07**	36.4
Thr11	0.08	**0.03**	30.0	0.07	29.6
Ala16	0.02	**-0.05**	21.8	0.00	21.6
Arg17	-0.13	**-0.22**	20.8	-0.17	20.8
Ile18	-0.08	**-0.08**	22.5	-0.01	22.4
Arg20	0.07	**-0.01**	25.5	0.06	25.4
Phe22	0.07	**0.03**	30.2	**0.07**	30.0
Ala25	0.06	**0.02**	38.3	**0.05**	37.9
Lys26	0.06	**0.02**	40.5	**0.04**	40.2
Ala27	0.04	**0.01**	40.5	**0.03**	40.3
Leu29	0.04	**0.02**	38.3	**0.04**	38.1
Thr32	0.02	**-0.00**	31.7	**0.03**	31.7
Phe33	0.05	**-0.00**	28.2	0.05	28.0
Arg39	0.35	**0.20**	23.1	0.31	22.4
Ala40	0.31	**0.17**	23.4	0.30	22.8
Lys41	0.18	**0.09**	27.2	0.17	26.6
Arg42	0.20	**0.08**	29.7	0.15	29.0
Phe45	0.16	**0.07**	26.3	0.15	25.9
Lys46	0.26	**0.12**	23.3	0.24	22.9
Ala48	0.11	**0.06**	27.1	0.11	27.0
Arg53	0.13	**0.06**	31.4	**0.10**	31.0
Thr54	0.15	**0.05**	31.5	**0.10**	31.0

[a] Samples contained either uniformly ^{15}N-labelled NS2B-NS3pro with unlabelled BPTI or unlabelled NS2B-NS3pro with selectively ^{15}N-labelled Ala, Arg, Lys, Thr, Leu, Phe, Ile labelled BPTI. In each case, NS2B-NS3pro was the S68C mutant tagged with either C2-Tb^{3+} (paramagnetic sample) or C2-Y^{3+} (diamagnetic reference). PCSs were measured as the chemical shifts observed with C2-Tb^{3+} minus the chemical shifts observed with C2-Y^{3+}.

[b] Fit 1 was obtained from fitting a $\Delta\chi$ tensor to the crystal structure 3U1J [2], using the PCS^{exp} values of backbone amide protons of NS3pro and NS2B (residues 55*–

65*). The Δχ tensor determined by Fit 1 was subsequently used to back-calculate PCS data, PCS^{cal}, for the same protein structure. Data in bold are PCSs back-calculated for nuclear spins, the PCS^{exp} data of which were not used for the tensor determination by Fit 1. To back-calculate PCS^{cal} values for NS2Bc residues without electron density in the structure 3U1J, their coordinates were generated by crafting the corresponding coordinates from 3U1I [2] onto 3U1J, following superimposition of the NS2Bn and NS3pro residues of both crystal structures.

[c] Fit 2 was performed like Fit 1, except using the PCS^{exp} data of NS2B (residues 55*–65*), NS3pro, and BPTI residues within 30 Å from the metal centre determined by Fit 1. PCS^{cal} values of NS2Bc and BPTI are highlighted in bold if they were not used in Fit 2.

[d] Distance of the nuclear spin from the metal centre determined by the Δχ-tensor determination of Fit 1.

[e] Distance of the nuclear spin from the metal centre determined by the Δχ-tensor determination of Fit 2.

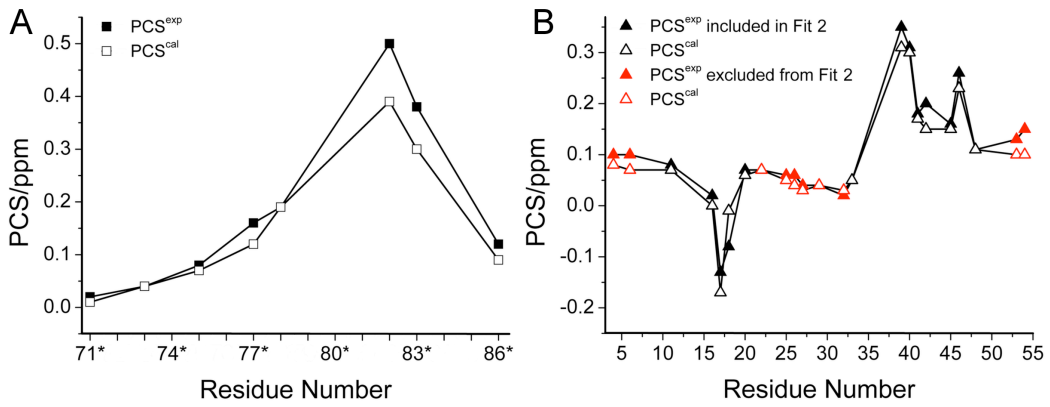


Fig. S3. Comparison of back-calculated and experimental PCSs for residues that were excluded from the Δχ-tensor fits. (A) Experimental PCSs and PCSs predicted by Fit 1 for backbone amide protons of NS2Bc. Filled and open squares indicate experimental and calculated PCSs, respectively. The PCS^{cal} values match the PCS^{exp} data less well than in Fig. 3D, because Fit 1 did not include PCSs from BPTI. This demonstrates the importance of determining the Δχ-tensor from PCS data of nuclei located near NS2Bc [3]. (B) Experimental PCSs (filled triangles) and PCSs predicted by Fit 2 (open triangles) for backbone amide protons of BPTI. Data of residues further than 30 Å from the metal center (determined by Fit 1) are highlighted in red; they were not used in the Δχ-tensor determination of Fit 2. By including a number of amides from BPTI in the Δχ-tensor fit, Fit 2 resulted in better agreement between predicted and experimental PCSs for BPTI than Fit 1 (Fig. 3C).

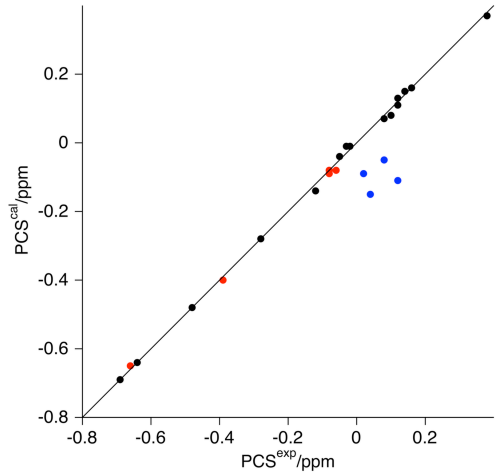


Fig. S4. Correlation between back-calculated and experimental PCSs for backbone amide protons, using the open conformation of Fig. 1A (2FOM [4]) and the PCSs of Table S1 to fit a $\Delta\chi$ tensor to NS2Bn and NS3pro. The data of NS2Bn (residues 55*–65*), NS2Bc (residues 66*–95*) and NS3pro are shown as red, blue and black points, respectively. The structure 2FOM contains electron density for only four of the NS2Bc residues, for which PCSs could be measured. For these residues, negative PCSs are predicted by the $\Delta\chi$ tensor but positive PCSs were measured, demonstrating that this or similar conformations of NS2Bc are not highly populated in solution in the complex with BPTI.

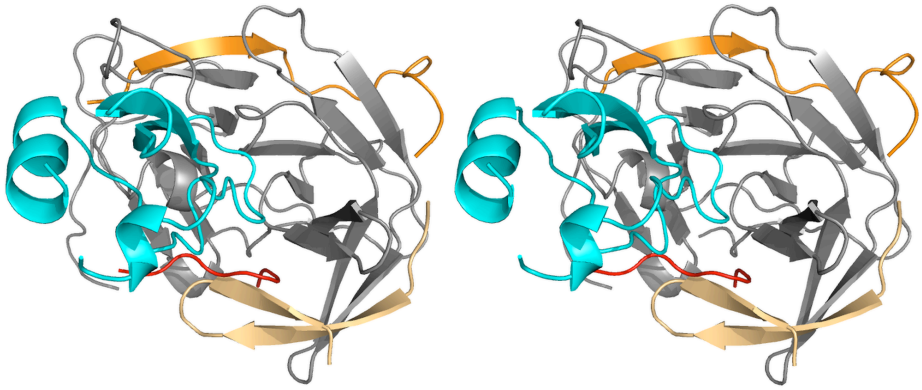


Fig. S5. Stereo view of crystal contacts at the NS2B binding site in the structure 3U1J. Grey: NS3pro; cyan: BPTI; orange: NS2B; light orange: NS2Bc copied from the crystal structure with the peptide inhibitor Bz-nKRR-H (3U1I, [2]); red: residues 6–13 from a neighboring NS3pro molecule in the crystal lattice of 3U1J, which displaces the β-hairpin of NS2Bc.

References

[1] de la Cruz, L., Nguyen, T.H., Ozawa, K., Shin, J., Graham, B., Huber, T. and Otting, G. (2011) Binding of low molecular weight inhibitors promotes large conformational changes in the dengue virus NS2B-NS3 protease: fold analysis by pseudocontact shifts. J. Am. Chem. Soc. 133, 19205–19215.

[2] Noble, C.G., Seh, C.C., Chao, A.T. and Shi, P.Y. (2012) Ligand-bound structures of the dengue virus protease reveal the active conformation. J. Virol. 86, 438–446.

[3] Shishmarev, D. and Otting, G. (2013) How reliable are pseudocontact shifts induced in proteins and ligands by mobile paramagnetic metal tags? A modelling study. J. Biomol. NMR 56, 203–216.

[4] Erbel, P., Schiering, N., D'Arcy, A., Renatus, M., Kroemer, M., Lim, S. P., Yin, Z., Keller, T.H., Vasudevan, S.G. and Hommel, U. (2006) Structural basis for the activation of flaviviral NS3 proteases from dengue and West Nile virus. Nat. Struct. Mol. Biol. 13, 372–373.

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

O-tert-Butyltyrosine, an NMR Tag for High-Molecular-Weight Systems and Measurements of Submicromolar Ligand Binding Affinities

Wan-Na Chen,[†] Kekini Vahini Kuppan,[†] Michael David Lee,[‡] Kristaps Jaudzems,[§] Thomas Huber,*[,†] and Gottfried Otting*[,†]

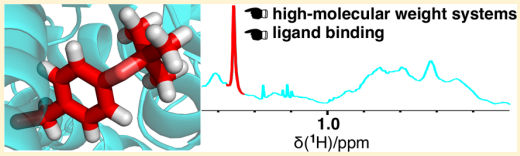
[†]Research School of Chemistry, Australian National University, Canberra, ACT 2601, Australia

[‡]Monash Institute of Pharmaceutical Sciences, Parkville, VIC 3052, Australia

[§]Latvian Institute of Organic Synthesis, Riga 1006, Latvia

Supporting Information

ABSTRACT: *O-tert*-Butyltyrosine (Tby) is an unnatural amino acid that can be site-specifically incorporated into proteins using established orthogonal aminoacyl-tRNA synthetase/tRNA systems. Here we show that the *tert*-butyl group presents an outstanding NMR tag that can readily be observed in one-dimensional ^{1}H NMR spectra without any isotope labeling. Owing to rapid bond rotations and the chemical equivalence of the protons of a solvent-exposed *tert*-butyl group from Tby, the singlet resonance from the *tert*-butyl group generates an easily detectable narrow signal in a spectral region with limited overlap with other methyl resonances. The potential of the *tert*-butyl ^{1}H NMR signal in protein research is illustrated by the observation and assignment of two resonances in the *Bacillus stearothermophilus* DnaB hexamer (320 kDa), demonstrating that this protein preferentially assumes a 3-fold rather than 6-fold symmetry in solution, and by the quantitative measurement of the submicromolar dissociation constant K_d (0.2 μM) of the complex between glutamate and the *Escherichia coli* aspartate/glutamate binding protein (DEBP, 32 kDa). The outstanding signal height of the ^{1}H NMR signal of the Tby *tert*-butyl group allows K_d measurements using less concentrated protein solutions than usual, providing access to K_d values 1 order of magnitude lower than established NMR methods that employ direct protein detection for K_d measurements.



INTRODUCTION

NMR spectroscopy offers structural information on biological macromolecules at a level of detail that is unsurpassed by any other solution method. The resolution of the solution NMR spectrum of a large protein, however, is compromised both by the number of NMR resonances as well as by line widths that increase with molecular mass.[1] These problems are routinely addressed by labeling with ^{15}N and ^{13}C to facilitate the recording of multidimensional NMR spectra. In addition, perdeuteration of the protein leads to narrower NMR signals of the remaining ^{1}H spins, permitting in exceptional cases two-dimensional (2D) NMR spectra to be recorded of protein multimers approaching total molecular masses of 500 kDa and beyond.[2−4] In particular, methyl groups are relatively easy to observe in such systems because methyl groups rotate rapidly, which results in chemical shift degeneracy of the three methyl protons and favorable ^{1}H relaxation properties.[5] It is a quite general phenomenon that enhanced mobility produces narrower, more easily observable NMR signals. For example, highly flexible polypeptide segments are typically found at the N- and C-termini of a protein and their NMR resonances can often be observed without perdeuteration even when the molecular weight of the system is large.[6]

These observations prompted us to investigate the potential of an unnatural amino acid containing a mobile *tert*-butyl group. Methods have recently been established for the genetic encoding of *O-tert*-butyltyrosine (Tby),[7,8] but NMR spectra of proteins with Tby incorporated in this way have not been reported, although the value of monitoring the NMR signal of a site-specifically incorporated chemical "spy" is well established. For example, Geierstanger and co-workers have shown that the NMR resonances of the methyl- and CF_3-groups of *p*-methoxyphenylalanine (OMePhe) and *p*-trifluoromethoxyphenylalanine (OCF_3Phe) can readily be observed after site-specific incorporation at different locations of a 33 kDa protein.[9] The signals of the newly introduced $^{13}CH_3$- and CF_3-groups were readily observed in two-dimensional (2D) ^{13}C HSQC and 1D ^{19}F NMR spectra, respectively. Similarly, Mehl and co-workers demonstrated that site-specific incorporation of trifluoromethylphenylalanine (CF_3Phe) produced 1D ^{19}F NMR spectra that allowed distinction of different ligand- and metal-bound states in 48 and 98 kDa homodimers.[10] While the ^{1}H NMR resonances of methyl groups are difficult to resolve without additional labeling with ^{13}C, the main drawback of ^{19}F NMR is the large chemical shift anisotropy (CSA) associated with ^{19}F spins which leads to broad NMR resonances. As CSA relaxation rates increase with the square of the magnetic field

Received: February 21, 2015
Published: March 19, 2015

DOI: 10.1021/jacs.5b01918
J. Am. Chem. Soc. 2015, 137, 4581−4586

strength, ^{19}F NMR spectra are best recorded at relatively low NMR frequencies, compromising sensitivity. In contrast, the *tert*-butyl group of Tby can be observed by ^{1}H NMR with increasing sensitivity at increasing magnetic fields. We hypothesized that, owing to the intensity of the methyl resonance (corresponding to nine degenerate ^{1}H spins) and the decreased relaxation rate associated with a mobile, solvent exposed *tert*-butyl group, its observation by high-field NMR should be straightforward without any stable isotope labeling.

In the following we show that Tby indeed produces a ^{1}H NMR signal that is readily observable in proteins without isotope labeling, including a 320 kDa system. Furthermore, we show that the outstanding sensitivity of the *tert*-butyl resonance allows its detection in dilute protein solutions, effectively reducing the protein demand by 1 order of magnitude. Finally, we show that Tby is an excellent probe for direct NMR measurements of submicromolar ligand binding affinities.

RESULTS

Incorporation of Tby. For site-specific incorporation of Tby into proteins, we chose the *p*-cyanophenylalanyl-tRNA synthetase (*p*CNF-RS) evolved from the *Methanocaldococcus jannaschii* tyrosyl-RS encoded in the pUltra system together with the requisite $tRNA_{CUA}$ for recognition of TAG stop codons.[11] *Bacillus stearothermophilus* (*Bst*) DnaB was produced in fusion with the solubility enhancer GB1 by in vivo expression in *Escherichia coli* (Supporting Information Figures S1 and S2A), whereas *Staphylococcus aureus* sortase A (SrtA) and the *E. coli* aspartate/glutamate binding protein (DEBP) were produced from linear PCR amplified DNA in a cell-free system based on an *E. coli* cell extract, from which the release factor 1 had been selectively removed by tagging with chitin-binding domains and filtration over chitin (Figure S2B).[12] The fusion with GB1 did not significantly change the 1D ^{1}H NMR spectrum of the *Bst* DnaB hexamer (Figure S3A).

NMR Spectroscopy of DnaB-Tby. Figure 1A compares the 1D ^{1}H NMR spectrum of the mutant Tyr104Tby with the corresponding spectrum of wild-type DnaB. As expected for a high-molecular-weight system, the proteins display very broad signals, whereas very narrow signals arise from buffer or low-molecular-weight impurities. Two new signals at about 1.2 ppm, however, are observed for DnaB-Tby but not wild-type DnaB. Their chemical shifts agree with the chemical shift of the *tert*-butyl group in free Tby (1.26 ppm in 25 mM Tris-HCl, pH 7.5). The observation of two resonances is in agreement with the two different chemical environments of the Tby site in crystal structures of the *Bst* DnaB hexamer in complex with DnaG (Figure 1B and C).[13] Previous electron microscopy studies of *E. coli* DnaB had reported hexamers with 3-fold and 6-fold symmetries.[14−16] The observation of two, rather than one, *tert*-butyl resonances identifies the 3-fold symmetry as the symmetry prevailing for the *Bst* helicase in solution.

To assign the local environments which the two resonances belong to, we introduced the additional mutations Cys253Val, Ala303Cys, and Cys305Val to enable site-specific ligation with a lanthanide tag at position 303. At the same time, the natural amino acids of the protein were perdeuterated to confirm the assignment of the *tert*-butyl signals. The mutations did not alter the appearance of the *tert*-butyl signals (Figure 1A). Following ligation of the cysteine residue with a C1-tag[17] loaded with Ce^{3+} (Figure S4A), one of the two Tby resonances broadened significantly, assigning it to the site, which is within 15 Å of Ala 303 in the crystal structure. This result presents further

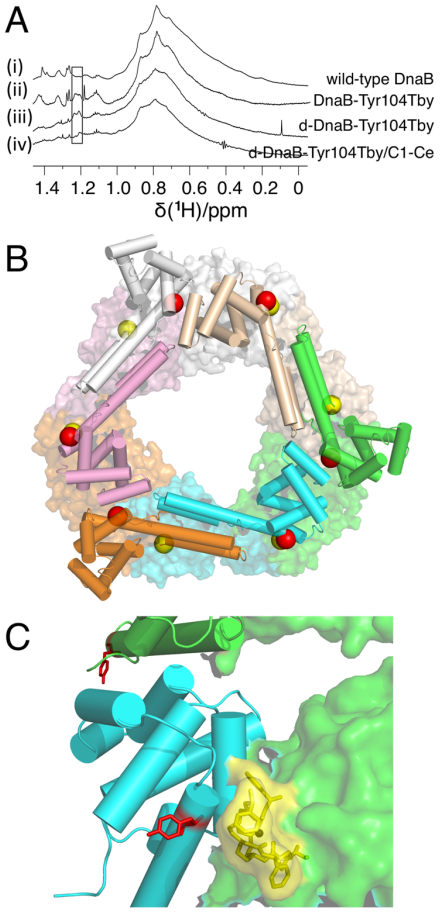


Figure 1. Conformational study of *Bst* DnaB helicase. (A) 1D ^{1}H NMR spectra of wild-type and mutant *Bst* DnaB. The spectra were recorded in a buffer of 25 mM Tris-HCl, pH 7.5, 150 mM NaCl at 25 °C, using a Bruker 600 MHz NMR spectrometer. Protein concentrations were about 20−40 μM and the total recording time per spectrum was 0.5−2 h. (i) Wild-type DnaB. (ii) DnaB with Tby at site 104. The *tert*-butyl group displays two peaks (highlighted by a frame). Unassigned narrow signals are from impurities in the buffer. (iii) Perdeuterated DnaB (70% deuteration) with Tby at site 104 and containing the additional mutations Cys253Val, Ala303Cys, and Cys305Val. The reduced intensity of the background confirms the identification of the Tby *tert*-butyl resonances. (iv) Same as (iii), except with a C1-Ce tag ligated to Cys 303. The downfield, but not the upfield, *tert*-butyl peak is affected, assigning the downfield peak to the Tby residue near residue 303 of the neighboring DnaB molecule, and confirming that residue 104 exists in two different chemical environments as expected for a hexamer with C3 symmetry. (B) View of the crystal structure of the DnaB hexamer along the C3 symmetry axis (PDB: 2R6A).[13] Each monomer is colored differently, showing the N-terminal domains in a cartoon representation, while the linker segments and the C-terminal domains are in a surface representation. The sites of Tyr 104, which was replaced by Tby in this work, are highlighted by red spheres. The sites of residue 303, mutated to a cysteine for subsequent ligation with the C1-Ce tag, are identified by yellow spheres. (C) The C1-Ce tag modeled at Cys 303 (yellow) is close to Tyr 104 in the blue monomer, but far from Tyr 104 in the green monomer in the crystal structure 2R6A. The N- and C-terminal domains are shown in cartoon and surface representations, respectively, as in (B). Tyr 104 is shown in red.

DOI: 10.1021/jacs.5b01918
J. Am. Chem. Soc. 2015, 137, 4581−4586

confirmation that the crystal structure of the *Bst* hexamer reflects its predominant three-dimensional (3D) structure in solution.

Most importantly, the result clearly demonstrates the possibility to detect the Tby *tert*-butyl resonance in systems of high molecular mass at low concentration without any isotope labeling. For comparison, we also incorporated a trifluoromethyl-phenylalanine residue in position 104 and recorded a ^{19}F NMR spectrum at a ^{1}H NMR frequency of 500 MHz. The line width of the ^{19}F NMR resonance was about 160 Hz at half height and thus too broad to resolve two different signals (Figure S5).

NMR Spectroscopy of Sortase A with Tby. Figure 2A and B compares the 1D ^{1}H NMR spectra obtained with the

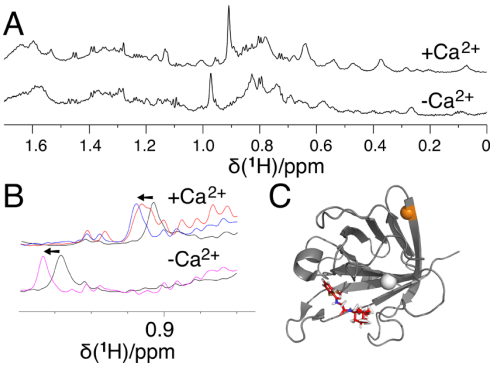


Figure 2. Ligand binding study of *S. aureus* sortase A. The spectra were recorded at 25 °C using a Bruker 800 MHz NMR spectrometer. (A) 1D ^{1}H NMR spectra of sortase A with Tby incorporated at position 150. Top trace: 20 μM protein in 20 mM MES buffer, pH 6.5, 50 mM NaCl, 5 mM $CaCl_2$. Bottom trace: same as above, but without $CaCl_2$ and with 1 mM EDTA. (B) 1D ^{1}H NMR spectra of a sortase A/inhibitor complex with (top) and without (bottom) calcium. The protein concentration was 20 μM. Black: inhibitor-free state. Red: half-saturation with inhibitor. Blue and magenta: inhibitor-bound state with 50% excess inhibitor. The inhibitor[18] was added from a 5 mM stock solution in DMSO-d_6. The final concentration of DMSO in the protein solution was always less than 1%. Arrows indicate the direction of chemical shift change upon addition of inhibitor. (C) Cartoon representation of sortase A (PDB: 1IJA[19]). Inhibitor, calcium, and the mutation site for Tby are shown as red sticks, a white sphere, and an orange sphere, respectively. The ligand and calcium coordinates were taken from the PDB files 2MLM[18] and 2KID,[20] respectively.

Thr150Tby mutant of *S. aureus* sortase A and the effects of calcium and inhibitor binding on the chemical shift of the *tert*-butyl resonance of Tby. The chemical shift of the *tert*-butyl group is unusually high-field, which can be attributed to ring current effects arising from its proximity to the side chain of Phe 144. Despite the distance of the Tby residue from the active site marked by Cys 184 (about 25 Å) and the calcium binding site (about 20 Å, Figure 2C), the chemical shift of the *tert*-butyl resonance was sensitive to the presence of Ca^{2+} or inhibitor, or both. Therefore, the Tby residue can be used for remote sensing of changes in the substrate binding site of the enzyme. Although the presence of calcium causes many spectral changes in the ^{1}H NMR spectrum of the protein, the *tert*-butyl resonance is easily picked out as the tallest resonance in the range between 0 and 1.7 ppm (Figure 2A). The size of the resonance greatly facilitates monitoring its chemical shift in titration experiments, and its narrow line width magnifies the significance of even small chemical shift changes. The inhibitor used here (Figure S4B) is known to bind irreversibly by formation of a covalent bond to the active-site cysteine.[18] As expected for a system with slow or no chemical exchange, the *tert*-butyl resonances of the inhibitor bound and unbound states of the protein appeared simultaneously at half-saturation (Figure 2B). A control experiment confirmed that the dimethyl sulfoxide (DMSO) of the inhibitor stock solution caused much smaller chemical shift changes than the ligand (Figure S3B).

Measuring Submicromolar Dissociation Constants. *E. coli* DEBP is a periplasmic transport protein that reversibly binds aspartate or glutamate. It binds glutamate with a binding affinity of 260 nM, as measured by isothermal calorimetry (ITC; B. Clifton and C. J. Jackson, personal communication). Despite the high molecular mass of the protein (32 kDa), the *tert*-butyl resonance is clearly visible in the 1D ^{1}H NMR spectrum of the mutant Val140Tby and over 1 order of magnitude taller than any other signal of the protein (Figure 3A). Therefore, NMR spectra of good sensitivity can be

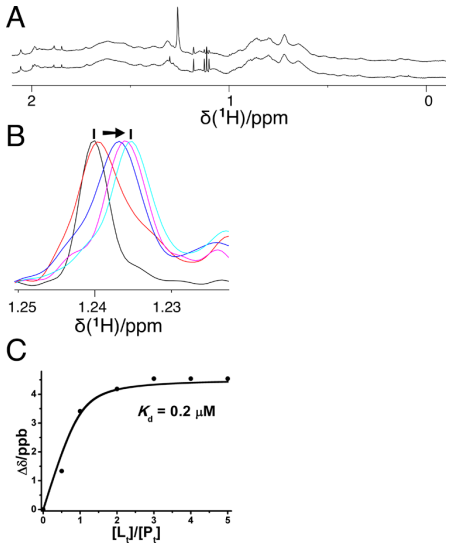


Figure 3. Binding affinity measurement of *E. coli* DEBP for glutamate. The 1D ^{1}H NMR spectra were recorded in 20 mM phosphate buffer, pH 7.5, 100 mM NaCl at 25 °C, using a Bruker 800 MHz NMR spectrometer. (A) NMR spectra of 50 μM solutions of DEBP. The top and bottom traces show the spectra of the mutant with Tby at position 140 and of wild-type DEBP, respectively. The signal of the *tert*-butyl group appears at about 1.24 ppm. Other narrow signals are from low-molecular weight impurities. (B) Titration of glutamate (1 mM stock solution in H_2O) into a 2 μM solution of DEBP. The *tert*-butyl resonance shifts upfield with increasing concentration of glutamate (indicated by the arrow). Each spectrum was acquired with a total recording time of 30 min on a 800 MHz NMR spectrometer equipped with a cryoprobe. (C) Fit of the binding affinity curve by eq 1. For best accuracy, the chemical shifts were determined by fitting Lorentzian line shapes to the *tert*-butyl resonances in (B). K_d was calculated as 0.20 μM with a standard error of 0.08 μM.

recorded at low protein concentrations, which is a prerequisite for determining small dissociation constants from chemical shift changes of protein resonances. Titration of glutamate into the protein solution indeed led to small changes in chemical shifts, indicating fast exchange between the bound and free state of the protein (Figure 3B).

DOI: 10.1021/jacs.5b01918
J. Am. Chem. Soc. 2015, 137, 4581−4586

The observed chemical shift changes can be described by eq 1:

$$\Delta_{obs} = \Delta_{max}\{(K_d + L_t + P_t) - [(K_d + L_t + P_t)^2 - 4P_tL_t]^{1/2}\}/(2P_t) \quad (1)$$

where Δ_{obs} is the observed change in chemical shift, Δ_{max} is the maximal chemical shift change observed when the protein is fully saturated with ligand, L_t and P_t are the total concentrations of ligand and protein, respectively, and K_d is the dissociation constant. The K_d value obtained by fitting the observed changes in chemical shifts by eq 1 was 0.20 μM (Figure 3C), in close agreement with the value determined by ITC measurement.

Near-Isotropic Bond Reorientation in Tby. The *tert*-butyl group of Tby produces an exceptionally narrow line because of its motional freedom. When it is fully solvent exposed, three different bond rotations contribute to the reorientation of the C−H bonds of the methyl groups. First, each methyl group rotates about its axis of symmetry. Second, the *tert*-butyl group reorientates by rotation about the C−O bond, that projects from the aromatic ring. Finally, the following O−C bond to the *tert*-butyl group is free to rotate. Crystal structures of compounds with *tert*-butyl-phenyl groups show dihedral angles of about 90° for the C−O bond, positioning the *tert*-butyl group outside the plane of the aromatic ring (Figure 4A). Assuming dihedral angles of 90° or 270° for the C−O bond, and 60°, −60°, or 180° for the O−C bond and the methyl group rotamers, an individual C−H group of a methyl can access 2 × 3 × 3 = 18 different rotameric states. Assuming equal probability of each dihedral angle combination owing to unrestricted motions of a solvent-exposed *tert*-butyl group, the C−H bond populates a remarkably isotropic distribution of orientations (Figure 4B). As a measure of the degree of isotropy, concatenation of all C−H bond vectors produced by the 18 rotameric states yields a sum vector of a length of only about 1 Å. Although each rotamer may be associated with different opportunities for dipole−dipole relaxation depending on the location of other protein protons in the vicinity, the remarkably slow relaxation of the *tert*-butyl resonance is fundamentally due to its exceptionally high degree of mobility.

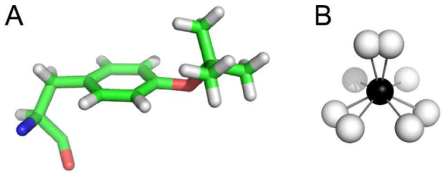


Figure 4. Methyl groups of O-*tert*-butyl-tyrosine can change their orientation in multiple directions. (A) Stick representation of the O-*tert*-butyl-tyrosine residue as determined by X-ray crystallography.[21] (B) Different orientations of a single methyl C−H bond produced by rotations around bonds of the O-*tert*-butyl group, including dihedral angles of 90° or 270° for the C−O bond, and 60°, −60°, or 180° for the O−C bond and the methyl groups. The variability in the resulting C−H bond orientations is highlighted by superimposing the carbon atom of each of the 18 rotameric states. The carbon and hydrogen atoms of the C−H group are represented by black and white spheres, respectively. Some of the rotamers produce degenerate orientations.

DISCUSSION

The present work shows that *tert*-butyl-tyrosine is an outstanding probe for the study of small and large proteins at natural isotopic abundance by NMR spectroscopy. The intense ^{1}H NMR signal of the *tert*-butyl group is easily detected in one-dimensional NMR spectra due to its narrow line width and chemical shift that is well separated from the NMR resonances of most organic buffers, the water resonance, and the usually dominant methyl signals of valine, leucine, and isoleucine. The system for genetic encoding of Tby is well established,[7] allowing its incorporation at any site of a protein. Importantly, the ^{1}H chemical shift of the *tert*-butyl resonance is still sensitive to changes in the chemical environment of the protein, even if the *tert*-butyl group is solvent exposed. Therefore, it can be used to monitor the state of a protein, which is of particular interest for studies of ligand binding.

Measurements of Dissociation Constants. The results obtained with DEBP demonstrate the outstanding advantage of Tby for ligand binding studies, in particular for the measurement of tight binding affinities. Measurements of small dissociation constants by direct observation of protein signals require a protein concentration in the order of the K_d value. The intensity of the *tert*-butyl resonance provides the sensitivity required to measure dilute solutions even of large proteins.

A graphical representation of titration curves for different K_d values illustrates the difficulty to measure dissociation constants that are much smaller than the protein concentration (Figure 5). To obtain K_d values with reasonable accuracy, a practical recommendation is to aim for protein concentrations that produce saturation ratios of ligand-bound protein to total protein between 0.2 and 0.8, when ligand and protein are present in 1:1 ratio.[22] Figure 5 shows that this condition is fulfilled for K_d values that are 10 times smaller than the protein concentration, but not if they are 100 or 1000 times smaller (see also Figure S6). Therefore, the measurement of submicromolar K_d values requires the ability to measure at protein concentrations in the low micromolar range. Such low concentrations pose a serious challenge in sensitivity for contemporary NMR spectrometers, especially if the NMR resonances of the protein are broad due to high molecular weight or chemical exchange broadening. As a result, NMR

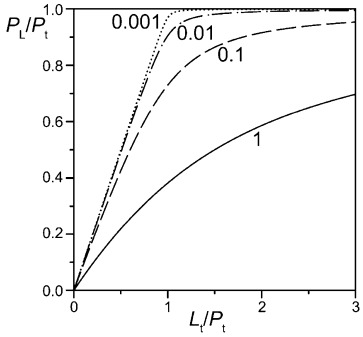


Figure 5. Titration curves calculated for different dissociation constants K_d. The curves were calculated using eq 1. The vertical axis plots the saturation of the protein with ligand (ratio of bound protein, P_L, relative to total protein, P_t). The horizontal axis plots the ratio of total ligand concentration, L_t, to total protein concentration. The curves are labeled with the respective dissociation constants normalized by the total protein concentration, K_d/P_t.

DOI: 10.1021/jacs.5b01918
J. Am. Chem. Soc. 2015, 137, 4581−4586

spectroscopy is traditionally regarded as a unique tool for detecting complexes with weakly binding ligands, whereas K_d values much smaller than about 10 μM are practically inaccessible by direct observation of protein resonances even in the case of small proteins,[23] let alone proteins of the molecular mass of DEBP. The exceptional peak height of the ^{1}H NMR resonance of the *tert*-butyl group of Tby changes this situation and lowers the detection limit by more than an order of magnitude. In principle, ligand detected competition experiments are a viable alternative for determining nanomolar binding affinities by NMR spectroscopy, but these experiments are less straightforward than the direct monitoring of an NMR resonance of the target protein in a protein observed experiment because they require the availability of an additional, different ligand, for which the dissociation constant must be determined independently.[24]

Tight binding affinities are invariably associated with slow dissociation rates, as the dissociation constant can be calculated from the ratio of dissociation versus association rates and diffusion poses an upper limit on the rate of complex formation. In the limit of slow chemical exchange between bound and free protein, separate NMR signals are observed and eq 1 no longer applies. Importantly, however, the limit between the fast and slow chemical exchange regimes depends on the frequency difference of the NMR resonances in the respective states, with small frequency differences extending the fast exchange regime to slower exchange rates. For the measurement of small dissociation constants, it is thus an advantage if the NMR signal which is monitored displays only small chemical shift changes upon titration with ligand, as in the case of a solvent-exposed *tert*-butyl group. As the examples of sortase A and DEBP show, the narrow line shape of the ^{1}H NMR resonance of the *tert*-butyl group of Tby permits accurate measurements also of small changes in chemical shift, allowing the positioning of the Tby probe sufficiently far from the ligand binding site to avoid any interference with ligand binding. In practice, we found that narrow NMR signals from buffer or low-molecular weight impurities provide excellent internal controls to confirm the accurate alignment of the NMR spectra recorded in a titration experiment.

Tby as a Probe in High-Molecular-Weight Systems. Established strategies for solution NMR spectroscopy of high-molecular weight systems employ isotope labeling strategies, in particular per-deuteration to reduce dipolar relaxation rates of the remaining ^{1}H spins.[25,26] Selective labeling or unlabeling of all amino acid residues of a certain type, however, usually labels more than a single site, making it challenging to obtain site-specific information. In contrast, the incorporation of a single unnatural amino acid naturally produces a label at a predetermined site, which is defined by the amber stop codon placed in the gene of interest.

In the case of *Bst* DnaB, the Tby probe positioned at a solvent exposed site of the N-terminal domain of the protein not only revealed clear ^{1}H NMR resonances of the *tert*-butyl group, but also provided clear evidence for the C3 symmetry of the hexamer as two different chemical environments of the *tert*-butyl group were apparent. Considering the molecular mass of the hexamer (320 kDa), it is remarkable that this could be achieved without isotope labeling, highlighting the substantial line narrowing associated with a mobile *tert*-butyl group of Tby. Once a site-specific signal can be observed, it offers a convenient probe for further structural analysis. In the case of *Bst* DnaB we were able to verify the crystal structure of the hexamer by the selective line broadening observed for one of the signals, when an intermonomer contact with a paramagnetic tag was produced in a cysteine mutant of the protein. The C1-tag is a cyclen-based lanthanide-binding tag designed to induce pseudocontact shifts (PCS). In the case of the *Bst* DnaB mutant with Tby residue and C1-Ce tag, the observation of line broadening rather than PCS can be attributed to paramagnetic relaxation enhancement, which becomes more severe with increasing molecular weight,[27] or to reduced motional freedom of the *tert*-butyl group due to interactions with the tag. For protein systems of lower molecular mass, where the *tert*-butyl signal is more readily observable, it will be possible to use a lanthanide tag to gain further structural information by inducing pseudocontact shifts in the *tert*-butyl resonances of Tby residues positioned at different sites of the protein.

■ CONCLUSION

In conclusion, the Tby probe offers superior sensitivity in ^{1}H NMR spectra, which makes larger and less soluble proteins amenable to NMR analysis than hitherto possible. The NMR signal of the *tert*-butyl group can be observed with minimal interference from the protein background. The robustness with which the *tert*-butyl group can be detected also in the presence of large concentrations of buffer and ligands could be enhanced further by ^{13}C labeling. While, in principle, isotope labels can be deployed to study wild-type proteins, the introduction of a Tby residue requires a point mutation. Notably, however, the capacity to study proteins without perdeuteration is important too, as perdeuterated proteins can be significantly less stable than their unlabeled forms, which can result in slightly different structures.[28,29] Unlabeled Tby is commercially available and inexpensive. We anticipate that the Tby probe will become an extremely valuable tool in structural biology and drug discovery.

■ ASSOCIATED CONTENT

Supporting Information

Figures illustrating the DnaB construct used, the production of full-length versus truncated protein, the effect of cleaving GB1 from DnaB, the effect of DMSO on *tert*-butyl group chemical shifts, the chemical structures of the C1-Ce tag and the sortase A inhibitor, 1D ^{19}F NMR spectra of trifluoromethyl-phenyl-alanine labeled DnaB, a glutamate binding affinity measurement performed with 10 μM DEBP, and the nucleotide sequence of the construct of *E. coli* DEBP. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Authors

*gottfried.otting@anu.edu.au
*t.huber@anu.edu.au

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

We thank Prof. Nicholas E. Dixon for helpful discussions on DnaB, Prof. Peter G. Schultz for the pEVOL and pUltra vector with the *p*-cyano-Phe-tRNA synthetase, Dr. Bim Graham for the C1-Ce tag, and Dr. Colin J. Jackson for an expression system of wild-type DEBP. W.-N.C. thanks the government of the People's Republic of China for a CSC scholarship. Financial support by the Australian Research Council is gratefully acknowledged.

DOI: 10.1021/jacs.5b01918
J. Am. Chem. Soc. 2015, 137, 4581−4586

REFERENCES

(1) Ryabov, Y. E.; Geraghty, C.; Varshney, A.; Fushman, D. *J. Am. Chem. Soc.* **2006**, *128*, 15432−15444.
(2) Fiaux, J.; Bertelsen, E. B.; Horwich, A. L.; Wüthrich, K. *Nature* **2002**, *418*, 207−211.
(3) Sprangers, R.; Kay, L. E. *Nature* **2007**, *445*, 618−622.
(4) Mas, G.; Crublet, E.; Hamelin, O.; Gans, P.; Boisbouvier, J. *J. Biomol. NMR* **2013**, *57*, 251−262.
(5) Rosenzweig, R.; Kay, L. E. *Annu. Rev. Biochem.* **2014**, *83*, 291−315.
(6) Ozawa, K.; Horan, N. P.; Robinson, A.; Yagi, H.; Hill, F. R.; Jergic, S.; Xu, Z.-Q.; Loscha, K. V.; Li, N.; Tehei, M.; Oakley, A. J.; Otting, G.; Huber, T.; Dixon, N. E. *Nucleic Acids Res.* **2013**, *41*, 5354−5367.
(7) Young, D. D.; Young, T. S.; Jahnz, M.; Ahmad, I.; Spraggon, G.; Schultz, P. G. *Biochemistry* **2011**, *50*, 1894−1900.
(8) Wang, Y.-S.; Fang, X.; Wallace, A. L.; Wu, B.; Liu, W. R. *J. Am. Chem. Soc.* **2012**, *134*, 2950−2953.
(9) Cellitti, S. E.; Jones, D. H.; Lagpacan, L.; Hao, X.; Zhang, Q.; Hu, H.; Brittain, S. M.; Brinker, A.; Caldwell, J.; Bursulaya, B.; Spraggon, G.; Brock, A.; Ryu, Y.; Uno, T.; Schultz, P. G.; Geierstanger, B. H. *J. Am. Chem. Soc.* **2008**, *130*, 9268−9281.
(10) Jackson, J. C.; Hammill, J. T.; Mehl, R. A. *J. Am. Chem. Soc.* **2007**, *129*, 1160−1166.
(11) Chatterjee, A.; Sun, S. B.; Furman, J. L.; Xiao, H.; Schultz, P. G. *Biochemistry* **2013**, *52*, 1828−1837.
(12) Loscha, K. V.; Herlt, A. J.; Qi, R.; Huber, T.; Ozawa, K.; Otting, G. *Angew. Chem., Int. Ed.* **2012**, *51*, 2243−2246.
(13) Bailey, S.; Eliason, W. K.; Steitz, T. A. *Science* **2007**, *318*, 459−463.
(14) San Martin, M. C.; Stamford, N. P. J.; Dammerova, N.; Dixon, N. E.; Carazo, J.-M. *J. Struct. Biol.* **1995**, *114*, 167−176.
(15) Yu, X.; Jezewska, M. J.; Bujalowski, W.; Egelman, E. H. *J. Mol. Biol.* **1996**, *259*, 7−14.
(16) San Martin, C.; Radermacher, M.; Wolpensinger, B.; Engel, A.; Miles, C. S.; Dixon, N. E.; Carazo, J.-M. *Structure* **1998**, *6*, 501−509.
(17) Graham, B.; Loh, C. T.; Swarbrick, J. D.; Ung, P.; Shin, J.; Yagi, H.; Jia, X.; Chhabra, S.; Pintacuda, G.; Huber, T.; Otting, G. *Bioconjugate Chem.* **2011**, *22*, 2118−2125.
(18) Zhulenkovs, D.; Rudevica, Z.; Jaudzems, K.; Turks, M.; Leonchiks, A. *Bioorg. Med. Chem.* **2014**, *22*, 5988−6003.
(19) Ilangovan, U.; Ton-That, H.; Iwahara, J.; Schneewind, O.; Clubb, R. T. *Proc. Natl. Acad. Sci. U. S. A.* **2001**, *98*, 6056−6061.
(20) Suree, N.; Liew, C. K.; Villareal, V. A.; Thieu, W.; Fadeev, E. A.; Clemens, J. J.; Jung, M. E.; Clubb, R. T. *J. Biol. Chem.* **2009**, *284*, 24465−24477.
(21) Goldup, S. M.; Pilkington, C. J.; White, A. J. P.; Burton, A.; Barrett, A. G. M. *J. Org. Chem.* **2006**, *71*, 6185−6191.
(22) Fielding, L. *Prog. NMR Spectrosc.* **2007**, *51*, 219−242.
(23) Marti, D. N.; Hu, C.-K.; An, S. S. A.; von Haller, P.; Schaller, J.; Llinás, M. *Biochemistry* **1997**, *36*, 11591−11604.
(24) Dalvit, C.; Fasolini, M.; Flocco, M.; Knapp, S.; Pevarello, P.; Veronesi, M. *J. Med. Chem.* **2002**, *45*, 2610−2614.
(25) Gardner, K. H.; Kay, L. E. *Annu. Rev. Biophys. Biomol. Struct.* **1998**, *27*, 357−406.
(26) Lu, K.; Miyazaki, Y.; Summers, M. F. *J. Biomol. NMR* **2010**, *46*, 113−125.
(27) Otting, G. *J. Biomol. NMR* **2008**, *42*, 1−9.
(28) Hattori, H.; Crespi, H. L.; Katz, J. J. *Biochemistry* **1965**, *4*, 1213−1225.
(29) Piszczek, G.; Lee, J. C.; Tjandra, N.; Lee, C.-R.; Seok, Y.-J.; Levine, R. L.; Peterkofsky, A. *Arch. Biochem. Biophys.* **2011**, *507*, 332−342.

DOI: 10.1021/jacs.5b01918
J. Am. Chem. Soc. 2015, 137, 4581−4586

Supporting Information

O-tert-Butyltyrosine, an NMR tag for high-molecular weight systems and measurements of submicromolar ligand binding affinities

Wan-Na Chen, Kekini Vahini Kuppan, Michael David Lee, Kristaps Jaudzems, Thomas Huber, Gottfried Otting

Expression constructs

The DNA nucleotide sequence of the *Bacillus stearothermophilus* (*Bst*) gene for DnaB was codon optimized for expression in *E. coli* using the program OPTIMISER.[1] The construct of the wild-type protein included a protein G B1 domain as a N-terminal solubility enhancement tag,[2] preceded by a MASMTG sequence (coded by 5' nucleotides of the bacteriophage T7 gene 10) and a His_6-tag. A tobacco etch virus (TEV) protease cleavage site was inserted between the GB1 domain and the DnaB (Figure S1).

The *S. aureus* sortase A (SrtA) construct used in this study is the catalytic domain (residues 60-206) of the full-length protein without the signal polypeptide. The catalytic core has been reported to retain the enzymatic activity.[3,4] The sequence was preceded by a MASMTG tag at the N-terminus, and contained a TEV cleavage site and a His_6-tag at the C-terminal end.

The expression construct for the wild-type *E. coli* aspartate/glutamate binding protein (DEBP) was preceded by a MHHHHHHMENLYFQG sequence, resulting in 291 residues.

The genes were cloned into the T7 expression plasmid pETMCSI,[5] which was transformed into the *E. coli* strain TOP10 for plasmid progression. Sites for the incorporation of *O-tert*-butyl-tyrosine (Tby) were selected by inspection of the three-dimensional protein structures to make sure that the *tert*-butyl group would be solvent exposed, using the structures with the PDB accession codes 2R6A for *Bst* DnaB,[6] 1IJA and 2KID for *S. aureus* SrtA,[4,7] and 2VHA for *E. coli* DEBP.[8] Site-directed mutations were generated by two pairs of primers (IDT-DNA Technologies) using Q5 DNA polymerase (New England Biolabs), following a modified overlap extension protocol as described.[9] The PCR amplified products were digested with the restriction enzymes *Nde*I and *Eco*RI for insertion into pETMCSI.

Protein expression and purification

Bst DnaB proteins containing Tby were produced by using a published pUltra vector[10] to express *p*-cyanophenylalanyl-tRNA synthetase (*p*CNF-RS) and the requisite $tRNA_{CUA}$. The pUltra vector was co-transformed into *E. coli* BL21 (DE3) together with the pETMCSI T7 vector harbouring the target *Bst* DnaB mutant. The proteins were produced in a high-cell density protocol to minimize the usage of unnatural amino acid.[11] After the 1L cell culture reached OD_{600} 0.5-1.0, the cells were spun down at 5,000 *g* at 10 ^{o}C and then resuspended in 250 mL fresh Luria-Bertani medium containing 1 mM Tby (Sigma-Aldrich). Following a recovery period of 1-2 h shaking at 37 ^{o}C, isopropyl-*β*-D-thiogalactopyranoside (IPTG) was added to a final concentration of 1 mM for protein overexpression. Perdeuterated DnaB samples were prepared in a similar way, except that the cells from a 1 L culture were washed twice using 25 mL minimal medium made from 70% D_2O and 99% D_2O prior to resuspension in minimal medium prepared in 99% deuterium oxide. Unlabeled ammonium chloride and glucose were used as nitrogen and carbon source, respectively. Following overnight expression at 25 ^{o}C, the cells were lysed by sonication and the lysate was loaded onto a 5 mL Ni-NTA column (GE Healthcare, USA). The target protein was then eluted using a gradient buffer mixture of buffer A (50 mM Tris-HCl, pH 7.5, 300 mM NaCl) and buffer B (50 mM Tris-HCl, pH 7.5, 300 mM NaCl, 300 mM imidazole). To separate the truncation product from the full-length DnaB, the protein was dialyzed into buffer C (25 mM Tris-HCl, pH 7.5, 300 mM NaCl, 1 mM DTT), concentrated to 500 µL using a centrifugal filter unit with molecular weight cut-off (MWCO) of 10 kDa (Amicon Ultra, Millipore, Billerica, USA), and loaded onto a Superdex 200 column (GE Healthcare, USA). The protein was eluted with buffer C. Finally, the buffer was exchanged to NMR buffer (25 mM Tris-HCl, pH 7.5, 150 mM NaCl) using a centrifugal filter device (MWCO 10 kDa).

To explore the effect of the GB1 domain fusion to DnaB on 1D ^{1}H NMR spectra, TEV protease was added after the gel filtration step at 1:100 weight ratio and the solution dialyzed into buffer D (25 mM Tris-HCl, pH 7.5, 300 mM NaCl, 1 mM 2-mercaptoethanol) at 4 ^{o}C for 14 h. The digested mixture was passed onto a 1 mL Ni-NTA gravity column (GE Healthcare, USA) and the flow-through fraction was exchanged into NMR buffer using a centrifugal filter unit (MWCO 10 kDa).

To minimize sample losses, the perdeuterated DnaB samples were applied to a centrifugal filter unit (MWCO 100 kDa) with NMR buffer instead of using gel filtration to separate the truncation product from the hexamer.

S. aureus SrtA and *E. coli* DEBP were produced by continuous exchange cell-free protein synthesis using PCR amplified DNA as template.[9,12-14] RF1-depleted S30 extract was prepared from an engineered *E. coli* BL21(DE3) strain.[13,15] The proteins were purified by diluting the reaction mixture with buffer A, loading the mixture onto a 1 mL Ni-NTA gravity column, washing with buffer A plus 30 mM imidazole, and eluting with buffer B. *S. aureus* SrtA was subsequently dialyzed into NMR buffer (20 mM MES buffer, pH 6.5, 50 mM NaCl). Following the Ni-NTA purification and prior to exchange into NMR buffer, *E. coli* DEBP was unfolded by incubation for 2 h at room temperature in the presence of 6 M guanidinium hydrochloride to release any bound substrate. Protein refolding and removal of denaturant were achieved by dialysis against NMR buffer (20 mM phosphate buffer, pH 7.5, 100 mM NaCl). Finally, the protein samples were concentrated using centrifugal filter units (MWCO 10 kDa).

Ligation with a lanthanide binding tag

The *Bst* DnaB Ala303Cys mutant was ligated with the C1 tag loaded with Ce^{3+} [16] by incubation for 14 h at room temperature with a three-fold excess of C1-Ce tag. Excess tag was removed by washing with NMR buffer using a centrifugal filter unit (MWCO 10 kDa).

NMR spectroscopy and measurements of K_d values

All 1D ^1H NMR spectra were recorded of 2-50 μM protein solutions in aqueous buffer containing 10% D_2O at 25 ^oC, using a Bruker 600 MHz or 800 MHz NMR spectrometers equipped with TCI cryoprobes. The spectra of DnaB were recorded using the jump-and-return pulse sequence,[17] while a double spin echo[18] was used to record the spectra of SrtA and DEBP. For accurate determination of peak positions in the titration experiment of DEBP with glutamate, line shape fits were performed for the *tert*-butyl resonance using TopSpin 2. The chemical shifts were fitted by the program Origin 8.0 using equation 1.

Figure S1. *Bst* DnaB expression construct used. The N-terminal fusions increased the molecular mass of the monomeric protein to 58.7 kDa. Due to the flexibility of the linker between GB1 domain and DnaB, this increase in molecular mass is not expected to influence the NMR spectrum of the hexamer very much.

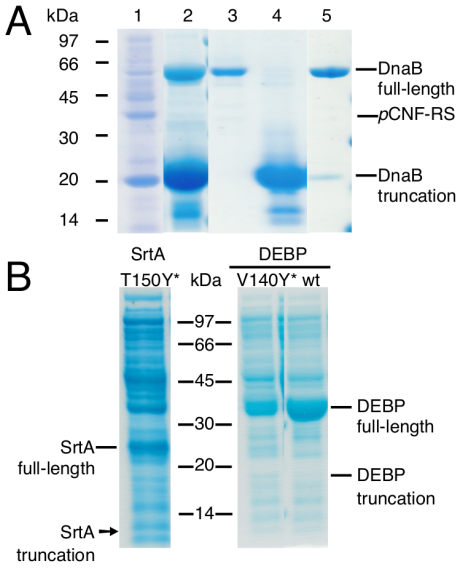


Figure S2. Protein production and purity analyzed by 15% SDS-polyacrylamide gel electrophoresis. The gels were stained with Coomassie blue. The positions of the expected full-length and truncated samples are indicated. (A) *Bst* DnaB produced *in vivo* with Tby incorporated at position 104. Lane 1: Crude protein fraction after IPTG induction. Lane 2: Elution from Ni-NTA affinity purification. Lane 3: Elution at 8 mL from Superdex 200 gel filtration. Lane 4: Elution at 12 mL from Superdex 200 gel filtration. Lane 5: Sample after Ni-NTA column and filtration with a centrifugal filter unit (MWCO 100 kDa), without using gel filtration. (B) Samples of SrtA and DEBP produced by cell-free synthesis. The expected positions of the truncation products are indicated. Left panel: cell-free reaction mixture containing SrtA with Tby (denoted Y*) at position 150. There is no evidence for truncation product. Right panels: DEBP with Tby at position 140 and wild-type protein. Notably, although the amber stop codon at site 140 was followed by a thymidine in the DNA sequence (Figure S7), the expression level of the Val140Tby mutant was not much reduced compared with the wild-type protein, with only little evidence of truncation in contrast to expectations based on previous literature.[19]

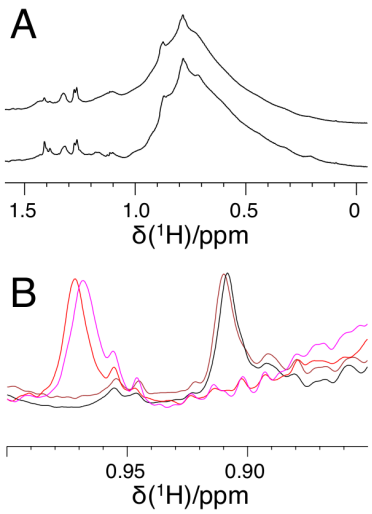


Figure S3. Control experiments to assess the effect of the fusion with GB1 on the structural integrity of wild-type *Bst* DnaB and the effect of DMSO on the chemical shift of the *tert*-butyl resonance in SrtA. (A) 40 µM solution of *Bst* DnaB in 25 mM Tris-HCl, pH 7.5, 150 mM NaCl. Top: after cleavage of the GB1 domain with TEV protease; bottom: before TEV protease cleavage. A few more signals can be observed in the uncleaved product, suggesting that the GB1 moiety is mobile relative to the DnaB. The conservation of the overall appearance of the spectrum suggests that the hexamer remains unaffected by the fusion with GB1. (B) *tert*-butyl resonance measured of a 20 µM solution of SrtA with Tby incorporated at position 150. Black: SrtA in 20 mM MES buffer, pH 6.5, 50 mM NaCl, 5 mM $CaCl_2$; brown: same as the black spectrum, but with 1% DMSO; red: SrtA in 20 mM MES buffer, pH 6.5, 50 mM NaCl, 1 mM EDTA; magenta: same as red spectrum, but with 1% DMSO. The chemical shift changes induced by DMSO are much smaller than the chemical shift changes generated by the inhibitor.

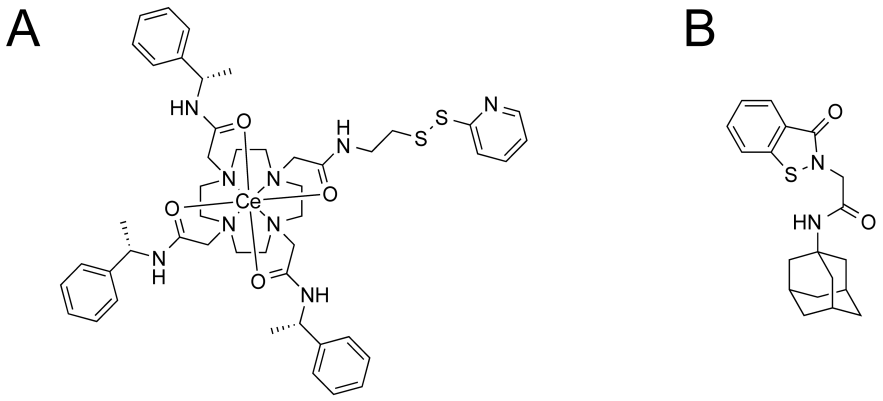


Figure S4. Chemical structures of compounds used. (A) C1-Ce tag.[16] (B) SrtA inhibitor N-(adamantan-1-yl)-2-(3-oxo- 2,3-dihydro-1,2-benzothiazol-2-yl)-acetamide.[20]

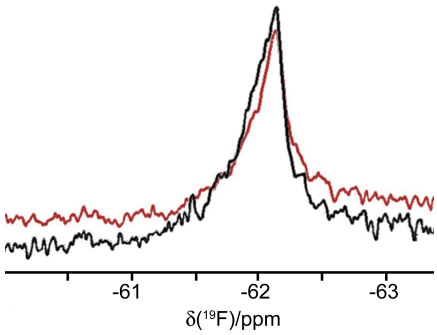


Figure S5. ^{19}F NMR spectra of 70 μM solutions of *Bst* DnaB with trifluoromethyl-phenylalanine (tfmF) incorporated at site 104. The samples were prepared using the same protocol as for the preparation of the Tby samples described above, except that Tby was replaced by 1 mM tfmF. The buffer used contained 50 mM Tris-HCl, pH 7.4, 1 mM EDTA, 1 mM DTT, and 200 mM NaCl. The spectra were recorded at room temperature on a Varian Inova 500 MHz NMR spectrometer in 4 h each. Spectrum calibration was relative to external tfmF (-62.1 ppm). Black: after TEV protease cleavage of the GB1 solubilization domain. Brown: before TEV protease cleavage.

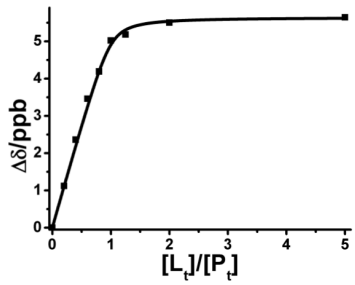


Figure S6. Binding affinity measurement of 10 μM *E. coli* DEBP for glutamate. The spectra were recorded in 20 mM phosphate buffer, pH 7.5, 100 mM NaCl at 25 ^{o}C, using a Bruker 600 MHz NMR spectrometer. The fit yielded a K_d value of 0.16 μM with a standard error of 0.04 μM. This result is unreliable because the K_d value is more than 10 times smaller than the protein concentration.

ATGCACCATCACCATCACCATATGGAAAACCTGTATTTTCAGGGCATGGCCGCAGGCAGCACGCTGGACAAAATCGCCAA
AAACGGTGTGATTGTCGTCGGTCACCGTGAATCTTCAGTGCCTTTCTCTTATTACGACAATCAGCAAAAAGTGGTGGGTT
ACTCGCAGGATTACTCCAACGCCATTGTTGAAGCAGTGAAAAAGAAACTCAACAAACCGGACTTGCAGGTAAAACTGATT
CCGATTACCTCACAAAACCGTATTCCACTGCTGCAAAACGGCACTTTCGATTTTGAATGTGGTTCTACCACCAACAACGT
CGAACGCCAAAAACAGGCGGCTTTCTCTGACACTATTTTCGTGGTCGGTACGCGCCTGTTGACCAAAAAGGGTGGCGATA
TCAAAGATTTTGCCAACCTGAAAGACAAAGCCGTAGTCGTCACTTCCGGCACTACCTCTGAA**GTT**TTGCTCAACAAACTG
AATGAAGAGCAAAAAATGAATATGCGCATCATCAGCGCCAAAGATCACGGTGACTCTTTCCGCACCCTGGAAAGCGGTCG
TGCCGTTGCCTTTATGATGGATGACGCTCTGCTGGCCGGTGAACGTGCGAAAGCGAAGAAACCAGACAACTGGGAAATCG
TCGGCAAGCCGCAGTCTCAGGAGGCCTACGGTTGTATGTTGCGTAAAGATGATCCGCAGTTCAAAAAGCTGATGGATGAC
ACCATCGCTCAGGTGCAGACCTCCGGTGAAGCGGAAAAATGGTTTGATAAGTGGTTCAAAAATCCAATTCCGCCGAAAAA
CCTGAACATGAATTTCGAACTGTCAGACGAAATGAAAGCACTGTTCAAAGAACCGAATGACAAGGCACTGAACTAA

Figure S7. Nucleotide sequence of the construct of *E. coli* DEBP used in the present work. The codon replaced by the amber stop codon is in bold and underlined.

References

(1) Puigbo, P.; Guzman, E.; Romeu, A.; Garcia-Vallve, S. *Nucleic Acids Res*. **2007**, *35* (Web Server issue): W126-131.

(2) Zhou, P.; Wagner G. *J. Biomol. NMR* **2010**, *46*, 23-31.

(3) Mazmanian, S. K.; Liu, G.; Ton-That, H.; Schneewind, O. *Science* **1999**, *285*, 760-763.

(4) Ilangovan, U.; Ton-That, H.; Iwahara, J.; Schneewind, O.; Clubb, R. T. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 6056-6061.

(5) Neylon, C.; Brown, S. E.; Kralicek, A. V.; Miles, C. S.; Love, C. A.; Dixon, N. E. *Biochemistry* **2000**, *39*, 11989-11999.

(6) Bailey, S.; Eliason, W. K.; Steitz, T. A. *Science* **2007**, *318*, 459-463.

(7) Suree, N.; Liew, C. K.; Villareal, V. A.; Thieu, W.; Fadeev, E. A.; Clemens, J. J.; Jung, M. E.; Clubb, R. T. *J. Biol. Chem*. **2009**, *284*, 24465-24477.

(8) Hu, Y. L.; Fan, C.-P.; Fu, G. S.; Zhu, D. Y.; Jin, Q.; Wang, D.-C. *J. Mol. Biol*. **2008**, *382*, 99-111.

(9) Wu, P. S. C.; Ozawa, K.; Lim, S. P.; Vasudevan, S. G.; Dixon, N. E.; Otting, G. *Angew. Chemie Int. Ed.* **2007**, *46*, 3356-3358.

(10) Chatterjee, A.; Sun, S. B.; Furman, J. L.; Xiao, H.; Schultz, P. G. *Biochemistry* **2013**, *52*, 1828-1837.

(11) Sivashanmugam, A.; Murray, V.; Cui, C. X.; Zhang, Y. H.; Wang, J. J.; Li, Q. Q. *Protein Sci*. **2009**, *18*, 936-948.

(12) Ozawa, K.; Headlam, M. J.; Schaeffer, P. M.; Henderson, B. R.; Dixon, N. E.; Otting, G. *Eur. J. Biochem.* **2004**, *271*, 4084-4093.

(13) Apponyi, M. A.; Ozawa, K.; Dixon, N. E.; Otting, G. *Methods Mol. Biol.* **2008**, *426*, 257-268.

(14) Jia, X.; Ozawa, K.; Loscha, K. V.; Otting, G. *J. Biomol. NMR* **2009**, *44*, 59-67.

(15) Loscha, K. V.; Herlt, A. J.; Qi, R.; Huber, T.; Ozawa, K.; Otting, G. *Angew. Chem. Int. Ed*. **2012**, *51*, 2243-2246.

(16) Graham, B.; Loh, C. T.; Swarbrick, J. D.; Ung, P.; Shin, J.; Yagi, H.; Jia, X., Chhabra, S.; Pintacuda, G.; Huber, T.; Otting, G. *Bioconjugate Chem*. **2011**, *22*, 2118-2125.

(17) Plateau, P.; Guéron, M. *J. Am. Chem. Soc*. **1982**, *104*, 7310-7311.

(18) Hwang, T. L.; Shaka, A. J. *J. Magn. Reson. A* **1995**, *112*, 275-279.

(19) Smolskaya, S.; Zhang, Z. J.; Alfonta, L. *PLoS One* **2013**, *8*, e68363.

(20) Zhulenkovs, D.; Rudevica, Z.; Jaudzems, K.; Turks, M.; Leonchiks, A. *Bioorg. Med. Chem*. **2014**, *22*, 5988-6003.

JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Sensitive NMR Approach for Determining the Binding Mode of Tightly Binding Ligand Molecules to Protein Targets

Wan-Na Chen,[†] Christoph Nitsche,[†,‡] Kala Bharath Pilla,[†] Bim Graham,[§] Thomas Huber,[†] Christian D. Klein,[‡] and Gottfried Otting*[,†]

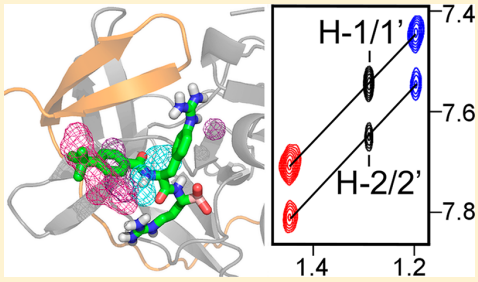
[†]Australian National University, Research School of Chemistry, Canberra, ACT 2601, Australia

[‡]Medicinal Chemistry, Institute of Pharmacy and Molecular Biotechnology, Heidelberg University, Im Neuenheimer Feld 364, 69120 Heidelberg, Germany

[§]Medicinal Chemistry, Monash Institute of Pharmaceutical Sciences, Monash University, Parkville, VIC 3052, Australia

Supporting Information

ABSTRACT: Structure-guided drug design relies on detailed structural knowledge of protein−ligand complexes, but crystallization of cocomplexes is not always possible. Here we present a sensitive nuclear magnetic resonance (NMR) approach to determine the binding mode of tightly binding lead compounds in complex with difficult target proteins. In contrast to established NMR methods, it does not depend on rapid exchange between bound and free ligand or on stable isotope labeling, relying instead on a *tert*-butyl group as a chemical label. *tert*-Butyl groups are found in numerous protein ligands and deliver an exceptionally narrow and tall ^{1}H NMR signal. We show that a *tert*-butyl group also produces outstandingly intense intra- and intermolecular NOESY cross-peaks. These enable measurements of pseudocontact shifts generated by lanthanide tags attached to the protein, which in turn allows positioning of the ligand on the protein. Once the ligand has been located, assignments of intermolecular NOEs become possible even without prior resonance assignments of protein side chains. The approach is demonstrated with the dengue virus NS2B-NS3 protease in complex with a high-affinity ligand containing a *tert*-butyl group.



INTRODUCTION

Structure-guided ligand design is a powerful approach in drug discovery. Usually, the approach relies on X-ray crystallography to obtain the 3D structural information on the protein−ligand complexes required to guide further ligand development. NMR spectroscopy is often used to aid this process when single crystals of the protein−ligand complexes cannot be obtained. As the NMR resonances of ligand molecules are often difficult to discriminate against the background of protein NMR signals, however, NMR spectroscopic methods tend to perform best for weakly binding and rapidly exchanging ligands that can be used in large excess over the protein.[1] In the case of tightly binding ligands, the ligand molecule can, for all intent and purpose, be considered to act like an integral part of the protein.

The most common NMR approaches to characterize the complexes between proteins and tightly binding ligands involve chemical shift perturbations[2] and nuclear Overhauser effects (NOE).[3] Both methods rely on the assignment of a large number of NMR signals. In particular, reliable assignments of intermolecular NOESY cross-peaks require either near-complete resonance assignments of the target protein, which are difficult to obtain for large target proteins irrespective of the availability of ^{15}N or ^{13}C-labeled protein samples, or information from independent data, such as site-directed mutagenesis experiments.[4] Detailed structural interpretation of chemical shift perturbations is difficult and requires measurements with a series of ligand derivatives.

In recent years, paramagnetic NMR methods have increasingly been recruited for structural analysis of protein−ligand complexes, in particular by making use of paramagnetic relaxation enhancements (PRE) and pseudocontact shifts (PCS) generated by site-specifically attached tags.[5] While much work has focused on these parameters for the analysis of weakly binding complexes,[6−12] the ligand PCSs are exceedingly small if a large excess of ligand is used. The reverse experiment, generating PCSs in the target protein by attaching a paramagnetic metal complex to the ligand, requires a significant chemical modification of the ligand and comprehensive resonance assignments of the protein for data interpretation.[13,14] Examples are rare where PCSs have been used to establish the structure and orientation of a tightly binding ligand at its binding site on the protein.[11,15−17] The fundamental problem in these cases arises from the difficulty in identifying the resonances of the tightly bound ligand. In the special situation of a protein−protein complex, isotope labeling allows selective observation of either protein component, which has been used for structure analysis of high-affinity protein−

Received: January 13, 2016
Published: March 14, 2016

DOI: 10.1021/jacs.6b00416
J. Am. Chem. Soc. 2016, 138, 4539−4546

protein complexes using PCSs induced by paramagnetic lanthanide ions.[15,17] In the case of a small-molecule ligand, however, synthesis of an isotope-labeled version would usually be prohibitively laborious. An alternative strategy employs selective suppression of the ^{1}H NMR spectrum of the protein by isotope labeling, either using ^{15}N/^{13}C-labeling in combination with half-filter pulse sequences,[18] or by perdeuteration of the protein.[19] Either approach is imperfect, as the half-filters achieve only incomplete suppression of the protein NMR signals, due to nonuniform one-bond coupling constants, and are associated with significant signal decay by relaxation, while very high deuteration levels are required to reduce the residual protein ^{1}H NMR signals to a level below those of the ligand. Instead of relying on costly isotope labeling, the present approach makes use of the unique NMR properties of *tert*-butyl groups.

The ^{1}H NMR signal of a *tert*-butyl group is an intense singlet with a chemical shift in the range of about 1.0 to 1.5 ppm depending on the chemical environment. Owing to rapid methyl rotation and methyl reorientation within this flexible group, the signal of a solvent-exposed *tert*-butyl group is much narrower than protein ^{1}H NMR signals, enabling its detection even in protein complexes of high molecular mass without any isotope labeling.[20] In experiments aimed at probing chemical exchange phenomena, we observed that the *tert*-butyl group also produces NOESY cross-peaks that are much more intense than anticipated and can be observed even under conditions where cross-peaks of the protein are barely detectable. This opens a new route to characterize the binding mode of a ligand labeled with a *tert*-butyl group. First, the NOESY cross-peaks with the *tert*-butyl signal enable measurements of PCSs induced by paramagnetic tags attached to the protein at different sites. These data position the *tert*-butyl group with respect to the protein. If intraligand NOEs with the *tert*-butyl group can be observed, the PCSs of additional ligand protons can be used to orient the ligand in its binding site. Finally, intermolecular NOEs with the *tert*-butyl group further pinpoint its location on the protein.

We demonstrate the approach with the structure determination of the complex between the two-component dengue virus NS2B-NS3 protease (NS2B-NS3pro) from serotype 2 (in the following referred to as DENpro) and a newly developed ligand that contains a boronic acid warhead, binds with high affinity, and contains a *tert*-butyl group (C. Nitsche and C. D. Klein, manuscript in preparation). Although a dengue vaccine has been developed that is effective against all serotypes, the vaccine offers only incomplete protection and is ineffective in children below the age of 6, who need it most.[21] Dengue causes severe complications and is spreading rapidly,[22] making DENpro a drug target of high priority.[23−25] Despite intense efforts over many years, however, crystallization of DENpro remains difficult and only a single structure has been published, which proved to be of an inactive conformation (PDB code 2FOM).[26] Co-crystal structures with inhibitors are available of serotype 3 (PDB code 3U1I and 3U1J)[27] and of the homologous West Nile virus protease,[26] which allow modeling of the structure of DENpro. In the absence of exact structural information, the development of high-affinity inhibitors has been slow. An effective inhibition constant of 43 nM was reported for the peptidic inhibitor Bz-Nle-Lys-Arg-Arg-B(OH)$_2$ owing to the presence of the electrophilic boronic acid warhead which is thought to form a covalent bond to Ser135 at the active site of DENpro.[28] A range of further inhibitors lacking a boronic acid warhead has been developed,[25,29,30] but their binding modes remain uncertain in the absence of experimental confirmation for any of the computational models built.

DENpro has a molecular mass of about 27 kDa and degrades readily in the absence of high-affinity ligands. Its NMR spectrum shows broad resonances and strong signal overlap, and varies between different constructs and inhibitors.[31−35] The well-resolved ^{15}N-HSQC cross-peaks comprise about half of the backbone amides (Figure S3). We identified three single-cysteine mutants of DENpro for the attachment of paramagnetic lanthanide tags, which generate pronounced PCSs in the NMR spectrum of the protein.[31] Using the resolved ^{15}N-HSQC cross-peaks, PCS measurements are straightforward and allowed characterization of the conformation of DENpro in isolation and in complex with a small-molecule and a protein inhibitor, respectively.[20,31] Assessing the binding mode of the synthetic nonprotein inhibitor, however, proved intractable because of the difficulty in observing and assigning intermolecular NOEs.

In the following we report the binding mode of the high-affinity ligand **1** (Figure 1) to DENpro. This is the first time that the complex between DENpro and a small-molecule ligand

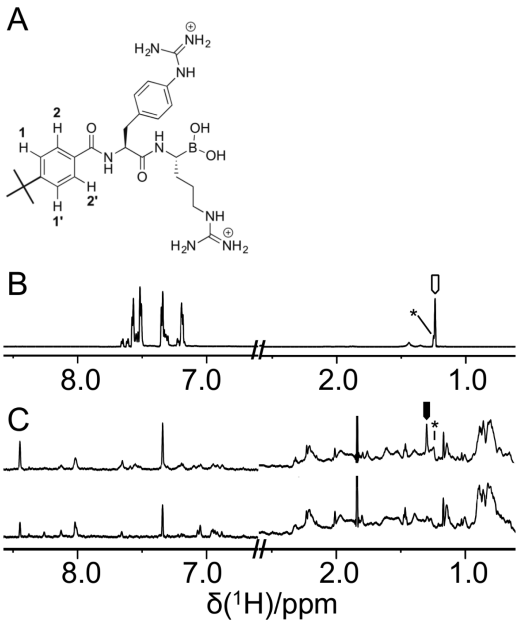


Figure 1. 1D ^{1}H NMR spectra of ligand **1** without and with DENpro. (A) Chemical structure of ligand **1**. Ligand **2** is structurally identical except for the absence of the *tert*-butyl group. (B) 1D ^{1}H NMR spectrum of ligand **1** in the absence of DENpro. The spectrum was recorded at 25 °C in NMR buffer (20 mM MES, pH 6.5, 50 mM NaCl) using a 800 MHz NMR spectrometer. The downfield region of the spectrum was scaled up 8-fold relative to the upfield part of the spectrum. The open arrow marks the ^{1}H NMR resonance of the *tert*-butyl group. The star identifies the signal from an impurity. (C) 1D ^{1}H NMR spectra of 50 μM solutions of DENpro in deuterated NMR buffer without ligand (bottom trace) and in a 1:1 complex with ligand **1** (top trace). The spectral appearance of the aromatic region was simplified by perdeuteration of Tyr, Phe, and Trp. Signals at 7.3 and 8.0 ppm are from the C-terminal His$_6$-tag. The low-field region was scaled up 2-fold relative to the high-field region. The resonance of the *tert*-butyl group of ligand **1** is marked by a filled arrow. The star identifies the signal of a nonbinding impurity.

DOI: 10.1021/jacs.6b00416
J. Am. Chem. Soc. 2016, 138, 4539−4546

has been modeled at atomic resolution using experimental NMR data. Importantly, the new protocol succeeded with a limited set of ^{15}N-HSQC cross-peaks of backbone amides, which are sufficiently well resolved to enable a straightforward transfer of their assignments between the free protein and complexes with different inhibitors.

MATERIALS AND METHODS

Sample Preparation. Three single-cysteine mutants of DENpro were prepared. The mutants were Ala57*Cys (site A; the asterisk marks residues in NS2B), Ser34Cys (site B), and Ser68Cys (site C). Uniformly ^{15}N-labeled protein samples were produced by over-expression in *E. coli* and selectively isotope-labeled samples by cell-free synthesis as described previously.[31] The samples were ligated with C2-Ln^{3+} tags (Figure S1) as described,[31,35,36] where Ln^{3+} is either Y^{3+}, Tm^{3+}, or Tb^{3+}.

Ligands **1** and **2** were synthesized as described elsewhere (C. Nitsche and C. D. Klein, manuscript in preparation).

NMR Measurements. All NMR spectra were measured on a Bruker 800 MHz NMR spectrometer equipped with a 5 mm TCI cryoprobe, using 3 mm NMR tubes. ^{15}N-HSQC spectra were recorded with t_{1max} = 26.2 ms and t_{2max} = 160 ms. NOESY spectra were recorded with a mixing time of 200 ms, t_{1max} = 58 ms, and t_{2max} = 232 ms. The total recording times of the NOESY spectra varied from 4 h for diamagnetic samples to 8 h for paramagnetic samples.

PCSs were measured as ^{1}H chemical shifts observed for DENpro tagged with C2−Tm^{3+} or C2−Tb^{3+} tags minus the corresponding chemical shifts measured with DENpro tagged with diamagnetic C2−Y^{3+}. PCSs are through-space effects that follow the equation[37]

$$\Delta\delta^{PCS} = 1/(12\pi r^3)[\Delta\chi_{ax}(3\cos^2\theta - 1) + 1.5\Delta\chi_{rh}\sin^2\theta\cos 2\varphi] \quad (1)$$

where $\Delta\delta^{PCS}$ is the PCS measured in ppm, r is the distance of the nuclear spin from the metal ion, $\Delta\chi_{ax}$ and $\Delta\chi_{rh}$ are the axial and rhombic components of the magnetic susceptibility anisotropy tensor $\Delta\chi$, and θ and φ are the polar angles describing the position of the nuclear spin with respect to the principal axes of the $\Delta\chi$ tensor. We used the PCSs of backbone amide protons (Table S1) to fit the $\Delta\chi$ tensors to the coordinates of DENpro models generated by Rosetta.[38]

PREs of the *tert*-butyl resonance were measured by comparing its ^{1}H NMR line widths in resolved NOESY cross-peaks recorded for samples tagged with paramagnetic and diamagnetic tags.

Modeling of the Protein−Ligand Complex. A structural model of DENpro in the catalytically active conformation was obtained using the program Modeller[39] and the crystal structure 3U1I[27] of the protease from serotype 3 as the template. A total of 100 models were generated following the advanced modeling protocol,[40] and the best model was selected based on the discrete optimized potential energy (DOPE) score.[41]

The program AutoDock-Vina[42] was used to generate 20 different poses of the docked ligand. Each pose was checked for agreement with the ligand PCSs using six different $\Delta\chi$ tensors, which had been determined from the PCSs of the backbone amides of DENpro with C2−Tm^{3+} and C2−Tb^{3+} tags at three sites. The best-fitting pose was selected for further refinement by the software suite Rosetta[43] to model a covalent bond between the boron and the side chain oxygen of Ser135. Further restraints were introduced to ensure tetrahedral coordination of the boron, as detected in boronic acid complexes of other serine proteases such as the hepatitis C virus protease,[44] and planarity of the guanidinum groups. The all-atom Relax module of Rosetta[43] was used to refine the amino acid side chain conformations in the substrate binding site, following the protocol available in the protocol demonstrations (demos directory) of the Rosetta software suite.[45] From the models generated by the all-atom Relax module, the model that best fulfilled the backbone PCSs of DENpro was selected as the final structure. New $\Delta\chi$ tensors were fitted to each model to find the structure with the lowest PCS score.[46]

RESULTS

^{1}H NMR Resonance of the *tert*-Butyl Group of Ligand 1 in the Absence and Presence of DENpro. In the DENpro construct used in the present work, the linker between NS2B and NS3 contained the natural recognition site EVKKQR, leading to autoproteolytic cleavage at this site. We have shown previously that the resulting product assumes the enzymatically active closed conformation of the protease.[35] The unlinked construct degraded during NMR measurements in less than a day, but was stable for months in the presence of ligand **1**.

The 1D ^{1}H NMR spectrum of the free ligand shows four prominent resonances for the aromatic rings between 7 and 8 ppm, and a particularly intense NMR signal for the *tert*-butyl group (Figure 1). The signal of the *tert*-butyl group is also easily discernible in the 1:1 complex with DENpro, whereas the aromatic resonances of the ligand are difficult to identify even in a sample prepared with a high level of perdeuteration of aromatic residues (Tyr, Phe, and Trp). The chemical shift of the *tert*-butyl group of ligand **1** in the bound state is slightly downfield compared with the chemical shift in the free state. In the presence of excess ligand, the signals of bound and free ligand were observed simultaneously, indicating that any chemical exchange is slow as expected for a tightly bound ligand (data not shown).

PCS Measurements. The strategy of the present work relies on PCS measurements of the protein (to determine the $\Delta\chi$ tensors) and of the ligand (to determine the position of the ligand relative to the $\Delta\chi$ tensors and thus relative to the protein). To account for possible conformational changes caused by ligand binding, PCSs of DENpro were measured in ^{15}N-HSQC spectra recorded of the complex between ^{15}N-labeled DENpro and ligand **1**, using C2−Tm^{3+} and C2−Tb^{3+} tags at sites A−C, with C2−Y^{3+} tagged samples as diamagnetic reference (Figures S2−S4). The PCSs were very similar, but not identical, to PCSs measured previously for samples without inhibitor[35] and with the protein inhibitor BPTI.[17] A total of 222 PCSs (Table S1) were measured and used to determine the $\Delta\chi$-tensor parameters (Table S3).

PCSs of the bound ligand **1** sometimes proved difficult to determine from simple 1D ^{1}H NMR spectra. The intense ^{1}H NMR resonance of the *tert*-butyl group was readily detected in the diamagnetic samples, but PREs combined with changes in the NMR spectrum of the protein can make it difficult to identify the resonance in the paramagnetically tagged samples (Figure S5).

In contrast, NOESY spectra turned out to yield intense cross-peaks with the *tert*-butyl resonance that could easily be followed even when the *tert*-butyl resonance was broadened by PRE. In addition, the NOESY spectra yield PCSs not only for the *tert*-butyl resonance itself but also for the protons involved in the NOEs with the *tert*-butyl group. Figure 2B illustrates the intensity of the *tert*-butyl cross-peaks with the wild-type protein. The cross-peaks are readily identified by their narrow line shape in the detection dimension. As expected, they were absent from the corresponding NOESY spectrum recorded with ligand **2**, which is the same as ligand **1** but without the *tert*-butyl group (Figure 2A). Particularly intense cross-peaks are observed with protons at about 7.5 ppm, which were assigned to the aromatic ligand protons in the positions *ortho* and *meta* to the *tert*-butyl group labeled 1, 1′, 2, and 2′ in Figure 1A. (Using a sample prepared with ^{15}N-labeled DENpro, a NOESY spectrum

DOI: 10.1021/jacs.6b00416
J. Am. Chem. Soc. 2016, 138, 4539−4546

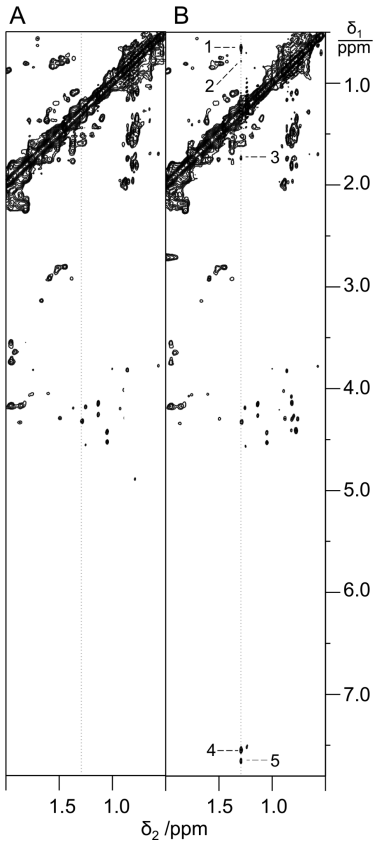


Figure 2. NOESY spectra of DENpro in complex with ligands **1** and **2**. The spectra were recorded at 25 °C of 150 μM solutions of unlabeled wild-type DENpro in NMR buffer containing 10% D_2O, using 3 mm NMR tubes in a 5 mm TCI cryoprobe. Experimental parameters: 200 ms mixing time, t_{1max} = 58 ms, t_{2max} = 232 ms, total recording time per spectrum 4 h. The spectral region shown captures the NOE cross-peaks of the *tert*-butyl resonance. (A) NOESY spectrum of the complex with ligand **2**. (B) NOESY spectrum of the complex with ligand **1**. Five cross-peaks with the *tert*-butyl resonance are labeled. Note that these cross-peaks were readily detected, while other cross-peaks, e.g., with amide protons, were much weaker or unobservable.

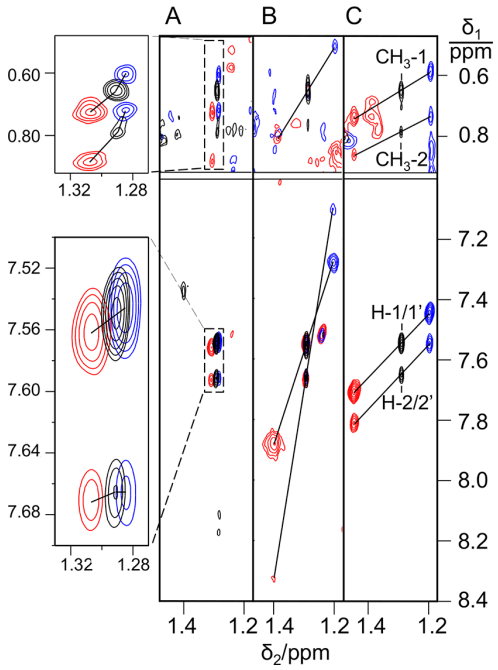


Figure 3. PCSs observed in 2D NOESY spectra of DENpro with C2 tags at three different sites. The spectra were recorded of 150 μM solutions of ^{15}N-labeled DENpro complexes with ligand **1** in NMR buffer at 25 °C. The figure shows superimpositions of NOESY spectra for samples tagged with C2−Y^{3+} (black), C2−Tm^{3+} (blue) and C2−Tb^{3+} (red). The spectral regions shown contain cross-peaks between the *tert*-butyl resonance and protons in the aromatic and methyl regions of the NMR spectrum. PCSs are identified by lines drawn between cross-peaks of the diamagnetic samples and their corresponding peaks in the paramagnetic samples. (A) NOESY spectra of the complex tagged at site A. (B) Same as (A), but for the complex tagged at site B. (C) Same as (A), but for the complex tagged at site C. NOEs with two initially unassigned methyl groups are labeled as CH_3-1 and CH_3-2. The NOEs with two aromatic protons were attributed to the protons 1/1′ and 2/2′ of the ligand (Figure 1A).

recorded without ^{15}N-decoupling confirmed that these NOEs are not with amide protons of the protein.) Additional cross-peaks at about 0.7 and 1.7 ppm cannot be explained as intraligand NOEs and, based on the chemical shifts, include NOEs with methyl groups.

Figure 3 illustrates the measurement of PCSs from the NOESY cross-peaks with the *tert*-butyl resonance. The cross-peaks were preserved between wild-type DENpro and DENpro tagged with diamagnetic C2−Y^{3+} tags, and were displaced by PCSs in the samples containing paramagnetic tags. Different to backbone amides, where PCSs displace ^{15}N-HSQC cross-peaks along mostly parallel lines owing to the close spatial proximity of ^{1}H and ^{15}N spins, the PCSs observed for the protons involved in NOEs with the *tert*-butyl resonance were mostly different from the PCS of the *tert*-butyl resonance itself, as expected for spins that are located at more disparate positions relative to the $\Delta\chi$ tensors of the paramagnetic tags. Importantly, the C2−Tm^{3+} and C2−Tb^{3+} tags produced PCSs of opposite sign[31,36] and the NOESY cross-peaks of the three samples tagged with C2−Tm^{3+}, C2−Y^{3+}, and C2−Tb^{3+} tags could be connected by almost straight lines. This greatly aids the identification of the paramagnetically shifted cross-peaks, even when the cross-peaks were significantly broadened by PREs, as for the paramagnetic samples tagged at site B (Figure 3B).

Altogether, the NOESY cross-peaks observed with the *tert*-butyl resonance delivered 18 ligand PCSs together with 14 protein PCSs (Table S2). The ligand PCSs immediately enable a qualitative structural interpretation regarding the orientation of the *tert*-butylphenyl group relative to the $\Delta\chi$ tensors. For example, the PCSs obtained with the tags at site B decrease in magnitude in moving from the protons at positions 2/2′ to those at positions 1/1′ to the *tert*-butyl group (Figure 3B), suggesting that the *tert*-butyl group faces away from the tag at site B. This conclusion is confirmed by the broad line widths of cross-peaks observed for the H-2/2′ resonances in the paramagnetic samples which can be attributed to PREs due to proximity to the tag.

PRE Measurement. As the *tert*-butyl resonance is a singlet, PREs can simply be measured as the difference in line width observed in paramagnetic and diamagnetic samples. The NOESY cross-peaks of Figure 3 clearly show that the *tert*-butyl resonance is broader in the samples with paramagnetic

DOI: 10.1021/jacs.6b00416
J. Am. Chem. Soc. 2016, 138, 4539−4546

versus diamagnetic tags, and broader with Tb^{3+} tags than with Tm^{3+} tags, as expected from the greater paramagnetic moment of Tb^{3+}.[47] It is difficult, however, to convert the PREs into quantitative distance restraints from the metal ions. First, the significant anisotropy of the magnetic susceptibility tensors associated with Tb^{3+} and Tm^{3+} needs to be taken into account.[48] Second, the PREs can be compromised by cross-correlation effects with nuclear chemical shift anisotropies.[49] Most importantly, however, the linker of the C2 tag is not entirely rigid, resulting in multiple possible metal positions. For a conservative estimate of the distance restraint indicated by a PRE, we calculated the largest distance compatible with the $\Delta\chi$ tensor anisotropies determined from the PCSs using the anisotropic Curie-spin relaxation equations[48] with a rotational correlation time of 10 ns. We only used the data from site B, which showed the greatest PRE effects (Figure 3). In this construct, the C2−Tb^{3+} and C2−Tm^{3+} tags broadened the *tert*-butyl signal by 22 and 13 Hz, respectively, relative to the diamagnetic sample with the C2−Y^{3+} tag. Both measurements correspond to a distance from the paramagnetic center of about 15.5 Å. To account for the uncertainty in metal position, we converted the PREs observed for mutant B into an upper distance restraint of 20 Å.

PCS-Guided Modeling of the Protein−Ligand Complex. As no crystal structure of DENpro in the active conformation is available and the amino acid sequences of NS2B-NS3pro of serotypes 2 and 3 are 66% sequence identical, we modeled the structure of DENpro on the crystal structure 3U1I of serotype 3.[27] Among 100 very similar homology models, we selected the best model based on the DOPE score,[41] which had a Cα-RMSD of 0.7 Å relative to the crystal structure 3U1I.

The $\Delta\chi$ tensors associated with the C2−Tm^{3+} and C2−Tb^{3+} tags at sites A−C produce different PCSs at different sites of the DENpro-ligand complex. Using the $\Delta\chi$ tensors determined for the best-fitting model of DENpro, we predicted the PCSs for every point of a dense grid placed over the structural model of DENpro to identify all the points that were in agreement with the PCSs generated by the C2−Tm^{3+} and C2−Tb^{3+} tags at sites A−C.

Allowing for an uncertainty range of ±0.02 ppm for the experimental PCSs, the PCSs of the *tert*-butyl group of the ligand identified a continuous space in the vicinity of NS2B, where the *tert*-butyl group could be located. We refer to this PCS-supported grid point volume as localization space. The localization space of the *tert*-butyl group was in full agreement with its expected position upon formation of a covalent bond between Ser135 and the boronic acid moiety of the ligand, as anticipated in the design of the ligand. The localization space shown in Figure 4 was also restricted by imposing the 20 Å distance restraint for the tertbutyl group from the metal position of the tag at site B, which had been derived from PREs.

To generate a complete model of the ligand, 20 poses of ligand **1** were generated on the structural model of DENpro using the program AutoDock-Vina.[42] The different poses produced dramatically different ligand structures and orientations (with a pairwise RMSD of about 10 Å) with very similar calculated binding energies (ranging from −7.5 to −7.9 kcal/mol). The two top-ranked and seven further poses approximately positioned the *tert*-butyl group in the localization space defined by the PCSs. Furthermore, the top-ranked pose positioned the boronic acid group in the vicinity of Ser135 and the direction of the *tert*-butylphenyl group was also in general agreement with the localization spaces of the 1/1′ and 2/2′ protons (Figure 4). To refine this model of the protein−ligand complex, we used the Rosetta software suite[38,43] to form a covalent bond between the boron of the ligand and the side chain oxygen of Ser135, and allowed small structural adjustments by the all-atom "Relax" refinement protocol of Rosetta. The protocol generated 1000 structures of the complex, all with conserved protein conformations and similar ligand conformations (over 90% within 1.3 Å of the structure with the lowest PCS energy, Figure S6). The protein structure with the lowest PCS score showed an excellent correlation between the experimental and back-calculated PCSs (Figure S7). Figure 4 shows the structure best fulfilling the PCSs of DENpro, which was chosen as the final representation of the DENpro-ligand **1** complex.

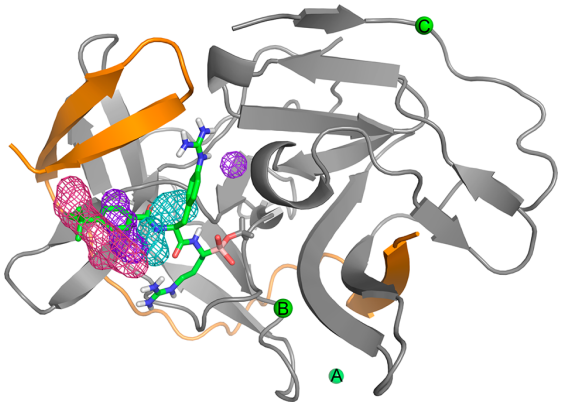


Figure 4. Model of the complex between DENpro and ligand **1**. NS2B (orange) and NS3pro (gray) are drawn in a cartoon representation, and ligand **1** is shown in a stick representation, displaying hydrogens only for amide and guanidinium groups. Green balls indicate the positions of the lanthanide ions as determined by PCSs for the tags at sites A−C. Meshes identify the localization spaces of the *tert*-butyl group (magenta), protons 1/1′ (purple), and protons 2/2′ (cyan) identified by PCSs from C2−Tm^{3+} and C2−Tb^{3+} tags at sites A−C. The meshes capture localization spaces with PCS deviations smaller than 0.02 ppm from the experimental PCSs.

The localization spaces of protons 1/1′ and 2/2′ (Figure 4) are progressively closer to the site B, reflecting the large PCSs induced by C2 tags at this site (Figure 3B). This provides strong support for the orientation of the *tert*-butyl phenyl group. The apparent discrepancies between experimental localization spaces and actual proton positions in the model of Figure 4 may arise from mobility of the tag at site B, which can increase the PCSs by transiently shortening the distance between metal and ligand. As PCSs present long-range restraints, experimental uncertainties translate into relatively large uncertainties for localization spaces. Therefore, we used the intermolecular NOEs observed with the *tert*-butyl group to verify the correctness of the model of the complex.

Verification of the Structural Model by Intermolecular NOEs. The model of the complex between DENpro and ligand **1** predicts intermolecular NOEs between the *tert*-butyl group and the methyl groups of Val154 and Val155. The chemical shifts of the NOEs observed with the signals of CH_3-1 and CH_3-2 (Figure 3) agree with expectations for valine methyl groups.

To support this assignment, we recorded a NOESY spectrum of the DENpro-ligand **1** complex in which wild-type DENpro

DOI: 10.1021/jacs.6b00416
J. Am. Chem. Soc. 2016, 138, 4539−4546

was selectively labeled with ^{13}C-valine. In the absence of ^{13}C decoupling, the NOESY cross-peaks labeled 1 and 2 in Figure 2 split by $^{1}J_{CH}$ coupling constants and the signal of 3 could no longer be resolved, showing that all three protons indeed belong to valine (Figure S8). To distinguish between Val154 and Val155, we measured a NOESY spectrum of the mutant Val155Ile selectively labeled with ^{13}C-valine. The NOESY cross-peaks observed were in agreement with an isoleucine residue and did not split in the absence of ^{13}C-decoupling, indicating that the NOEs between the *tert*-butyl group of ligand **1** and wild-type DENpro are with Val155 (Figure S9).

DISCUSSION

The surprisingly high signal-to-noise ratio of NOESY cross-peaks with a *tert*-butyl group can be attributed to three factors. Most importantly, its ^{1}H NMR resonance is intrinsically narrow because of short effective rotational correlation times arising from fast methyl group rotation and rotation of the *tert*-butyl group. Second, the signal stands out because it corresponds to nine protons. Finally, due to the very strong distance dependence of NOEs, a small population of a short ^{1}H−^{1}H distance can lead to a stronger NOE than a larger population of a slightly longer ^{1}H−^{1}H distance. Therefore, motions of the *tert*-butyl group that transiently shorten the ^{1}H−^{1}H distance can strongly enhance the overall NOE. In this way, NOEs can be observed for proteins of increased molecular weight, which usually display broad ^{1}H NMR lines. In fact, the broad widths of the protein resonances are an advantage in our approach, as this decreases the background of NOESY cross-peaks between other protons. The high sensitivity of the NOESY cross-peaks with the *tert*-butyl resonance enabled their detection in a few hours, using about 100 μM solutions of the complex in 3 mm NMR tubes.

The attempt to use the prominence of the *tert*-butyl resonance in 1D ^{1}H NMR spectra for PCS measurements was compromised by too many spectral changes caused by the paramagnetic tags, including a much broader *tert*-butyl resonance (Figure S5). In contrast, the line-up of NOESY cross-peaks with the same *tert*-butyl signal delivered safe assignments of its PCSs. It is an advantage of the NOESY spectrum that different cross-peaks have different intensities, as this can be used to track the cross-peaks in the paramagnetic samples, as in the example of the aromatic peaks in Figure 3B. This makes the assignment of PCSs easier in 2D NOESY spectra than in 2D ^{15}N-HSQC spectra.

It is an important additional benefit that NOESY cross-peaks also deliver the PCSs of the interacting proton spins. While the PCSs of a ^{15}N−^{1}H group tend to be very similar for the ^{15}N and ^{1}H spins due to their close spatial proximity, the ^{1}H spins involved in a NOE are further apart and thus more likely to display PCSs of different magnitude. This presents highly useful information for orienting the internuclear vector relative to the $\Delta\chi$ tensor. In the case of the *tert*-butyl phenyl group of ligand **1**, PCSs were available from three different tagging sites, which unambiguously determined its orientation relative to the protein.

As is common with drug targets of moderate complexity, resonance assignments of DENpro could be obtained with reasonable effort for spectrally resolved ^{15}N-HSQC cross-peaks, whereas near-complete resonance assignments, as required for conventional NOESY cross-peak assignments, would be prohibitively laborious to establish for every complex with a new ligand. In this situation, PCSs present a much more efficient route to structural information, as $\Delta\chi$ tensors can be determined from PCSs of a small number of resolved backbone amide peaks, which can be assigned by, e.g., systematic site-directed mutagenesis of selectively isotope-labeled samples.[50,51] While the accuracy of PCSs may be insufficient to predict close intermolecular protein−ligand contacts, the localization spaces defined by PCSs from paramagnetic tags at multiple sites greatly limit the number of possible assignments for intermolecular NOEs observed with the *tert*-butyl resonance. As shown in the present work, the residue type of the protein protons involved in the intermolecular NOEs can be ascertained by chemical shifts as well as selective isotope labeling.

In many instances, it is easier to synthesize a ligand molecule with a *tert*-butyl group than with a stable isotope, in particular as *tert*-butyl groups are part of common synthetic building blocks and protection strategies. Consequently, *tert*-butyl groups are also frequently found in fragment libraries and series of inhibitor candidates during the early phases of the drug discovery process. The *tert*-butyl moiety can establish hydrophobic protein−ligand interactions and therefore often increases ligand affinity. It is also present in recently approved drugs, such as the HCV protease inhibitors boceprevir and telaprevir.[52] If a *tert*-butyl group is highly solvent exposed and flexible, it will produce less intense NOEs, but its ^{1}H NMR signal will be narrower, which facilitates identification of its localization space as a taller signal makes it easier to extract PCSs from 1D NMR spectra.

Ligands containing fluorine have found widespread application as NMR probes for protein−ligand studies.[16,53] In principle, the combined PCS and NOE strategy outlined in the present work could be extended to ^{1}H−^{19}F HOESY spectra, except that sensitivity would be compromised by the much broader line width of ^{19}F combined with typically fewer ^{19}F spins in the ligand molecules.

In conclusion, the combined PCS and NOE approach of the present work opens a practical and expedient route to elucidate the binding mode of high-affinity ligands in 1:1 complexes of protein targets. In contrast to studies with weakly binding ligands, no excess of ligand is required, which is an important advantage in the case of poorly soluble compounds. Detailed structural information about the protein−ligand complex can be obtained without costly perdeuteration, and dilute solutions of protein−ligand complexes with highly incomplete NMR resonance assignments can be investigated. While multiple samples with different paramagnetic tags are required to measure the PCSs and NOEs, each NMR experiment is quick and uses protein sparingly, making it practical to study a protein with a set of different ligands. The structural data that can be obtained with this strategy open a new and unique window of opportunity for the rational development of high-affinity lead compounds in drug development projects, when cocrystal structures of the protein−ligand complex cannot be obtained with acceptable effort.

ASSOCIATED CONTENT

Supporting Information

The Supporting Information is available free of charge on the ACS Publications website at DOI: 10.1021/jacs.6b00416.

Tables of PCSs; table of fitted $\Delta\chi$ tensor parameters; chemical structure of the C2-Ln^{3+} tag; comparison of ^{15}N-HSQC spectra of DENpro with and without ligand

DOI: 10.1021/jacs.6b00416
J. Am. Chem. Soc. 2016, 138, 4539−4546

1; ^{15}N-HSQC spectrum with resonance assignments; ^{15}N-HSQC spectra of DENpro tagged with C2−Y^{3+}, C2−Tm^{3+}, and C2−Tb^{3+} at different sites; 1D ^{1}H NMR spectra showing PCSs of the *tert*-butyl group of ligand **1** in complex with DENpro; RMSDs of ligand poses after Rosetta refinement; correlation between experimental and back-calculated PCSs for the final model structure of DENpro; NOESY spectra recorded with and without ^{13}C-decoupling; NOESY spectrum of DENpro Val55Ile; comparison between the structural model of the complex between DENpro and ligand **1** of the present work with the crystal structure 3U1I. (PDF)

AUTHOR INFORMATION

Corresponding Author

*gottfried.otting@anu.edu.au

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

Financial support by the Australian Research Council is gratefully acknowledged. C. D. K. appreciates financial support by the Deutsche Forschungsgemeinschaft (grant KL-1356/3-1). W.-N. C. thanks the government of the People's Republic of China for a CSC scholarship. C. N. thanks the Alexander von Humboldt Foundation for a Feodor Lynen fellowship.

REFERENCES

(1) (a) Barile, E.; Pellecchia, M. *Chem. Rev.* **2014**, *114*, 4749−4763. (b) Harner, M. J.; Frank, A. O.; Fesik, S. W. *J. Biomol. NMR* **2013**, *56*, 65−75. (c) Pellecchia, M.; Sem, D. S.; Wüthrich, K. *Nat. Rev. Drug Discovery* **2002**, *1*, 211−219. (d) Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Q. Rev. Biophys.* **1999**, *32*, 211−240. (e) Otting, G. *Curr. Opin. Struct. Biol.* **1993**, *3*, 760−768.

(2) (a) Shuker, S. B.; Hajduk, P. J.; Meadows, R. P.; Fesik, S. W. *Science* **1996**, *274*, 1531−1534. (b) Medek, A.; Hajduk, P. J.; Mack, J.; Fesik, S. W. *J. Am. Chem. Soc.* **2000**, *122*, 1241−1242.

(3) (a) Breeze, A. L. *Prog. Nucl. Magn. Reson. Spectrosc.* **2000**, *36*, 323−372. (b) Eichmüller, C.; Tollinger, M.; Kräutler, B.; Konrat, R. *J. Biomol. NMR* **2001**, *20*, 195−202.

(4) Pellecchia, M.; Meininger, D.; Dong, Q.; Chang, E.; Jack, R.; Sem, D. S. *J. Biomol. NMR* **2002**, *22*, 165−173.

(5) (a) Liu, W. M.; Overhand, M.; Ubbink, M. *Coord. Chem. Rev.* **2014**, *273*−*274*, 2−12. (b) Otting, G. *Annu. Rev. Biophys.* **2010**, *39*, 387−405. (c) Otting, G. *J. Biomol. NMR* **2008**, *42*, 1−9.

(6) John, M.; Pintacuda, G.; Park, A. Y.; Dixon, N. E.; Otting, G. *J. Am. Chem. Soc.* **2006**, *128*, 12910−12916.

(7) Zhuang, T.; Lee, H.-S.; Imperiali, B.; Prestegard, J. H. *Protein Sci.* **2008**, *17*, 1220−1231.

(8) Balogh, E.; Wu, D.; Zhou, G.; Gochin, M. *J. Am. Chem. Soc.* **2009**, *131*, 2821−2823.

(9) Keizers, P. H. J.; Mersinli, B.; Reinle, W.; Donauer, J.; Hiruma, Y.; Hannemann, F.; Overhand, M.; Bernhardt, R.; Ubbink, M. *Biochemistry* **2010**, *49*, 6846−6855.

(10) Saio, T.; Ogura, K.; Shimizu, K.; Yokochi, M.; Burke, T. R.; Inagaki, F. *J. Biomol. NMR* **2011**, *51*, 395−408.

(11) Guan, J.; Keizers, P. H. J.; Liu, W.; Löhr, F.; Skinner, S. P.; Heeneman, E. A.; Schwalbe, H.; Ubbink, M.; Siegal, G. *J. Am. Chem. Soc.* **2013**, *135*, 5859−5868.

(12) Saio, T.; Yokochi, M.; Kumeta, H.; Inagaki, F. *J. Biomol. NMR* **2010**, *46*, 271−280.

(13) Canales, A.; Mallagaray, A.; Berbís, M. A.; Navarro-Vázquez, A.; Domínguez, G.; Cañada, F. J.; André, S.; Gabius, H. J.; Pérez-Castells, J.; Jiménez-Barbero, J. *J. Am. Chem. Soc.* **2014**, *136*, 8011−8017.

(14) Brath, U.; Swamy, S. I.; Veiga, Z. X.; Tung, C.-C.; Van Petegem, F.; Erdélyi, M. *J. Am. Chem. Soc.* **2015**, *137*, 11391−11398.

(15) Pintacuda, G.; Park, A. Y.; Keniry, M. A.; Dixon, N. E.; Otting, G. *J. Am. Chem. Soc.* **2006**, *128*, 3696−3702.

(16) Poppe, L.; Tegley, C. M.; Li, V.; Lewis, J.; Zondlo, J.; Yang, E.; Kurzeja, R. J.; Syed, R. *J. Am. Chem. Soc.* **2009**, *131*, 16654−16655.

(17) Chen, W. N.; Loscha, K. V.; Nitsche, C.; Graham, B.; Otting, G. *FEBS Lett.* **2014**, *588*, 2206−2211.

(18) (a) Otting, G.; Wüthrich, K. *Q. Rev. Biophys.* **1990**, *23*, 39−96. (b) Zwahlen, C.; Legault, P.; Vincent, S. J. F.; Greenblatt, J.; Konrat, R.; Kay, L. E. *J. Am. Chem. Soc.* **1997**, *119*, 6711−6721. (c) Whitehead, B.; Tessari, M.; Düx, P.; Boelens, R.; Kaptein, R.; Vuister, G. W. *J. Biomol. NMR* **1997**, *9*, 313−316. (d) Sattler, M.; Schleucher, J.; Griesinger, C. *Prog. Nucl. Magn. Reson. Spectrosc.* **1999**, *34*, 93−158. (e) Andersson, P.; Otting, G. *J. Magn. Reson.* **2000**, *144*, 168−170.

(19) Hsu, V. L.; Armitage, I. M. *Biochemistry* **1992**, *31*, 12778−12784.

(20) Chen, W.-N.; Kuppan, K. V.; Lee, M. D.; Jaudzems, K.; Huber, T.; Otting, G. *J. Am. Chem. Soc.* **2015**, *137*, 4581−4586.

(21) Capeding, M. R.; Tran, N. H.; Hadinegoro, S. R.; Ismail, H. I.; Chotpitayasunondh, T.; Chua, M. N.; Luong, C. Q.; Rusmil, K.; Wirawan, D. N.; Nallusamy, R.; Pitisuttithum, P.; Thisyakorn, U.; Yoon, I. K.; van der Vliet, D.; Langevin, E.; Laot, T.; Hutagalung, Y.; Frago, C.; Boaz, M.; Wartel, T. A.; Tornieporth, N. G.; Saville, M.; Bouckenooghe, A. *Lancet* **2014**, *384*, 1358−1365.

(22) (a) Messina, J. P.; Brady, O. J.; Scott, T. W.; Zou, C.; Pigott, D. M.; Duda, K. A.; Bhatt, S.; Katzelnick, L.; Howes, R. E.; Battle, K. E.; Simmons, C. P.; Hay, S. I. *Trends Microbiol.* **2014**, *22*, 138−146. (b) Screaton, G.; Mongkolsapaya, J.; Yacoub, S.; Roberts, C. *Nat. Rev. Immunol.* **2015**, *15*, 745−759.

(23) Clum, S.; Ebner, K. E.; Padmanabhan, R. *J. Biol. Chem.* **1997**, *272*, 30715−30723.

(24) Lescar, J.; Luo, D.; Xu, T.; Sampath, A.; Lim, S.; Canard, B.; Vasudevan, S. *Antiviral Res.* **2008**, *80*, 94−101.

(25) (a) Nitsche, C.; Holloway, S.; Schirmeister, T.; Klein, C. D. *Chem. Rev.* **2014**, *114*, 11348−11381. (b) Behnam, M. A. M.; Nitsche, C.; Boldescu, V.; Klein, C. D. *J. Med. Chem.* **2016**, DOI: 10.1021/acs.jmedchem.5b01653.

(26) Erbel, P.; Schiering, N.; D'Arcy, A.; Renatus, M.; Kroemer, M.; Lim, S. P.; Yin, Z.; Keller, T. H.; Vasudevan, S. G.; Hommel, U. *Nat. Struct. Mol. Biol.* **2006**, *13*, 372−373.

(27) Noble, C. G.; Seh, C. C.; Chao, A. T.; Shi, P. Y. *J. Virol.* **2012**, *86*, 438−446.

(28) Yin, Z.; Patel, S. J.; Wang, W.-L.; Wang, G.; Chan, W.-L.; Rao, K. R. R.; Alam, J.; Jeyaraj, D. A.; Ngew, X.; Patel, V.; Beer, D.; Lim, S. P.; Vasudevan, S. G.; Keller, T. H. *Bioorg. Med. Chem. Lett.* **2006**, *16*, 36−39.

(29) Nitsche, C.; Schreier, V. N.; Behnam, M. a M.; Kumar, A.; Bartenschlager, R.; Klein, C. D. *J. Med. Chem.* **2013**, *56*, 8389−8403.

(30) Lim, S. P.; Wang, Q. Y.; Noble, C. G.; Chen, Y. L.; Dong, H.; Zou, B.; Yokokawa, F.; Nilar, S.; Smith, P.; Beer, D.; Lescar, J.; Shi, P. Y. *Antiviral Res.* **2013**, *100*, 500−519.

(31) de la Cruz, L.; Nguyen, T. H. D.; Ozawa, K.; Shin, J.; Graham, B.; Huber, T.; Otting, G. *J. Am. Chem. Soc.* **2011**, *133*, 19205−19215.

(32) Kim, Y. M.; Gayen, S.; Kang, C.; Joy, J.; Huang, Q.; Chen, A. S.; Wee, J. L.; Ang, M. J.; Lim, H. A.; Hung, A. W.; Li, R.; Noble, C. G.; Lee, L. T.; Yip, A.; Wang, Q.-Y.; Chia, C. S. B.; Hill, J.; Shi, P.-Y.; Keller, T. H. *J. Biol. Chem.* **2013**, *288*, 12891−12900.

(33) Bi, Y.; Zhu, L.; Li, H.; Wu, B.; Liu, J.; Wang. *Biomol. NMR Assignments* **2013**, *7*, 137−139.

(34) Li, H.; Zhu, L.; Hou, S.; Yang, J.; Wang, J.; Liu, J. *FEBS Lett.* **2014**, *588*, 2794−2799.

(35) de la Cruz, L.; Chen, W. N.; Graham, B.; Otting, G. *FEBS J.* **2014**, *281*, 1517−1533.

(36) Graham, B.; Loh, C. T.; Swarbrick, J. D.; Ung, P.; Shin, J.; Yagi, H.; Jia, X.; Chhabra, S.; Barlow, N.; Pintacuda, G.; Huber, T.; Otting, G. *Bioconjugate Chem.* **2011**, *22*, 2118−2125.

(37) Bertini, I.; Luchinat, C.; Parigi, G. *Prog. Nucl. Magn. Reson. Spectrosc.* **2002**, *40*, 249−273.

DOI: 10.1021/jacs.6b00416
J. Am. Chem. Soc. 2016, 138, 4539−4546

(38) Pilla, K. B.; Leman, J. K.; Otting, G.; Huber, T. *PLoS One* **2015**, *10*, e0127053.
(39) Šali, A.; Blundell, T. L. *J. Mol. Biol.* **1993**, *234*, 779−815.
(40) Webb, B.; Sali, A. In *Protein Structure Prediction*; Springer, 2014; pp 1−15.
(41) Shen, M.; Sali, A. *Protein Sci.* **2006**, *15*, 2507−2524.
(42) Trott, O.; Olson, A. J. *J. Comput. Chem.* **2010**, *31*, 455−461.
(43) Conway, P.; Tyka, M. D.; DiMaio, F.; Konerding, D. E.; Baker, D. *Protein Sci.* **2014**, *23*, 47−55.
(44) Venkatraman, S.; Wu, W.; Prongay, A.; Girijavallabhan, V.; Njoroge, F. G. *Bioorg. Med. Chem. Lett.* **2009**, *19*, 180−183.
(45) www.rosettacommons.org/demos/latest/public/relax_around_chemically_bound_ligand/README, accessed 1 Nov. 2014.
(46) Schmitz, C.; Vernon, R.; Otting, G.; Baker, D.; Huber, T. *J. Mol. Biol.* **2012**, *416*, 668−677.
(47) Bleaney, B. *J. Magn. Reson.* **1972**, *8*, 91−100.
(48) Vega, A. J.; Fiat, D. *Mol. Phys.* **1976**, *31*, 347−355.
(49) Pintacuda, G.; Kaikkonen, A.; Otting, G. *J. Magn. Reson.* **2004**, *171*, 233−243.
(50) Wu, P. S. C.; Ozawa, K.; Lim, S. P.; Vasudevan, S. G.; Dixon, N. E.; Otting, G. *Angew. Chem., Int. Ed.* **2007**, *46*, 3356−3358.
(51) Abdelkader, E. H.; Yao, X.; Feintuch, A.; Adams, L. A.; Aurelio, L.; Graham, B.; Goldfarb, D.; Otting, G. *J. Biomol. NMR* **2016**, *64*, 39−51.
(52) De Luca, A.; Bianco, C.; Rossetti, B. *Curr. Opin. Pharmacol.* **2014**, *18*, 9−17.
(53) Vulpetti, A.; Dalvit, C. *Drug Discovery Today* **2012**, *17*, 890−897.

DOI: 10.1021/jacs.6b00416
J. Am. Chem. Soc. 2016, 138, 4539−4546

Supporting information

Sensitive NMR Approach for Determining the Binding Mode of Tightly Binding Ligand Molecules to Protein Targets

Wan-Na Chen, Christoph Nitsche, Kala Bharath Pilla, Bim Graham, Thomas Huber, Christian D. Klein, Gottfried Otting*

Table S1. PCSs measured of backbone amide protons of DENpro in complex with ligand **1**.[a]

Residue	Site A		Site B		Site C	
	Tm^{3+}	Tb^{3+}	Tm^{3+}	Tb^{3+}	Tm^{3+}	Tb^{3+}
NS2B						
Ser70*	-0.35	0.46	-0.06	0.08	-0.01	0
Ser71*	-0.26	0.33	-0.05	0.06	-0.02	0.02
Ile73*	-	-	-0.05	0.07	-0.08	0.09
Leu74*	-0.12	0.18	-0.06	0.09	-0.05	0.08
Ser75*	-0.09	0.11	-0.06	0.08	-0.06	0.09
Ile78*	-0.07	-	-0.11	0.17	-0.09	0.20
Ser79*	-0.01	0.02	-0.21	0.25	-0.23	0.34
Asp81*	-	-	-	0.33	-0.34	-0.54
Gly82*	-	-	-0.26	0.33	-0.37	0.58
Ser83*	-	-	-0.34	0.42	-0.31	0.48
Ser85*	-0.03	0.05	-0.16	0.20	-0.14	0.21
Ile86*	-0.05	-	-0.11	0.12	-0.09	0.14
Lys87*	-0.05	0.08	-0.07	0.08	-0.08	0.11
NS3pro						
Gly21	0.58	-	0.16	-0.19	-	-
Ala22	-	-	0.18	-0.22	-	-
Tyr23	-	-	0.32	-0.38	-	-
Lys28	-	-	-	-	0.31	-0.53
Ser34	-	-	-	-	0.21	-

Gly39	0.83	-0.99	0.37	-0.44	0.49	-0.76
Tyr41	0.59	-0.72	0.20	-0.23	0.56	-0.84
Lys42	0.20	-0.25	0.15	-0.17	-	-
Glu43	0.04	-0.06	0.10	-0.12	-	-
Gly44	0.14	-0.19	0.12	-0.16	-	-
Phe46	0.26	-0.30	0.20	-0.23	-	-
Trp50	0.10	-0.12	-	-	-	-
His51	0.12	-0.14	-	-	-0.32	0.05
Val52	0.19	-0.23	-	-	-0.03	-
Thr53	0.24	-0.29	-	-	0.14	-0.26
Gly55	0.19	-0.23	-	-	-	-
Gly62	0.47	-0.58	0.38	-0.44	-	-1.26
Glu66	0.23	-	0.24	-	-	-
Ala70	0.05	-0.05	-0.05	0.05	-	-
Ser78	-	-	0.06	-0.07	-	-
Gly81	0.14	-0.15	0.12	-0.16	-	-
Gly82	0.05	-0.06	0.07	-0.08	-	-
Leu85	-	0.10	0.05	-0.04	-	-
Trp89	-	-	-0.03	0.03	-0.06	0.08
Gly92	-0.28	0.37	-	-	-0.03	0.04
Gln96	-	-	-	-	0.04	-0.07
Leu98	-1.15	-	-	-	0.10	-0.17
Leu100	-	-	-	-	0.17	-0.26

Gly103	-	-	-	-	0.16	-0.26
Gln110	-	-	-0.03	0.05	0.03	-0.06
Thr111	-	1.18	-0.04	0.06	-	-
Gly114	-0.28	-	-0.08	0.08	-	-
Phe116	-0.15	0.21	-0.09	0.10	-0.08	0.11
Lys117	-	-	-0.08	0.09	-0.07	0.11
Thr120	-0.06	0.08	-0.07	0.09	-0.12	0.19
Gly121	-0.07	0.10	-0.06	0.08	-0.11	0.15
Thr122	-	-	-0.05	0.07	-0.10	0.14
Ile123	-	-	-	-	-0.12	0.11
Ser127	-0.36	-	-0.09	0.11	-	-
Val140	-	-	-	-	0.06	-0.10
Gly148	-0.38	0.49	-	-	0.02	-0.06
Ala160	-0.11	0.15	-0.04	0.05	-0.03	0.05
Ala164	-	-	-0.22	0.28	-0.08	0.12
Ile165	-	-	-0.12	0.14	-	-
Ala166	-	-	-0.08	-	-0.08	0.13

[a] PCSs were generated by C2-Tm^{3+} and C2-Tb^{3+} tags at sites A-C. Corresponding samples with C2-Y^{3+} tags were used as the diamagnetic reference. PCSs are reported in ppm as the ^{1}H chemical shift measured in the presence of paramagnetic tag minus the ^{1}H chemical shift measured of the diamagnetic reference.

Table S2. PCSs observed in NOESY spectra of the complex between ligand **1** and DENpro with C2-Ln^{3+} tags[a]

Tag site and metal	Ligand **1**			Protein		
	tert-butyl group (ppm)	H-1/1' (ppm)	H-2/2' (ppm)	CH_3-1 (ppm)	CH_3-2 (ppm)	H-3 (ppm)
A-Tm^{3+}	-0.01	0.00	0.00	-0.05	-0.06	-0.05
A-Tb^{3+}	0.02	0.01	0.01	0.07	0.09	0.08
B-Tm^{3+}	-0.09	-0.27	-0.56	-0.14	n.d.	n.d.
B-Tb^{3+}	0.10	0.33	0.67	0.15	n.d.	n.d.
C-Tm^{3+}	-0.10	-0.10	-0.10	-0.06	-0.05	-0.06
C-Tb^{3+}	0.16	0.16	0.16	0.09	0.07	0.10
Chemical shift[a]	1.30	7.54	7.66	0.67	0.78	1.73

[a] Measured for the complex of ligand **1** with wild-type DENpro without tag. See Figures 1A and 2B for the numbering of the proton resonances. n.d., not detected.

Table S3. $\Delta\chi$ tensor parameters fitted to the homology model built of DENpro and refined by Rosetta[a]

Tagging site and metal	$\Delta\chi_{ax}$	$\Delta\chi_{rh}$	x	y	z	α	β	γ
A-Tm^{3+}	17.1	6.7	16.415	-18.172	36.099	140.1	139.5	38.5
A-Tb^{3+}	-21.1	-8.3	16.415	-18.172	36.099	141.4	139.8	42.7
B-Tm^{3+}	8.7	2.0	8.634	-21.149	7.976	102.9	54.8	43.0
B-Tb^{3+}	-10.5	-2.7	8.634	-21.149	7.976	102.5	56.8	41.8
C-Tm^{3+}	17.7	3.5	13.935	4.362	4.462	49.3	147.9	82.2
C-Tb^{3+}	-27.4	-6.9	13.935	4.362	4.462	49.3	147.9	82.2

[a] The tensor parameters are reported with respect to the final Rosetta-refined model, which is available from http://rsc.anu.edu.au/~go/coordinates/. The axial and rhombic components of the $\Delta\chi$ tensors are given in 10^{-32} m^3 and the Euler angles in degrees, using the zyz convention and unique tensor representation.[1] $\Delta\chi$ tensor fits were performed by restricting Tb^{3+} and Tm^{3+} to the same site.

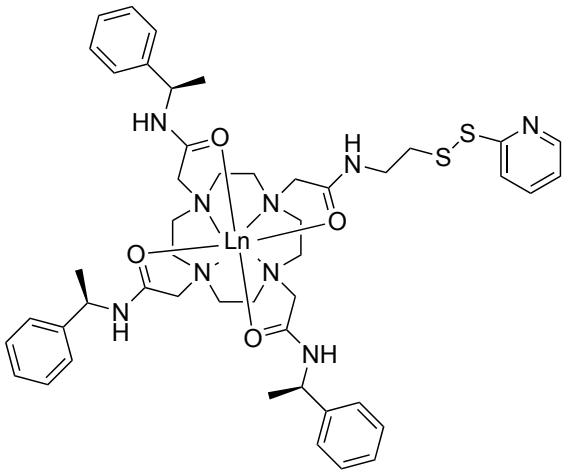


Figure S1. Chemical structure of the C2-Ln^{3+} tag.[2]

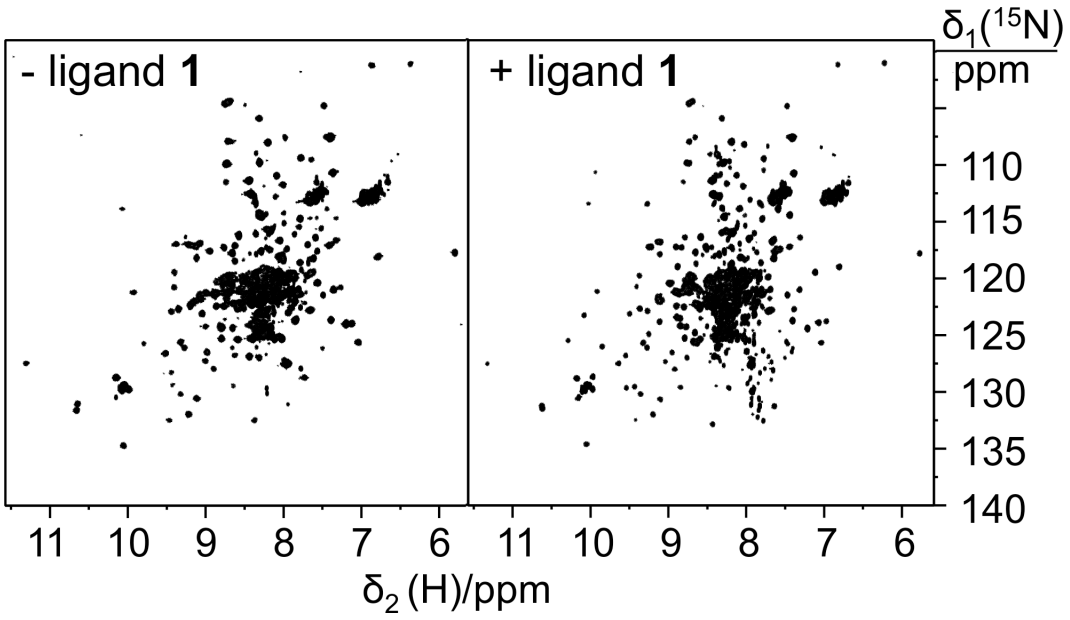


Figure S2. ^{15}N-HSQC spectra of DENpro without (left spectrum) and with ligand **1** (right spectrum). The spectra were recorded of 150 μM solutions of uniformly ^{15}N-labeled DENpro tagged at site B with a diamagnetic C2-Y^{3+} tag. The comparison shows that many of the well-resolved peaks can be tracked despite some changes in chemical shifts.

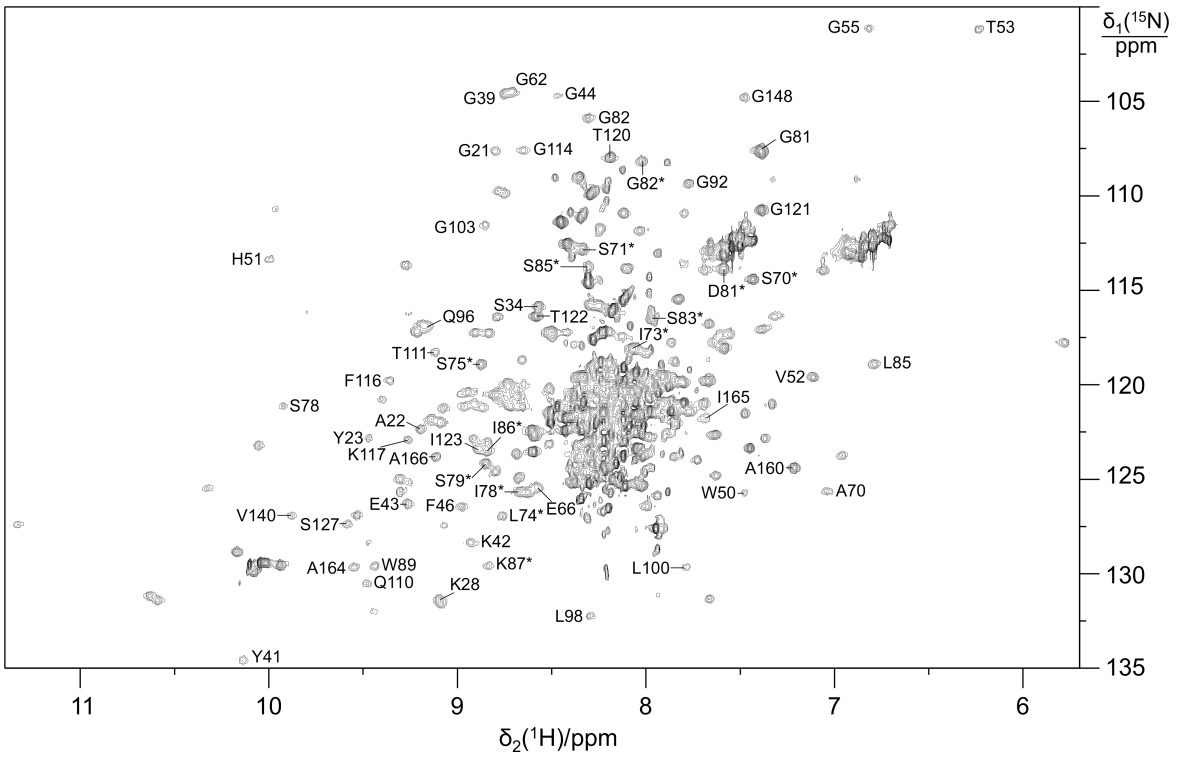


Figure S3. ^{15}N-HSQC spectrum recorded of a 150 μM solution of uniformly ^{15}N-labeled DENpro tagged at site A with a diamagnetic C2-Y^{3+} tag and in complex with ligand **1**. Solution conditions: 20 mM MES, pH 6.5, 50 mM NaCl, 25 ^{o}C. The spectrum was recorded using a 3 mm NMR tube on a Bruker 800 MHz NMR spectrometer equipped with a TCI cryoprobe. The cross-peaks used as diamagnetic reference for PCS measurements (Table S1) are marked with their sequence-specific residue assignments.

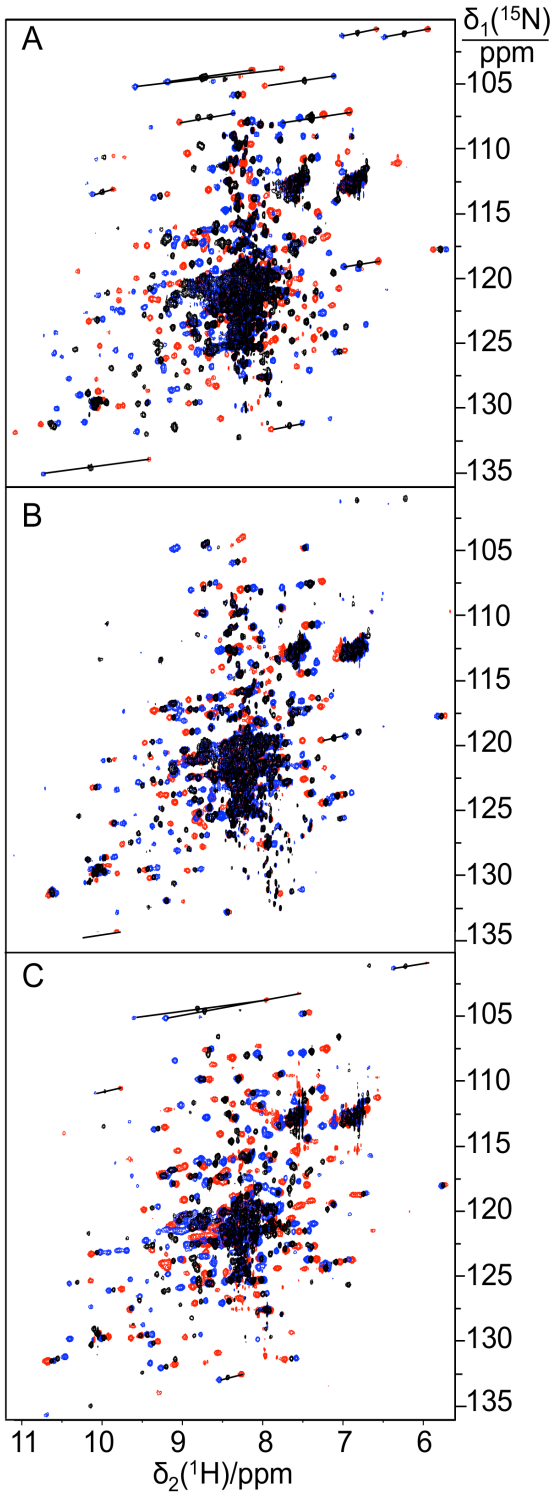


Figure S4. PCS measurements of backbone amide protons. The figure shows superimpositions of ^{15}N-HSQC spectra recorded of uniformly ^{15}N-labeled DENpro in complex with ligand **1** with

C2-Ln^{3+} tags at different sites of DENpro. Panels A-C show the spectra with tags at sites A, B, and C, respectively. Spectra of DENpro with diamagnetic tag (C2-Y^{3+}) are shown in black, and the corresponding spectra with paramagnetic tags are plotted in red (C2-Tm^{3+}) and blue (C2-Tb^{3+}). Selected cross-peaks are linked by lines indicating the pseudocontact shifts. The spectra were recorded of 150 μM solutions of DENpro-ligand **1** complexes in NMR buffer (pH 6.5) at 25 ^{o}C, using a Bruker 800 MHz NMR spectrometer. Based on the cross-peaks observed in these ^{15}N-HSQC spectra, the tag ligation yields at sites A-C were about 80%, 70%, and 90%, respectively.

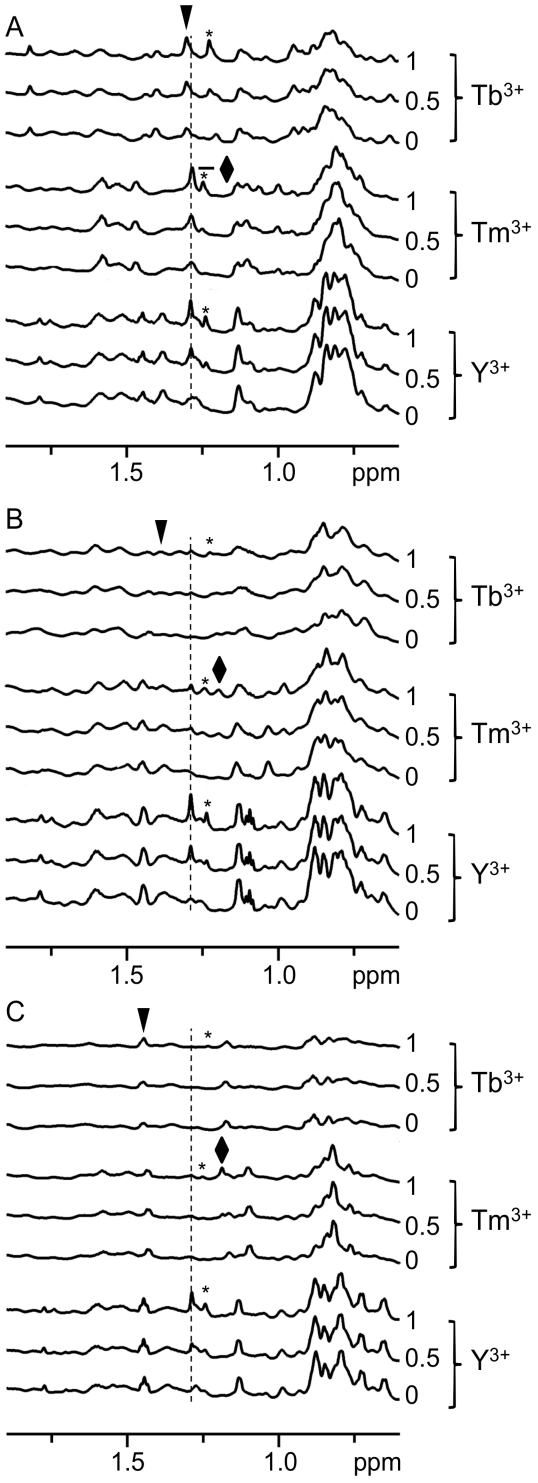


Figure S5. PCSs of the *tert*-butyl signal observed in 1D ^{1}H NMR spectra measured of ligand **1** in complex with DENpro with different C2-Ln^{3+} tags loaded with different metal ions as indicated.

The panels A-C show the data obtained with C2-Ln^{3+} tags at sites A, B, and C, respectively. The spectra were recorded of ca. 150 μM solutions of uniformly ^{15}N labeled DENpro samples in NMR buffer containing 10% D_2O at 25 ^{o}C. Each panel contains three groups of three spectra, where each group shows the result of titration of the tagged protein with increasing amounts of ligand **1** from bottom to top in stoichiometric ratios as indicated. The dashed line identifies the chemical shift of the *tert*-butyl signal of the bound ligand in the diamagnetic samples. This chemical shift is independent of the site of the diamagnetic tag and the same as in wild-type protein. Signals from a non-binding impurity in the ligand stock are labeled with a star. Filled rhombuses and triangles mark the chemical shifts of the *tert*-butyl group in samples tagged with C2-Tm^{3+} and C2-Tb^{3+}, respectively. Note the increase of the *tert*-butyl resonance of bound ligand **1** with increasing ligand concentrations, but also that the *tert*-butyl resonance is difficult to assign in panels A and B due to spectral changes induced in the protein NMR spectrum upon ligand binding as well as PREs associated with paramagnetic metal ions.

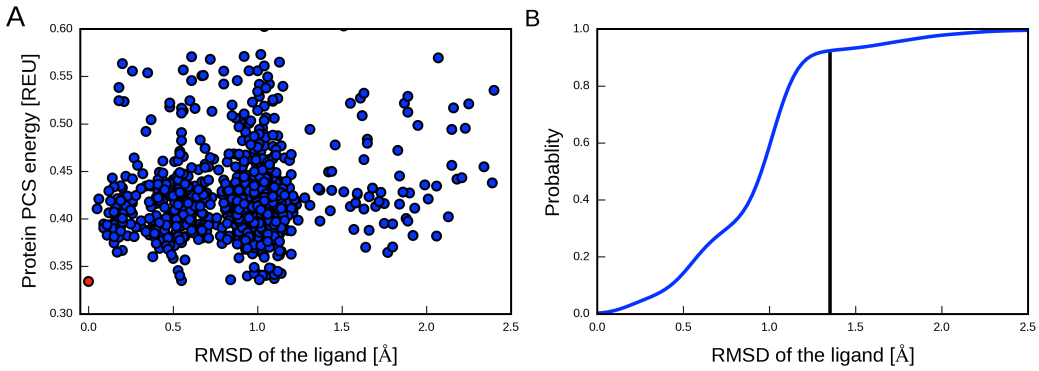


Figure S6. PCS energies and RMSDs of the ligand poses after all-atom Rosetta refinement. (A) PCS energy of the protein (in Rosetta energy units) vs RMSD of the ligand to the final pose (highlighted in red). 1000 structures were produced by the Rosetta refinement protocol, which were all very similar and showed very similar PCS energies. The final ligand pose and protein structure was chosen as the protein-ligand complex with the lowest PCS energy of the protein. The RMSD was calculated for heavy atoms of the ligand only. All ligand poses were within 2.5 Å from the best pose. (B) Probability density plot of the ligand RMSDs with respect to the final docked pose of the ligand. The vertical bar shows that about 92% of the ligand poses are within 1.3 Å from the best pose.

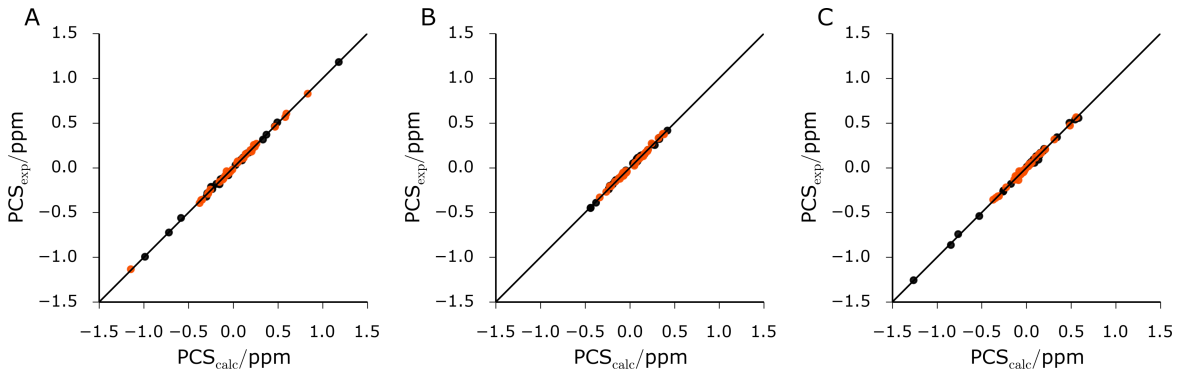


Figure S7. Correlation between experimental and back-calculated PCSs fitted on the model structure. The data from C2-Tm^{3+} and C2-Tb^{3+} tags are shown as red and black points, respectively. (A)-(C) are the correlation plots drawn for the three different mutants A, B, and C, respectively. The $\Delta\chi$ tensor parameters are reported in Table S3.

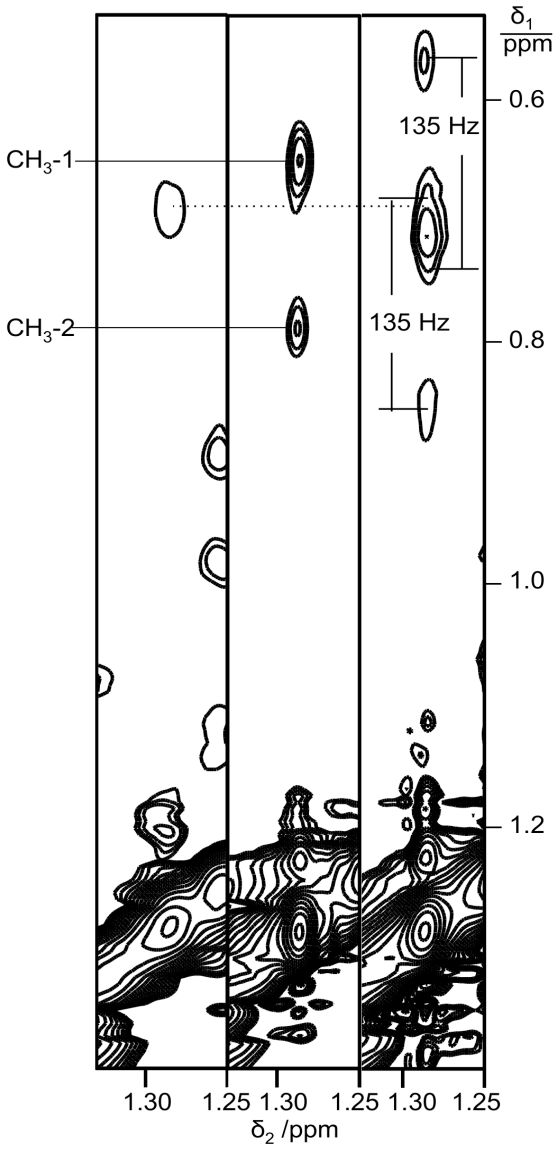


Figure S8. Selected spectral region of NOESY spectra of DENpro in complex with ligand **1**, showing cross-peaks between the methyl groups of Val and the *tert*-butyl resonance of the ligand. The left-most panel shows the spectrum recorded with ligand **2**, which is devoid of a *tert*-butyl group. The center and right panels show the spectra of complexes of ligand **1** with unlabeled DENpro and DENpro prepared with $^{13}C/^{15}N$-labeled valine, respectively. The cross-peaks are clearly split by $^{1}J_{CH}$ couplings in the spectrum with isotope-labeled valine. The chemical shifts of the cross-peaks in the indirect dimensions identify the cross-peaks as NOEs with methyl groups of valine.

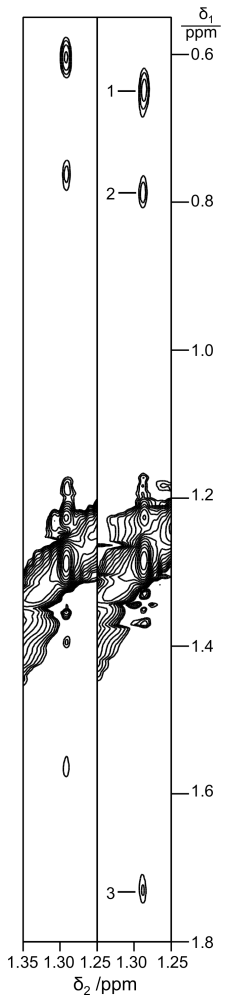


Figure S9. Assignment of NOEs between methyl groups of protein side chains and the *tert*-butyl group of ligand **1**. Right panel: zoom into the spectral region from the NOESY spectrum shown in Figure 2B of the main text. Cross-peaks with the *tert*-butyl resonance of ligand **1** are numbered as in Figure 2B. Left panel: same as the right panel, except for the complex of ligand **1** with the mutant Val155Ile of DENpro rather than wild-type DENpro. Note that the cross-peaks of the Val155Ile mutant display no splitting by $^{1}J_{HC}$ couplings, although the protein had been prepared with ^{13}C-labeled valine and the spectrum was recorded without ^{13}C-decoupling. This result indicates that the cross-peaks 1–3 in wild-type DENpro are with Val155 rather than Val154, and that the corresponding cross-peaks observed in the Val155Ile mutant are with isoleucine. The localization space of the *tert*-butyl group shown in Figure 4 of the main text is not near any other methyl group of DENpro. Also note that all NOESY cross-peaks with the *tert*-butyl group are slightly shifted because of a small change in its chemical shift, providing an independent means of identification beyond their narrow line width in the detection dimension.

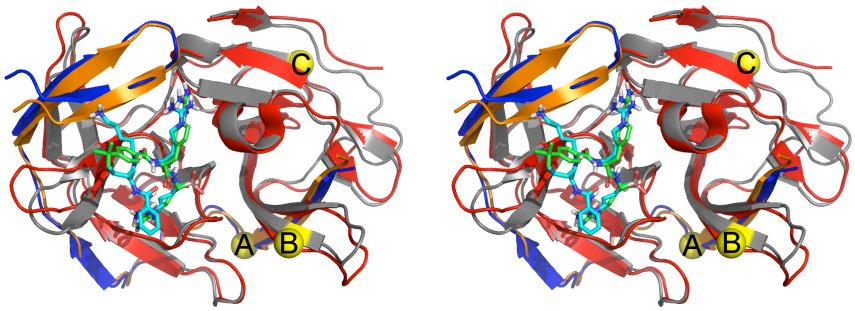


Figure S10. Stereo view comparing the crystal structure of serotype 3 with peptide inhibitor (PDB code 3U1I),[3] and the Rosetta-refined model structure of DENpro in complex with ligand **1** determined in the present work. The Rosetta all-atom refinement adjusted the conformation of the backbone atoms around the ligand binding site both for NS2B and NS3pro. NS2B and NS3pro of the crystal structure are shown in blue and red, respectively, with the peptide ligand shown in a stick representation with a cyan backbone. NS2B and NS3pro of the model are shown in orange and grey, with ligand **1** in a stick representation with green backbone atoms. Nitrogen and oxygen atoms are shown in dark blue and red, respectively. Yellow balls mark the sites A-C mutated to cysteine for attachment of C2-Ln^{3+} tags. The backbone of ligand **1** closely aligns with the backbone of the peptide ligand in the crystal structure. The Rosetta all-atom refinement resulted in a small shift of the C-terminal two-stranded β-sheet of NS3pro (residues 154-164) and of the C-terminal β-strand of NS2B (residues 73*-83*) towards the ligand.

References

(1) Schmitz, C.; Stanton-Cook, M. J.; Su, X.-C.; Otting, G.; Huber, T. *J. Biomol. NMR* **2008**, *41*, 179–189.

(2) Graham, B.; Loh, C. T.; Swarbrick, J. D.; Ung, P.; Shin, J.; Yagi, H.; Jia, X.; Chhabra, S.; Barlow, N.; Pintacuda, G.; Huber, T.; Otting, G. *Bioconjugate Chem.* **2011**, *22*, 2118–2125.

(3) Noble, C. G.; Seh, C. C.; Chao, A. T.; Shi, P. Y. *J. Virol.* **2012**, *86*, 438–446.