

# Global Dynamics as Communication Sensors in Peptide Synthetase Cyclization Domains.

Subrata H. Mishra<sup>1†</sup>, Aswani K. Kancherla<sup>1#</sup>, Kenneth A. Marincin<sup>1#</sup>, Guillaume Bouvignies<sup>2</sup>, Santrupti Nerli<sup>3</sup>, Nikolaos Sgourakis<sup>4</sup>, Daniel P. Dowling<sup>5</sup>, Dominique P. Frueh.<sup>1\*</sup>

<sup>1</sup>Department of Biophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, 725 N Wolfe St Room 701 Hunterian, Baltimore, MD 21205, USA.

<sup>2</sup>Department of Chemistry, ENS - Laboratoire des BioMolécules, 24 rue Lhomond, 75005 Paris, France.

<sup>3</sup>Department of Biomolecular Engineering, University of California, Santa Cruz, Santa Cruz, CA, 95064, USA.

<sup>4</sup>Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, 19104, USA.

<sup>5</sup>Department of Chemistry, University of Massachusetts Boston, Boston, Massachusetts, 02125, USA

# These authors contributed equally to this work

\* **corresponding author: [dfrueh1@jhmi.edu](mailto:dfrueh1@jhmi.edu)**

**THIS WORK HAS NOT BEEN PEER-REVIEWED.**

† Present address: Reference Standard Laboratory, United States Pharmacopeial Convention, 12601 Twinbrook Pkwy, Rockville, MD, 20852, USA

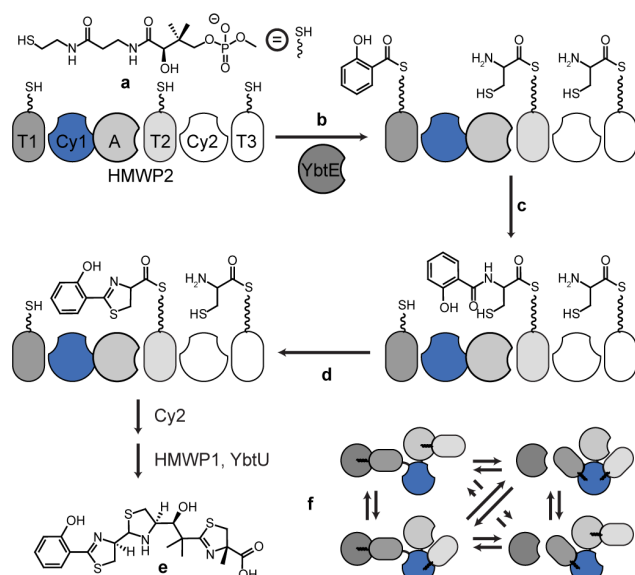
## ABSTRACT

Structural biology is the foundation for deriving molecular mechanisms, where snapshots of macromolecules and binding partners inform on mutations that test or modify function. However, frequently, the impact of mutations violates the underpinnings of structural models, and mechanisms become cryptic. This conundrum applies to multidomain enzymatic systems called nonribosomal peptide synthetases (NRPSs), which assemble simple substrates into complex metabolites often with pharmaceutical properties. Engineering NRPSs can generate new pharmaceuticals<sup>1-3</sup> but a dynamic domain organization challenges rational design.<sup>4-8</sup> Using nuclear magnetic resonance (NMR), we determined the solution structure of a 52 kDa cyclization domain and demonstrate that global intra-domain dynamics enable sensing of substrates tethered to partner domains and draw an allosteric response encompassing the enzyme's buried active site and two binding sites 40 Å apart. We show that a point-site mutation that impedes the domain dynamics globally hampers the allosteric response. We demonstrate this mechanism through NMR experiments that provide atomic-level read-outs of allosteric responses during biochemical transformations *in situ*. Our results establish global structural dynamics as sensors of molecular events that can remodel domain interactions and illustrate the need for integrating structural dynamics explicitly when deriving molecular mechanisms through mutagenesis and structural biology.

## INTRODUCTION

Structural dynamics are increasingly recognized as a component of biological activity yet determining their presence and their function remains challenging. Structural dynamics play a role in enzymatic activity,<sup>9-12</sup> allosteric communication,<sup>13-15</sup> and molecular binding.<sup>16</sup> However, detecting dynamics and determining their function is experimentally demanding, often because

establishing correlations between complementary experiments is hindered by experimental precision. Notably, although NMR can provide valuable information, the method is challenged by molecular size. Yet, when dynamics go undetected, kinetic models are incomplete, allostery may go unnoticed, and structure-based mechanisms often need revisions. The latter is true for large enzymatic systems called nonribosomal peptide synthetases, and we sought to determine whether structural dynamics may resolve gaps in understanding these fascinating enzymes.



**Figure 1. HMWP2 and NRPS synthesis.** **a**, holo T-domains harbor phosphopantetheines. **b**, Adenylation domains (*in cis*, A, and *in trans*, YbtE) load substrates onto T-domains. **c**, C- (not shown) and Cy-domains catalyze peptide bond formation between substrates of T domains. **d**, Cy-domains additionally catalyze cyclodehydration. **e**, Cy2, HMWP1, and YbtU finalize the synthesis of yersiniabactin. **f**, Selected domain interactions during NRPS synthesis involving T, A and Cy/C domains.

NRPSs are microbial molecular factories that employ a dynamic multi-domain architecture to assemble simple substrates into complex natural products, including pharmaceuticals such as antibiotics (bacitracin), anticancer agents (epothilones), or immunosuppressants (cyclosporins).<sup>17</sup> Adenylation (A) domains attach substrates to 20 Å phosphopantetheine moieties (Fig. 1a) of holo thiolation (T) domains through thioester bonds (Fig. 1b). Condensation (C) domains then catalyse the peptide bond formation between substrates of upstream and downstream T-domains,

leading to a downstream acceptor harbouring an extended intermediate and an upstream donor restored to its holo form (Fig. 1c). As part of the C-domain family, cyclization (Cy) domains further catalyse cyclodehydration (Fig. 1d). Iteration leads to chain elongation before product release, e.g. through thioesterases. Our model system, HMWP2, combines salicylate with two cysteines to synthesize a precursor of yersiniabactin (Fig. 1e), a virulence factor of *Yersinia*

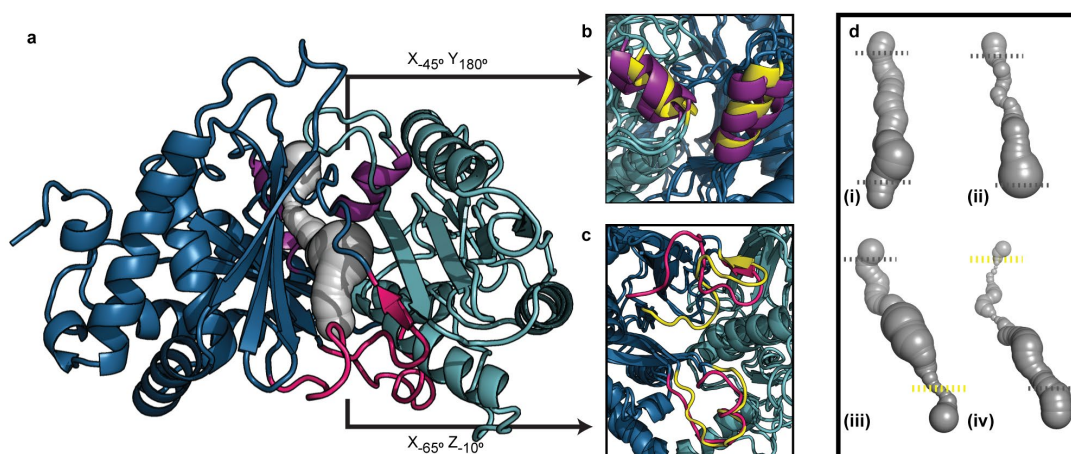
*pestis*, the causative agent of the bubonic plague. Such modular organization has stimulated efforts to engineer NRPSs and control substrate incorporation to produce novel pharmaceuticals<sup>1-3</sup>, but a dynamic domain architecture<sup>4-8</sup> hampers rational design. Notably, T-domains visit catalytic domains through sequential, transient interactions (Fig. 1f), and it is unclear whether these interactions are random or if chemistry promotes interactions in line with synthesis. Further, intra-domain structural dynamics are often necessary for function,<sup>18,19</sup> although dynamics within condensation domains have proven puzzling. Schmeing and Co. noted that open and closed conformations observed in C-domains may be needed at different stages of synthesis,<sup>20</sup> but the conformations of C-domains in crystal structures are unaltered in the presence of other domains or substrates (Supporting Information (SI) Fig. 1). The mechanisms of peptide bond formation and heterocyclization are under frequent revisions and mutagenesis impacts function in sometimes unexpected manner.<sup>21</sup> Thus, the C-domain family appears to bear the hallmarks of structural dynamics, and we strived to determine them experimentally and establish their function using the domain Cy1 of HMWP2.

## RESULTS AND DISCUSSION

### Structural fluctuations in Cy1

The solution structure of Cy1 displays the conserved C-domain family fold<sup>21</sup>, but its NMR ensemble reflects malleability (SI Table 1, Fig. 2, SI Fig. 2). N- and C-terminal residues display distinct connected regions, with C-terminal residues (Fig. 2a, dark blue) spatially crossing over twice into the N-terminal region (Fig. 2a, c). Both crossovers include loops (L16 and L20, SI Fig. 3) that interact with the donor T-domain. L20 leads to an extended region that occasionally pairs with  $\beta$ -strand  $\beta$ 2 of the N-terminal region in the NMR ensemble. Like all Cy- and C-domains, Cy1 displays a tunnel between both regions, which harbours the active sites for peptide bond formation<sup>21</sup> and cyclodehydration.<sup>22,23</sup> In multi-domain NRPS structures, donor and

acceptor T-domains funnel their phosphopantetheine arms from opposite ends of this 40 Å tunnel (Fig. 2a, pink, purple) to reach the buried active site.<sup>8,24,25</sup> NMR provides ensemble of



**Figure 2. Solution structure of Cy1 and transient tunnel accessibility.** **a**, medoid NMR conformer displayed with N- and C-subdomains in light and dark blue respectively. T domain sites in pink (donor) and purple (acceptor). **b**, **c**, open (purple, pink) and closed (yellow) conformations at acceptor (**b**) and donor (**c**) sites. **d**, tunnels in conformers are found fully open (i), constricted at the center (ii), closed at donor (iii), and closed at acceptor (iv) sites. Dashes approximate tunnel entrances (grey: open; yellow: constricted).

conformers (SI Fig. 2) that may reflect dynamics. Most notably, Cy1's tunnel is constricted in several conformers, suggestive of transient interactions within the tunnel (Fig. 2d, SI Fig. 2c). Further, the tunnel entrances at the acceptor (Fig. 2b) and donor (Fig. 2c) sites display structural heterogeneity that occasionally prevent substrate access. This malleability recalls that seen in a distant member of the C-domain family, in which accelerated molecular dynamics revealed the closing of an otherwise open channel replacing the tunnel of this domains.<sup>26</sup> Our results may reflect structural fluctuations, in which access to the tunnel is intermittent, a gating mechanism that can be harnessed for substrate recognition and that we investigated further.

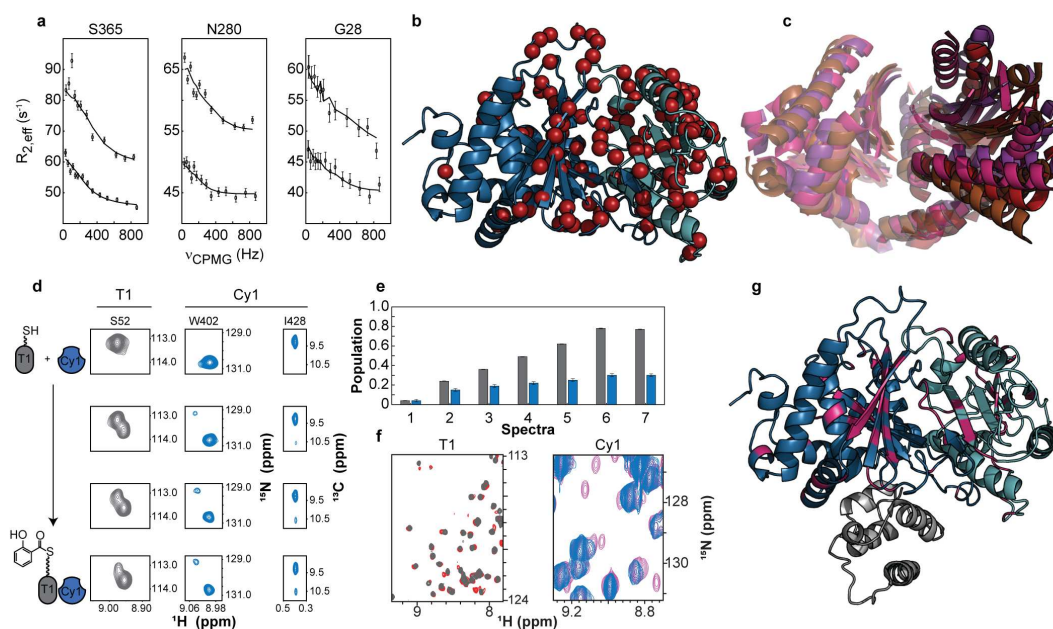
Cy1 is subject to global structural fluctuations at  $\mu$ s-ms time-scales, which mirror conformational changes in the C-domain family. We used NMR relaxation dispersion<sup>27</sup> to identify residues that sample different chemical environments (Fig. 3a-b). We found that both T-

domain binding sites are dynamic (Fig. 3a), consistent with structural fluctuations seen in our NMR ensemble and in crystallographic studies (SI Fig. 4) thus strengthening a gating mechanism where structural fluctuations modulate entry of the tethered substrates to access the active site. However, we also observed wide-spread dynamics involving 131 residues, with the N-terminal lobe displaying a dynamic footprint reminiscent of conformational changes reported by crystallography (Fig. 3b, c).<sup>20</sup> Hence, we used multi-field relaxation dispersion to determine whether motions in Cy1 reflected global changes seen in C domains. The majority of dynamic residues (126/131) fit a global two-conformer model, with a minor conformer populated at 3% and an exchange rate of 1,522 s<sup>-1</sup>. Chemical shift analysis demonstrates that we are not probing unfolding of Cy1 (SI Fig. 4e, f). The function of this malleability appears to include a gating mechanism for substrate recognition, but the global nature of Cy1 dynamics also points at allostery, with a network of dynamic residues linking the remote T-domain binding sites. Thus, Cy1 dynamics may serve as a conduit to propagate the impact of T-domain binding from one site to the other remote site. We proceeded to demonstrate this hypothesis experimentally.

### **Cy1 responds allosterically to its partner substrate loading**

We found that Cy1 responds to its donor thiolation domain only when it holds a substrate. Structures in the C-domain family appear unaffected by partner T-domains or substrates (SI Fig. 1), and substrates rapidly fall off of T domains, hampering traditional structure-function approaches to investigate allostery. We designed an experiment that builds upon our previous method<sup>28</sup> to overcome these obstacles. We first presented Cy1 with its donor T-domain (T1) in holo form and observed no spectroscopic perturbation (SI Fig. 6a), suggestive of a dissociation constant larger than several mM. We then loaded T1 with its salicylate substrate *in situ* and used a combination of isotope labelling and tailored pulse sequences<sup>29,30</sup> to monitor the spectra of each protein simultaneously. New Cy1 signals appeared as T1 was loaded with salicylate, denoting strong interaction between Cy1 and loaded T1 (Fig. 3d-f, SI Fig. 5). These observations

corroborate recent studies where substrate-loaded donor T domains were needed to detect binding with a C-domain.<sup>31</sup> Our results are remarkable as one would expect substrate recognition to take place at the active site near the centre of the tunnel, but our data indicates that the



**Figure 3. Dynamics and allosteric response in Cy1.** **a**, Representative relaxation dispersion profiles of residues at the donor site (S365), within the tunnel (N280), and at the acceptor site (G28). **b**, Cy1 residues (red spheres) exhibiting dynamics on the ms- $\mu$ s timescale. **c**, Crystal structures of C and Cy domains aligned by the C-terminal subdomain. (PDB: 5T3D, 4JN3, 6P1J, and 1L5A). For clarity, loops have been hidden. **d**, Time-shared NMR data showing concomitant loading of T1 (S52, grey) and emergence of a minor Cy1 population (W402 amide, I428 methyl, blue). **e**, Population of loading for T1 (grey) and of Cy1 minor conformation (blue). **f**, IDIS-NMR of T1 (holo: grey, loaded: red) and Cy1 spectra (free: blue, with T1 loaded: pink) allow for complete analysis of each protein without spectral crowding. **g**, Cy1 residues (pink) exhibiting a response to substrate-loaded T1.

phosphopantetheine arm of holo-T1 does not go into the tunnel for substrate probing. Instead, our results point at substrate recognition at the surface of Cy1 paired with a dynamic gating mechanism, where dynamics respond to the presence of substrate to promote engagement.

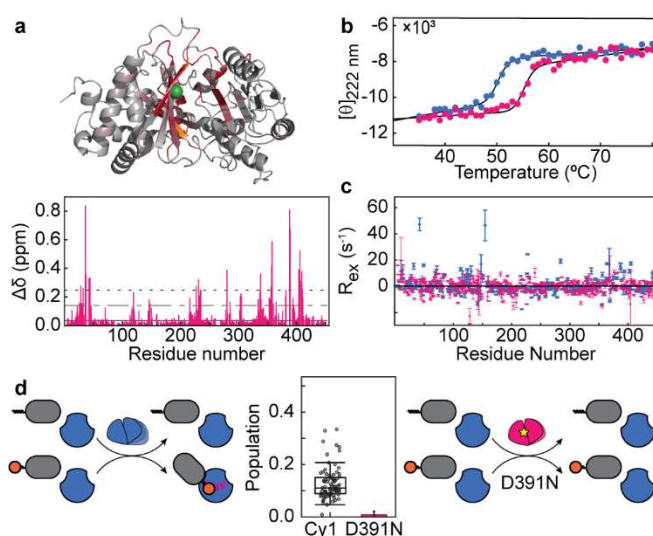
Substrate recognition at the donor site induces allosteric remodelling of Cy1's active site, tunnel, and remote acceptor binding site. New signals observed in Cy1 identify residues with altered environments when Cy1 binds to T1. Mapping affected residues (Fig. 3f, Cy1) on our



structure (Fig. 3g) reveals not only changes in the donor site but depicts a conformer with an average population of approximately 11% (SI Fig. 5c) with changes encompassing the active site and the remote acceptor site. This conformer appears not to be selected from conformations of free Cy1 (SI Fig. 6). These observations demonstrate that, upon substrate recognition at the donor site, Cy1 responds allosterically to propagate changes to all sites involved in synthesis. Such a global reaction is reminiscent of Cy1's global dynamics, and we next used mutagenesis to demonstrate that dynamics mediate this allosteric response.

### Impeding dynamics hinders recognition

Mutating a residue at the centre of the tunnel impairs both Cy1 dynamics and its response



**Figure 4. A Cy1 point-site mutation impedes dynamics and allosteric response.** (a) Chemical shift perturbations ( $\Delta\delta$ ) between Cy1 and D391N mapped onto the Cy1 structure (top panel) using a color gradient where bright red indicates a CSP two standard deviations above the median (dotted line in the bottom panel). Residues that were uniquely assigned in D391N but not in wild type are in orange and D391 is shown as a green sphere (top panel). (b) Thermal melt of Cy1 (blue) and Cy1 D391N (pink) monitored by circular dichroism at 222 nm. (c) Dynamics within Cy1 (blue) and D391N (pink) as determined by Hartman-Hahn  $R_{ex}$ . (d) Population of minor conformers in Cy1 and D391N. A dynamic Cy1 engages with a loaded partner with an allosteric response (left cartoon) whereas a rigidified Cy1 is severely impeded (right cartoon).

towards its substrate loaded partner. To demonstrate that dynamics facilitate Cy1's allosteric response, we mutated a residue distant from the donor site, such that the mutation could not impact binding through direct interactions. The conserved aspartate D391 was shown by two teams to govern heterocyclization through interactions with the acceptor substrate.<sup>22,23</sup>

Unexpectedly, in Cy1, mutating this residue to an asparagine (D391N) induces a spectacular global response that reaches both binding sites (Fig. 4a, SI Fig. 7a). Such changes could indicate assaults to structural integrity



but instead this mutation stabilizes Cy1's structure (Fig. 4b, SI Fig. 8a-c) and impairs structural fluctuations, as evidenced by Hartman-Hahn  $R_{ex}$  profiles<sup>32</sup> (Fig. 4c, SI Fig. 8d). The global response imparted by this mutation does not reflect conformational selection of a minor Cy1 conformation (SI Figs. 7b, c). Finally, we repeated our experiment to monitor allostery and, although a higher population of loaded T1 should have favoured complex formation for D391N (SI Fig. 9c), we observed a severely hampered response of D391N to loaded T1, with only five signals detected for D391N versus 77 for wild type Cy1 and a dramatic reduction in the apparent population of molecules responding (Fig. 4d). D391N's mitigated response to loaded T1 indicates that, while the two domains do still interact, the formation of a minor state with an allosteric response is now hindered. These experiments establish dynamics as a mediator of the substrate-specific allosteric response we uncovered in the cyclization domain.

## CONCLUSION

Overall, our studies establish global structural dynamics as universal sensors of molecular events and highlight challenges in interpreting outcomes of mutagenesis in dynamic proteins. Thus, structural fluctuations within a protein enable sensing of post-translational modifications within partners and can promote interactions together with remodelling of distant sites. Here, a point-site mutation cannot be interpreted solely through local effects, e.g. by considering the position of the mutation in a binding site (SI Discussion and SI Fig. 10). Instead, the mutation throws sand into intricate, concerted dynamics such that it affects binding events on remote sites, thus confounding interpretation through local criteria. As a corollary, our results indicate that intra-domain dynamics in Cy or C domains should be preserved in engineered NRPSs as they enable to remodel the inter-domain NRPS dynamics by promoting engagement with loaded partners and preventing unproductive interactions with holo domains (SI Discussion). The conventional wisdom is that structure is function, and our results show that function includes

structural fluctuations. As structures become easier to determine, it is time to establish methods and models to routinely integrate dynamics to molecular mechanisms.

## METHODS

**Chemicals and reagents.** Routine chemicals used for protein expression and purifications were purchased from Sigma, Fisher Scientific, VWR, Research Products International (RPI), or MP Biomedicals. NMR isotopes were purchased from Sigma (glucose and D<sub>2</sub>O for 4 samples used for NOESY experiments) or Cambridge Isotope Laboratories (all other samples). Salicylate, ATP, and DTT were purchased from Sigma, GoldBio, and VWR, respectively. Unless otherwise noted, *E. coli* strains for expression were purchased from Novagen. *E. coli* ΔEntD cell lines used in the expression of T1 were courtesy of Drs. Chalut and Guilhot (CNRS, Toulouse, France). Cloning and mutagenesis enzymes and primers were purchased from New England BioLabs (NEB) and Integrated DNA Technologies (IDT), respectively. The adenylation domain YbtE and thioesterase SrfAD were prepared as described in,<sup>28</sup> with vectors provided by the C. T. Walsh lab (previously at Harvard Medical School). The precursors to prepare stereospecific <sup>1</sup>H-<sup>13</sup>C-Me samples were a gift from Dr. Haribabu Arthanari (Dana-Farber Cancer Institute).

**Protein expression, purification, and *in situ* reactions.** Holo-T1<sup>33</sup> and Cy1<sup>34</sup> were produced as previously described with variations in labelling schemes detailed in the SI Methods. The D391N variant was obtained through mutagenesis using PCR guided site-directed mutagenesis with primers 5'-gtctggataaatcatctggcggttcgagcatcacggcgaggtc-3' (forward primer) and 5'-cagatgatttatccagacctgcggcggtttgcgagatgccccattc-3' (reverse primer). Expression and purification protocols followed those used for wild-type Cy1 (SI Methods). Incomplete factorial design<sup>35</sup> revealed conditions to optimize Cy1 stability. Following purification, Cy1 is stored in 20 mM sodium phosphate, 100 mM NaCl, 1 mM ethylenediamine tetraacetate (EDTA), 5 mM dithiothreitol (DTT), pH 8 at 4 °C. For deuterated samples, Cy1 was incubated at room

temperature for 4 to 5 days in 50 mM Tris, 10 mM NaCl, 1 mM EDTA, 10 mM DTT, pH 8.5 to accelerate back-exchange of solvent-protected amides. The buffer for all NMR experiments, except for *in situ* experiments (*vide infra*), was 20 mM sodium phosphate, 10 mM NaCl, 1 mM EDTA, 5 mM DTT, 0.05 % w/v sodium azide, pH 7.0 at 25°C. D<sub>2</sub>O was added at 10 % v/v and sodium trimethylsilylpropanesulfonate (DSS) was added to reach defined concentrations in the range 0.2-0.5 mM (DSS protons). Cy1 is sensitive to oxidation and all buffers were freshly prepared with fresh DTT and degassed. NMR samples were degassed in Shigemi tubes, bubbled with argon, and sealed with glass plunger and parafilm.

Complexes of Cy1 or D391N with holo-T1 or loaded-T1 were prepared by co-concentration of individual stocks (prepared in identical buffers) to reach the desired concentrations. Here, 1D-NMR isotope-edited spectra were recorded for isolated proteins and for proteins in complexes to calculate the final experimental concentrations (details in SI Methods) by scaling concentrations initially determined using U.V absorption, with absorbance coefficients ( $\epsilon$ ) of 20,970 and 88,265 M<sup>-1</sup> cm<sup>-1</sup> for T1 and Cy1, respectively.

Holo-T1 was loaded with salicylate with a protocol modified from that of Goodrich et al.<sup>33</sup> with the buffer 50 mM ACES, 10 mM NaCl, 2 mM MgCl<sub>2</sub>, 1 mM TCEP, pH 7 at room temperature. All NMR samples used in the *in situ* assay contained 10% D<sub>2</sub>O and DSS (0.2 mM DSS protons) for referencing. Loaded-T1 was presented four times to wild-type Cy1 and once to D391N. In each case, the following controls were performed. 2D-HN-TROSY<sup>36</sup> were recorded for holo-T1, Cy1, and for holo-T1 with Cy1. 2D-IDIS-TROSYs<sup>30</sup> were recorded for the latter, as well as after addition of salicylate to 1 mM and ATP to 2 mM. Loading of holo-T1 with salicylate were initiated by addition of the adenylation domain YbtE and monitored *in situ* through the emergence of signals characteristic of loaded-T1<sup>28</sup> in HN-TROSY experiments (Figure 3, SI Figures 5 and 6). As discussed in the SI Discussion, Cy1 apparently accelerates hydrolysis of the thioester bond tethering salicylate to T1, suggesting that the reported hydrolase activity of C-domains<sup>37</sup> extends to cyclization domains. We did not focus on this aspect other than varying the

concentrations of YbtE, T1, and Cy1 such that we could detect Cy1 signals with sufficient signal-to-noise whilst controlling the amount of loaded T1. Depletion of ATP was monitored by 1D NMR and the integrity of loaded T1 was monitored by 2D HN-TROSYs. ATP needed to be replenished around 3 days after T1 was loaded to a stable amount. A first sample was prepared for quantifying the population of the minor conformer of Cy1 using holo- $^1\text{H}$ - $^{15}\text{N}$ - $^{13}\text{C}$  T1 and  $^2\text{H}$ - $^{15}\text{N}$ - $^{13}\text{C}$  Cy1 (Fig. 3 and SI Figs. 5, 6, and 9). Here, a 3D-HNCO was recorded for free Cy1 as well as before loading holo-T1 to verify that no spectroscopic changes occurred when it is presented to Cy1 (SI Fig. 6). T1 was then loaded to 58 % *in situ* and a second HNCO was immediately recorded. A second sample was prepared to help assign signals of the minor conformer of Cy1 using HNCO and HNCA recorded with  $^2\text{H}$ - $^{15}\text{N}$ - $^{13}\text{C}$  T1 loaded to 60% and  $^1\text{H}$ -Me-ILV- $^2\text{H}$ - $^{15}\text{N}$ - $^{13}\text{C}$  Cy1. To correlate the population of loaded T1 with that of the minor conformer of Cy1 (Fig. 3e), a series of 190 time-shared HN-TROSY/HC-HSQC (each 7 minutes long) were collected at 5 nM YbtE, 50 nM YbtE, and 200 nM YbtE. Spectra were then summed in groups of 15, starting from the first spectrum recorded at 5 nM, with the next group starting from the third spectrum, and the last group containing the last 15 spectra recorded at 200 nM YbtE. Here, T1 was loaded to 80% in the final spectra recorded at 50 nM YbtE, with no additional loading observed at 200 nM YbtE.

A 3D HNCO was recorded when D391N was presented with loaded-T1 using an *in situ* protocol, with controls first performed, and holo-T1 loaded upon addition of YbtE to 100 nM. D391N and T1 concentrations were similar to that used for WT HNCO and HNCA (SI Methods). This time, T1 was loaded to 77%. We did not reduce the concentration of YbtE to reach the same conversion rate as for wild-type as the larger population of loaded-T1 strengthened our observation that D391N does not respond to loaded-T1 as efficiently as the wild-type Cy1 does. It is tempting to extrapolate that the mutation impacts the hydrolase activity of Cy1, but we had not conducted controls with free holo-T1 during that set of experiments, and future studies should be performed to validate this observation (SI Discussion).

To demonstrate that the minor conformer of Cy1 (WT) does not result from the build-up of AMP or Sal-AMP during the *in situ* reaction, we added the thioesterase SrfAD to 500 nM and verified that T1 was restored to its holo form. A 3D-HNCO was then recorded to confirm that the signals of the minor conformer disappeared (SI Fig. 6), thus confirming that they reflected specifically a response of Cy1 to salicylate attached to T1 through a thioester bond.

**NMR data acquisition and processing.** All experiments were conducted at 25 °C on a 600 MHz AVANCEIII Bruker spectrometer (Hopkins School of Medicine) equipped with a QCI cryoprobe, an 800 MHz Varian Unity+ spectrometer equipped with a chiliprobe (Hopkins Arts and Science), and a 950 MHz AVANCEIII Bruker spectrometer (University of Maryland School of Medicine) equipped with a TCI cryoprobe. *In situ* experiments and Hahn-Echo experiments were recorded at 600 MHz, experiments to assign resonances were recorded at 600 and 800 MHz, distance constraints were measured at 600, 800, and 950 MHz, and relaxation dispersion profiles were recorded at 600 MHz and 950 MHz (*vide infra*). A list of experiments featuring acquisition parameters, samples, and fields is provided in the SI Methods. All experiments used for Cy1 used TROSY<sup>36</sup> to minimize losses due to relaxation. All spectra were processed with NMRPipe<sup>38</sup> and analysed with CARA<sup>39</sup> (assignments) or Sparky<sup>40</sup> (chemical shift perturbations).

**Assignment of backbone resonances.** Details of acquisition parameters, experiments, and samples are provided in the SI Methods. Notable strategies to overcome spectral crowding included using the TROSY-hNCA<sup>41</sup> and TROSY-hNcaNH<sup>42</sup> experiments as well as calculating 4D covariance correlation maps with non-uniformly sampled HNCO, HN(CA)CO, HN(CO)CA, HNCA, HN(COCA)CB, and HN(CA)CB. In the end, 93% of Cy1 backbone resonances were assigned. The mutant D391N was assigned through transposition of assignments from wild-type Cy1 using HNCA and HNCO. 82% of the mutant backbone resonances were assigned. Assignments of T1 were reported before.<sup>33</sup>

**Assignment of side-chain resonances.** The majority of Cyl methyl resonances were assigned using covariance maps calculated through HMCM(CG)CB, HMCM(CGCB)CA, HNCA, and HN(CA)CB spectra<sup>34</sup> collected for a  $^1\text{H}$ -Me-ILV- $^2\text{H}$ - $^{15}\text{N}$ - $^{13}\text{C}$  sample. HCCONH<sup>43</sup> experiments were recorded on the same sample together with an HCCH-TOCSY. 98 % of Cyl methyls (ILV) were assigned. 42% of aromatic (tryptophan, phenylalanine, and tyrosine) sidechain proton signals were assigned following initial structure calculations, using NOESY spectra and  $\text{H}^\alpha$  assignments from MQ-HACACO<sup>44</sup> as check points.

**NMR constraints.** Torsion angle constraints were obtained with the program TALOS+.<sup>45</sup> The majority of the distance constraints were determined with a time-shared 3D HN-TROSY/HC-HSQC-NOESY<sup>46</sup> with a mixing time ( $\tau_m$ ) of 40 ms collected at 950 MHz on a  $^1\text{H}^{13}\text{C}$ - $\delta\text{I}$ - $\delta\text{L}$ - $\gamma\text{V}$ - $^2\text{H}$ - $^{15}\text{N}$  sample.<sup>47</sup> A second spectrum was collected for long-range constraints ( $\tau_m$  150 ms). Assignments were facilitated through a non-uniformly sampled 4D HN-TROSY/HC-HSQC-NOESY-HN-TROSY/HC-HSQC experiment<sup>48,49</sup> ( $\tau_m$  200 ms), recorded at 800 MHz on a  $^1\text{H}^{13}\text{C}$ - $\delta\text{I}$ -LV- $^2\text{H}$ - $^{15}\text{N}$  sample that also contained protonated tyrosine and phenylalanine.<sup>50</sup> A 3D NOESY-HN-TROSY/HC-HSQC was also recorded to provide constraints with aromatic moieties. Other NOESYs were collected on Cyl samples during method developments, and distance constraints were retrieved as described in the next section and the SI Methods.

$^1\text{H}$ - $^{15}\text{N}$  residual dipolar couplings were obtained by recording TROSY and  $^{15}\text{N}$  decoupled 3D HNCOs in an interleaved manner in both isotropic (NMR buffer) and aligned media. Alignment was achieved using 12 mg/mL of Pfl phage (deuterium quadrupolar splitting of 11.4 Hz). Residues displaying relaxation dispersion or with  $^{15}\text{N}$  order parameters smaller than 0.7 (obtained from TALOS+<sup>45</sup>) were excluded from the restraint list.

**Structure calculation.** Details in the SI Methods. Structures were calculated using CYANA 3.98<sup>51,52</sup> and refined using CNS.<sup>53</sup> 810 dihedral angle restraints were prepared using a



consensus between TALOS+<sup>54</sup> and CSI 3.0.<sup>55</sup> 293 hydrogen bond constraints were introduced based on secondary structure prediction through TALOS+ and NOE patterns. A structural model calculated with RASREC CS-Rosetta<sup>56</sup> (SI Methods) was used to provide hydrogen bond and dihedral angle constraints for unassigned residues or residues whose signals were too weak to provide constraints; 7 residues in  $\alpha 4$  and 5 in  $\beta 11$ . First, a structural ensemble (root-mean-square deviation (RMSD) of 1.96 Å over residues in secondary structured elements) was calculated using 1187 accurate distances obtained from time-shared NOESY spectra collected on a sample with stereospecific protonation of Leu and Val methyls at  $\delta 2$  and  $\gamma 2$ , respectively, and with a mixing time of 40 ms, i.e. with minimal spin diffusion. A representative conformer was then identified from the resulting structural ensemble. To include long-range constraints, an additional 1,002 distance restraints were selected from a set of NOESY spectra at longer mixing times (excluding those at 40 ms, see SI Methods) and binned into groups with upper limits at 3, 5, 7, and 9 Å such that they were in agreement with the initial representative model. 176 residual dipolar couplings were then incorporated, excluding those of residues showing dispersion or possessing low order parameters (estimated through TALOS+). The deposited bundle features no violations larger than 0.5 Å and exhibits an RMSD of 1.2 Å for residues in secondary structured elements (SI Table S1).

**Experiments to determine protein dynamics.** Relaxation dispersion experiments at 600 and 950 MHz were recorded with TROSY and relaxation compensation of constant-time Carr-Purcell-Meiboom-Gill pulse sequences,<sup>57,58</sup> using a 600  $\mu$ M sample of Cy1 with  $^1\text{H}^{13}\text{C}$ - $\delta\text{I}$ - $\delta 2\text{L}\gamma 2\text{V}$ - $^2\text{H}$ - $^{15}\text{N}$  labelling. The complete set of relaxation dispersions were recorded in an interleaved manner such that all CPMG frequencies were sampled before indirect nitrogen dimensions were encoded. The CPMG frequencies were acquired in random order. Intensities were estimated through line-shape fitting with the nLinLS module of NMRPipe.<sup>38</sup> The relaxation dispersion profiles were fitted using ChemEx,<sup>59</sup> which integrates the Bloch-McConnell equation over the NMR pulse sequence. The fit was run in three consecutive steps. First, a subset of

isolated signals showing dispersions and good signal-to-noise were selected and fit together to obtain an estimate of the population and exchange rate for a two-site model. Second, all profiles were fit with exchange rates and populations fixed to the values obtained in the first step. Third, residues for which the change in chemical shift was smaller than the estimated uncertainty were kept fixed at a value of 0.0 during a final fit. Of 298 residues integrated, 178 could be fit. The profiles and their fits were subject to further statistical analysis following data fitting. Dispersion was considered significant only if the following two criteria were fulfilled. 1) Their data could not be fitted by a single average value,<sup>60,61</sup> as measured by the following  $\chi^2$ :

$$\chi^2_{<R_{2,eff}>} = \sum \frac{(R_{2,eff}^i - \langle R_{2,eff} \rangle)^2}{(\Delta R_{2,eff}^i)^2}$$

where  $R_{2,eff}^i$  denotes the effective transverse relaxation rate for a given CPMG frequency  $\nu_i$ ,<sup>57</sup>  $\langle \rangle$  denotes the mean value over all frequencies, and  $\Delta R_{2,eff}^i$  denotes the uncertainty as described in <sup>60</sup>. 2) The magnitude of the fitted dispersion had to exceed the average residual of the fit. 52 residues were used for quantitative comparisons of chemical shifts based on low  $\chi^2$  at both 600 MHz and 950 MHz, another 74 residues showed unambiguous dispersion, and 5 residues clearly demonstrated departure from the global fit.

Structural fluctuations over a wider  $\mu$ s-10's of ms timescale were detected through the Hahn-Echo method with a 600  $\mu$ M sample of Cy1 WT (CDN) and a 122  $\mu$ M sample of Cy1 D391N (CDN) using the analysis described in <sup>32</sup> and a pulse sequence modified to account for <sup>13</sup>C labelling. See SI Methods for more details.

**Calculation of Chemical Shift Perturbations (CSPs,  $\Delta\delta$ ).** For all CSPs, samples were prepared in an identical buffer, spectra were referenced through DSS,<sup>62</sup> and temperatures were calibrated. CSPs were calculated through the relation:<sup>63</sup>

$$\Delta\delta = \sqrt{\left[\frac{1}{3}\left\{\Delta\delta(^1H^N)^2 + \Delta\delta(^{15}N)^2 + \Delta\delta(^{13}CO)^2\right\}\right]}$$

where  $\delta$  is the chemical shift difference in ppm for a given nucleus between the two sets of signals that are compared (e.g. Cy1 WT and Cy1 D391N or Cy1 major conformer and Cy1 minor conformer). For figures where the CSPs are mapped onto the medoid structure of Cy1, residues are colored according to their CSP value with a gradient ranging from grey to red, where red indicates a CSP that is at or above two standard-deviations from the median value of CSPs. See SI Methods for more details.

**Cy1 Tunnel Calculations.** All conformers from the Cy1 NMR ensemble were analysed using the CAVER 3.0.3 Pymol plugin.<sup>64</sup> For each conformer, tunnels were generated by optimizing parameters such that a tunnel connecting the putative donor and acceptor thiolation domain entrances in the structure were found (details in SI Methods). To identify closed tunnels, where donor or acceptor entrances were inaccessible, the probe radius used in the calculation was decreased and the obstructed tunnels were identified from a set of unrealistic tunnels by comparison with open tunnels in other conformations.

**Cy1 WT and D391N Thermal Stability.** Unlabelled wild-type Cy1 (WT) and D391N samples were spectroscopically characterized using circular dichroism (CD). Cy1 WT and Cy1 D391N at concentrations of 0.035 and 0.028 mg/mL, respectively, were thermally unfolded in NMR buffer (20 mM sodium phosphate, 10 mM NaCl, 1 mM EDTA, pH 7) with 5  $\mu$ M TCEP and 0.0025% w/v NaN<sub>3</sub>. For CD measurements, a wavelength of 222 nm was monitored over the course of thermal unfolding. Both Cy1 WT and D391N exhibited an irreversible, two-state unfolding transition and were fit to a two-state model for determination of melting temperatures (see SI Methods).

**Statistics and Error Analysis.** See also details of experimental procedures. The experimental error bars in the relaxation dispersion profiles (Fig. 3) are calculated as described in

<sup>60</sup> following integration through nLinLS with NMRPipe.<sup>65</sup> The error on reported chemical shifts (SI Fig. 4, 6, and 7) were estimated using the covariance matrix approach.<sup>66</sup> When determining the percentage of loaded-T1 and minor peaks of Cy1 as T1 was loaded *in situ* (Fig. 3e), populations were calculated as:

$$Population(loaded\ T1) = \frac{I_{loaded}}{I_{holo} + I_{loaded}}$$

and

$$Population(Cy1_{minor}) = \frac{I_{minor}}{I_{major} + I_{minor}}$$

The values reported are means over two signals (T1) and six signals (Cy1) detected in time-shared experiments and the error bars were calculated via error propagation where the noise of the associated spectrum is the uncertainty of measurement of each signal intensity. In other experiments, loading of T1 was quantified through five signals resolved in 2D IDIS-TROSY<sup>30</sup> data. Means were calculated and the error bars in SI Figs. 5 and 9 represent standard deviations to the mean. The populations of Cy1 minor residues are reported as a scatter plot in Fig. 4d, with a box-and-whisker in which the central line denotes the median and the edges the 25<sup>th</sup> and 75<sup>th</sup> percentiles. The limits of the whiskers identify populations with more than 1.5 times the interquartile range. Only 5 signals were detected for the minor conformer of D391N, and they could not be assigned. For all other residues, the noise amplitude defines a threshold for detection that was used to calculate the population, which lead to a 10-fold reduction in population when compared to wild-type Cy1. For Cy1 thermal melt analyses (Fig. 4b and SI Fig. 8), the error on the signals was computed using the standard deviation of the signal averaged over one second on the CD instrument. Error bars in the determination of Hartmann-Hahn  $R_{ex}$  profiles for Cy1 WT and D391N (Fig. 4c) were determined using Monte Carlo error estimates over 300 iterations, with the noise used as the variance for the distribution.

## Data Availability:

The data that support the findings of this study are available from the corresponding author upon reasonable request. NMR chemical shifts have been deposited in the Biological Magnetic Resonance Data Bank under the following accession codes: 30943 for wild-type Cy1 and 51105 for mutant D391N. The structure of Cy1 has been deposited to the Protein Data Bank (PDB) under accession code 7RY6.

## REFERENCES

- 1 Kaniusaite, M., Goode, R. J. A., Tailhades, J., Schittenhelm, R. B. & Cryle, M. J. Exploring modular reengineering strategies to redesign the teicoplanin non-ribosomal peptide synthetase. *Chemical Science* **11**, 9443-9458, doi:10.1039/d0sc03483e (2020).
- 2 Niquille, D. L. *et al.* Nonribosomal biosynthesis of backbone-modified peptides. *Nature Chemistry* **10**, 282-287, doi:10.1038/nchem.2891 (2018).
- 3 Bozhüyük, K. A. J. *et al.* Modification and de novo design of non-ribosomal peptide synthetases using specific assembly points within condensation domains. *Nature Chemistry* **11**, 653-661, doi:10.1038/s41557-019-0276-z (2019).
- 4 Frueh, D. P. *et al.* Dynamic thiolation-thioesterase structure of a non-ribosomal peptide synthetase. *Nature* **454**, 903-906, doi:10.1038/nature07162 (2008).
- 5 Tanovic, A., Samel, S. A., Essen, L. O. & Marahiel, M. A. Crystal structure of the termination module of a nonribosomal peptide synthetase. *Science* **321**, 659-663 (2008).
- 6 Drake, E. J. *et al.* Structures of two distinct conformations of holo-non-ribosomal peptide synthetases. *Nature* **529**, 235-238, doi:10.1038/nature16163(2016).
- 7 Reimer, J. M., Aloise, M. N., Harrison, P. M. & Martin Schmeing, T. Synthetic cycle of the initiation module of a formylating nonribosomal peptide synthetase. *Nature* **529**, 239-242, doi:10.1038/nature16503 (2016).
- 8 Reimer, J. M. *et al.* Structures of a dimodular nonribosomal peptide synthetase reveal conformational flexibility. *Science* **366**, doi:10.1126/science.aaw4388 (2019).
- 9 Bhabha, G. *et al.* A Dynamic Knockout Reveals That Conformational Fluctuations Influence the Chemical Step of Enzyme Catalysis. *Science* **332**, 234-238, doi:10.1126/science.1198542 (2011).
- 10 Henzler-Wildman, K. A. *et al.* A hierarchy of timescales in protein dynamics is linked to enzyme catalysis. *Nature* **450**, 913-916, doi:10.1038/nature06407 (2007).

- 11 Whittier, S. K., Hengge, A. C. & Loria, J. P. Conformational motions regulate phosphoryl transfer in related protein tyrosine phosphatases. *Science* **341**, 899-903, doi:10.1126/science.1241735 (2013).
- 12 Xie, T., Saleh, T., Rossi, P. & Kalodimos, C. G. Conformational states dynamically populated by a kinase determine its function. *Science* **370**, doi:10.1126/science.abc2754 (2020).
- 13 Oyen, D. *et al.* Defining the Structural Basis for Allosteric Product Release from E. coli Dihydrofolate Reductase Using NMR Relaxation Dispersion. *Journal of the American Chemical Society* **139**, 11233-11240, doi:10.1021/jacs.7b05958 (2017).
- 14 Saavedra, H. G., Wrabl, J. O., Anderson, J. A., Li, J. & Hilser, V. J. Dynamic allostery can drive cold adaptation in enzymes. *Nature* **558**, 324-328, doi:10.1038/s41586-018-0183-2 (2018).
- 15 Hadzipasic, A. *et al.* Ancient origins of allosteric activation in a Ser-Thr kinase. *Science* **367**, 912-917, doi:10.1126/science.aay9959 (2020).
- 16 Smock, R. G. & Gierasch, L. M. Sending Signals Dynamically. *Science* **324**, 198-203, doi:10.1126/science.1169377 (2009).
- 17 Süssmuth, R. D. & Mainz, A. Nonribosomal Peptide Synthesis—Principles and Prospects. *Angewandte Chemie International Edition* **56**, 3770-3821, doi:10.1002/anie.201609079 (2017).
- 18 Gulick, A. M. Conformational dynamics in the Acyl-CoA synthetases, adenylation domains of non-ribosomal peptide synthetases, and firefly luciferase. *ACS Chem Biol* **4**, 811-827, doi:10.1021/cb900156h (2009).
- 19 Huguenin-Dezot, N. *et al.* Trapping biosynthetic acyl-enzyme intermediates with encoded 2,3-diaminopropionic acid. *Nature* **565**, 112-117, doi:10.1038/s41586-018-0781-z (2019).
- 20 Bloudoff, K., Rodionov, D. & Schmeing, T. M. Crystal structures of the first condensation domain of CDA synthetase suggest conformational changes during the synthetic cycle of nonribosomal peptide synthetases. *J Mol Biol* **425**, 3137-3150, doi:10.1016/j.jmb.2013.06.003 (2013).
- 21 Bloudoff, K. & Schmeing, T. M. Structural and functional aspects of the nonribosomal peptide synthetase condensation domain superfamily: discovery, dissection and diversity. *Bba-Proteins Proteom* **1865**, 1587-1604, doi:10.1016/j.bbapap.2017.05.010 (2017).
- 22 Bloudoff, K., Fage, C. D., Marahiel, M. A. & Schmeing, T. M. Structural and mutational analysis of the nonribosomal peptide synthetase heterocyclization domain provides insight into catalysis. *Proc Natl Acad Sci U S A* **114**, 95-100, doi:10.1073/pnas.1614191114 (2017).
- 23 Dowling, D. P. *et al.* Structural elements of an NRPS cyclization domain and its intermodule docking domain. *Proc Natl Acad Sci U S A* **113**, 12432-12437, doi:10.1073/pnas.1608615113 (2016).
- 24 Drake, E. J. *et al.* Structures of two distinct conformations of holo-non-ribosomal peptide synthetases. *Nature* **529**, 235-238, doi:10.1038/nature16163 (2016).
- 25 Izoré, T. *et al.* Structures of a non-ribosomal peptide synthetase condensation domain suggest the basis of substrate selectivity. *Nature Communications* **12**, 2511, doi:10.1038/s41467-021-22623-0 (2021).
- 26 Kosol, S. *et al.* Structural basis for chain release from the enacyloxin polyketide synthase. *Nature chemistry* **11**, 913-923, doi:10.1038/s41557-019-0335-5 (2019).



- 27 Palmer, A. G. & Koss, H. in *Methods in Enzymology* Vol. 615 (ed A. Joshua Wand) 177-236 (Academic Press, 2019).
- 28 Goodrich, A. C. & Frueh, D. P. A nuclear magnetic resonance method for probing molecular influences of substrate loading in nonribosomal peptide synthetase carrier proteins. *Biochemistry* **54**, 1154-1156, doi:10.1021/bi501433r (2015).
- 29 Frueh, D. P., Vosburg, D. A., Walsh, C. T. & Wagner, G. Determination of all nOes in 1H-13C-Me-ILV-U-2H-15N proteins with two time-shared experiments. *J Biomol NMR* **34**, 31-40, doi:10.1007/s10858-005-5338-4 (2006).
- 30 Golovanov, A. P., Blankley, R. T., Avis, J. M. & Bermel, W. Isotopically discriminated NMR spectroscopy: a tool for investigating complex protein interactions in vitro. *J Am Chem Soc* **129**, 6528-6535 (2007).
- 31 Shi, C., Miller, B. R., Alexander, E. M., Gulick, A. M. & Aldrich, C. C. Design, Synthesis, and Biophysical Evaluation of Mechanism-Based Probes for Condensation Domains of Nonribosomal Peptide Synthetases. *ACS Chem Biol* **15**, 1813-1819, doi:10.1021/acscchembio.0c00411 (2020).
- 32 Wang, C., Rance, M. & Palmer, A. G., 3rd. Mapping chemical exchange in proteins with MW > 50 kD. *Journal of the American Chemical Society* **125**, 8968-8969, doi:10.1021/ja035139z (2003).
- 33 Goodrich, A. C., Harden, B. J. & Frueh, D. P. Solution Structure of a Nonribosomal Peptide Synthetase Carrier Protein Loaded with Its Substrate Reveals Transient, Well-Defined Contacts. *Journal of the American Chemical Society* **137**, 12100-12109, doi:10.1021/jacs.5b07772 (2015).
- 34 Mishra, S. H. & Frueh, D. P. Assignment of methyl NMR resonances of a 52 kDa protein with residue-specific 4D correlation maps. *J Biomol NMR* **62**, 281-290, doi:10.1007/s10858-015-9943-6 (2015).
- 35 Ducat, T., Declerck, N., Gostan, T., Kochoyan, M. & Demene, H. Rapid determination of protein solubility and stability conditions for NMR studies using incomplete factorial design. *J Biomol NMR* **34**, 137-151 (2006).
- 36 Pervushin, K., Riek, R., Wider, G. & Wüthrich, K. Attenuated T2 relaxation by mutual cancellation of dipole-dipole coupling and chemical shift anisotropy indicates an avenue to NMR structures of very large biological macromolecules in solution. *Proc Natl Acad Sci U S A* **94**, 12366-12371 (1997).
- 37 Schoppet, M. *et al.* The biosynthetic implications of late-stage condensation domain selectivity during glycopeptide antibiotic biosynthesis. *Chemical science* **10**, 118-133, doi:10.1039/c8sc03530j (2019).
- 38 Delaglio, F. *et al.* NMRPipe a Multidimensional Spectra Processing System Based on UNIX Pipes. *J Biomol NMR* **6**, 277-293 (1995).
- 39 Keller, R. L. J. *The Computer Aided Resonance Assignment Tutorial*. (Cantina Verlag, 2004).
- 40 SPARKY 3 (Univeristy of California, San Fransisco).
- 41 Frueh, D. P., Arthanari, H., Koglin, A., Walsh, C. T. & Wagner, G. A double TROSY hNCaNH experiment for efficient assignment of large and challenging proteins. *Journal of the American Chemical Society* **131**, 12880-12881, doi:10.1021/ja9046685 (2009).
- 42 Frueh, D. P. *et al.* Non-uniformly sampled double-TROSY hNcaNH experiments for NMR sequential assignments of large proteins. *Journal of the American Chemical Society* **128**, 5757-5763, doi:10.1021/ja0584222 (2006).

- 43 Grzesiek, S., Anglister, J. & Bax, A. Correlation of Backbone Amide and Aliphatic Side-Chain Resonances in  $^{13}\text{C}/^{15}\text{N}$ -Enriched Proteins by Isotropic Mixing of  $^{13}\text{C}$  Magnetization. *J Magn Reson B* **101**, 114-119 (1992).
- 44 Pervushin, K. & Eletsky, A. A new strategy for backbone resonance assignment in large proteins using a MQ-HACACO experiment. *J Biomol NMR* **25**, 147-152 (2003).
- 45 Shen, Y., Delaglio, F., Cornilescu, G. & Bax, A. TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts. *J Biomol NMR* **44**, 213-223, doi:10.1007/s10858-009-9333-z (2009).
- 46 Mishra, S. H., Harden, B. J. & Frueh, D. P. A 3D time-shared NOESY experiment designed to provide optimal resolution for accurate assignment of NMR distance restraints in large proteins. *J Biomol NMR* **60**, 265-274, doi:10.1007/s10858-014-9873-8 (2014).
- 47 Gans, P. *et al.* Stereospecific isotopic labeling of methyl groups for NMR spectroscopic studies of high-molecular-weight proteins. *Angew Chem Int Ed Engl* **49**, 1958-1962, doi:10.1002/anie.200905660 (2010).
- 48 Frueh, D. P., Vosburg, D. A., Walsh, C. T. & Wagner, G. Determination of all nOes in  $^1\text{H}$ - $^{13}\text{C}$ -Me-ILV-U- $^2\text{H}$ - $^{15}\text{N}$  proteins with two time-shared experiments. *J Biomol NMR* **34**, 31-40 (2006).
- 49 Wurtz, P., Aitio, O., Hellman, M. & Permi, P. Simultaneous detection of amide and methyl correlations using a time shared NMR experiment: application to binding epitope mapping. *J Biomol NMR* **39**, 97-105 (2007).
- 50 Medek, A., Olejniczak, E. T., Meadows, R. P. & Fesik, S. W. An approach for high-throughput structure determination of proteins by NMR spectroscopy. *J. Biomol. NMR* **18**, 229-238. (2000).
- 51 Güntert, P., Mumenthaler, C. & Wüthrich, K. Torsion angle dynamics for NMR structure calculation with the new program DYANA. *J. Mol. Biol.* **273**, 283-298 (1997).
- 52 Schmidt, E. & Guntert, P. Automated structure determination from NMR spectra. *Methods Mol Biol* **1261**, 303-329, doi:10.1007/978-1-4939-2230-7\_16 (2015).
- 53 Brunger, A. T. Version 1.2 of the Crystallography and NMR system. *Nat Protoc* **2**, 2728-2733, doi:10.1038/nprot.2007.406 (2007).
- 54 Shen, Y. & Bax, A. Protein backbone and sidechain torsion angles predicted from NMR chemical shifts using artificial neural networks. *J Biomol NMR* **56**, 227-241, doi:10.1007/s10858-013-9741-y (2013).
- 55 Hafsa, N. E., Arndt, D. & Wishart, D. S. CSI 3.0: a web server for identifying secondary and super-secondary structure in proteins using NMR chemical shifts. *Nucleic Acids Res* **43**, W370-377, doi:10.1093/nar/gkv494 (2015).
- 56 Lange, O. F. *et al.* Determination of solution structures of proteins up to 40 kDa using CS-Rosetta with sparse NMR data from deuterated samples. *Proc Natl Acad Sci U S A* **109**, 10873-10878, doi:10.1073/pnas.1203013109 (2012).
- 57 Mulder, F. A., Skrynnikov, N. R., Hon, B., Dahlquist, F. W. & Kay, L. E. Measurement of slow (micros-ms) time scale dynamics in protein side chains by  $(^{15}\text{N})$  relaxation dispersion NMR spectroscopy: application to Asn and Gln residues in a cavity mutant of T4 lysozyme. *Journal of the American Chemical Society* **123**, 967-975 (2001).
- 58 Loria, J. P., Rance, M. & Palmer, A. G., 3rd. A TROSY CPMG sequence for characterizing chemical exchange in large proteins. *J Biomol NMR* **15**, 151-155, doi:10.1023/a:1008355631073 (1999).

- 59 Vallurupalli, P., Bouvignies, G. & Kay, L. E. Studying "invisible" excited protein states in slow exchange with a major state conformation. *Journal of the American Chemical Society* **134**, 8148-8161, doi:10.1021/ja3001419 (2012).
- 60 Ishima, R. in *Protein Dynamics: Methods and Protocols* (ed Dennis R. Livesay) 29-49 (Humana Press, 2014).
- 61 Sauerwein, A. C. & Hansen, D. F. in *Protein NMR Vol. 32 Biological Magnetic Resonance* (ed Lawrence Berliner) Ch. 3, 75-132 (Springer US, 2015).
- 62 Wishart, D. S. *et al.* <sup>1</sup>H, <sup>13</sup>C and <sup>15</sup>N chemical shift referencing in biomolecular NMR. *J Biomol NMR* **6**, 135-140, doi:10.1007/BF00211777 (1995).
- 63 Terada, T. *et al.* Nuclear magnetic resonance and molecular dynamics studies on the interactions of the Ras-binding domain of Raf-1 with wild-type and mutant Ras proteins. *J Mol Biol* **286**, 219-232, doi:10.1006/jmbi.1998.2472 (1999).
- 64 Chovancova, E. *et al.* CAVER 3.0: a tool for the analysis of transport pathways in dynamic protein structures. *PLoS Comput Biol* **8**, e1002708, doi:10.1371/journal.pcbi.1002708 (2012).
- 65 Delaglio, F. *et al.* NMRPipe a Multidimensional Spectra Processing System Based on UNIX Pipes. *JBNMR* **6**, 277-293 (1995).
- 66 Press, W. H., Teukolsky, S. A., Vetterling, W. T. & Flannery, B. P. *Numerical Recipes 3rd Edition: The Art of Scientific Computing*. (Cambridge University Press, 2007).

**Acknowledgements** We thank Drs A. Majumdar (Johns Hopkins Biomolecular NMR Center), Katie Tripp (Hopkins Center for Molecular Biophysics), David Weber and Kristen Varney (University of Maryland School of Medicine), and Bob Cole (Hopkins Mass Spectrometry and Proteomics Facility) for their assistance. This work was supported by NIH grant GM104257 to D.P.F.; N.S. acknowledges funding from NIAID grant 5R01AI143997 and NIGMS grant 5R35GM125034; D.P.D acknowledges funding from NIH grant 1R15GM123425-01.

**Author Contributions** S.H.M., A.K.K., K.A.M. and D.P.F. designed research and wrote the manuscript. S.H.M. assigned the resonances of Cy1 and its distance constraints, with help from D.D.'s EpoB crystal structure model, determined the initial Cy1 NMR structure model, performed *in situ* reactions with A.K.K. and K.A.M., and collected data for Hahn-Echo and relaxation dispersion for Cy1. In addition, A.K. determined the structure of Cy1 and analysed RDCs. K.A.M. collected (assisted by S.H.M. and A.K.K.) and analysed all data for D391N, collected and analysed CD data for Cy1, performed Cy1 tunnel calculations, and determined

minor Cyl populations. S.N. and N.S. calculated a RASREC model of Cyl's C-terminal domain, and G.B. analysed the relaxation dispersion data.

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

**Supporting Information** is available separately.

**THIS WORK HAS NOT BEEN PEER-REVIEWED.**