Published in final edited form as:

Chembiochem. 2017 April 04; 18(7): 629–632. doi:10.1002/cbic.201700030.

Molecular Cross-Talk between Nonribosomal Peptide Synthetase Carrier Proteins and Unstructured Linker Regions

Bradley J. Harden^a and Dominique P. Frueh^{*,a}

^aBiophysics and Biophysical Chemistry, Johns Hopkins University School of Medicine, 725 N. Wolfe Street, Baltimore, MD, 21205

Abstract

Nonribosomal peptide synthetases (NRPSs) employ multiple domains separated by linker regions to incorporate substrates into natural products. During synthesis, substrates are covalently tethered to carrier proteins that translocate between catalytic partner domains. The molecular parameters that govern translocation and associated linker remodeling remain unknown. Here, we used NMR to characterize the structure, dynamics, and invisible states of a peptidyl carrier protein flanked by its linkers. We show that the N-terminal linker stabilizes and interacts with the protein core while modulating dynamics at specific sites involved in post-translational modifications and/or domain interactions. The results detail the molecular communication between peptidyl carrier proteins and their linkers and may guide efforts in engineering NRPSs to obtain new pharmaceuticals.

TOC image

Linkers that matter: The solution structure of a peptidyl carrier protein flanked by its unstructured linkers reveals contacts between the core and linkers. The interaction is shown to both stabilize the protein core and modulate dynamics in a manner suggestive of allosteric communication.

* dfrueh1@jhmi.edu.

Supporting information for this article is given via <http://onlinelibrary.wiley.com/doi/10.1002/cbic.201700030/full>

Experimental Section

See supporting information.

Keywords

allostery; nonribosomal peptide synthetase; Nuclear Magnetic Resonance; protein dynamics; protein folding

Biological systems often rely on protein domains separated by linker regions to partition or multiply functionality. For example, nonribosomal peptide synthetases (NRPSs) are microbial enzymatic systems that employ a modular, multi-domain architecture to incorporate substrates into secondary metabolites^[1]. This assembly-line strategy, however, does not operate through a rigid molecular assembly but instead relies on structural fluctuations both within and between individual domains connected by linkers^[2,3]. Importantly, carrier protein domains covalently tether substrates through a phosphopantetheinyl (PP) arm introduced on a conserved serine, and they were found to shuttle substrates between catalytic domains during synthesis^[4,5]. We recently provided evidence of molecular cross-talk between a carrier protein and its tethered substrate^[6]. Here, we show that similar communication exists between the core of a carrier protein and its inter-domain linker regions. Using NMR we show that residues in the N-terminal linker interact with the domain core and modulate its dynamics. Further, by monitoring the protein's invisible unfolded state, we show that this linker region stabilizes the carrier protein fold.

We determined the solution structure (PDB 5U3H) and fast dynamics (ps-ns) of PCP1_{ybt}, a peptidyl carrier protein (PCP) from Yersiniabactin Synthetase^[7], including all of the flanking residues linking PCP1_{ybt} to its upstream adenylation and downstream cyclization domains (Figure 1). The core displays a four helix bundle commonly observed in carrier proteins^[8] and features an occasionally observed single-turn helix, α I, that lies between α 1 and α 2. The active serine S1439 is found at the N-terminus of α 2 (* in all Figures). Model Free analysis of ¹⁵N NMR relaxation^[9,10] provided us with order parameters S^2 to characterize ps-ns fluctuations of amide bonds. S^2 range from 0 to 1, and are sensitive to

both the amplitudes and orientations of motional processes. A value of 0 indicates an unrestricted motion and 1 denotes either a motion coaxial with the amide bond or rigidity.

The pattern of flexible residues observed in PCP1_{ybt} (Figure 1C, D) is reminiscent of those reported in aryl and acyl carrier proteins^[6]. Loop2 and the N-terminus of loop1 are malleable, while the four helices are relatively rigid. Of particular interest, the residues preceding the active serine (1436–1438) display alternately low and high values of S^2 , with the S^2 of G1437 10% and 15% higher than those of its predecessor and successor, respectively (D). This behavior may reflect the sensitivity of order parameters to anisotropic motions rather than an alternation of rigid and flexible residues. Acyl and aryl carrier proteins also display dynamics around the PP site, and malleability here is likely needed during the many molecular events involving these residues, including post-translational modifications^[6]. Our results indicate that peptidyl carrier proteins also possess malleability in this region.

Although most of the N- and C-terminal linker regions of PCP1_{ybt} are strongly disordered (Figures 1A and S5), there are three exceptions. In the C-terminal linker, which has recently been shown to be essential for communication between NRPS modules^[11], five residues immediately following $\alpha 4$ interact with the folded core along loop2. In the N-terminal linker, a short region approximately 20 residues before $\alpha 1$ displays helical character, as evidenced by chemical shift indexing and weak nuclear Overhauser effect (nOe) cross-peaks. A recent study assigned these residues to the preceding adenylation domain^[12], where they are expected to form a helix, and our results confirm this prediction. This region does not interact with the PCP core; its low S^2 indicates independent molecular tumbling (Figure S9), and no nOe cross-peaks to the core are observed. In contrast, a third region immediately prior to $\alpha 1$ displays extensive nOe's to the folded core, and its fast time-scale dynamics resemble those seen in loops rather than those seen in disordered linkers (Figure 1D). Consequently, we refer to this region as loop0. Contacts between loop0 and the core occur principally through the proline residue P1405. Notably, proline is the most common amino acid at this position^[12], and other PCP structures substantiate that observation. Out of 12 different Type I PCPs selected from the PDB, 11 contain a proline residue at this position and 9 of these display contacts between the proline and the core (Figure S4).

In PCP1_{ybt}, loop0 makes contacts to $\alpha 1$, loop3, and loop1. Loop1 plays an important role in communication with partner enzymatic domains^[8], prompting us to further characterize the influence of loop0 on the structure and dynamics of the folded core. Thus, we developed a second construct (blue in secondary structure diagrams of Figures 1, 2, and 4) that removes both N- and C-terminal linkers and excludes P1405, breaking the interaction between loop0 and the core.

Backbone and C β chemical shift perturbations (CSPs) between the full-length and short constructs localize primarily to the sites of truncation and to regions that contact loop0 (Figure 2), indicating that the structure of the core is largely unaltered. Only residue H1464, at the N-terminus of loop3, stands apart since it is distant from the contact point between loop3 and loop0. Notably, only the chemical shift of the carbonyl carbon is strongly affected

(Figure S5), which may indicate (transient) hydrogen bond formation upon subtle rearrangements of loop3.

Investigation of slow time-scale dynamics reveals that truncation of loop0 significantly destabilizes the folded core. We used ^{15}N Chemical Exchange Saturation Transfer (CEST) experiments^[13] to probe for “invisible” signals arising from secondary conformational states, and we found many in both the full length and short constructs (Figure 3).

Their chemical shifts indicate an unstructured environment (Figure 3C, D) and kinetic analysis reveals a single exchange process (SI), which together point to an unfolding process as seen in acyl carrier proteins^[14]. Removal of loop0 increases the population of the unfolded state to nearly 5%, up from 0.8% in the full length protein. Accordingly, the melting temperature of the short construct decreases relative to the full-length construct (Figure S7). Thus, loop0 contributes to the stability of the folded core.

Investigation of structural fluctuations in the folded PCP core using relaxation dispersion experiments (RD)^[15] did not reveal conformational exchange processes beyond the aforementioned unfolding. When compared to CEST experiments, RD experiments provide less direct evidence of exchange at millisecond time-scales but can characterize faster exchange processes, up to microsecond time-scales. However, we could not identify any dispersion that could not be accounted for by the parameters obtained with CEST (Figure S8). Consequently, within the range of time-scales probed by RD and CEST experiments, we could only detect the unfolding exchange process. Similarly, contributions from unfolding dominate in the exchange parameters obtained by Model Free analysis, complicating interpretation (Figure S12). Conformational fluctuations of the ground state may exist, but within the limits of our experimental procedures, they are masked by the unfolding process or are outside the time-scales covered.

To probe for allosteric communication between loop0 and the protein core at faster time-scales (ps-ns), we also performed Model Free analysis of the short construct of PCP1_{ybt}. The rotational correlation time (τ_c) for this construct falls exactly in line with the value predicted from its molecular weight (SI). However, this was not the case for the full length construct, which tumbles 30% more slowly than anticipated (SI). This behavior is presumably caused by the flexible regions of the linkers, in agreement with a study on such effects^[16]. Indeed, size exclusion chromatography indicates that full length PCP1_{ybt} has an unusually large hydrodynamic volume (Figure S13), yet multi-angle light scattering reveals a monomeric protein in solution (Figure S14). Thus, to prevent confounding hydrodynamic effects when interpreting changes in S^2 with and without loop0 contacts, we prepared a third construct of PCP1_{ybt}. This construct is identical to the short construct discussed throughout the text except that loop0 has been restored by extending the linker by four residues (beige in secondary structure diagrams of Figures 1, 2, and 4). Model Free analysis of the new construct confirms that its rotational correlation time agrees well with prediction (SI). As a result, comparisons of S^2 between the two short constructs faithfully report on modulations in fast time-scale dynamics due to contact between loop0 and the protein core.

Several differences emerge when evaluating the change in S^2 (S^2) between the two short constructs. Figure 4 maps the absolute value of S^2 to the ribbon width of the PCP1_{ybt} structure. Using the absolute value of S^2 emphasizes that changes in S^2 do not necessarily represent an increase or decrease in flexibility (see above). When loop 0 is shortened, we observe large changes in S^2 at the site of truncation and at the contact points between loop 0 and the core.

However, additional sites throughout the protein appear to be affected as well. Changes in S^2 within loop3 may reflect the subtle rearrangements invoked to explain the CSP at H1464. Changes in S^2 in the C-terminal tail contacting loop2 denote communication between both ends of the protein, which may occur indirectly through $\alpha 3$ and $\alpha 4$ via loop3, which contacts loop0. Further changes in S^2 are apparent at the N-terminus of loop1, which has been implicated in binding between related aryl carrier proteins and their associated adenylation domains^[17]. Finally, the region immediately preceding the post-translationally modified serine, residues 1436–1438, is of particular interest. The S^2 of G1437 is substantially larger than that of its neighbors when loop0 contacts the core (see above and SI) but the pattern is inverted in the shorter construct, denoting a substantial change in the amplitude and/or orientation of the motion. In any case, the conformational landscape in this region is modified, likely impacting communication between PCP1_{ybt} and other, enzymatic domains during synthesis. Critically, this dynamic region belongs to the major conserved motif of carrier proteins^[18].

In summary, we have determined the solution structure of PCP1_{ybt}, an NRPS peptidyl carrier protein from Yersiniabactin Synthetase, and we have identified contacts between its linker residues and its folded core. We found that a region of the N-terminal linker, loop0, provides a six-fold increase in fold stability and influences the ps-ns time-scale dynamics of residues throughout the protein, particularly those directly preceding the post-translationally modified serine, which are conserved in sequence and structure. This region has been shown to be dynamic in other carrier proteins, and it is affected by post-translational modifications^[6]. Loop0 may act as an allosteric sensor between the N-terminal linker and the protein core, propagating molecular events affecting the core to linkers in a manner that ensures their remodeling during sequential NRPS domain interactions. These findings underline the importance of unstructured regions in multi-domain proteins, and highlight the role of protein dynamics in molecular communication.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This work was supported by the National Institute of Health, grant R01 GM 104257. We thank Drs. A. Ringel and M. Morgan for assistance with SEC-MALS measurements. We also thank the JHU Center for Molecular Biophysics for their assistance with CD experiments.

References

1. Marahiel MA, Stachelhaus T, Mootz HD. Chem Rev. 1997; 97:2651–2674. [PubMed: 11851476]

2. Frueh D, Arthanari H, Koglin A, Vosburg D. *Nature*. 2008; 454doi: 10.1038/nature07162
3. Strieker M, Tanovi A, Marahiel MA. *Curr Opin Struct Biol*. 2010; 20:234–240. [PubMed: 20153164]
4. Drake EJ, Miller BR, Shi C, Tarrasch JT, Sundlov JA, Leigh Allen C, Skiniotis G, Aldrich CC, Gulick AM. *Nature*. 2016; 529:235–238. [PubMed: 26762461]
5. Reimer JM, Aloise MN, Harrison PM, Schmeing TM. *Nature*. 2016; 529:239–242. [PubMed: 26762462]
6. Goodrich AC, Harden BJ, Frueh DP. *J Am Chem Soc*. 2015; 137:12100–12109. [PubMed: 26334259]
7. Gehring AM, DeMoll E, Fetherston JD, Mori I, Mayhew GF, Blattner FR, Walsh CT, Perry RD. *Chem Biol*. 1998; 5:573–86. [PubMed: 9818149]
8. Miller BR, Gulick AM. *Methods Mol Biol*. 2016; 1401:3–29. [PubMed: 26831698]
9. Lipari G, Szabo A. *J Am Chem Soc*. 1982; 104:4546–4559.
10. Berlin K, Longhini A, Dayie TK, Fushman D. *J Biomol NMR*. 2013; 57:333–352. [PubMed: 24170368]
11. Yu D, Xu F, Gage D, Zhan J. *Chem Commun*. 2013; 49:6176–6178.
12. Miller BR, Sundlov JA, Drake EJ, Makin TA, Gulick AM. *Proteins*. 2014:1–12.
13. Vallurupalli P, Bouvignies G, Kay LE. *J Am Chem Soc*. 2012; 134:8148–8161. [PubMed: 22554188]
14. Lim J, Xiao T, Fan J, Yang D. *Angew Chemie – Int Ed*. 2014; 53:2358–2361.
15. Palmer AG. *J Magn Reson*. 2014; 241:3–17. [PubMed: 24656076]
16. Bae SH, Dyson HJ, Wright PE. *J Am Chem Soc*. 2009; 131:6814–6821. [PubMed: 19391622]
17. Sundlov JA, Shi C, Wilson DJ, Aldrich CC, Gulick AM. *Chem Biol*. 2012; 19:188–198. [PubMed: 22365602]
18. Weber T, Marahiel MA. *Structure*. 2001; 9:3–9.

Figure 1.

Structure and dynamics of PCP1_{ybt} including flanking linker regions (1381–1491). A) NMR bundle displaying residues 1400 to 1478. B) Alternative view, also showing the disordered N- and C-termini residues in linkers. C) 3D representation of PCP1_{ybt} dynamics. A thicker width indicates a lower order parameter and increased ps-ns dynamics. The orientation and range of residues are as in A. D) Order parameters used in C and secondary structure. Colors in A, C, D are consistent and help define the constructs used in this work. Blue: 1406–1482 (protein core and emerging linkers), beige: 1402–1405 (contact region of loop0) and black: disordered regions in linkers.

Figure 2.

Chemical shift perturbations between the full length (1383–1491) and truncated (1406–1482) constructs of PCP1_{ybt}. The perturbations are largely limited to contact points between loop0 and the core, which therefore maintains the same fold. Ω is defined in the supporting information. Orientation as in 1A and 1C.

Figure 3.

CEST reveals unfolded states. A,B) CEST profiles of residue E1429 using nominal B_1 field strengths of 12.5 Hz (red) and 25 Hz (blue) for the full-length (A) and short (B) constructs. C) Comparison of predicted and experimental changes in ^{15}N backbone amide chemical shifts between the unfolded state ω_U and the folded state ω_F . D) Comparison of the ^{15}N backbone amide chemical shifts of the unfolded state ω_U in the full length and short constructs. The same unfolded state is observed in both constructs.

Figure 4.

Modulation of fast (ps-ns) dynamics in PCP1_{ybt} upon contact with loop0. A thicker width denotes a larger absolute change in S^2 ($|\Delta S^2|$) between constructs when truncating loop0, up to a cap defined by the dashed line. The residues of loop0 that are removed are represented by the beige region in the secondary structure diagram.