Biophysical Chemistry 159 (2011) 152–161

Contents lists available at ScienceDirect

Biophysical Chemistry

journal homepage: http://www.elsevier.com/locate/biophyschem

Deletion of internal structured repeats increases the stability of a leucine-rich repeat protein, YopM

Ellen F. Vieux [a], Doug Barrick [b,*]

[a] Department of Biology, Massachusetts Institute of Technology, 77 Massachusetts Ave., Cambridge, MA 02139
[b] T.C. Jenkins Department of Biophysics, Johns Hopkins University, 3400 N. Charles St., Baltimore, MD 21218, USA

ARTICLE INFO

Article history:
Received 1 April 2011
Received in revised form 7 June 2011
Accepted 7 June 2011
Available online 22 June 2011

Keywords:
LRR
Leucine-rich repeat
Cooperativity
Interface stability
Repeat protein
YopM

ABSTRACT

Mapping the stability distributions of proteins in their native folded states provides a critical link between structure, thermodynamics, and function. Linear repeat proteins have proven more amenable to this kind of mapping than globular proteins. C-terminal deletion studies of YopM, a large, linear leucine-rich repeat (LRR) protein, show that stability is distributed quite heterogeneously, yet a high level of cooperativity is maintained [1]. Key components of this distribution are three interfaces that strongly stabilize adjacent sequences, thereby maintaining structural integrity and promoting cooperativity.

To better understand the distribution of interaction energy around these critical interfaces, we studied internal (rather than terminal) deletions of three LRRs in this region, including one of these stabilizing interfaces. Contrary to our expectation that deletion of structured repeats should be destabilizing, we find that internal deletion of folded repeats can actually *stabilize* the native state, suggesting that these repeats are *destabilizing*, although paradoxically, they are folded in the native state. We identified two residues within this destabilizing segment that deviate from the consensus sequence at a position that normally forms a stacked leucine ladder in the hydrophobic core. Replacement of these nonconsensus residues with leucine is stabilizing. This stability enhancement can be reproduced in the context of nonnative interfaces, but it requires an extended hydrophobic core. Our results demonstrate that different LRRs vary widely in their contribution to stability, and that this variation is context-dependent. These two factors are likely to determine the types of rearrangements that lead to folded, functional proteins, and in turn, are likely to restrict the pathways available for the evolution of linear repeat proteins.

1. Introduction

Determining the contribution of discrete structural units to the overall stability of globular proteins poses a significant challenge. Elongated repeat proteins have simpler linear architectures than globular proteins, allowing the stability contribution of individual supersecondary structural units to be determined through removal and addition of whole repeats [2–4] (Fig. 1). If the removal of repeating units from either end ("terminal" deletion, Fig. 1A) does not disrupt the structure of adjacent repeats, the stability change from deletion provides an estimate of the thermodynamic stability of the deleted repeat, including the single interface it forms with the neighboring repeat [2,3,5]. If the removal of repeating units from the interior ("internal" deletion, Fig. 1B) results in formation of a new interface, the stability change from deletion provides an estimate of the thermodynamic stability of the deleted repeat and the two interfaces it forms with the N- and C-terminal neighboring repeats, offset by the contribution from the new interface. In the Notch ankyrin repeat domain [3], terminal deletions are modestly destabilizing, whereas internal deletions are strongly destabilizing [6].

In *Yersinia pestis* outer protein M (YopM), a repeat protein containing 15.5 leucine-rich repeats (LRRs) from the bacterial subfamily (Fig. 2A), the removal of the C-terminal β-strand cap results in the partial unfolding of adjacent repeats [1]. Further C-terminal deletions demonstrate that internal LRRs 11 and 6 are necessary for structural integrity in adjacent LRRs [1]. YopM sequences from different species and strains of *Yersinia* contain duplications and deletions of LRRs. As a result, these YopM sequences vary in length, from 13 to 21 LRRs [7]. This variability in LRR number indicates that some interfaces are interchangeable in LRR proteins. This finding is consistent with the prevalence of expansion of LRR proteins through internal duplication [8], and with the conservation of terminal regions in variable lymphocyte receptors, which undergo somatic recombination to produce LRR arrays of variable length and composition [9]. However, given the tight packing interactions seen crystallographically between the LRRs of YopM from *Y. pestis* strain KIM (Fig. 2B) [10], such duplications and/or deletions are likely to be selected for stabilizing (or against grossly destabilizing) interfaces.

Here we have created internal deletion constructs of YopM from *Y. pestis* strain KIM to study the relative stability of repeats and their

* Corresponding author. Tel.: +1 410 516 0409.
E-mail address: barrick@jhu.edu (D. Barrick).

0301-4622/$ – see front matter © 2011 Elsevier B.V. All rights reserved.
doi:10.1016/j.bpc.2011.06.004

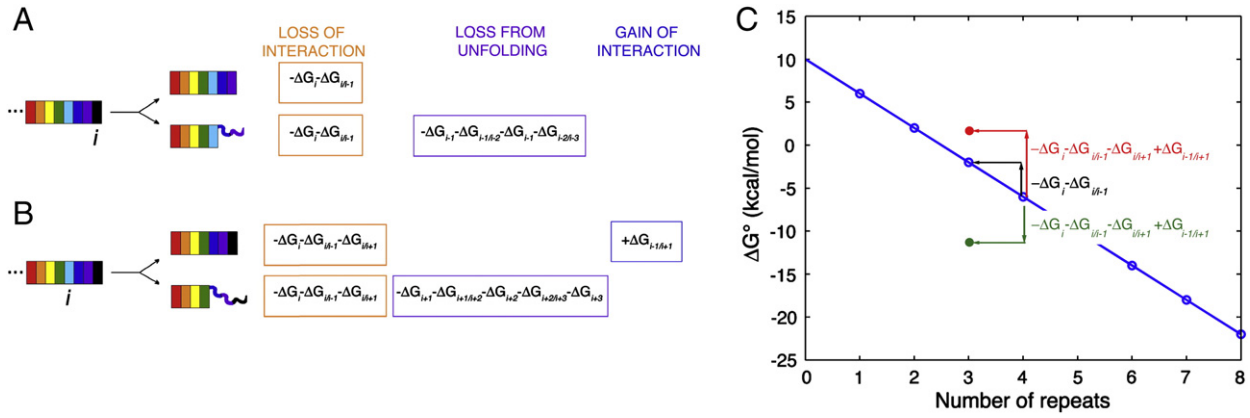
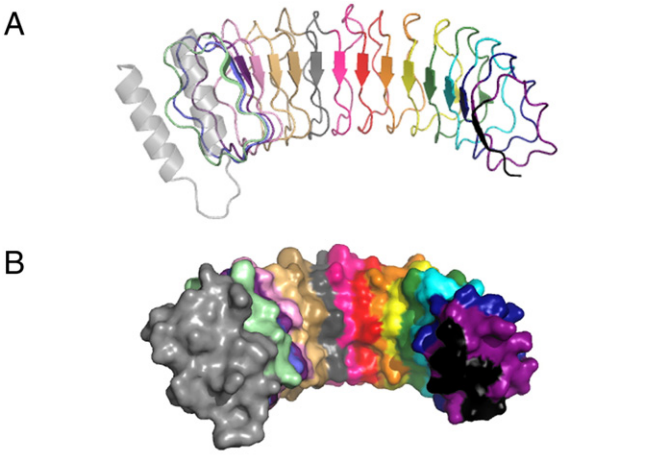


Fig. 1. Thermodynamics of deleting terminal versus internal repeats. (A, B) Schematic diagram comparing (A) terminal versus (B) internal deletion of whole repeats. Changes in folding free energy are on the right, with intrinsic and interfacial terms designated as ΔG_i and $\Delta G_{i/i+1}$, respectively. Energy terms resulting directly from loss of interactions in the fully folded state are listed in orange; terms resulting from partial unfolding (lower lines, A and B) are listed in purple; and the term resulting from forming new non-adjacent interaction ($\Delta G_{i-1/i+1}$) is listed in blue (internal deletion, no partial unfolding). (C) Folding free energy is depicted as a function of repeat number. For the purpose of illustration, each repeat is assumed to have the same intrinsic and nearest-neighbor free energy (blue line). For each terminally deleted repeat, the free energy change is equal to the lost intrinsic ($-\Delta G_i$) and native interfacial energies ($-\Delta G_{i/i+1}$), as indicated with the black arrow. For internal deletion, if the new non-native interface ($\Delta G_{i-1/i+1}$) is less stable than the interface that is lost, the resulting construct is less stable than that created by terminal deletion (red arrow). To get a stability *enhancement* from internal deletion (green arrow), the stability of the new interface ($\Delta G_{i-1/i+1}$) must exceed the sum of the stabilities of all three of the interactions lost (ΔG_i, $\Delta G°_{i/i-1}$, and $\Delta G°_{i/i+1}$).



C

Segment name	Sequence	Segment terminal residue number
LRR 1	HELELNNLGLSSLPEL--PPHL	93
LRR 2	ESLVASCNSLTELPEL--PQSL	113
LRR 3	KSLLVDNNNLKALSDL--PPLL	133
LRR 4	EYLGVSNNQLEKLPELQNSSFL	155
LRR 5	KIIDVDNNSLKKLPDL--PPSL	175
LRR 6	EFIAAGNNQLEELPELQNLPFL	197
LRR 7	TAIYADNNSLKKLPDL--PLSL	217
LRR 8	ESIVAGNNILEELPELQNLPFL	239
LRR 9	TTIYADNNLLKTLPDL--PPSL	259
LRR 10	EALNVRDNYLTDLPEL--PQSL	279
LRR 11	TFLDVSENIFSGLSEL--PPNL	299
LRR 12	YYLNASSNEIRSLCDL--PPSL	319
LRR 13	EELNVSNNKLIELPAL--PPRL	339
LRR 14	ERLIASFNHLAEVPEL--PQNL	359
LRR 15	KQLHVEYNPLREFPDI--PESV	379
β-cap	EDLRMN	385

Fig. 2. Three-dimensional surface and sequence of LRRs in YopM. (A) Ribbon representation of YopM. (B) Surface representation of N- and C-terminal caps and LRRs of YopM (PDB ID: 1JL5 [10]). From left to right the different colors represent different structural units: α-helical cap (gray), LRR 1 (light green), LRR 2 (light blue), LRR 3 (light purple), LRR 4 (light pink), LRR 5 (tan), LRR 6 (brown), LRR 7 (gray), LRR 8 (pink), LRR 9 (red), LRR 10 (orange), LRR 11 (yellow), LRR 12 (green), LRR 13 (teal), LRR 14 (blue), LRR 15 (purple), and β-strand cap (black). (C) Sequence of LRRs in YopM; residues are shaded from light to dark with increasing identity. The solvent accessible surface area in (B) was generated using default settings in Pymol [31].

adjoining interfaces, and to delineate rules that restrict deletion and duplication events leading to natural diversity in LRR arrays. We evaluated the ability of new interfaces to retain structure by circular dichroism (CD) spectroscopy and limited proteolysis, and obtained thermodynamic information from urea-induced unfolding transitions. We find the effects of different internal deletions to be quite variable. Surprisingly, the internal deletion of repeats 11 and 12 actually *increases* stability. To probe the sequence determinants underlying this stability enhancement, substitutions of two non-consensus residues were made in these two repeats (Fig. 2C). These consensus point-substitutions show stability enhancements similar to internal deletion, suggesting that the interaction of these non-consensus residues with their neighboring repeats contributes to the unexpected stability enhancement of internal repeat deletion. Moreover, stability changes resulting from these substitutions underscore the importance of the stacked hydrophobic core for stability, interface compatibility, and for the cooperative folding of YopM.

The broader goals of this research are to contribute to understanding the extent to which biomolecular structure can be resolved into thermodynamic components, to discover the fundamental rules that connect macromolecular structure to thermodynamics (both through generalities and exceptions), and to use these findings to gain insight into biological function, adaptation, and evolution. Answering these broader questions has been a major intellectual driving force for the Gibbs Conferences in Biothermodynamics over the last 25 years. Based on the substantial progress made over the last quarter century by the research groups the Gibbs conference comprises, we expect that answers to these questions will soon be in hand.

2. Results

We have previously mapped the stability contributions of LRRs of YopM by deleting repeats one at a time from the C-terminus [1]. This study showed the stability contributions of individual LRRs to be highly heterogeneous. For example, C-terminal deletion of LRR 13 (as judged by comparing a construct from the N-cap to repeat 13 [N-13] with a construct spanning from the N-cap to repeat 12 [N-12]) has no effect on stability. In contrast, C-terminal deletion of LRR 12 (comparing N-12 with N-11) modestly decreases stability, and C-terminal deletion of LRR 11 (comparing N-11 with N-10) dramatically decreases stability (and results in partial unfolding of adjacent repeats [1]). Here we create internal deletion constructs to probe the relative

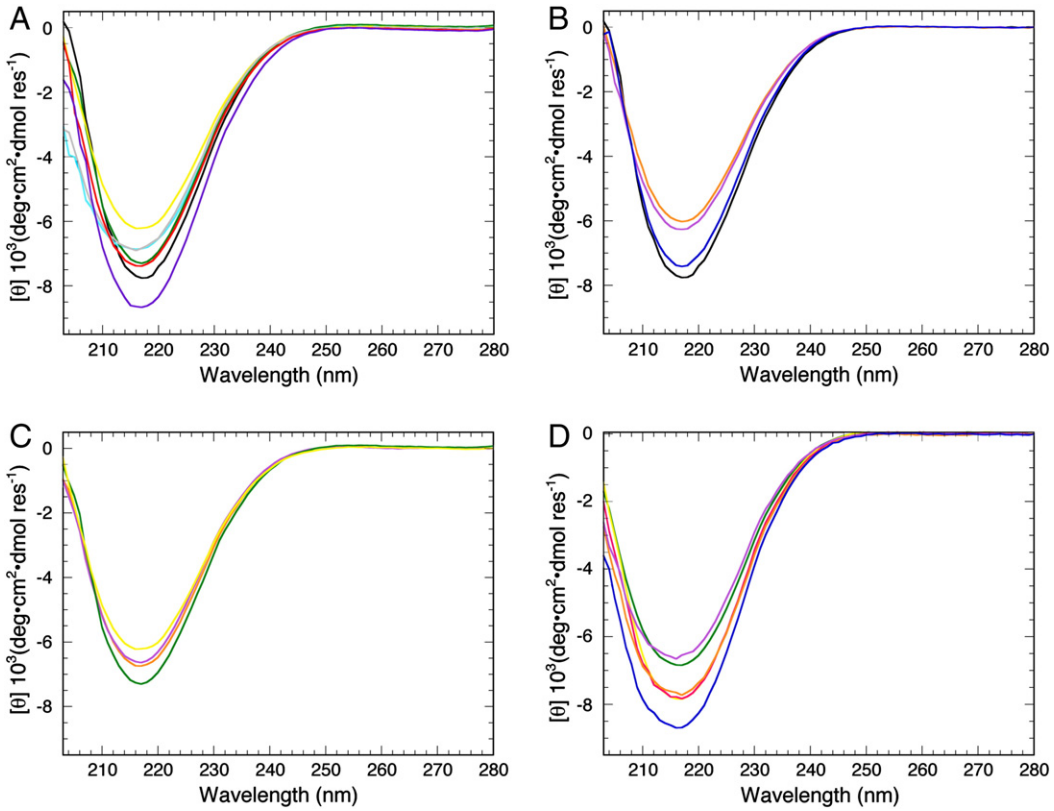


Fig. 3. Far-UV CD spectra of internal deletion and consensus substitution constructs. (A) Far-UV CD spectra of internal deletions: full-length YopM (black), Δ13 (aqua), Δ12 (dark green), Δ11 (yellow), Δ12Δ13 (gray), Δ11Δ12 (red), Δ11Δ12Δ13 (purple). (B) Far-UV CD spectra of consensus substitutions: full-length YopM (black), F289L (orange), I309L (light purple), F289L/I309L (blue). (C) Far-UV CD spectra of consensus substitutions in background of internal deletions: Δ12 (dark green), Δ11 (yellow), Δ12 F289L (orange), Δ11 I309L (light purple). (D) Far-UV CD spectra of consensus substitutions in the background of C-terminal deletion constructs: N-11 (yellow), N-11 F289L (pink), N-12 (dark green), N-12 F289L (orange), N-12 I309L (light purple), N-12 F289L/I309L (blue).

interface stability between LRRs. Single LRRs were deleted internally to create constructs Δ11 (deletion of LRR 11), Δ12 (deletion of LRR 12), and Δ13 (deletion of LRR 13). In addition, adjacent LRRs were deleted internally to create constructs Δ12Δ13 (deletion of LRRs 12 and 13), Δ11Δ12 (deletion of LRRs 11 and 12), and Δ11Δ12Δ13 (deletion of LRRs 11, 12, and 13).

Each single-repeat internal deletion construct removes two interfaces (e.g., Δ11 removes the interface between LRR 10 and 11, and the interface between LRR 11 and 12). Each two-repeat internal deletion removes three interfaces (e.g., Δ11Δ12 removes the interfaces between LRRs 10 and 11, 11 and 12, and between 12 and 13). In addition, internal deletion has the potential to either form a new, non-native interface, or cause partial unfolding involving adjacent repeats. Thus, internal deletion can be viewed as a larger perturbation than terminal deletion, probing all of the interactions affected by terminal deletion, along with several new interactions (Fig. 1).

2.1. Structural consequences of internal deletion

To monitor the secondary structure formation of the internal deletion constructs, far-UV CD spectra were obtained (Fig. 3A). As with full-length YopM, all internal deletion constructs have a single ellipticity minimum near 217 nm, characteristic of proteins containing β-strand secondary structure [11], thus, we conclude that none of these deletions result in large-scale unfolding. Moreover, CD spectra of internal deletion constructs Δ11, Δ12, Δ11Δ12, and Δ11Δ12Δ13 have nearly identical shapes, consistent with a similar secondary structure content for these constructs. In contrast, spectra of Δ13 and Δ12Δ13 have broader signals around the minima, consistent with some unstructured polypeptide [11], suggesting that deletion of repeat 13 may be structurally disruptive, rather than supporting non-native interface formation. Although Δ11 also displays lower ellipticity, its high sensitivity to urea and resistance to proteolysis (see below) suggest that it retains structure at a level near to that for wild-type YopM.[1]

2.2. Information available from equilibrium unfolding studies of repeat proteins

The two parameters that are most directly accessible from the urea-induced protein unfolding transitions are the transition midpoint (C_m) and the steepness of the transition, traditionally represented by the m-value ($d\Delta G^\circ/d[\text{urea}]$) [12–16]. Provided the unfolding transition conforms to a two-state mechanism, these two parameters can be combined to determine the free energy of folding in the absence of denaturant (ΔG°_{H2O}). Although ΔG°_{H2O} provides a fundamental representation of the thermodynamics of unfolding, it is inappropriate for multistate transitions.[2] Here we use both ΔG°_{H2O} and C_m to evaluate stability changes, emphasizing the former when the data support a two-state mechanism, and the latter when data indicate a more complex mechanism. Although C_m is not a direct measure of folding free energy, it provides a quantitative index of structural stability near to the unfolding transition.

[1] We attribute much of the variation in magnitude of the CD signal to uncertainty in concentration determination. Accurate concentration determination of YopM constructs is particularly difficult owing to its net negative charge (42 glu and asp residues, 17 arg and lys residues). This strong negative charge makes it difficult to purify away from trace nucleic acids, which have a high extinction coefficient.

[2] In addition to inadequately representing the complexity of a multistate mechanism, using a two-state ΔG°_{H2O} parameter is likely to underestimate conformational stability, owing to the large extrapolations that are required.

One way to assess the extent to which equilibrium unfolding adheres to a two-state mechanism is to compare the *m*-value to that of other proteins of similar size. For wild-type YopM, which adheres to an equilibrium two-state mechanism, the *m*-value obtained for urea-induced unfolding [17] matches that predicted from chain length [14]. Since our deletion constructs are significantly shorter than the parent YopM construct, we expect the *m*-value to decrease linearly with the number of repeats deleted (black line, Fig. 5). Additional *m*-value decreases beyond this chain-length effect would suggest either a multistate unfolding mechanism, in which the native state is converted to a partly folded intermediate at moderate urea concentration, or that the structure of the native state is disrupted, resulting in a partly folded state in the absence of denaturant. In such a partly folded state, unfolding would be expected between the site of the deletion and the terminus (in this case the C-terminus, since we have previously shown N-terminal fragments to be folded [1]).

For constructs with conspicuously low *m*-values, the spectroscopic test helps to resolve whether unfolding proceeds from the native state by a multistate folding mechanism or whether unfolding proceeds from a partly folded state by a two-state reaction mechanism. Deviation of local and global spectroscopic probes indicates a multistate unfolding mechanism, whereas coincidence suggests two-state unfolding of a partly disordered native state. In addition, limited proteolysis can help in interpreting the origins of decreased *m*-values (see below). Constructs that are partly unfolded under native conditions should display increased susceptibility to proteolysis. We have previously combined limited proteolysis with mass spectrometry to show that structurally disruptive perturbations in the C-terminus of YopM increase proteolytic susceptibility, and that enhanced proteolysis sites map to the C-terminus [1].

2.3. Thermodynamic consequences of internal deletion

To determine the effects of internal deletions on stability, urea-induced unfolding transitions were monitored by far-UV CD (a probe of secondary structure formation over the entire molecule) and tryptophan fluorescence (a probe of the environment surrounding the two tryptophans in the N-terminal α-helical cap) [17]. For all internal deletion constructs, full sigmoidal unfolding transitions were observed. CD and fluorescence transitions for these deletion constructs are coincident, consistent with a simple two-state (all or none) equilibrium unfolding mechanism. Each of these transitions was well fitted by a two-state unfolding model in which the free energy of folding depends linearly on urea concentration (Fig. 4A).

The C_m values for Δ12 and Δ13 decrease substantially compared to full-length YopM (Fig. 4A, Table 1), suggesting that internal deletion of these repeats decreases stability. Given the low apparent stability contribution of these repeats suggested from C-terminal deletion studies [1], this decrease in C_m is somewhat surprising, and emphasizes the contribution of interfaces to the folding of LRRs 12 and 13. For Δ13, the *m*-value also decreases substantially (Table 1), and both the C_m and *m*-values match those of a partially unfolded C-terminal deletion construct N-15 (Table 1; [1]). This decrease suggests that internal deletion of LRR 13 may disrupt LRRs 14, 15, and the β-strand cap, resulting in partial structure formation through LRR 12. In

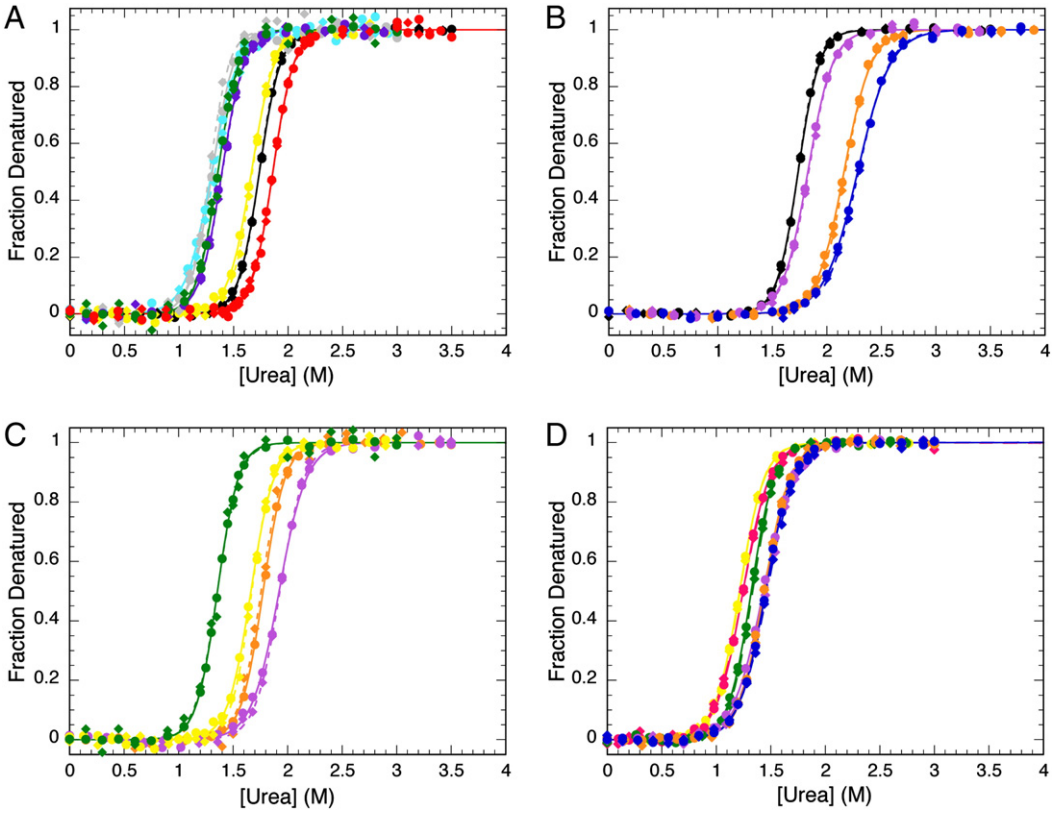


Fig. 4. Urea-induced unfolding transitions of internal deletion and consensus substitution constructs. Normalized urea-induced unfolding transitions monitored by far-UV CD (filled circles, solid lines) and tryptophan fluorescence (filled diamonds, dashed lines). (A) Internal deletions: full-length YopM (solid black line), Δ13 (aqua), Δ12 (dark green), Δ11 (yellow), Δ12Δ13 (gray), Δ11Δ12 (red), Δ11Δ12Δ13 (purple). (B) Consensus substitutions: full-length YopM (solid black line), F289L (orange), I309L (light purple), F289L/I309L (blue). (C) Consensus substitutions in the background of internal deletion constructs: Δ12 (dark green), Δ11 (yellow), Δ12 F289L (orange), Δ11 I309L (light purple), (D) Consensus substitutions in the background of C-terminal deletion constructs: N-11 (yellow), N-11 F289L (pink), N-12 (dark green), N-12 F289L (orange), N-12 I309L (light purple), N-12 F289L/I309L (blue). Lines are the result of fitting a two-state unfolding model to individual transitions.

Table 1
Thermodynamic parameters for YopM internal deletion and consensus substitution constructs.

		ΔG°_{H2O} (kcal·mol^{-1})	m (kcal·mol^{-1}·M^{-1})	C_m (M)
Full-length YopM		−10.7 ± 0.2	6.2 ± 0.03	1.73 ± 0.02
	F[a]	−11.0 ± 0.2	6.3 ± 0.1	1.74 ± 0.03
Internal deletions				
Δ13		−6.7 ± 0.2	5.0 ± 0.1	1.34 ± 0.02
Δ12		−8.3 ± 0.3	6.1 ± 0.3	1.36 ± 0.02
Δ11		−10.5 ± 0.3	6.0 ± 0.2	1.76 ± 0.04
Δ12Δ13		−7.6 ± 0.1	5.6 ± 0.1	1.37 ± 0.02
Δ11Δ12		−11.1 ± 0.2	6.0 ± 0.1	1.86 ± 0.01
Δ11Δ12Δ13		−7.9 ± 0.2	5.5 ± 0.2	1.43 ± 0.02
N-15		−7.3 ± 0.1	5.2 ± 0.1	1.39 ± 0.02
Consensus substitutions				
Full-length[b]				
F289L		−10.0 ± 0.2	4.7 ± 0.04	2.12 ± 0.02
F289L	F[a]	−10.7 ± 0.2	5.0 ± 0.03	2.13 ± 0.02
I309L		−10.5 ± 0.5	5.9 ± 0.3	1.79 ± 0.02
F289L/I309L		−8.7 ± 0.2	3.7 ± 0.1	2.34 ± 0.06
F289L/I309L	F[a]	−9.4 ± 0.2	4.0 ± 0.1	2.33 ± 0.04
Internal Deletions[b]				
Δ11 I309L		−9.8 ± 0.2	5.2 ± 0.1	1.89 ± 0.01
Δ12 F289L		−9.7 ± 0.4	5.6 ± 0.2	1.73 ± 0.02
Terminal deletions[b]				
N-11		−6.5 ± 0.1	5.2 ± 0.2	1.25 ± 0.03
N-11 F289L		−6.1 ± 0.1	4.8 ± 0.1	1.26 ± 0.01
N-12		−7.5 ± 0.1	5.5 ± 0.1	1.35 ± 0.02
N-12 F289L		−7.1 ± 0.02	5.0 ± 0.1	1.41 ± 0.02
N-12 I309L		−6.3 ± 0.1	4.4 ± 0.1	1.43 ± 0.01
N-12 F289L/I309L		−7.2 ± 0.1	5.0 ± 0.1	1.44 ± 0.02

Parameters obtained from urea-induced unfolding transitions monitored by CD, unless indicated by an F. Parameters are averages of three or more independent urea-induced unfolding transitions; uncertainties represent the standard error on the mean. Conditions: 200 mM NaCl, 20 mM Tris–HCl pH 7.6, 0.01 mM TCEP.
[a] When indicated by an F, unfolding transitions were monitored by fluorescence.
[b] Indicates the background in which consensus substitutions are made.

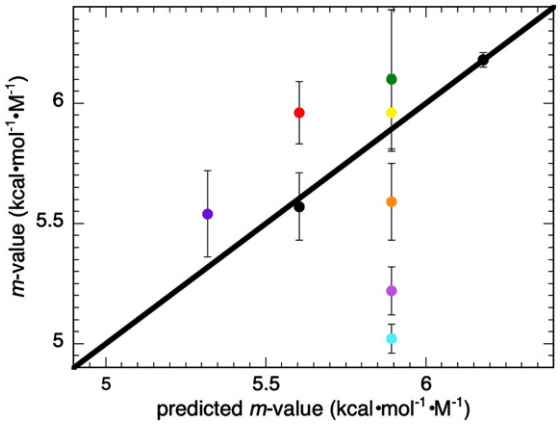


Fig. 5. Comparison of measured and predicted m-values. Predicted m-values for deletion constructs were calculated by scaling the experimentally determined full-length YopM m-value based on relative chain length. Δ13 (aqua), Δ12 (dark green), Δ11 (yellow), Δ12Δ13 (gray), Δ11Δ12 (red), Δ11Δ12Δ13 (purple), Δ12 F289L (orange), Δ11 I309L (light purple), and full-length YopM (black). The black line is of unit slope and zero intercept, and indicates the expected m-value for a two-state transition spanning the entire construct.

contrast, although the C_m decreases for Δ12, the m-value remains high, suggesting that the C-terminus remains fully structured.

Construct Δ12Δ13 shows a decrease in C_m similar to Δ13 (Fig. 4A; Tables 1 and S-1). The m-value of Δ12Δ13 is decreased compared to Δ12 and wild-type YopM, but is larger than the m-value of Δ13. The modest decrease in m-value for the Δ12Δ13 construct is consistent with that expected from deletion of two folded repeats from an otherwise native state (Fig. 5). These results suggest that LRR11 can form a non-native interface with LRR13 (based on results for Δ12) and LRR14 (based on results from Δ12Δ13), although this comes at a significant thermodynamic cost ($\Delta\Delta G^{\circ} = +2.4$ and $+3.1$ kcal/mol, respectively).

Unlike Δ12 and Δ13, Δ11 has a C_m value similar to full-length YopM (Fig. 4A). This apparent insensitivity of stability to internal deletion of LRR 11 is unexpected, based on the destabilizing effect of the removing LRR 11 from the C-terminus [1]. Δ11 shows only a slight decrease in m-value (Table 1), which matches that expected for deletion of 20 residues (Fig. 5). These observations suggest LRRs 10 and 12 can form a non-native interface that is stable enough to drive folding of the C-terminal repeats with a very modest cost to stability.

Surprisingly the simultaneous removal of LRRs 11 and 12 (Δ11Δ12) results in an increase in C_m, compared to both single deletions (Δ11 and Δ12) and to full-length YopM (Fig. 4A). This shift in C_m is consistent with an *increase* in stability resulting from internal deletion. Compared to wild-type YopM, the C_m of Δ11Δ12 increases by 0.13 M, corresponding to a decrease in the folding free energy (ΔG°_{H2O}) of 0.4 kcal mol^{-1}. An even more striking way to represent this stability enhancement is as the effect of internal deletion of LRR 11 in a background in which LRR 12 is missing. Compared to Δ12, the C_m of Δ11Δ12 increases by 0.5 M urea, corresponding to a decrease in ΔG°_{H2O} of 2.8 kcal mol^{-1}. Both of these constructs have identical m-values, consistent with a large cooperative unit and a two-state folding mechanism. A similar, albeit somewhat less dramatic effect can be seen when LRR 11 is deleted in a construct that is missing both LRRs 12 and 13. Taken together these observations may reflect the formation of highly stabilizing non-native interfaces between LRR 10 and more C-terminal LRRs (13 and 14), which are of greater stability than the native interface between LRRs 11 and 12.

2.4. Structural and thermodynamic consequences of consensus substitutions in the stacked hydrophobic core

To explore sequence determinants that underlie the relative instability of the interface between LRRs 11 and 12 (and contribute to stability increase upon deletion of these repeats), we compared the conservation of residues at internal positions within each LRR (Fig. 2C). Numerous studies have shown that consensus sequences in α-helical repeat proteins are highly stabilizing [5,18–20]. Position 10 deviates uniquely in LRRs 11 and 12 (F289 and I309, respectively, Fig. 2C) from the otherwise completely conserved leucine. To test whether these sequence variations destabilize the interfaces involving LRRs 11 and 12, substitutions that restore consensus at these sites (F289L and I309L) were made in full-length YopM.

Far-UV CD spectra for all consensus substitutions are similar, showing a single ellipticity minimum at 217 nm (Fig. 3B), indicating that the secondary structures of F289L, I309L, and F289L/I309L remain intact. Moreover, the urea-induced unfolding transitions of these three single-residue variants are well-fitted by a two-state (all-or-none) model for unfolding (Fig. 4B). The C_m values increase, indicating that indeed, these two non-consensus residues in LRRs 11 and 12 significantly destabilize wild-type YopM. The m-value of I309L is close to that of wild-type, consistent with a highly cooperative transition (Table 1). However, the m-value of F289L decrease is significantly lower than that of wild-type. Moreover, the m-value decreases further upon combination with I309L, suggesting a multistate unfolding mechanism (structural disruption in the native state would be inconsistent with increased C_m values. For both F289L and F289L/I309L, the fitted free energies of folding and m-values are

larger for transitions monitored by fluorescence than those monitored by CD (Table 1). The data indicate that single substitutions disrupt the high degree of coupling observed in YopM [17].

2.5. Structural and thermodynamic consequences of consensus substitutions in internal deletion constructs

If deviations from consensus have a general destabilizing effect on interfaces, the F289L and I309L consensus substitutions should increase the stability of internal deletion constructs by strengthening the non-native interfaces, as they do in wild-type YopM. To test whether the stabilizing effect of consensus substitution is generic, we examined the effects of the F289L and I309L consensus substitutions in combination with single internal deletions, creating Δ11 I309L and Δ12 F289L. Gross secondary structure is maintained for both constructs, as judged by far-UV CD spectra (Fig. 3C). Both Δ11 I309L and Δ12 F289L show stabilization by the single consensus substitutions (Fig. 4C), based on increased C_m values compared to internal deletion constructs (Table 1). As in full-length YopM, consensus substitutions in Δ11 and Δ12 decrease the *m*-value, suggesting a loss of cooperativity.

2.6. Structural and thermodynamic consequences of consensus point substitutions in C-terminal deletion constructs

We further explored the sequence determinants of interface formation and the impact of non-consensus residues on the intrinsic stability of LRRs by making consensus substitutions in C-terminal deletion constructs N-11 and N-12 [1]. Gross secondary structure is maintained for all constructs, as judged by far-UV CD spectra (Fig. 3D). The urea-induced unfolding transition of N-11 F289L is nearly identical to N-11 (Fig. 4D; Table 1). This observation indicates that neither LRR 11 nor its interface with LRR 10 are stabilized by the consensus substitution F289L, and that C-terminal repeats are required for stabilization by this consensus point substitution. Although this same substitution is modestly stabilizing in the N-12 C-terminal deletion construct, this effect is much smaller, based on C_m shift, than the same substitution in wild-type YopM (Fig. 4B vs. D, Table 1). This observation suggests that the interface between LRRs 11 and 12 is not as well formed in N-12 as in wild-type YopM (consistent with an earlier C-terminal deletion study [1]). In contrast, the I309L consensus point substitution stabilizes the N-12 terminal deletion construct to a similar extent to wild-type YopM, although for both substitutions the effect is modest (Fig. 4D, Table 1).

2.7. Limited proteolysis of internal deletion and consensus substitution constructs

YopM is resistant to trypsin proteolysis when the structural integrity of the C-terminal LRRs is maintained [1]. We used limited proteolysis to probe the retention of structure in internal deletion and consensus substitution constructs. After 180 seconds of exposure to trypsin, most constructs show a modest decrease in mass (compared to the untreated control; Fig. 6A). This decrease results from the removal of the N-terminal histidine tag and N-terminal leader sequence (Fig. 6A, B [1]). Constructs Δ13 and Δ12Δ13 show a larger decrease in mass (Fig. 6A and B). The major proteolytic product in these constructs is close in size to the 34 kDa trypsin proteolysis products with a partially unfolded C terminus [1]. These results indicate that the C-terminal repeats (LRRs 12–15) have increased lability, consistent with the large destabilization seen for these constructs.

The Δ12 construct develops a faint band at this size, consistent with the intermediate destabilization seen for this construct. Moreover, this band is suppressed by combination of F289L with the Δ12 deletion construct. This indicates that an interface may form between LRRs 11 and 13, but is weaker than the interface between LRRs 10 and 13, due in part to the nonconsensus Ile at position 289. Based on resistance to proteolysis, all other interfaces appear to be stable, most notably those resulting from internal deletion of LRR 11, which is neutral to stabilizing, depending on the context (Fig. 6).

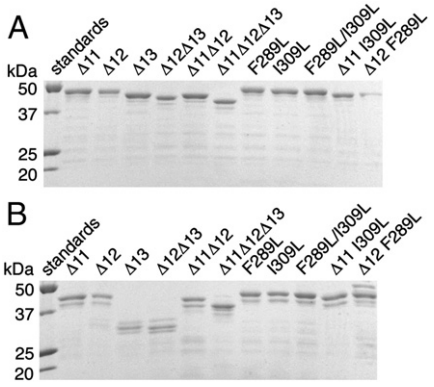


Fig. 6. Limited trypsin proteolysis of internal deletion and consensus substitution constructs. (A) SDS-PAGE of untreated constructs, and (B) reacted with trypsin for 180 seconds at 25 °C.

3. Discussion

Duplications and deletions of whole repeats occur in many repeat proteins, and are implicated both as a mechanism for protein evolution and as a way to obtain complexity and sequence variation [8,21]. *YopM* is located on a virulence plasmid that is highly conserved among *Yersinia* species (97.5% average pairwise identity among the other six *yop* genes [7]). However, YopM shows variation in length, ranging from 13 to 21 LRRs, making it significantly more variable than its surroundings. Here we find that specific interactions in the stacked hydrophobic core influence interfacial stability substantially, such that in some circumstances, non-native interfaces appear to be more stabilizing than native interfaces. This observation suggests that, based on sequence variation from repeat to repeat, there should be energetically favored routes to generate genetic variation by whole-repeat insertion/deletion events.

3.1. Internal deletion of repeats can be destabilizing

Based on studies of terminal deletions in the Notch ankyrin domain [3] and YopM [1], and internal deletions in the Notch ankyrin domain [6], internal deletions in YopM are expected to be destabilizing. In addition to a potential decrease in stability from loss of favorable intrinsic (local) interactions, stability is expected to decrease from the replacement of stable native interfaces with a non-native interfaces. If a new non-native interface is of sufficiently low stability (high free energy), the repeats adjacent to the site of deletion may unfold rather than forming a structured interface.

The above results show that when LRRs 12 and 13 are internally deleted (Δ12, Δ13, and Δ12Δ13) the C_m decreases. Moreover, when LRR 13 is deleted, the *m*-values decrease substantially, indicating that the size of the cooperative unit is smaller than would be expected for the loss of one repeat (Table 1, Fig. 5), suggesting that the adjacent C-terminal LRRs do not retain structure. Partial unfolding of C-terminal LRRs is supported by broadening of the far-UV CD spectrum (Fig. 3A), and by increased susceptibility to proteolysis (Fig. 6). These observations suggest that non-native interfaces of LRR 12 and LRR 14 are not stable enough to drive structure formation at the C terminus (Fig. 7). In contrast, when LRR 12 is deleted on its own (Δ12), the *m*-value remains relatively high, suggesting that the adjacent C-terminal LRRs retain structure in this construct. Thus, although deletion of LRR 12 decreases stability based on C_m, the

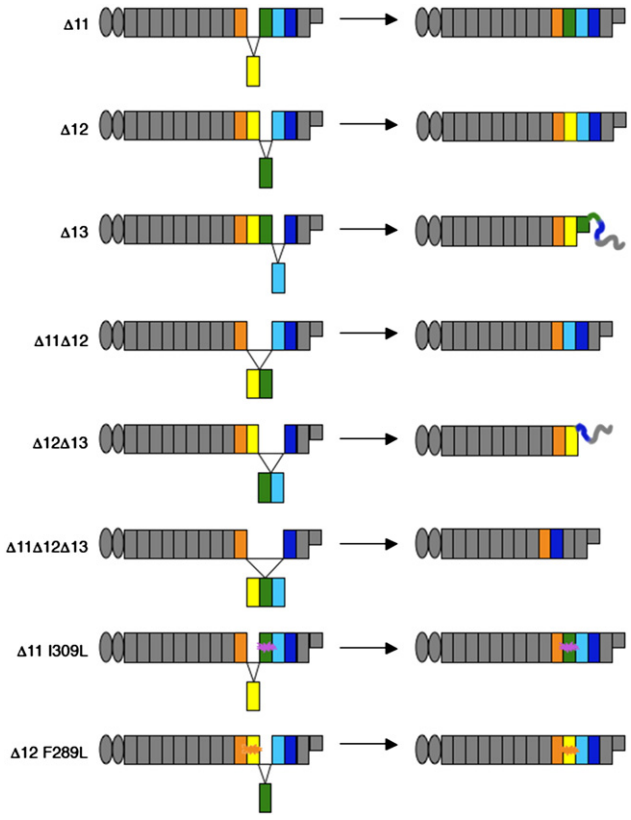


Fig. 7. Structural interpretation of internal deletions and consensus substitutions. Internal deletions of LRR 11, LRR 12, LRRs 11–12, and LRRs 11–13 create new interfaces between previously nonadjacent repeats, and retain structure through the C-terminus. Internal deletion of LRR 13 results in the unfolding of the C-terminus, unless deleted simultaneously with LRR 11. The N-terminus remains folded in all constructs. LRR 10 (orange), LRR 11 (yellow), LRR 12 (green), LRR 13 (aqua), LRR 14 (blue), F289L (orange), and I309L (purple).

resulting non-native interface between LRRs 11 and 13 appears to be stable enough to drive folding of C-terminal repeats (Fig. 7).

3.2. Internal deletion of repeats can be neutral

Results here show that internal deletion of the single LRR 11 from wild-type YopM has no effect on stability. The C_m does not change when LRR 11 is deleted internally (construct Δ11), compared to full-length YopM (Fig. 4A), and the cooperative unit remains relatively large, based on its *m*-value (Table 1 and Fig. 5). Adjacent structure is maintained, as judged by limited trypsin proteolysis (Fig. 6). This finding is unexpected, given that LRR 11 appears to be strongly stabilizing in the C-terminal deletion format [1]. The net neutrality of internally deleting LRR 11 implies that the formation of the non-native interface between LRRs 10 and 12 can fully compensate for the loss of native interfaces between LRRs 10 and 11 and between LRRs 11 and 12, and for the intrinsic folding of LRR 11. Although these latter components may be relatively unstable, they must be stable enough to maintain a single, cooperatively folded unit in YopM that spans LRR 11.

3.3. Internal deletion of repeats can be stabilizing

Perhaps the most surprising stability change we have identified with internal deletion is the *increase* in stability that results from deletion of LRRs 11 and 12. The unfolding transition of Δ11Δ12 retains a high level of cooperativity, based on its *m*-value, which is higher than expected given the removal of 40 residues (Table 1, Fig. 5). The structure of C-terminal LRRs is maintained in Δ11Δ12 (Fig. 7), as shown by the far-UV CD (Fig. 3A) and protection from trypsin proteolysis (Fig. 6). These observations demonstrate that a highly stabilizing non-native interface forms (between LRRs 10 and 13) that more than compensates for the removal of three interfaces (between LRRs 10 and 11, 11 and 12, and 12 and 13), and loss of intrinsic folding units (LRRs 11 and 12). Enhancement of stability as a result of internal deletion of LRR 11 is even more dramatically demonstrated using a construct missing LRR 12 as a background (Δ12 versus Δ11Δ12): deletion of LRR 11 increases stability by almost 3 kcal mol^{-1} (Fig. 4, Table 1).

The observation that a folded structural element destabilizes its context is rather counterintuitive. The second law of thermodynamics maintains that processes that increase free energy do not occur spontaneously. For example, a ligand that binds specifically to a folded protein structure can (and will) preferentially stabilize that structure (decreasing overall free energy), whereas ligands that would destabilize a folded protein (increasing free energy) in a bound state simply do not bind (they are repelled, compared to the bulk phase). Following this line of reasoning, it might be expected that the free energy of full-length YopM would be minimized by local unfolding of a "destabilizing" element of structure such as LRR 11–12, allowing the stable interface between LRR 10 and 13 to form. However, this would require the ends of the unfolded segment spanning LRRs 11 and 12 to be close together. In addition to resulting in significant steric repulsion, the formation of a closed loop in a 40 residue segment would result in a significant entropic destabilization. We propose that in wild-type YopM, the energy penalty for forming this looped structure must exceed the enhanced stabilization provided by the non-native interface between LRRs 10 and 13.

3.4. Two non-consensus residues contribute to the relative instability of LRRs 11 and 12

Interfaces formed between LRRs are critical for thermodynamic cooperativity and structural integrity in YopM [1]. The internal deletions described here suggest that the interface formed between LRRs 11 and 12 is less stable than the other interfaces. Sequence analysis of LRRs 11 and 12 shows a deviation from consensus at position 10. Leucine is strongly preferred at this position in the bacterial LRR subfamily to which YopM belongs, forming a regular stacked array in the hydrophobic core in other repeats.

Single site consensus substitutions in LRR11 and 12 increase stability, based on increases in C_m values (0.39 ± 0.03 M and 0.06 ± 0.03 M units, respectively). Combining the two consensus substitutions (F289L/I309L) further increases stability, with an increase in C_m (0.62 ± 0.07 M units) that exceeds the sum expected from both single substitutions. Based on C_m values alone, these stability increases are as large as those produced by deletion of LRRs 11 and 12 (0.13 ± 0.03 M units). However, *m*-values of the two variants that include the F289L substitution in LRR11 are decreased relative to wild-type YopM (Table 1; this effect is most pronounced for the combined substitutions F289L/I309L). Comparisons of fluorescence and CD-monitored unfolding transitions are consistent with the formation of intermediates through the unfolding transition region, rather than with the partial unfolding of the native state. This interpretation is supported by trypsin digestion (Fig. 6). This result calls into question the necessity of the α-helical domain at the N terminus as a template for folding of the entire LRR domain [10], suggesting instead that if the C-terminus is sufficiently stabilized, it can remain folded even as the N-terminus begins to unfold.

Although it is not possible to demonstrate that the relative instability of repeats 11 and 12 is *entirely* due to the nonconsensus residues at position 10, two observations suggest a major role of the nonconsensus residues. First, we see a similar level of stabilization when consensus point substitutions are made in a non-native context,

i.e., Δ12 F289L, and Δ11 I309L, indicating the stabilizing effect of consensus substitution does not depend on local sequence details. Second, taking F298L/I309L as a full-length background with a better match to the LRR consensus, internal deletion of LRRs 11 and 12, both singly and in combination, significantly decreases the C_m by −0.45 to −0.6 M units (Table 1), as would be expected from internal deletion of stably folded repeats.

3.5. Interface stability is highly dependent on an extended hydrophobic core

Combining consensus substitutions in LRRs 11 and 12 demonstrates that consensus stabilization is independent of the identity (and thus sequence) of the surrounding LRRs. The C_m increases for both Δ11 I309L and Δ12 F289L (compared to wild-type YopM) are equivalent to the sums of the increases in C_m values of the single internal deletions and consensus substitutions (Table 1). Thus, the stabilizing effect of F289L and I309L substitution appears to be independent of the specific sequence context of LRRs 12 and 11, respectively. This conclusion is consistent with an increased protection from trypsin digestion for Δ12 F289L, compared to Δ12 (Fig. 6).

However, consensus stabilization appears to require both N- and C-terminal flanking repeats. In N-11, F289L shows only a slight increase of 0.02 ± 0.01 M, compared to the increase of 0.39 ± 0.03 M units in full-length YopM, indicating that the stabilization associated with F289L requires a C-terminal interface. In N-12, F289L is slightly more stabilizing, with an increase in C_m of 0.06 ± 0.09 M, although again this is significantly less than the increase seen in full-length YopM (~0.4 M). I309L shows a similar, modest increase in C_m (0.08 M) in N-12. When F289L and I309L are combined in N-12, the same modest increase in C_m is produced (0.09 M) as for F289L alone in N-12, in contrast to the strong nonadditive effect of these two substitutions on the C_m of full-length YopM. Thus, it appears the increase in stability requires C-terminal repeats to form buried interfaces, providing an extended packed hydrophobic core. Although the N-12 construct provides a repeat (LRR12) C-terminal to the F289L substitution, C-terminal deletion studies revealed partial unfolding of LRR 12.

3.6. Implications for evolution

The *yopM* genes of various *Yersinia* strains range in repeat number from 13 to 21 LRRs. This level of variation is significantly higher than that of the other *yop* genes on the virulence plasmids. Based on sequence comparisons of other types of repeat proteins, through duplication of whole (or multiple) repeats seems to be a facile and common mechanism for generating diversity [8,21–23]. Presumably, such expansion mechanisms are balanced by internal deletions to loosely maintain overall length and stability.

The *yopM* gene appears to be a critical component for virulence [24]. It gains access to intracellular compartments of the host cell, where it interacts with two protein kinases [25]. As all known protein–protein interactions involving LRR proteins use the *folded* state, we assume YopM must remain folded to function as a virulence factor. Thus, our findings that some whole-repeat rearrangements are structurally disruptive but others are actually *stabilizing* suggest that sequence-specific thermodynamic factors restrict which whole-repeat rearrangements will be viable.

To get clues as to how this variation may have arisen, given our findings regarding repeat deletion and insertion, we created a phylogenetic tree describing a plausible evolutionary relationship between five strains of *Yersinia* (Fig. 8). Distances between strains were determined using sequence analysis of the entire virulence plasmids, and duplication, insertion, and deletion events were reconstructed based on sequence similarity between LRRs, and by minimizing the number of recombination events needed to generate these five YopM sequences. Our reconstruction does not consider horizontal gene transfer.

Within the *Y. pestis* species (Fig. 8, left), strain 91001 has only 13 LRRs compared to 15 found in strains KIM and CO92. In our proposed reconstruction, deletion of LRRS 7 and 8 created a new interface between LRRs 6 and 9. More striking is the closely related *Yersinia pseudotuberculosis* IP32953 *yopM gene, which encodes* 21 LRRs. This could have arisen from the deletion of LRRs 8 and 9, and two duplications of the segment containing LRRs 4, 5, 6, and 7, and would have created new interfaces between repeats 7 and 10 (by deletion) and between repeats 7 and 4 (by duplication).

Based on comparison of virulence plasmids, the *Yersinia enterocolitica* strains are more distantly related to *Y. pestis* strains than is *Y. pseudotuberculosis* (Fig. 8, right). Whereas the *yopM* gene of *Y. enterocolitica* strain 8081 is similar to the proposed common ancestor, *Y. enterocolitica* strain A127/90 contains 20 LRRs, which could have arisen from three duplications of the segment containing LRRs 12, 13, and 14, along with a deletion of LRRs 5 and 6. This rearrangement would have created new interfaces between repeats 14 and 12 (by duplication), and between repeats 4 and 9.

Sequence comparison shows that the LRR on the N-terminal side of all of the new interfaces depicted in Fig. 8 retain the consensus leucine at position 10. Interestingly, in the two YopMs of greater length (*Y. pseudotuberculosis* IP32953 and *Y. enterocolitica* strain A127/90), the position 10 isoleucine in LRR 12 is retained, but the phenylalanine in the flanking LRR 11 has been replaced by an isoleucine and a leucine, respectively. With the increased length and higher stability, some sequence variation may have been tolerated. This natural variation has no known effect on the virulence of the bacteria [7]. It would be interesting to test if these proteins maintain a cooperative folding mechanism.

4. Materials and methods

4.1. Sub-cloning, protein expression, and purification

Internal deletion constructs of *yopM* in the pET15b expression vector (Novagen, Madison, WI) were created using a sub-cloning strategy previously described by Noriega et al. [26]. Consensus substitutions were created using a Quikchange kit (Stratagene, La Jolla, CA). All polypeptides were expressed in *Escherichia coli* (BL21 [DE3]) overnight at 20 °C in TB [27], and were purified using a nickel-NTA (Qiagen, Valencia, CA) or HisTrap HP column on an AKTA purifier (GE Healthcare Bio-Sciences Corp, Piscataway, NJ). Fractions containing YopM were dialyzed to 20 mM NaCl, 20 mM Tris–HCl pH 7.6, 2 mM EDTA, 0.1 mM DTT, and were run on a HiTrap Q ion exchange column. YopM-containing fractions of greater than 95% purity, as judged by SDS-PAGE, were dialyzed into storage buffer (200 mM NaCl, 20 mM Tris–HCl pH 7.6, 0.1 mM Tris(2-Carboxyethyl)-phosphine (TCEP)), concentrated using a VivaSpin 500 microconcentrator (Sartorius; Aubagne, France), passed through an 0.22 μm filter, flash-frozen, and stored at −80 °C. Protein concentrations were determined as described by Edelhoch [28].

4.2. CD spectroscopy and urea-induced unfolding transitions

Far-UV CD spectra were collected on an Aviv 62A DS spectropolarimeter (Lakewood, NJ). Far-UV CD spectra were collected in 0.1 cm cuvettes with protein concentrations of 10–20 μM in storage buffer at 25 °C. Spectra were obtained by signal averaging for 20 seconds every 1.0 nm.

Urea-induced unfolding transitions were monitored by CD at 217 nm and by tryptophan fluorescence using an excitation wavelength of 280 nm with an emission filter at 320 nm on Aviv 62A DS and 410 CD spectropolarimeters. Urea purchased from Amresco (Solon, OH) was stirred with mixed-bed resin (Bio-Rad; Hercules,

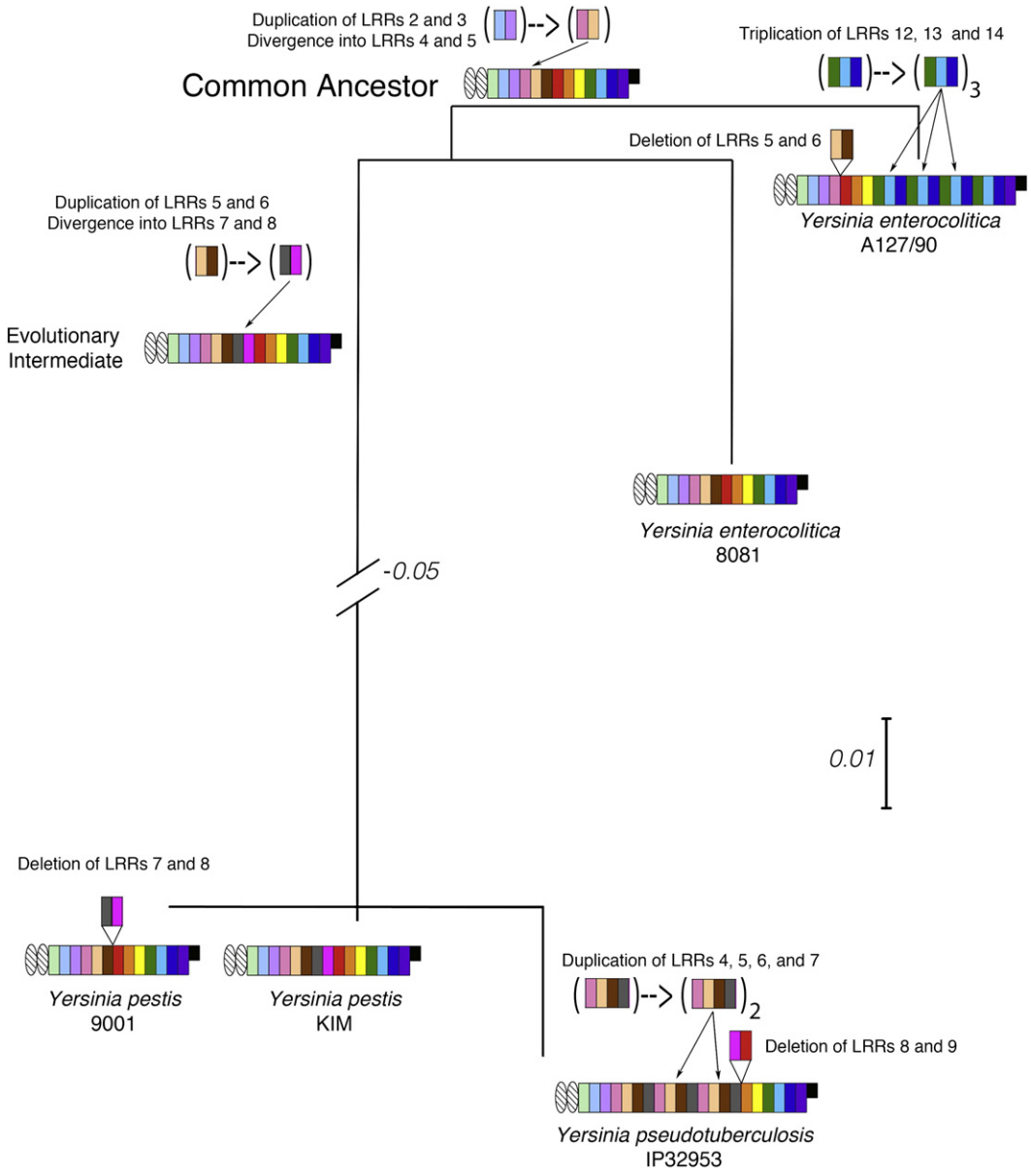


Fig. 8. A potential phylogenetic relationship between *Yersinia* species, and likely LRR duplication and deletion events in the *yopM* gene. The relationship between LRRs was determined using BLASTP [32] to align each repeat against the *Y. pestis* KIM YopM sequence. Relatedness of aligned repeats was evaluated from the default substitution matrix (BLOSUM 62), and was combined with phylogenetic relationships among *Yersinia* strains (determined by analysis of virulence plasmid sequences using the program MUAVE) to map out a parsimonious history of duplication and deletion events. The phylogenetic guide tree was constructed using the phylogram setting in SplitsTrees 4 [33].

CA) as described in Street et al. [29]. Urea concentration was determined by refractometry [30]. Urea titrations were carried out automatically using a computer controlled Microlab syringe titrator (Hamilton Company, Reno, NV) to deliver buffered urea solutions (containing protein) into a buffered protein sample. At each urea concentration, samples were equilibrated for 5 minutes at 25 °C, and CD and fluorescence signals were averaged for 30 seconds. For each titration, protein stocks were diluted with fresh storage buffer to a concentration of 1–3.2 μM.

4.3. Two-state analysis of equilibrium urea-induced unfolding transitions

Equilibrium urea-induced unfolding transitions were related to the equilibrium constant for unfolding using a two-state model, assuming a population-weighted average of the signal of the native (Y_N) and denatured states (Y_D):

$$Y_{obs} = f_N Y_N + f_D Y_D = \left(\frac{K_f}{1+K_f}\right) Y_N + \left(\frac{1}{1+K_f}\right) Y_D \quad (1)$$

where f_N and f_D are the fraction of native and denatured protein, and K_f represents the equilibrium constant for unfolding. Y_N and Y_D are assumed to have linear dependence on urea and are represented as $Y_N = a_N + b_N[\text{urea}]$ and $Y_D = a_D + b_D[\text{urea}]$. The equilibrium constant for unfolding is related to the reaction free energy by the standard formula:

$$K_u = exp\left(-\frac{\Delta G_u^\circ}{RT}\right) \quad (2)$$

The two-state model was fitted to data from CD and fluorescence signals by combining Eqs. (1) and (2):

$$Y_{obs} = \frac{(a_N + b_N[\text{urea}]) + (a_D + b_D[\text{urea}])\exp\left(-\Delta G_u^\circ / RT\right)}{1 + \exp\left(-\Delta G_u^\circ / RT\right)} \quad (3)$$

Thermodynamic parameters were estimated using the linear extrapolation method [13,30], in which the free energy of unfolding varies linearly with urea concentration[13,15,30]:

$$\Delta G_u^\circ = \Delta G_{H_2O}^\circ + m[\text{urea}] \quad (4)$$

Eq. (4) was inserted into Eq. (3), and was then fitted to unfolding transitions using the nonlinear least-squares tool of Kaleidagraph 3.4 (Synergy Software, Reading, PA).

4.4. Limited trypsin proteolysis

YopM internal deletion constructs and variants with consensus substitutions were diluted with storage buffer to 10 μM. Trypsin (Sigma-Aldrich, St. Louis, MO) was dissolved in storage buffer to a final concentration of 0.96 μM as determined gravimetrically. Reactions were incubated at room temperature in storage buffer. Aliquots of samples were removed at various time points and quenched with a final concentration of 25 mM PMSF in 25% isopropanol. All samples were immediately flash frozen and stored at −80 °C. Digests were electrophoresed on pre-cast Tris–HCl gels (10–20% acrylamide; Bio-Rad, Hercules, CA).

Acknowledgments

This work was stimulated by many conversations that DB and EV have had with members of the Gibbs community, both at numerous annual meetings at Touch of Nature, and in conversations carried on throughout the years. Some of the data above was presented as preliminary results by EV at a student talk at Gibbs 2007, and ensuing conversations with Wayne Bolen, Bertrand Garcia Moreno, Terry Oas, Rohit Pappu, Patricia Clark, Chiwook Park, and Vince Hilser helped us accept the idea that stabilization through deletion of folded segments of structure is not in violation of the laws of thermodynamics. We thank members of the Barrick lab for helpful comments and scientific discussion throughout the work. We thank David Waugh for providing a YopM expression vector. This work was supported by NIH grant GM 068462 to DB. EV was supported by an NIH doctoral training grant (2T32 GM008403) from NIGMS.

References

[1] E. Kloss, D. Barrick, C-terminal deletion of leucine-rich repeats from YopM reveals a heterogeneous distribution of stability in a cooperatively folded protein, Protein Sci. 18 (9) (2009) 1948–1960.
[2] E. Kloss, N. Courtemanche, D. Barrick, Repeat-protein folding: new insights into origins of cooperativity, stability, and topology, Arch. Biochem. Biophys. 469 (1) (2008) 83–99.
[3] C.C. Mello, D. Barrick, An experimentally determined protein folding energy landscape, Proc. Natl. Acad. Sci. U. S. A. 101 (39) (2004) 14102–14107.
[4] E.R. Main, et al., Local and long-range stability in tandemly arrayed tetratricopeptide repeats, Proc. Natl. Acad. Sci. U. S. A. 102 (16) (2005) 5721–5726.
[5] T. Kajander, et al., A new folding paradigm for repeat proteins, J. Am. Chem. Soc. 127 (29) (2005) 10188–10190.
[6] K.W. Tripp, D. Barrick, The tolerance of a modular protein to duplication and deletion of internal repeats, J. Mol. Biol. 344 (1) (2004) 169–178.
[7] A. Boland, S. Havaux, G.R. Cornelis, Heterogeneity of the Yersinia YopM protein, Microb. Pathog. 25 (6) (1998) 343–348.
[8] A.K. Bjorklund, D. Ekman, A. Elofsson, Expansion of protein domain repeats, PLoS Comput. Biol. 2 (8) (2006) e114.
[9] Z. Pancer, et al., Somatic diversification of variable lymphocyte receptors in the agnathan sea lamprey, Nature 430 (6996) (2004) 174–180.
[10] A.G. Evdokimov, et al., Unusual molecular architecture of the *Yersinia pestis* cytotoxin YopM: a leucine-rich repeat protein with the shortest repeating unit, J. Mol. Biol. 312 (4) (2001) 807–821.
[11] W.C. Johnson Jr., Protein secondary structure and circular dichroism: a practical guide, Proteins 7 (3) (1990) 205–214.
[12] D.W. Bolen, M.M. Santoro, Unfolding free energy changes determined by the linear extrapolation method. 2. Incorporation of delta G degrees N-U values in a thermodynamic cycle, Biochemistry 27 (21) (1988) 8069–8074.
[13] R.F. Greene Jr., C.N. Pace, Urea and guanidine hydrochloride denaturation of ribonuclease, lysozyme, alpha-chymotrypsin, and beta-lactoglobulin, J. Biol. Chem. 249 (17) (1974) 5388–5393.
[14] J.K. Myers, C.N. Pace, J.M. Scholtz, Denaturant m values and heat capacity changes: relation to changes in accessible surface areas of protein unfolding, Protein Sci. 4 (10) (1995) 2138–2148.
[15] M.M. Santoro, D.W. Bolen, Unfolding free energy changes determined by the linear extrapolation method. 1. Unfolding of phenylmethanesulfonyl alpha-chymotrypsin using different denaturants, Biochemistry 27 (21) (1988) 8063–8068.
[16] J.A. Schellman, Fifty years of solvent denaturation, Biophys. Chem. 96 (2–3) (2002) 91–101.
[17] E. Kloss, D. Barrick, Thermodynamics, kinetics, and salt dependence of folding of YopM, a large leucine-rich repeat protein, J. Mol. Biol. 383 (5) (2008) 1195–1209.
[18] K.W. Tripp, D. Barrick, Enhancing the stability and folding rate of a repeat protein through the addition of consensus repeats, J. Mol. Biol. 365 (4) (2007) 1187–1200.
[19] S.K. Wetzel, et al., Folding and unfolding mechanism of highly stable full-consensus ankyrin repeat proteins, J. Mol. Biol. 376 (1) (2008) 241–257.
[20] V.S. Devi, et al., Folding of a designed simple ankyrin repeat protein, Protein Sci. 13 (11) (2004) 2864–2870.
[21] T.O. Street, G.D. Rose, D. Barrick, The role of introns in repeat protein gene formation, J. Mol. Biol. 360 (2) (2006) 258–266.
[22] G. Apic, J. Gough, S.A. Teichmann, Domain combinations in archaeal, eubacterial and eukaryotic proteomes, J. Mol. Biol. 310 (2) (2001) 311–325.
[23] N. Takahashi, Y. Yoko Takahashi, F.W. Putnam, Periodicity of leucine and tandem repetition of a 24-amino acid segment in the primary structure of leucine-rich alpha2-glycoprotein of human serum, Proc. Natl. Acad. Sci. U. S. A. 82 (1985) 1906–1910.
[24] K. Trulzsch, et al., Contribution of the major secreted yops of *Yersinia enterocolitica* O: 8 to pathogenicity in the mouse infection model, Infect. Immun. 72 (9) (2004) 5227–5234.
[25] J.B. McPhee, P. Mena, J.B. Bliska, Delineation of regions of the Yersinia YopM protein required for interaction with the RSK1 and PRK2 host kinases and their requirement for interleukin-10 production and virulence, Infect. Immun. 78 (8) (2010) 3529–3539.
[26] F.R. Noriega, et al., Construction and characterization of attenuated delta aroA delta virG *Shigella flexneri* 2a strain CVD 1203, a prototype live oral vaccine, Infect. Immun. 62 (11) (1994) 5168–5172.
[27] Tartoff, Hobbs, Bethesda Research Laboratories Focus, 9(12), 1987.
[28] H. Edelhoch, Spectroscopic determination of tryptophan and tyrosine in proteins, Biochemistry 6 (7) (1967) 1948–1954.
[29] T.O. Street, N. Courtemanche, D. Barrick, Protein folding and stability using denaturants, Methods Cell. Biol. 84 (2008) 295–325.
[30] C.N. Pace, Determination and analysis of urea and guanidine hydrochloride denaturation curves, Methods Enzymol. 131 (1986) 266–280.
[31] W.L. De Lano, The PyMOL Molecular Graphics System, DeLano Scientific, Palo Alto, California, USA, 2008.
[32] S.F. Altschul, et al., Basic local alignment search tool, J. Mol. Biol. 215 (3) (1990) 403–410.
[33] D.H. Huson, D. Bryant, Application of phylogenetic networks in evolutionary studies, Mol. Biol. Evol. 23 (2) (2006) 254–267.