

Interaction of the Disordered *Yersinia* Effector Protein YopE with Its Cognate Chaperone SycE[†]

Xin Hu,[‡] Michael S. Lee,^{‡,§} and Anders Wallqvist^{*,‡}

[‡]*Biotechnology HPC Software Applications Institute, Telemedicine and Advanced Technology Research Center, US Army Medical Research and Materiel Command, Fort Detrick, Maryland 21702 and* [§]*Computational Sciences and Engineering Branch, US Army Research Laboratory, Aberdeen Proving Ground, Maryland 21005*

Received October 7, 2009; Revised Manuscript Received October 29, 2009

ABSTRACT: We describe an efficient approach to model the binding interaction of the disordered effector protein to its cognate chaperone in the type III secretion system (T3SS). Starting from de novo models, we generated ensembles of unfolded conformations of the *Yersinia* effector YopE using REMD simulations and docked them to the chaperone SycE using a multistep protein docking strategy. The predicted YopE/SycE complex was in good agreement with the experimental structure. The ability of our computational protocol to mimic the structural transition upon chaperone binding opens up the possibility of studying the underlying specificity of chaperone/effector interactions and devising strategies for interfering with T3SS transport.

The type III secretion system (T3SS) utilized by many Gram-negative bacteria plays an important role in the pathogenic invasion of the host cell by delivering effector proteins directly into the cytosol of the host (1, 2). The delivery mechanism depends on an initial association step of the effector protein with specific cognate chaperones in the bacterial cytosol (3). This association is characterized by a binding mode whereby the unfolded effector chaperone binding domain (CBD) wraps around its chaperone in an extended, nonglobular form (4). Though this binding is typically thought of as disordered, structural analysis of a number of effector proteins with secretion chaperones across many different species has revealed a common structural binding pattern, termed the β -motif (5). The nature of the binding process itself with the chaperone is typically postulated as an order-to-disorder transition since the CBD is initially required to unfold in order to effectively bind to the chaperone (6). An alternative disorder-to-order binding mechanism has been proposed that the native CBD exists in an intrinsically disordered state but transitions into an ordered state upon binding to its cognate chaperone (7).

The dynamics of the disordered effector protein and structural transitions associated with chaperone binding are challenging problems to study both experimentally and computationally (8). Though a large number of T3SS effector/chaperone complexes have been identified, only a few have been structurally determined so far (5). Here, we explore computational approaches to model the disordered effector proteins and predict the binding complex with their cognate chaperones. Our focus is on the

specific interactions that are required between the CBD of the *Yersinia* effector protein YopE (YopE_{CBD}) and its cognate chaperone SycE to form a T3SS-competent complex. The structures of the unbound chaperone SycE as well as the bound YopE_{CBD}/SycE complex have been experimentally solved (9, 10), whereas the free CBD of YopE (residues 20–78) is considered to be unstructured in solution (7). Given that the unbound effector is in a globular, yet partially unfolded state, is it possible to computationally sample the unfolded conformations sufficiently and obtain the near-native conformers suitable for complex prediction?

To address this issue, we generated three-dimensional structures of unbound YopE_{CBD} using the Rosetta de novo fragment assembly program (11). The structural models were refined by energy minimization, rescored with an implicit solvent model, and clustered using a pairwise hierarchical method with a structural similarity metric in order to identify a small set of unique models (12, 13). Analysis showed that the predicted solution structures of YopE_{CBD} were molten globules with a hydrophobic core. Molecular dynamics (MD) simulations indicated that the structure remained compact at room temperature. We also calculated the NOE intensities of these structural models. The results confirmed that these conformations were qualitatively compatible with the partially disordered structure indicated by the experimental NOE data (7) (Figure S1, Supporting Information).

To explore the unfolded conformational space of YopE_{CBD}, we performed replica exchange MD (REMD) simulations for the modeled structures in implicit solvent. The best predicted models with the lowest scores in the top two major clusters were selected as representatives for REMD. Ten replicas were used with temperature windows ranging from 270 to 600 K. The results showed that the globular conformation denatured above 400 K, resulting in a wide range of unfolded conformations (Figure S2, Supporting Information). These temperature-denatured structures constituted a large ensemble of unfolded conformations of YopE_{CBD} that likely include conformations similar to those known to bind the chaperone. To examine this hypothesis, we measured the root-mean square deviation (rmsd) compared to the experimental structure of SycE-bound YopE_{CBD} and the radius of gyration (R_g) of the denatured conformations (Figure 1). Remarkably, a substantial number of structural models obtained at higher temperatures were found to closely resemble the wrapped YopE_{CBD} conformation of the crystal structure.

The existence of near-native conformations of unfolded YopE_{CBD} from REMD indicated that temperature denaturation alone can produce conformations commensurate with the known

[†]This work was sponsored by the U.S. Department of Defense High Performance Computing Modernization Program (HPCMP), under the High Performance Computing Software Applications Institutes (HSAI) initiative.

^{*}Corresponding author. Phone: (301) 619-1989 Fax: (301) 619-1983. E-mail: awallqvist@bioanalysis.org.

chaperone-bound structure. These results suggest that REMD is a promising adjunct to protein–protein docking algorithms that can be used for predicting the interactions of the denatured effector protein binding to its chaperone (14). However, the probability of occurrence of the near-native conformations in the simulations was low (<0.1%). Given the large number of denatured conformations, the challenge remains as to whether we can discriminate near-native binding conformations from nonbinders.

Hence, we developed a stepwise hierarchical protein ensemble docking strategy to predict the disordered CBD/chaperone binding complex. The overall computational procedure is summarized in Figure 2 (for further details, see Supporting Information). First, we performed a structural clustering analysis to extract representative protein structures from the multiple REMD trajectories. The representative structures of each cluster

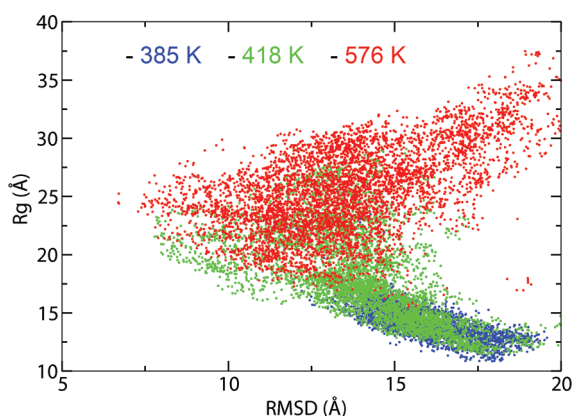


FIGURE 1: Scatter plot of the radius of gyration (R_g) vs rmsd for the unfolded conformations of YopE_{CBD} generated at different temperatures. The rmsd's (backbone) were calculated with respect to the X-ray structure of SycE-bound YopE_{CBD}.

were then docked to the unbound form of chaperone SycE using ZDOCK, followed by rescoring with ZRANK (15). To improve hit selection, the top ZRANK-scoring decoys were assessed on the basis of the presence of the conserved β -motif interaction (5). Two β -motif interacting residues (hotspots) from the CBD-binding pocket of chaperone SycE were selected (Leu15 and Val31 of SycE subunits A and B). The presence of an interaction between these residues and the YopE CBD in the docked complex was based on distance. Only docked decoys that have these interactions were retained. This hotspot filtering was required to identify the near-native binding conformations from the large pool of docked decoys because the scoring function alone was insufficient. Consequently, a number of decoys that lacked hotspot interactions were filtered out, leaving a smaller set of representative YopE_{CBD} clusters in the top hit list (Figure 2D).

In the second step, we selected all structural members associated with the identified top cluster representatives and docked them to the chaperone. A total of 350 ensemble conformations of YopE_{CBD} were docked to SycE using a more extensive searching by ZDOCK, followed by ZRANK rescoring and hotspot filtering. Inspection of the top decoys showed that the unfolded effector YopE_{CBD} bound to chaperone SycE in the same manner as that observed in the crystal structure. The top-scoring binding conformation of YopE_{CBD} had the lowest rmsd of 7.6 Å compared to that of the experimental structure. Interestingly, a number of binding complexes adopted a symmetric binding mode in which the effector wrapped the dimer chaperone from the back of the binding surface (Figure 2E). This binding model is essentially the same as the front binding of the experimental structure because the two SycE subunits A and B in the dimer are symmetric. In addition, a reverse binding mode was also found in which the effector wrapped around the same chaperone surface, but the N-terminus bound to the B subunit instead of the A subunit (Figure 2E). Binding free energy calculations indicated

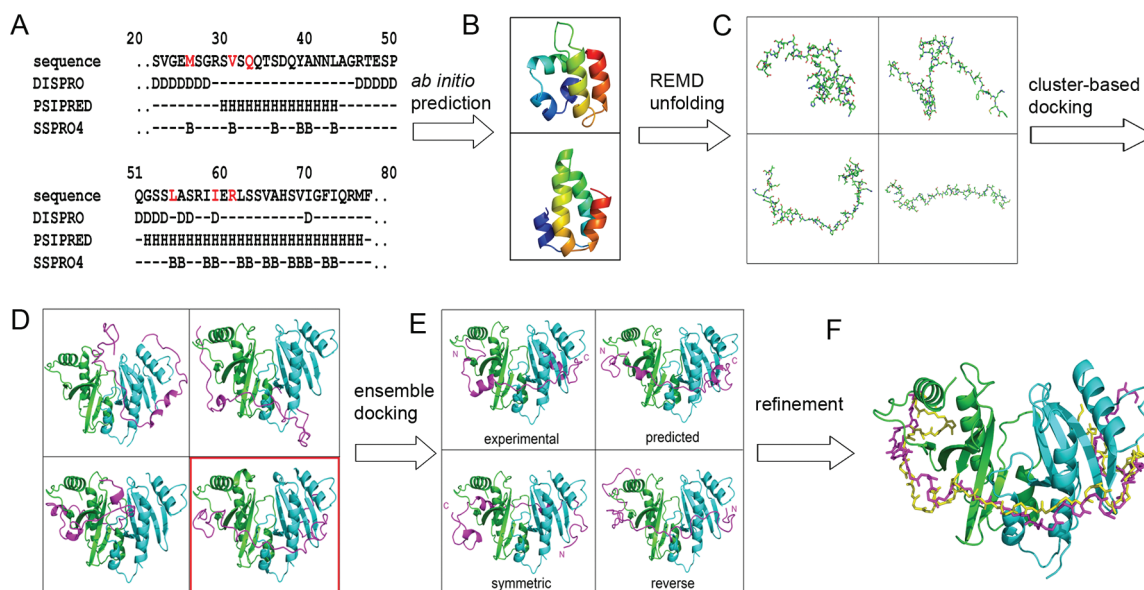


FIGURE 2: REMD-based protein ensemble docking approach for effector–chaperone binding prediction. (A) Sequence of the N-terminal chaperone binding domain (CBD) of *Yersinia* YopE and predicted secondary structure properties. (B) 3D structures of YopE_{CBD} from the ab initio prediction. (C) Unfolded conformation of YopE_{CBD} generated from REMD simulations. (D) Docked complexes of YopE_{CBD}/SycE obtained from cluster-based docking. The chaperone is shown as a ribbon in green (subunit A) and cyan (subunit B). The hit selected from hotspot filtering is shown in the red box. (E) Representative binding modes of YopE_{CBD}/SycE complexes obtained from ensemble docking. (F) Refined binding complex of YopE_{CBD}/SycE from MD simulations. The predicted conformation of YopE_{CBD} is shown in magenta, and the X-ray structure is shown in yellow.

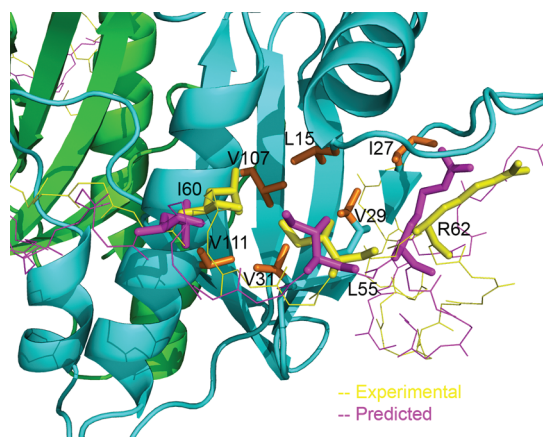


FIGURE 3: Interactions of the conserved β -motif of *Yersinia* effector YopE (stick representation in magenta and yellow) with its cognate chaperone SycE (ribbon in cyan and green). The binding residues of the chaperone hydrophobic pocket are shown as sticks in orange.

that this reverse binding interaction was less favored because of different binding interface properties across the chaperone (Figure S3, Supporting Information).

Finally, we refined the predicted binding complexes with MD simulations. This allowed for local side chain and backbone rearrangements to improve the binding interface contacts. The best predicted binding conformation of YopE_{CBD} was refined to a rmsd of ~ 5 Å (Figure 2F). The most significant variation was located at the terminal regions, which exhibited the largest fluctuations in the simulations. In fact, if the flexible termini were excluded, the binding conformation of YopE_{CBD} had a rmsd of 2 Å. Most importantly, Figure 3 shows that the conserved CBD binding motif interactions were well reproduced.

Our REMD-based conformational sampling and protein ensemble docking approach were able to mimic the disorder-to-order transition seen between an effector protein and its cognate chaperone. This transition requires that the disordered hydrophobic core of the compact CBD structure ruptures to form an extended contact state with the cognate chaperone, rendering the effector protein transportable through the T3SS. From the theory of conformational selection upon binding (16), it is likely that the free CBD of the effector protein exists as an ensemble of partially disordered conformations in dynamic equilibrium. The cognate chaperone binds preferentially to weakly populated, higher-energy conformations, shifting the equilibrium in favor of effector conformations conducive to complex formation. Exactly how this unfolding and binding process occurs in nature cannot be addressed here, but it is worth noting that the REMD-based temperature denaturation did generate the near-native conformations of effector protein that can be docked to the chaperone to closely reproduce the experimentally determined complex. An important concern is the general applicability of the proposed approach since naively, the possibility of an unfolded conformation from the REMD

simulation matching the bound conformation seems remote. Furthermore, the number of conformations that might need to be examined by protein–protein docking would appear to be incomprehensibly large. While extensive testing on a broad range of protein–protein complexes will definitively answer this question, there are two theoretical arguments which can provide some optimism. First, the relative free energy of the correct unfolded conformation cannot be larger than the compensating free energy due to complex formation. Second, the conformations actually sampled by an unfolding protein at low temperatures must be greatly reduced compared to the space of all possible conformations (i.e., nature’s solution to Levinthal’s paradox). Nevertheless, the REMD-based protein docking approach opens up the possibility of computationally mimicking and probing interactions between bacterial effector proteins and their cognate chaperones to understand their underlying specificity and to devise possible strategies for interfering with T3SS transport.

ACKNOWLEDGMENT

Computational time was provided by the U.S. Army Research Laboratory Department of Defense (DoD) Supercomputing Resource Center and the Maui High Performance Computing Center.

SUPPORTING INFORMATION AVAILABLE

Detailed computational methods, REMD simulations, and docking results. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES

- Galan, J. E., and Collmer, A. (1999) *Science* 284, 1322–1328.
- Cornelis, G. R., and Van Gijsegem, F. (2000) *Annu. Rev. Microbiol.* 54, 735–774.
- Matsumoto, H., and Young, G. M. (2009) *Curr. Opin. Microbiol.* 12, 94–100.
- Ghosh, P. (2004) *Microbiol. Mol. Biol. Rev.* 68, 771–795.
- Lilic, M., Vujanac, M., and Stebbins, C. E. (2006) *Mol. Cell* 21, 653–664.
- Stebbins, C. E., and Galan, J. E. (2001) *Nature* 414, 77–81.
- Rodgers, L., Gamez, A., Riek, R., and Ghosh, P. (2008) *J. Biol. Chem.* 283, 20857–20863.
- Thirumalai, D., and Lorimer, G. H. (2001) *Annu. Rev. Biophys. Biomol. Struct.* 30, 245–269.
- Birtalan, S., and Ghosh, P. (2001) *Nat. Struct. Biol.* 8, 974–978.
- Birtalan, S. C., Phillips, R. M., and Ghosh, P. (2002) *Mol. Cell* 9, 971–980.
- Simons, K. T., Kooperberg, C., Huang, E., and Baker, D. (1997) *J. Mol. Biol.* 268, 209–225.
- Feig, M., Karanicolas, J., and Brooks, C. L. (2004) *J. Mol. Graphics Modell.* 22, 377–395.
- Lee, M. S., Bondugula, R., Desai, V., Zavaljevski, N., Yeh, I. C., Wallqvist, A., and Reifman, J. (2009) *PLoS One* 4, e6254.
- Keskin, O., GURSOY, A., Ma, B., and Nussinov, R. (2008) *Chem. Rev.* 108, 1225–1244.
- Pierce, B., and Weng, Z. P. (2007) *Proteins* 67, 1078–1086.
- Tsai, C. J., Ma, B., and Nussinov, R. (1999) *Proc. Natl. Acad. Sci. U.S.A.* 96, 9970–9972.