RECOLLECTIONS

Mutagenic dissection of the sequence determinants of protein folding, recognition, and machine function

Robert T. Sauer*

Department of Biology, Massachusetts Institute of Technology, Cambridge, Massachusetts, 02139

Received 31 July 2013; Accepted 12 August 2013

DOI: 10.1002/pro.2334

Published online 21 August 2013 proteinscience.org

Robert Sauer is the recipient of the Protein Society 2013 Stein & Moore Award.

Abstract: Understanding the relationship between the amino-acid sequence of a protein and its ability to fold and to function is one of the major challenges of protein science. Here, cases are reviewed in which mutagenesis, biochemistry, structure determination, protein engineering, and single-molecule biophysics have illuminated the sequence determinants of folding, binding specificity, and biological function for DNA-binding proteins and ATP-fueled machines that forcibly unfold native proteins as a prelude to degradation. In addition to structure-function relationships, these studies provide information about folding intermediates, mutations that accelerate folding, slow unfolding, and stabilize proteins against denaturation, show how new binding specificities and folds can evolve, and reveal strategies that proteolytic machines use to recognize, unfold, and degrade thousands of distinct substrates.

Keywords: lambda repressor; Arc repressor; tmRNA tagging; intracellular degradation; ClpXP; ATP-dependent unfolding

Introduction

Bill Stein and Stanford Moore were pioneers in probing the relationship between protein sequence and function, and I am honored to be the recipient of an award that bears their names. This review has two major goals. My primary aim is to illustrate how biochemical, biophysical, structural, and biological studies of mutant proteins can reveal how the amino-acid sequence of a protein dictates its ability to fold, to recognize specific macromolecular part-

ners, and to carry out complex biological functions. My second goal is to recount a personal journey that began with determining amino-acid sequences using classical methods of protein chemistry and progressed to probing the mechanisms that ATP-fueled molecular machines use to forcibly unfold native proteins as a prelude to proteolytic destruction.

Protein Sequencing the Hard Way

My scientific career began in the summer of 1968. John Potts, who was the newly appointed Chief of the Endocrine Unit at Mass General Hospital and had trained with Chris Anfinsen, hired me as a research technician. One of my first jobs was operating an amino-acid analyzer, an instrument invented by Bill Stein and Stanford Moore. I also learned to

Grant sponsor: National Institutes of Health; Grant numbers: AI-15706; AI-16892; GM-101988.

*Correspondence to: Robert T. Sauer, Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139. E-mail: bobsauer@mit.edu

sequence proteins using a method developed by Pehr Edman, who had trained my immediate supervisor, Hugh Niall. I remember Hugh telling me about a breakfast for young protein chemists that Stanford Moore would host at the Federation Meetings in Atlantic City each year. Thus, I had somewhat tenuous connections to Stein, Moore, and Anfinsen, who would share the Nobel Prize in Chemistry in 1972. Sequencing proteins the classical way was like solving a puzzle, but one that could take years to finish, giving me plenty of time to memorize the sequences I was deciphering and to wonder how those strings of amino acids encoded biological function.

λ Repressor

I applied to two grad schools in 1973. MIT rejected me, citing the recent discontinuation of NIH training grants and too many qualified applicants, but Harvard offered me a fellowship. As a grad student, I worked for Mark Ptashne, who was studying the bacteriophage λ cI repressor protein and how its binding to operator-DNA sites controlled viral development.¹ This was an exciting time in molecular biology when methods of cloning and plasmid over-expression were first being developed. Part of my thesis was determining the sequence of λ repressor. I deduced most of the sequence the classical way, but couldn't resolve whether residue 52 was an asparagine or an aspartic acid (Fig. 1). Allan Maxam and Wally Gilbert, on the floor below, had developed a method to sequence DNA,² which I used to confirm that the questionable residue was Asn.^{52,3} It turned out that getting this residue right was important, as we later found that an Asn⁵² $\rightarrow$ Asp substitution eliminated repressor activity. The article reporting the sequence was mentioned in a commentary enti-

Figure 1. The amino-acid sequence of a portion of λ repressor was determined by sequential Edman degradation (arrows) of repressor fragments, including one generated by hydroxylamine cleavage, but whether residue 52 was Asn or Asp could not be determined. DNA sequencing of the repressor gene established that this position 52 was Asn.

tled "*The decline and fall of protein chemistry*", which noted that nucleic-acid sequencing would soon be the preferred method of determining most protein sequences.⁴ Carl Pabo, a fellow grad student, and I isolated λ repressor domains by limited proteolysis and found that the N domain bound operator DNA, although weakly compared to intact repressor, whereas the C domain formed stable dimers and larger oligomers.^{5,6} In subsequent years, Carl, Mitch Lewis, and colleagues would solve the crystal structures of the individual domains in various oligomeric forms, as well as an N-domain dimer and an intact repressor dimer bound to an operator-DNA site.⁷⁻¹¹

While I was finishing my Ph.D., the MIT Biology Department offered me a position as an Assistant Professor. They were looking for someone to run a departmental protein-sequencing facility, liked the fact that I knew molecular biology, were willing to count my time as a technician in lieu of a postdoc, and probably felt bad about turning me down for grad school. I accepted the position and walked into a completely empty lab on July 1, 1978. One of my goals was to understand how the amino-acid sequence of λ repressor encoded its functional properties. At the time, site-directed mutagenesis was not possible. Hillary Nelson and Michael Hecht, who were among my first grad students, developed ways to isolate, sequence, and characterize repressor mutations with strongly defective phenotypes.¹²⁻¹⁴ Biochemical and genetic studies revealed that some of these mutations prevented stable folding of the N domain, others appeared to define the contact surface with operator DNA, which was confirmed when the co-crystal structure was solved several years later,⁸ and one mutation affected the dimerization interface of the N domain, which was formed by symmetric packing between α -helix 5 of each subunit (Fig. 2). It had been established that each 17 base-pair λ operator site had an approximately two-fold symmetric DNA sequence and bound a single repressor dimer.¹ Surprisingly, although the K_D for N-domain dimerization was only 100 μ M,¹⁵ the mutation in helix 5 was as strong as any DNA-interface mutation. In fact, we never detected monomer binding to an operator half site. These results taught us that "weak" and "distant" didn't mean unimportant and that coupling two sets of DNA-binding interactions via dimerization was critical for recognition. One wouldn't have guessed this looking at the structure. By isolating suppressors of some of the weaker defective mutations, Hillary and Michael also identified amino-acid changes that enhanced DNA-binding affinity and specificity and/or increased N-domain stability.¹⁶⁻¹⁸

By the mid to late 1980's, the automation of DNA synthesis allowed directed-mutagenesis experiments. To study the information content of a region of sequence, John Reidhaar-Olson randomized codons in the helix-5 region of λ repressor synthetically, selected

Figure 2. Mutations that eliminate λ -repressor function. (A) Cartoon of the secondary structure of the N domain and positions of severe mutations. Mutations that affected DNA binding directly (red or magenta squares) were found in the N-terminal arm, in the helix-turn-helix motif, and in the dimerization helix. Mutations that affected DNA binding indirectly by disrupting or destabilizing the structure (blue squares, hydrophobic-core substitutions; green triangles, Xxx $\rightarrow$ Pro mutations) were distributed throughout the sequence. (B) Positions of mutations that affect function directly were mapped onto the cocrystal structure of the N domain bound to an operator-DNA site.

functional variants, and sequenced the “neutral” mutations.^{19,20} He found that many amino-acid substitutions were tolerated at surface positions, whereas the repertoire of allowable side chains was more restricted at buried and partially buried positions (Fig. 3). Indeed, only the wild-type isoleucine appeared to be tolerated at position 84, where we had found that a serine substitution eliminated DNA binding. Wendell Lim applied this method of combinatorial mutagenesis and selection to residue positions that defined the N-domain hydrophobic core and found that many different combinations of hydrophobic residues were tolerated (Fig. 4), albeit with some restrictions on volume and stereochemistry.²¹ Subsequent studies of some of these variants revealed the energetic and structural consequences of core substitutions.^{22–24}

Arc and Mnt Repressors

During the late 1970’s and early 1980’s, we worked on a number of different bacteriophage transcription factors. Sequence and structural homology suggested that many of these proteins as well as some bacte-

rial repressors were similar to λ repressor in using a helix-turn-helix motif for DNA recognition.²⁵ However, the Arc and Mnt repressors from bacteriophage P22, which shared strong sequence homology with each other, did not have an obvious helix-turn-helix motif.^{26,27} Phil Youderian in Mimi Susskind’s lab, Drew Vershon, and Ken Knight performed genetic, biochemical, and altered-specificity experiments that suggested that the N-terminal residues of Arc and Mnt were responsible for operator-DNA recognition.^{28–32} Using combinatorial mutagenesis, Jim Bowie found that the pattern of obligatory hydrophobic positions in the 53-residue Arc protein strongly predicted a β - α - α fold,³³ which was subsequently confirmed when Robert Kaptein’s group solved the solution structure of the Arc dimer.³⁴ In this dimer, many of the residues implicated in DNA recognition were part of an anti-parallel β -sheet.

Bronwen Brown discovered that Arc bound to its 21 base-pair operator site as a tetramer, with each dimer binding with nanomolar affinity to a half site and strong positive cooperativity between the two DNA-bound dimers [Fig. 5(A)].^{35,36} Dimer binding to half sites was surprising, as there was little symmetry in the sequence of these sites. Brigitte Raumann solved the crystal structure of the Arc tetramer bound to operator DNA and found that all of the major-groove contacts were made by residues from the antiparallel β -sheet [Fig. 5(B)].³⁷ The structure revealed striking asymmetry in the DNA contacts made by equivalent side chains in the dimer [Fig. 5(C)]. For example, one Asn¹¹ side chain made hydrogen bonds with a cytosine base, whereas the other Asn¹¹ made hydrogen bonds with an adjacent thymine on the same DNA strand. These results made it clear that there was no simple “code” of DNA recognition. Another surprise came from probing the importance of Phe¹⁰, a residue that packed into the core of the free dimer but rotated out to contact the phosphate backbone in the operator complex. Joel Schildbach and Wali Karzai solved the crystal structure of the F10V mutant bound to operator DNA and found that it was essentially identical to the wild-type complex except for the expected missing backbone contacts.³⁸ Nevertheless, F10V Arc had lost the ability to discriminate between operator and non-operator sites. These results implied that the Phe¹⁰-phosphate contacts contribute to specific binding, either because contacts between Arc and bases in the major groove trigger the Phe¹⁰ conformational change or because this backbone contact cannot be made unless there is a tight fit of the rest of the protein into the major groove, which only occurs for the cognate DNA site.

The N domain of Mnt had the same β - α - α fold as Arc, whereas the C domain formed a stable coiled-coil tetramer.^{39,40} Although the 21-base-pair *mnt* operator differed at most positions from the *arc* operator,

Figure 3. Substitutions compatible with biological function in α -helix 5 of λ repressor, which mediates dimerization, were identified by combinatorial mutagenesis and selection. Hydrophobic, polar, and charged residues are colored blue, green, and red, respectively. Substitutions shown below the wild-type sequence result in inactive repressor variants. The relative solvent accessibilities of the wild-type side chains are shown in the histogram. In general, surface positions tolerate many substitutions, whereas buried positions are more restricted.

Brigitte Raumann and Ken Knight discovered that a few residue substitutions were sufficient to alter binding specificity for these sites! For example, a hybrid protein containing the arm and β -sheet residues of Arc and body of Mnt bound tightly and specifically to

the *arc* operator, changing a single arm residue in the hybrid resulted in tight affinity for both the *arc* and *mnt* operators, and changing one additional β -sheet side chain resulted in a protein that bound tightly and specifically to the *mnt* operator [Fig. 5(D)].^{30,41} Before these studies, I would never have imagined that such a dramatic change in binding specificity could be mediated by so few residue changes. The existence of an intermediate that both retained the original specificity and had acquired a new specificity was also provocative in terms of possible evolutionary trajectories.

Studies of Arc folding and stability resulted in several surprises. For example, Jim Bowie showed that Arc dimers equilibrated with unfolded monomers, resulting in concentration-dependent stability with a dimerization/folding constant of ~ 10 nM.⁴² Moreover, Marcos Milla and Thor Jonsson found that the kinetics of dimer dissociation and unfolding were remarkably fast, with a half life of ~ 7 s under native conditions.^{43,44} Despite this rapid unfolding, refolding was sufficiently fast to ensure that 99.9% of the Arc molecules in the population were native at a concentration of 10 μ M. The small size of Arc (53 residues) made it attractive for determining how each residue contributed to folding and stability. Marcos and Bronwen Brown constructed and purified a complete set of Arc variants with single alanine substitutions and

Figure 4. Tolerance of the hydrophobic core to amino-acid substitutions. (A) The positions of Leu¹⁸, Val³⁶, Met⁴⁰, Val⁴⁷, Phe⁵¹, Leu⁵⁷, and Leu⁶⁵, which form most of the hydrophobic core of the globular region of the λ -repressor N domain, are shown in stick representation and dot representation. (B) Core sequences in functional variants identified by combinatorial mutagenesis and selection revealed variations in total core volume and packing but the major requirement was maintenance of hydrophobic side chains at these positions.

Figure 5. Binding of Arc repressor to operator DNA. (A) Single Arc dimers bind with similar affinities to the left and right half sites of the 21-base-pair *arc* operator. When one half site is occupied, however, the second dimer binds with 6000-fold higher affinity. This strong positive cooperativity results from favorable dimer-dimer contacts. (B) Co-crystal structure of two Arc dimers bound to *arc*-operator DNA. Side chains from the β sheet of each dimer make all of the major-groove contacts. The boxed portions of the *arc* operator sequence show the bases identified by mutagenesis as being important for recognition. Although the Arc dimer is essentially two-fold symmetric, there is little symmetry in the DNA sequence of each half site. (C) Arrows indicate contacts between β -sheet side chains and DNA bases. The Gln⁹ side chains make symmetric contacts but the Asn¹¹ and Arg¹³ side chains make asymmetric contacts. (D) The *mnt* operator differs from the *arc* operator at most positions and all of the important base pairs, which are boxed. A hybrid protein with the 13 N-terminal residues of Arc and body of Mnt (hy1) binds tightly and specifically to the *arc* operator. In this context, the single Arg5→Ser change allows tight binding of hy2 to both operators, whereas the additional Gln⁹→His substitution restores specific binding of hy3 to the *mnt* operator.

measured equilibrium stability and the kinetics of unfolding and refolding.^{45–47} One mutant was more stable than wild type, ~50% had stabilities similar to wild type, ~40% were destabilized, and ~10% were completely unfolded under native conditions. Most mutations that destabilized Arc increased the unfolding/dissociation rate and decreased the folding/dimerization rate, suggesting that the native interactions affected by these mutations are partially formed in the transition state, with ~80% of energetic effects of this information being consolidated between the transition and native states. Almost all of the alanine-substitution mutations that had minimal stability effects altered surface residues. To test if *en masse* removal of the wild-type side chains at these surface positions might disrupt folding, we made mutants with multiple alanine substitutions.⁴⁸ One of these variants, containing 15 alanine substitutions per subunit, displayed ~20-fold faster refolding, ~100-fold slower unfolding, and displayed nuclear magnetic resonance (NMR) chemical dispersion in the amide-aromatic and methyl regions expected for a well-folded protein. For stable folding, these results,

once again, clearly emphasized the importance of core side chains and relative unimportance of surface side chains.

Alanine-substitution mutations at five core positions of Arc prevented folding.⁴⁶ One of these residues, Glu³⁶, formed partially buried salt bridges and hydrogen bonds with the Arg³¹ and Arg⁴⁰ side chains in the same subunit [Fig. 6(A)]. At the time, substantial evidence supported the importance of salt-bridge networks for protein stability. Indeed, alanine substitutions for any of the salt-bridge residues were destabilizing by 2–6 kcal/mol.⁵⁰ However, computational studies had predicted that such networks might make proteins less stable compared to hydrophobic packing.⁴⁹ To test this possibility, Carey Waldburger randomized Arc residues 31, 36, and 40, selected functional variants, and found that mutants with non-polar side chains at these positions were indeed more stable than wild-type Arc.^{50,51} For example, the Met³¹/Tyr³⁶/Leu⁴⁰ variant had a crystal structure like wild type but was ~4 kcal/mol more stable, in part because of substantially faster folding and dimerization [Fig. 6(A)]. Changes in the position of the

Figure 6. Replacing a salt-bridge triad with hydrophobic interactions stabilizes Arc and accelerates folding. (A) Salt-bridge interactions between the side chains of Arg³¹, Glu³⁶, and Arg⁴⁰ in wild-type Arc were replaced by hydrophobic interactions between Met³¹, Tyr³⁶, and Leu⁴⁰ in a fully functional variant identified by combinatorial mutagenesis and selection. Compared to wild-type Arc, the mutant was 3.9 kcal/mol more stable, folded 37-fold more rapidly, and unfolded 12-fold more slowly. (B) Arc folding and unfolding proceed through a dimeric intermediate (I₂). For the Met³¹/Tyr³⁶/Leu⁴⁰ variant, formation of I₂ is the rate-limiting step in folding. For wild-type Arc, the I₂ → N₂ transition is the rate-limiting step in folding, suggesting that formation of the Arg³¹-Glu³⁶-Arg⁴⁰ salt bridge is also a late step in folding and early step in unfolding.

transition state and the viscosity dependence of folding/dimerization suggested that denatured monomers of this mutant interact at close to the diffusion-controlled limit to form a partially folded dimer (I₂), which rapidly progressed to the native dimer [Fig. 6(B)]. An I₂ intermediate probably also exists in wild-type Arc folding but dissociates faster than forming N₂, thereby slowing the folding/dimerization rate. Additional experiments provided evidence for at least one additional intermediate between I₂ and N₂ with buried surface area very similar to native Arc.⁴⁴ Interestingly, some alanine-substitution mutations were less destabilizing in Met³¹/Tyr³⁶/Leu⁴⁰ Arc than in the wild-type background and *vice versa*.⁵²

Although many mutations in Arc and in λ repressor affected protein stability, the native structures of these variants appeared to be similar to the wild-type protein as judged by spectroscopy, biological function, and crystallography. Matt Cordes discovered a striking exception to this rule when he swapped the side chains of Asn¹¹ and Leu¹² in Arc. In the NMR

structure of this mutant, two 3₁₀ helices replaced the wild-type β sheet.^{53,54} Remarkably, an Arc variant containing just the N11L mutation served as an evolutionary bridge between the two alternative structures, switching between the β sheet and 3₁₀ helices at a rate much faster than dimer dissociation/unfolding.⁵⁵ These results suggested that the β sheet is normally the last element of secondary structure to form during folding and break during unfolding, a result consistent with the effects of glycine-substitution mutations in different regions of Arc.⁵⁶

Degradation Tags

Our interest in intracellular protein degradation was initially piqued in studies of λ Cro mutants, when Andy Pakula found that mutations expected to destabilize the protein resulted in rapid intracellular degradation.⁵⁷ This correlation had not been observed for destabilizing mutations in full-length λ repressor, but studies of variants corresponding to just the N domain revealed that decreased thermodynamic stability increased the rate of intracellular degradation.⁵⁸ Two additional studies convinced us that factors in addition to thermodynamic stability must play important roles in determining degradation rates in *Escherichia coli*. Jim Bowie selected revertants of a destabilizing mutation in Arc, expecting to find suppressor mutations that increased thermodynamic stability.⁵⁹ Instead, the suppressors slowed intracellular degradation. One suppressor changed the C-terminal sequence from Gly-Ala to Asp-Ala and others were frame-shifts near the C terminus that resulted in addition of tails that terminated with polar amino acids. Dawn Parsell and Karen Silber randomized an unstructured region at the C terminus of the λ -repressor N domain, screened for proteins resistant or sensitive to degradation, and sequenced variants from both groups.⁶⁰ All of the rapidly degraded proteins had relatively non-polar C-terminal sequences (e.g., Ala-Ala), whereas all of the proteins resistant to degradation had more polar tails (e.g., Asn-Arg). Clearly, residues at the C terminus were having major effects on the rate of intracellular degradation.

In 1995, Richard Simpson's group reported that expression of recombinant IL6 in *E. coli* resulted in a heterogeneous subset of insoluble truncated proteins that had acquired an 11-residue C-terminal tag.⁶¹ The last 10 residues of this tag were encoded by a small open-reading-frame in an RNA molecule called 10Sa or SsrA, which had some tRNA-like features including charging with alanine.⁶² The "ssrA tag" terminated with Ala-Ala, and thus we hypothesized that it probably targeted proteins for degradation. However, it was not clear how "tagging" occurred? Ken Keiler reasoned that if SsrA functioned in protein quality control, it might bind ribosomes unable to complete protein synthesis because the A site was unoccupied,

act as a tRNA to add the charged alanine to the nascent chain, and then replace the original message and act as an mRNA to complete translation of the tag (Fig. 7). The net effect of these reactions would be to rescue the stalled ribosome and mark the polypeptide for proteolysis. Ken and Patrick Waller tested this model by engineering mRNAs without in-frame stop codons on which ribosomes would stall at the 3' end. As predicted, model proteins expressed from nonstop mRNAs terminated with the *ssrA* tag and were rapidly degraded unless specific proteases were also eliminated.⁶³ By engineering tmRNA variants that encoded non-degradable tags, Eric Roche and Sean Moore demonstrated that ~0.5% of translation events in *E. coli* normally terminate with tmRNA tagging, producing a heterogeneous mix of more than 1000 tagged proteins, most with molecular weights sufficiently large to encode folded proteins or domains^{64,65}

ATP-Fueled Proteolytic Machines

We found that *ssrA*-tagged proteins in the cytoplasm of *E. coli* were degraded by a protease called ClpXP.^{66,67} Tania Baker, my colleague and neighbor at MIT, also worked on ClpXP, and our labs began a collaboration to determine how *ssrA*-tagged proteins were recognized and degraded by this ATP-dependent enzyme. Work from numerous labs had established that the proteolytic active sites of ClpP₁₄ resided within a hollow double-ring chamber accessible through small pores that excluded native proteins, whereas ClpX₆, a member of the AAA+ family of ATPases, was a ring-shaped homohexamer that could interact with ClpP (for review, see Ref. 68). The roles

of ClpX in protein degradation appeared to include recognition of appropriate substrates, ATP-dependent unfolding of these proteins by an unknown mechanism, and translocation of the unfolded polypeptide through an axial pore in the ClpX ring and into ClpP (Fig. 8). Studies by David Wah, Samia Siddiqui, and Andy Martin established that the *ssrA* tag bound in the axial pore of the ClpX ring with an affinity similar to the K_M for degradation of *ssrA*-tagged substrates, indicating that the tag simply provides a way for ClpX to grip the substrate prior to initiating unfolding and translocation.^{69–71}

Did ClpXP actively unfold substrates or simply trap spontaneous unfolding? Yong-In Kim and Randy Burton showed that *ssrA*-tagged green fluorescent protein (GFP) unfolded on a timescale of years in solution but was degraded in ~1 min by ClpXP and could also be unfolded, although slightly more slowly, by ClpX alone.⁷² By contrast, untagged GFP was neither degraded nor unfolded. These results established that recognition requires the *ssrA* tag and strongly suggested that ClpX actively unfolds protein substrates. Supporting this model, Randy found that ClpXP degraded a set of *ssrA*-tagged variants of Arc at rates differing by less than 2 fold despite the fact that the most and least stable of these mutants differed by 15 kcal/mol in thermodynamic stability and by 10⁸-fold in kinetic stability.⁷³

To address the rate limiting steps and energetic cost of degradation, Jon Kenniston used *ssrA*-tagged variants of the I27 domain of human titin.⁷⁴ He found that alkylation of cysteines normally buried in the titin^{I27} core produced an unfolded but highly

Figure 7. tmRNA model for co-translational addition of the *ssrA* tag. A tmRNA molecule, charged with alanine, recognizes ribosomes stalled at an empty A site. The charged alanine (yellow) is added to the nascent chain by transpeptidation, the original mRNA is replaced by the open-reading frame of tmRNA (red), and translation of this sequence adds the rest of the *ssrA* tag to the nascent chain and allows normal terminal of protein synthesis. The *ssrA*-tagged protein is then degraded by cellular proteases.

Figure 8. The ClpXP protease degrades ssrA-tagged proteins. ClpXP, shown as a cut-away cartoon, consists of the AAA+ ClpX hexamer and the self-compartmentalized ClpP peptidase. The ssrA tag binds in the axial pore of ClpX. Conformational changes in ClpX, fueled by ATP binding and hydrolysis, power unfolding of the ssrA-tagged protein and translocation of the unfolded polypeptide into the degradation chamber of ClpP.

soluble protein. ClpXP degraded unfolded titin^{I27}-ssrA substantially faster than native titin^{I27}-ssrA, establishing that unfolding was the rate-limiting step in degradation of the latter protein. The rate of ATP hydrolysis by ClpXP during unfolding of native titin^{I27}-ssrA substrates was ~4-fold slower than during translocation, as expected for a motor working under an increased mechanical load. During the time required for degradation of one unfolded substrate, ClpXP hydrolyzed an average of ~100 ATPs. By contrast, this value was ~600 ATPs for native titin^{I27}-ssrA, indicating that ~500 ATPs were consumed before successful denaturation. Apparently, most unfolding attempts fail and thus repeated applications of force are required to ensure eventual success. Indeed, Jon showed that ClpXP often releases the engaged substrate when denaturation fails.⁷⁵ ClpXP unfolding of native titin^{I27}-ssrA variants with mutations in the β sheet proximal to the ssrA tag required hydrolysis of as few as 20 ATPs, supporting a model in which unfolding is promoted by enzymatic pulling against and disruption of elements of structure attached to the degradation tag, a model originally proposed by Andreas Matouschek and colleagues.⁷⁶ The rates of ClpXP degradation of a large number of ssrA-tagged substrates show some correlation with global thermodynamic and kinetic stability, as unfolded and metastable substrates were generally degraded faster than very stable substrates (Fig. 9). However, as noted above for Arc variants, substrates degraded at similar rates by ClpXP often differed enormously in global stability and proteins with comparable stabilities were often degraded at very different rates. These results suggest that the local mechanical stability of the structural elements

attached to the degradation tag appears to dictate resistance to enzymatic unfolding.⁷⁴

The stochastic nature of ClpXP unfolding was illustrated elegantly by Marie Aubin-Tam and Adrian Olivares in single-molecule experiments using a dual-laser optical trap [Fig. 10(A,B)].⁷⁷ These studies, which emerged from a long-standing collaboration between Matt Lang's lab, Tania's lab, and my lab, showed that the dwell times preceding ClpXP unfolding of the filamin-A domains or HaloTag domain of an engineered substrate were exponentially distributed, as expected if both transient thermal destabilization of the domain and simultaneous pulling by ClpXP were required for successful unfolding. This work also demonstrated that ClpXP could work against forces up to 20 pN and that denaturation, once initiated, was very fast and highly cooperative. The cooperativity of protein unfolding by ClpXP depends on the substrate. In population studies of the unfolding of GFP-ssrA, for example, Andy Martin and Drew Nager showed that ClpXP initially extracts the C-terminal β strand but the remaining portions of the GFP β barrel remain folded, requiring additional ClpXP translocation and unfolding attempts to ensure global denaturation and degradation.^{78,79}

What chemical features of a polypeptide chain does ClpX recognize during translocation? Sarah Barlow and Igor Levchenko found that neither side-chain identity nor chirality were important and demonstrated that ClpX could translocate synthetic ssrA-tagged peptides in which 10 additional methylene groups were inserted between successive amide bonds in a guest region [Fig. 10(C)].⁸⁰ Thus, ClpX appears to be able to translocate almost any molecule attached to an ssrA tag. Dan Bolon engineered a substrate in

Figure 9. Structural determinants of degradation. Rates of ClpXP degradation of ssrA-tagged variants of Arc, titin¹²⁷, RNase H, and GFP are plotted as a function of the global thermodynamic stabilities (left panel) or the global kinetic stabilities (right panel) of these proteins. For a given protein family, less stable proteins are generally degraded more rapidly. When comparing different families, however, there are wide variations. For example, some variants are degraded at comparable rates despite having thermodynamic stabilities that differ by 10 kcal/mol and kinetic stabilities that differ by seven orders of magnitude.

Figure 10. ClpXP unfolding and translocation. (A) Laser-trap single-molecule assay of ClpXP unfolding and translocation. ClpXP is attached to one laser-trapped bead and a multidomain ssrA-tagged substrate is attached to another laser-trapped bead via a DNA linker. (B) ClpXP unfolding of each domain results in a sudden increase in bead-bead distance (marked by arrows). Following an unfolding event, ClpXP translocation results in a gradual decrease in bead-bead distance until the unfolded chain has been completely translocated. Upon completion of translocation, there is variable dwell time during which ClpXP hydrolyzes ATP but fails to unfold the next domain. Successful unfolding requires coincidence between a ClpXP power stroke and transient thermal destabilization of the local structure that the enzyme pulls against. (C) Polypeptide translocation by ClpXP occurs in steps of ~ 5 – 7 residues, can proceed in an C→N or N→ direction, and is not dependent upon side-chain identity, chirality, or the spacing between peptide bonds. Moreover, ClpXP can simultaneously translocate multiple polypeptide chains.

which the *ssrA* tag was disulfide bonded to an adaptor protein, with the expectation that ClpX would not be able to unfold or translocate this substrate.⁸¹ To our surprise, this crosslinked complex was degraded by ClpXP, showing that three polypeptide chains can simultaneously pass through the ClpX pore. In combination, these results suggest that the axial pore must be highly elastic, with the ability to widen to accommodate multiple chains or bulky regions of polypeptide and to contract to grip skeletal sequences, such as polyglycine. How this feat is accomplished in structural terms remains to be determined.

Andy Martin pioneered the use of genetically encoded single-chain ClpX hexamers, which allowed mutations to be introduced into a subset of specific subunits, and demonstrated that variants with just one or two subunits active in ATP hydrolysis could still function in protein unfolding and degradation.⁸² His results demonstrated that a ClpX power stroke could be generated by ATP hydrolysis in a single subunit, ruled out concerted and strictly sequential models of function, and provided evidence for a probabilistic mechanism that might prevent stalling if one subunit failed to bind or hydrolyze ATP. Andy also found that mutating the axial-pore loops in

even a few ClpX subunits resulted in substrate slippage, less efficient unfolding, and an increase in the energetic cost of degradation, supporting a model in which ATP-dependent conformational changes in these pore loops transmit force and drive substrate translocation and unfolding.⁸³

When Steve Glynn solved the crystal structure of the ClpX hexamer, he found that ATP bound between the large and small AAA+ domains of some subunits [Fig. 11(A,B)], as expected from related enzymes, but each binding site was slightly different because of changes in the relative orientations of the large and small domains.⁸⁴ Moreover, the nucleotide-binding site was destroyed in some subunits by a substantial rotation between the small and large domains. These structural studies helped explain Greg Hersch's discovery that some subunits in the ClpX ring did not bind ATP and the remaining subunits fell into different classes with respect to nucleotide affinity.⁸⁵ Despite the dramatic structural asymmetry, the small AAA+ domain of each subunit packed against the large AAA+ domain of its neighbor in a rigid-body fashion [Fig. 11(C)], suggesting that changes in ring conformation arise from nucleotide-dependent changes in the conformation of the short hinge that

Figure 11. ClpX structure. (A) In “loadable” or “L” subunits of the ClpX hexamer, the orientation between the large and small AAA+ domains creates a binding site for ATP or ADP. In “unloadable” or “U” subunits (not shown), major changes in hinge conformation rotate the small domain $\sim 80^\circ$ and destroy the nucleotide-binding site. (B) In crystal structures, the ClpX ring contains four L subunits and two U subunits. (C) Despite nucleotide-binding asymmetry, each small domain packs against the large domain of the counterclockwise subunit in a rigid-body manner. Thus, nucleotide-dependent changes in the conformation of a single hinge could propagate structural changes around the ClpX ring to drive unfolding and translocation. (D) ATP binding to a 4L:2U hexamer drives structural changes that facilitate ATP hydrolysis and mechanical function. Subunit switching between L and U conformations is required for ClpX unfolding and translocation.

links the large and small domains. To test this model, Steve connected each part of the rigid-body unit with designed disulfide bonds and created a covalently closed ring in which the hinges were the only source of substantial conformational flexibility.⁸⁶ This disulfide-bonded variant retained robust degradation activity, but a single-residue insertion in the hinge, which forms part of the nucleotide-binding site, uncoupled ATP hydrolysis from productive unfolding. Thus, ATP hydrolysis appears to drive change in the conformation of one hinge and its flanking domains that are propagated around the AAA+ ring through the topologically constrained set of rigid-body units and hinges to produce coupled motions that power substrate unfolding and translocation.

We call ClpX subunits that can bind nucleotide “loadable” and those that cannot “unloadable.” Do loadable and unloadable subunits maintain their identities during degradation or does switching occur? Using mutational and disulfide-crosslinking experiments as well as novel fluorescent methods that allowed assays of ATP binding and conformation of single subunits in the ClpX ring, Ben Stinson and Drew Nager provided strong evidence that dynamic interconversion between loadable and unloadable conformations is required to couple ATP hydrolysis to mechanical work [Fig. 11(D)].⁸⁷ ATP binding was found to initially drive staged allosteric changes that set the conformation of the ClpX ring to allow hydrolysis and linked mechanical steps. Switching of subunits between loadable and unloadable conformations then resets or isomerizes the configuration of the nucleotide-loaded ring and changes the roles of individual subunits, which may be part of a fail-safe mechanism that prevents ClpX stalling after failed unfolding attempts. Experiments to test these ideas and to define more completely the mechanism by which ClpX unfolds and translocates proteins are in progress.

Summary

In 1963, long before the molecular-biology revolution and the development of most of the techniques that we use today, Bill Stein and Stanford Moore used chemical modification to establish that His¹² and His¹¹⁹ were close together and part of the active site of ribonuclease A.⁸⁸ It is worth reading this early structure-function paper to appreciate how dramatically the methods used to study proteins have changed over the last 50 years. More importantly, however, the biochemical principles and logic of interrogating protein structure and function have remained remarkably constant.

Acknowledgments

I would like to thank all of the graduate students, post docs, collaborators, and other scientists who

contributed to the work mentioned here and to many additional studies that could not be included.

References

1. Ptashne M (1986) A genetic switch: gene control and phage λ . Palo Alto, CA: Blackwell Scientific Publications.
2. Maxam AM, Gilbert W (1977) A new method for sequencing DNA. *Proc Natl Acad Sci USA* 4:560–564.
3. Sauer RT, Anderegg R (1978) Primary structure of the lambda repressor. *Biochemistry* 17:1092–1100.
4. Malcolm ADB (1978) The decline and fall of protein chemistry? *Nature* 275:90–91.
5. Pabo CO, Sauer RT, Sturtevant JM, Ptashne M (1979) The lambda repressor contains two domains. *Proc Natl Acad Sci USA* 76:1608–1612.
6. Sauer RT, Pabo CO, Meyer BJ, Ptashne M, Backman KC (1979) Regulatory functions of the lambda repressor reside in the amino-terminal domain. *Nature* 279:396–400.
7. Pabo CO, Lewis M (1982) The operator-binding domain of lambda repressor: structure and DNA recognition. *Nature* 298:443–447.
8. Jordan SR, Pabo CO (1988) Structure of the lambda complex at 2.5 Å resolution: details of the repressor-operator interactions. *Science* 242:893–899.
9. Bell CE, Frescura P, Hochschild A, Lewis M (2000) Crystal structure of the lambda repressor C-terminal domain provides a model for cooperative operator binding. *Cell* 101:801–811.
10. Bell CE, Lewis M (2001) Crystal structure of the lambda repressor C-terminal domain octamer. *J Mol Biol* 314:1127–1136.
11. Stayrook S, Jaru-Ampornpan P, Ni J, Hochschild A, Lewis M (2008) Crystal structure of the lambda repressor and a model for pairwise cooperative operator binding. *Nature* 452:1022–1025.
12. Nelson HC, Hecht MH, Sauer RT (1983) Mutations defining the operator-binding sites of bacteriophage lambda repressor. *Cold Spring Harb Symp Quant Biol* 47:441–449.
13. Hecht MH, Nelson HC, Sauer RT (1983) Mutations in lambda repressor's amino-terminal domain: implications for protein stability and DNA binding. *Proc Natl Acad Sci USA* 80:2676–2680.
14. Hecht MH, Sturtevant JM, Sauer RT (1984) Effect of single amino acid replacements on the thermal stability of the NH₂-terminal domain of phage lambda repressor. *Proc Natl Acad Sci USA* 81:5685–5689.
15. Weiss MA, Pabo CO, Karplus M, Sauer RT (1987) Dimerization of the operator binding domain of phage lambda repressor. *Biochemistry* 26:897–904.
16. Hecht MH, Hehir KM, Nelson HC, Sturtevant JM, Sauer RT (1985) Increasing and decreasing protein stability: effects of revertant substitutions on the thermal denaturation of phage lambda repressor. *J Cell Biochem* 29:217–224.
17. Nelson HC, Sauer RT (1985) Lambda repressor mutations that increase the affinity and specificity of operator binding. *Cell* 42:549–558.
18. Hecht MH, Sauer RT (1985) Phage lambda repressor revertants. Amino acid substitutions that restore activity to mutant proteins. *J Mol Biol* 186:53–63.
19. Reidhaar-Olson JF, Sauer RT (1988) Combinatorial cassette mutagenesis as a probe of the informational content of protein sequences. *Science* 241:53–57.
20. Reidhaar-Olson JF, Sauer RT (1990) Functionally acceptable substitutions in two alpha-helical regions of lambda repressor. *Proteins* 7:306–316.

21. Lim WA, Sauer RT (1989) Alternative packing arrangements in the hydrophobic core of lambda repressor. *Nature* 339:31–36.
22. Lim WA, Sauer RT (1991) The role of internal packing interactions in determining the structure and stability of a protein. *J Mol Biol* 219:359–376.
23. Lim WA, Farruggio DC, Sauer RT (1992) Structural and energetic consequences of disruptive mutations in a protein core. *Biochemistry* 31:4324–4333.
24. Lim WA, Hodel A, Sauer RT, Richards FM (1994) The crystal structure of a mutant protein with altered but improved hydrophobic core packing. *Proc Natl Acad Sci USA* 91:423–427.
25. Sauer RT, Yocum RR, Doolittle RF, Lewis M, Pabo CO (1982) Homology among DNA-binding proteins suggests use of a conserved super-secondary structure. *Nature* 298:447–451.
26. Sauer RT, Krovatin W, DeAnda J, Youderian P, Susskind MM (1983) Primary structure of the *immI* immunity region of bacteriophage P22. *J Mol Biol* 168:699–713.
27. Vershon AK, Youderian P, Susskind MM, Sauer RT (1985) The bacteriophage P22 arc and mnt repressors. Overproduction, purification, and properties. *J Biol Chem* 260:12124–12129.
28. Youderian P, Vershon A, Bouvier S, Sauer RT, Susskind MM (1983) Changing the DNA-binding specificity of a repressor. *Cell* 3:777–783.
29. Vershon AK, Bowie JU, Karplus TM, Sauer RT (1986) Isolation and analysis of Arc repressor mutants: evidence for an unusual mechanism of DNA binding. *Proteins* 1:302–311.
30. Knight KL, Sauer RT (1989) DNA binding specificity of the Arc and Mnt repressors is determined by a short region of N-terminal residues. *Proc Natl Acad Sci USA* 86:797–801.
31. Knight KL, Sauer RT (1989) Identification of functionally important residues in the DNA binding region of the mnt repressor. *J Biol Chem* 264:13706–13710.
32. Knight KL, Sauer RT (1992) Biochemical and genetic analysis of operator contacts made by residues within the beta-sheet DNA binding motif of Mnt repressor. *EMBO J* 11:215–223.
33. Bowie JU, Sauer RT (1989) Identifying determinants of folding and activity for a protein of unknown structure. *Proc Natl Acad Sci USA* 86:2152–2156.
34. Breg JN, van Opheusden JH, Burgering MJ, Boelens R, Kaptein R (1990) Structure of Arc repressor in solution: evidence for a family of beta-sheet DNA-binding proteins. *Nature* 346:586–589.
35. Brown BM, Bowie JU, Sauer RT (1990) Arc repressor is tetrameric when bound to operator DNA. *Biochemistry* 29:11189–11195.
36. Brown BM, Sauer RT (1993) Assembly of the Arc repressor-operator complex: cooperative interactions between DNA-bound dimers. *Biochemistry* 32:1354–1363.
37. Raumann BE, Rould MA, Pabo CO, Sauer RT (1994) DNA recognition by beta-sheets in the Arc repressor-operator crystal structure. *Nature* 367:754–757.
38. Schildbach JF, Karzai AW, Raumann BE, Sauer RT (1999) Origins of DNA-binding specificity: role of protein contacts with the DNA backbone. *Proc Natl Acad Sci USA* 96:811–817.
39. Burgering MJ, Boelens R, Gilbert DE, Breg JN, Knight KL, Sauer RT, Kaptein R (1994) Solution structure of dimeric Mnt repressor (1–76). *Biochemistry* 33:15036–15045.
40. Nooren IM, Kaptein R, Sauer RT, Boelens R (1999) The tetramerization domain of the Mnt repressor consists of two right-handed coiled coils. *Nat Struct Biol* 6:755–759.
41. Raumann BE, Knight KL, Sauer RT (1995) Dramatic changes in DNA-binding specificity caused by single residue substitutions in an Arc/Mnt hybrid repressor. *Nat Struct Biol* 2:1115–1122.
42. Bowie JU, Sauer RT (1989) Equilibrium dissociation and unfolding of the Arc repressor dimer. *Biochemistry* 28:7139–7143.
43. Milla ME, Sauer RT (1994) P22 Arc repressor: folding kinetics of a single-domain, dimeric protein. *Biochemistry* 33:1125–1133.
44. Jonsson T, Waldburger CD, Sauer RT (1996) Nonlinear free energy relationships in Arc repressor unfolding imply the existence of unstable, native-like folding intermediates. *Biochemistry* 35:4795–4802.
45. Milla ME, Brown BM, Sauer RT (1994) Protein stability effects of a complete set of alanine substitutions in Arc repressor. *Nat Struct Biol* 1:518–523.
46. Milla ME, Sauer RT (1995) Critical side-chain interactions at a subunit interface in the Arc repressor dimer. *Biochemistry* 34:3344–3351.
47. Milla ME, Brown BM, Waldburger CD, Sauer RT (1995) P22 Arc repressor: transition state properties inferred from mutational effects on the rates of protein unfolding and refolding. *Biochemistry* 34:13914–13919.
48. Brown BM, Sauer RT (1999) Tolerance of Arc repressor to multiple-alanine substitutions. *Proc Natl Acad Sci USA* 96:1983–1988.
49. Hendsch ZS, Tidor B (1994) Do salt bridges stabilize proteins? A continuum electrostatic analysis. *Protein Sci* 3:211–226.
50. Waldburger CD, Schildbach JF, Sauer RT (1995) Are buried salt bridges important for protein stability and conformational specificity? *Nat Struct Biol* 2:122–128.
51. Waldburger CD, Jonsson T, Sauer RT (1996) Barriers to protein folding: formation of buried polar interactions is a slow step in acquisition of structure. *Proc Natl Acad Sci USA* 93:2629–2634.
52. Srivastava AK, Sauer RT (2002) Mutational studies of protein stability and folding of the hyperstable MYL Arc repressor variant. *Biophys Chem* 101–102:35–42.
53. Cordes MH, Walsh NP, McKnight CJ, Sauer RT (1999) Evolution of a protein fold *in vitro*. *Science* 284:325–328.
54. Cordes MH, Walsh NP, McKnight CJ, Sauer RT (2003) Solution structure of switch Arc, a mutant with 3₁₀ helices replacing a wild-type beta-ribbon. *J Mol Biol* 326:899–909.
55. Cordes MH, Burton RE, Walsh NP, McKnight CJ, Sauer RT (2000) An evolutionary bridge to a new protein fold. *Nat Struct Biol* 7:1129–1132.
56. Srivastava AK, Sauer RT (2000) Evidence for partial secondary structure formation in the transition state for Arc repressor refolding and dimerization. *Biochemistry* 39:8308–8314.
57. Pakula AA, Young VB, Sauer RT (1986) Bacteriophage lambda Cro mutations: effects on activity and intracellular degradation. *Proc Natl Acad Sci USA* 83:8829–8833.
58. Parsell DA, Sauer RT (1989) The structural stability of a protein is an important determinant of its proteolytic susceptibility in *Escherichia coli*. *J Biol Chem* 264:7590–7595.
59. Bowie JU, Sauer RT (1989) Identification of C-terminal extensions that protect proteins from intracellular proteolysis. *J Biol Chem* 264:7596–7602.
60. Parsell DA, Silber KR, Sauer RT (1990) Carboxy-terminal determinants of intracellular protein degradation. *Genes Dev* 4:277–286.

61. Tu GF, Reid GE, Zhang JG, Moritz RL, Simpson RJ (1995) C-terminal extension of truncated recombinant proteins in *Escherichia coli* with a 10Sa RNA decapeptide. *J Biol Chem* 270:9322–9326.
62. Komine Y, Kitabatake M, Yokogawa T, Nishikawa K, Inokuchi H (1994) A tRNA-like structure is present in 10Sa RNA, a small stable RNA from *Escherichia coli*. *Proc Natl Acad Sci USA* 91:9223–9227.
63. Keiler KC, Waller PR, Sauer RT (1996) Role of a peptide tagging system in degradation of proteins synthesized from damaged messenger RNA. *Science* 271:990–993.
64. Roche ED, Sauer RT (2001) Identification of endogenous SsrA-tagged proteins reveals tagging at positions corresponding to stop codons. *J Biol Chem* 276:28509–28515.
65. Moore SD, Sauer RT (2005) Ribosome rescue: tmRNA tagging activity and capacity in *Escherichia coli*. *Mol Microbiol* 58:456–466.
66. Gottesman S, Roche E, Zhou Y, Sauer RT (1998) The ClpXP and ClpAP proteases degrade proteins with carboxy-terminal peptide tails added by the SsrA-tagging system. *Genes Dev* 12:1338–1347.
67. Farrell CM, Grossman AD, Sauer RT (2005) Cytoplasmic degradation of ssrA-tagged proteins. *Mol Microbiol* 57:1750–1761.
68. Baker TA, Sauer RT (2012) ClpXP, an ATP-powered unfolding and protein-degradation machine. *Biochim Biophys Acta* 1823:15–28.
69. Wah DA, Levchenko I, Baker TA, Sauer RT (2002) Characterization of a specificity factor for an AAA+ ATPase: assembly of SspB dimers with ssrA-tagged proteins and the ClpX hexamer. *Chem Biol* 9:1237–1245.
70. Siddiqui SM, Sauer RT, Baker TA (2004) Role of the processing pore of the ClpX AAA+ ATPase in the recognition and engagement of specific protein substrates. *Genes Dev* 18:369–374.
71. Martin A, Baker TA, Sauer RT (2008) Diverse pore loops of the AAA+ ClpX machine mediate unassisted and adaptor-dependent recognition of ssrA-tagged substrates. *Mol Cell* 29:441–450.
72. Kim YI, Burton RE, Burton BM, Sauer RT, Baker TA (2000) Dynamics of substrate denaturation and translocation by the ClpXP degradation machine. *Mol Cell* 5:639–648.
73. Burton RE, Siddiqui SM, Kim YI, Baker TA, Sauer RT (2001) Effects of protein stability and structure on substrate processing by the ClpXP unfolding and degradation machine. *EMBO J* 20:3092–3100.
74. Kenniston JA, Baker TA, Fernandez JM, Sauer RT (2003) Linkage between ATP consumption and mechanical unfolding during the protein processing reactions of an AAA+ degradation machine. *Cell* 114:511–520.
75. Kenniston JA, Baker TA, Sauer RT (2005) Partitioning between unfolding and release of native domains during ClpXP degradation determines substrate selectivity and partial processing. *Proc Natl Acad Sci USA* 102:1390–1395.
76. Lee C, Schwartz MP, Prakash S, Iwakura M, Matouschek A (2001) ATP-dependent proteases degrade their substrates by processively unraveling them from the degradation signal. *Mol Cell* 7:627–637.
77. Aubin-Tam ME, Olivares AO, Sauer RT, Baker TA, Lang MJ (2001) Single-molecule protein unfolding and translocation by an ATP-fueled proteolytic machine. *Cell* 145:257–267.
78. Martin A, Baker TA, Sauer RT (2008) Protein unfolding by a AAA+ protease is dependent on ATP-hydrolysis rates and substrate energy landscapes. *Nat Struct Mol Biol* 15:139–145.
79. Nager AR, Baker TA, Sauer RT (2011) Stepwise unfolding of a β barrel protein by the AAA+ ClpXP protease. *J Mol Biol* 413:4–16.
80. Barkow SR, Levchenko I, Baker TA, Sauer RT (2009) Polypeptide translocation by the AAA+ ClpXP protease machine. *Chem Biol* 16:605–162.
81. Bolon DN, Grant RA, Baker TA, Sauer RT (2009) Nucleotide-dependent substrate handoff from the SspB adaptor to the AAA+ ClpXP protease. *Mol Cell* 16:343–350.
82. Martin A, Baker TA, Sauer RT (2005) Rebuilt AAA+ motors reveal operating principles for ATP-fueled machines. *Nature* 437:1115–1120.
83. Martin A, Baker TA, Sauer RT (2008) Pore loops of the AAA+ ClpX machine grip substrates to drive translocation and unfolding. *Nat Struct Mol Biol* 15:1147–1151.
84. Glynn SE, Martin A, Nager AR, Baker TA, Sauer RT (2009) Structures of asymmetric ClpX hexamers reveal nucleotide-dependent motions in a AAA+ protein-unfolding machine. *Cell* 139:744–756.
85. Hersch GL, Burton RE, Bolon DN, Baker TA, Sauer RT (2005) Asymmetric interactions of ATP with the AAA+ ClpX6 unfoldase: allosteric control of a protein machine. *Cell* 121:1017–1027.
86. Glynn SE, Nager AR, Baker TA, Sauer RT (2012) Dynamic and static components power unfolding in topologically closed rings of a AAA+ proteolytic machine. *Nat Struct Mol Biol* 19:616–622.
87. Stinson BM, Nager AR, Glynn SE, Schmitz KR, Baker TA, Sauer RT (2013) Nucleotide binding and conformational switching in the hexameric ring of a AAA+ machine. *Cell* 153:628–639.
88. Crestfield AM, Stein WH, Moore S (1963) Properties and conformation of the histidine residues at the active site of ribonuclease. *J Biol Chem* 238:2421–2428.