The Kinetic and Molecular Basis for the Interaction of LexA and Activated RecA Revealed by a Fluorescent Amino Acid Probe

Zachary M. Hostetler¹, Michael B. Cory², Chloe M. Jones², E. James Petersson^{3,4,*}, Rahul M. Kohli^{4,5,*}

¹Graduate Group in Cell and Molecular Biology, University of Pennsylvania, Philadelphia, PA 19104, United States ²Graduate Group in Biochemistry and Molecular Biophysics, University of Pennsylvania, Philadelphia, PA 19104, United States 3Department of Chemistry, University of Pennsylvania, Philadelphia, PA 19104, United States 4Department of Biochemistry and Biophysics, University of Pennsylvania, Philadelphia, PA 19104, United States 5Department of Medicine, University of Pennsylvania, Philadelphia, PA 19104, United States

Supporting Information

Methods

Supplemental Figures

Supplemental Tables

Recombinant gene cloning and protein purification

Incorporation of acridonylalanine (Acd or δ) into overexpressed LexA variants was accomplished as previously described,¹ with the following notes and changes. *E. coli* strains used for overexpression were transformed with two plasmids, one containing the recombinant *lexA* gene bearing the amber stop codon at the desired position for Acd incorporation, and the other encoding the aminoacyl tRNA synthetase with improved specificity for Acd incorporation.² To scale up expressions, cultures were grown in 400 mL of MDA-5052 autoinduction media³ in 4 L baffled flasks with solubilized Acd added to 1 mM. After growth for 24 hours, cells were harvested and stored at −80 °C. Cells were resuspended in buffer containing 20 mM sodium phosphate pH 7.2, 500 mM NaCl, and 25 mM imidazole; cells were then lysed using a benchtop microfluidizer (Microfluidics). Lysates were centrifuged for 40 min at 20,000 rpm. Clarified lysates were affinity purified with HisPur resin (Thermo Fisher) using a 25 to 400 mM imidazole gradient. Fractions containing purified Acd-labeled LexA eluted in the second half of the gradient. Purified LexA was dialyzed into buffer containing 50 mM Tris-HCl pH 7.6 (at 4 °C), 200 mM NaCl, 0.5 mM EDTA, and 10% glycerol and then stored at −80 °C.

Plasmids bearing the genes for N- or C-terminally HIS-tagged LexA were used as the starting point for generating catalytically inactive unlabeled LexA truncations, proteolytic fragments, or point mutants. Oligonucleotides used in Phusion polymerase site-directed mutagenesis or in Gibson assembly were ordered from Integrated DNA Technologies (IDT) and the resulting plasmids are available upon request. Successful cloning was confirmed through restriction mapping of the plasmids and sequencing the insert gene (GeneWiz). For protein overexpression, plasmids were transformed into BL21(DE3) cells. Overnight cultures of transformed cells were used to inoculate 400 mL of LB media with 30 μg/mL kanamycin. Cultures were grown at 37 °C to an OD595 of 0.4–0.7 and induced with IPTG to a concentration of 1 mM. For LexA variants that aggregated under these expression conditions, they were instead switched to 16 °C after reaching the critical OD595 and induced with 0.1 mM IPTG and grown overnight. Cells were harvested and lysed similarly to the procedures for Acdlabeled proteins described above, with the exception that these were purified in batches according to recommended protocol (Thermo Fisher). Following dialysis, HIS-tags were removed via TEV protease cleavage, and remaining contaminants were removed by flowing the sample back over HisPur resin. Alkaline proteolytic fragments were similarly purified following the proteolysis step. Proteins were stored as above.

A plasmid encoding an N-terminally HIS-tagged wild type RecA gene¹ was transformed into BLR(DE3) cells (BL21(DE3) with *recA[−]* genotype). Cultures were grown in 2 L of LB media plus 30 μg/mL kanamycin at 37 °C for 2 hours and then harvested and stored at −80 °C. Purification of HIS-RecA was carried out as previously described for wild type RecA.⁴ Notably, HIS-RecA eluted from hydroxyapatite resin at significantly higher sodium phosphate concentration than that reported for wild type RecA. Following hydroxyapatite chromatography, RecA was deemed sufficiently pure for biochemical assays, and was dialyzed into storage buffer and stored as previously described.⁴

Mass spectrometry analysis of LexΑ^δ

20 μL of 22 μM LexA^δ was dissolved in sterile water and immediately analyzed by LC-MS on a Waters Acquity UPLC-TUV-SQD UPLCMS with ESI. LC separation for LexA^δ was carried out on a Waters Acquity UPLC BEH

2

C8 (1.7 μ m, 2.1x50 mm) column using 0.1% formic acid in H₂O as solvent A and 0.1% formic acid in acetonitrile as solvent B with a gradient of 5–95% solvent B over ten minutes. For the first two minutes, the column flow through was diverted to waste, the rest of the sample was then directed to the MS and analyzed in positive mode ion mode with m/z of 100–2000 in continuum mode. The parent protein mass was calculated from the observed multiply charged species using MassLynx 4.1 software (Waters).

RecA activation reactions

RecA* activation reactions were performed as follows, unless otherwise indicated in the text. A 2x concentrated reaction was prepared in reaction buffer containing 70 mM Tris-HCl pH 7.5, 10 mM MgCl₂, 150 mM NaCl, 5 mM DTT, 50 μg/mL BSA with 0.25 mM ATPγS, 16 μM poly(dT) ssDNA, and 4 μM HIS-RecA. The reactions were incubated overnight on ice, a condition previously described to result in optimal rates of LexA cleavage.⁵ For experiments involving concentration series of RecA*, the ssDNA substrate was maintained at a 4:1 ratio of nucleotides to RecA monomers. Non-activated RecA controls were prepared similarly, with the exclusion of ATPγS. Upon mixing with an equal volume of sample in the stopped-flow reaction chamber, the final concentration of RecA* was 2 μM (unless otherwise indicated in the text).

Alkaline-mediated and RecA*-mediated LexA auto-proteolysis assays

To assess the ability of catalytically active LexA^{δ} (LexA δ with S119A reverted to wild type; LexA-Q161 δ -HIS) to undergo auto-proteolysis, gel-based assays were used to examine LexA auto-proteolysis in the presence of activated RecA (RecA*) or alkaline pH. For the RecA*-mediated auto-proteolysis assay, RecA was activated as described above, with the following changes. 8 μM HIS-RecA was combined with a 30-mer poly-dT ssDNA substrate at a 1:4 molar ratio of RecA protein to nucleotides in activation buffer (70 mM Tris-HCl, pH 7.5, 10 mM MgCl2, 5 mM DTT, +/− 250 µM ATPγS) and incubated for 2 hours at room temperature. Activated RecA* was then added 1:1 with 16 μM HIS-LexA or LexA-Q161δ-HIS in activation buffer to start the reaction. At time zero, 5 μL RecA* and LexA were independently added to 10 μL of a stop mixture containing 2x Laemmli buffer (Bio-Rad) and boiled at 95 °C for 5 minutes. At the end of the reaction, 10 μL of reaction mixture was added to 10 μL of 2X Laemmli buffer and boiled at 95 °C for 5 minutes. 10 μL of each sample was loaded onto a 15% SDS-PAGE gel and ran at 200 V for 55 minutes and the gel was stained with Coomassie Brilliant Blue R-250 (Bio-Rad) to visualize cleavage bands.

For the alkaline-mediated auto-proteolysis assay, 8 μM of HIS-LexA or LexA-Q161δ-HIS was mixed with an equal volume of alkaline pH reaction buffer (100 mM Tris-Glycine-CAPS pH 10.0, 300 mM NaCl) and incubated at room temperature. Samples for each indicated time point were processed as above. Following gel image acquisition, protein bands were quantified using ImageJ (NIH). Molecular weight-based correction factor multipliers were applied to the higher molecular weight cleavage fragments (the C-terminal LexA fragment) to account for mass-based differences in Coomassie staining. The percentage of cleaved LexA was calculated as a ratio of the full-length LexA band over the sum of the full-length and C-terminal fragment bands. Data were fit to single exponential decay curves from which intrinsic rate constants were extracted.

3

Fluorescence anisotropy measurements

All stopped-flow measurements were made in a KinTek AutoSF-120 stopped-flow spectrofluorometer, with the following settings. A Xenon/Mercury lamp (Hamamatsu) with a monochromator (Oriel) was set to emit light through a fiber optic at 385 nm with a 3.16 mm diffraction grating. Incident light was plane polarized using a Glen-Taylor polarized optic and passed into the observation chamber. Emitted light was filtered with two 440 +/− 40 nm emission filters (Edmund Optics) with polarized films in the vertical or horizontal direction for either photomultiplier tube (PMT). PMT voltages were calibrated so that a fluorescence reading of 5.0 was achieved, and a grating factor (G-factor) was obtained at the start of each instrument run, per instrument protocol. Buffermatched samples for each syringe were prepared at twice the final concentration indicated in the text prior to mixing. Sample measurements were collected for the indicated times, and anisotropy was reported as follows:

$$
r = \frac{I_{polarized}}{I_{total}} = \frac{I_{\parallel} - GI_{\perp}}{I_{\parallel} + 2GI_{\perp}}
$$

Here, r represents the fluorescence anisotropy, I the fluorescence intensity, I_{\parallel} the parallel polarized fluorescence intensity, I_1 the perpendicular polarized fluorescence intensity, and G the G-factor (the correction factor for differences in instrument sensitivity for parallel polarized versus perpendicular polarized light).

Measuring LexA:RecA* association rates and fitting a binding model

To track association rates, mixtures of LexA δ at 250 nM and RecA* at concentrations indicated in the text were monitored for changes in fluorescence anisotropy for 60 seconds. Following exclusion of low-quality traces, the remaining curves were averaged to find means and standard errors at each time point. The means and standard errors (sigma) at all times for each curve were imported into KinTek Explorer global fitting software, a program that simulates reactions from a user-specified model.⁶ Parameters in a one-step reversible binding model with LexA and RecA in a 1:1 stoichiometry were manually adjusted until the simulated curves approximately fit the imported data by visual inspection. Fits were optimized using built-in regression analysis in the software. Briefly, a set of parameters for the inputted model are identified that minimize Chi², defined as the sum of the residuals squared divided by the standard error (sigma) at each time point:

$$
\chi^2 = \sum_{i=0}^{N-1} \left(\frac{y_i - y(x_i)}{\sigma_i} \right)^2
$$

Next, the best-fit parameters are systematically varied to determine a range of values yielding acceptable fits based on the Chi² analysis and to reveal how well the model is constrained by the data.⁷ A Chi² threshold of 0.70 was selected as a moderately stringent boundary cut-off to allow for determination of upper and lower limits for each parameter. At this Chi², a high proportion of experimental data points fall within the predicted upper and lower boundaries for each curve by visual inspection. More complex binding models (additional binding steps, alternative stoichiometric ratios) were tested and excluded because these models were no longer wellconstrained by the data.

Measuring LexA:RecA* dissociation rates

To monitor the dissociation of LexA from RecA*, labeled LexA was incubated with non-activated or activated RecA to give final concentrations of 250 nM LexA and 2 μM RecA or RecA*, respectively. An equal volume of sample containing a 44mer dsDNA substrate⁸ containing the consensus operator sequence or a scrambled version of that operator sequence were used as quench molecules for LexA that dissociates from RecA. The dsDNA substrates were prepared in a buffer-matched mixture at a twenty-fold higher molar concentration than LexA. Reactions were mixed and fluorescence anisotropy monitored for the indicated time. An estimate of the dissociation rate from RecA* was generated by fitting the appropriate curve to a single exponential decay equation.

Competition binding assays with unlabeled LexA variants

A concentration series of unlabeled LexA variants was prepared for each experiment, using the range of concentrations indicated in the text. Each unlabeled LexA variant was prepared at four times (4x) the final concentration. Activated RecA reactions were also prepared at 4x the final concentration. Equal volumes of the 4x unlabeled LexA variants and 4x RecA* samples were mixed and incubated for 1 hour so that an equilibrium was reached. These mixtures, with both components now at twice (2x) the final concentration, were mixed with an equal volume of 2x LexA-S119A-Q161δ on the stopped flow spectrofluorometer.

Individual traces from competition binding assay were examined and low-quality traces were discarded. For each trace at a particular concentration of unlabeled LexA, a smoothed spline function representing the data is computed using a Gaussian Process smoothing function from the "mgcv" package for R. To approximate the initial rate for each smoothed curve, the instantaneous rate at time = 0 sec was estimated using a finite difference interpolation method. The instantaneous rate for Lex A^{δ} binding in the absence of unlabeled LexA was selected as the theoretical maximal rate; calculated instantaneous rates were normalized to this theoretical maximal rate and normalized values exceeding 105% of the theoretical maximal rate were discarded as outliers. Normalized rates were mathematically converted to percentages describing the degree of inhibition of LexA δ binding to RecA^{*}, where 100% inhibition represents no observed LexA δ binding.

To approximate the relative binding affinity of different unlabeled LexA variants, the percentage inhibition values were fit to variable slope sigmoidal dose-response curves as a function of unlabeled LexA variant concentrations. This permitted calculation of the half maximal effective concentration for any given unlabeled LexA variant, a value that reflects the concentration at which the binding of LexA δ is 50% inhibited. The data were fitted using the nls.multstart non-linear regression package in R. 95% confidence intervals for the best fit curves and model parameter estimates were obtained by bootstrapping the data in R and obtaining the 2.5 and 97.5 percentiles of each value. At least 1000 iterations were performed for each round of bootstrapping.

Figure S1: Identifying an Acd-labeled LexA variant suitable for a RecA* binding assay

- a) Cartoon model of the LexA dimer (PDB: 1JSO). One monomer (light gray) includes selected residues highlighted in space-filling representation (light blue); we previously demonstrated that residues at these positions tolerate incorporation of the fluorescent unnatural amino acid acridonylalanine (Acd or $δ$).¹ Shown in the inset is the chemical structure of Acd.
- b) SDS-PAGE gels of Acd-containing LexA constructs. For each construct, samples were mixed with an equal volume of 2x Laemmli gel loading buffer and run for 55 minutes at 200 V. From left-to-right, for each construct, the lanes include: (S) soluble fraction of cell extract, (FT) flow-through from HisPur affinity chromatography, (E) imidazole elution fraction. The molecular weights of bands in the protein ladder are indicated on the left of each gel.
- c) Schema depicting the reactions used to identify a signal change that can be attributed to specific binding of LexA to RecA*. LexA constructs containing Acd at one of each of the positions indicated in (a) are mixed with buffer alone, inactive RecA, or activated RecA* to determine baseline anisotropy signal, signal change due to nonspecific binding, and signal change due to specific binding, respectively.
- d) Plot of fluorescence anisotropy values (λ_{ex} 386, λ_{em} 440) for each LexA construct in (a) after allowing mixtures described in (b) to reach equilibrium.

Figure S2: Large scale expression and purification of LexA-S119A-Q161δ (LexAδ)

- a) SDS-PAGE gel of over-expressed and purified LexA δ (arrowhead), visualized by Coomassie staining. Lanes from left-to-right: (W) whole cell extract, (P) insoluble pellet from cell extract, (S) soluble fraction of cell extract, (F) HisPur column flow-through fraction, (wash) fractions from a low imidazole wash, (elution) fractions from high imidazole elution step.
- b) The same SDS-PAGE gel of over-expressed and purified LexA δ (arrowhead), visualized by UV illumination to examine the fluorescence of acridone.

Figure S3: Mass Spectrometry Analysis of LexAδ

- a) Electrospray ionization (ESI) of LexA^δ. (Top) Mass spectrum of 4.4 µM LexA^δ in the range between 515 and 1100 *m/z* at a retention time of 4.753 minutes. Identified *m/z* values are listed in the table to the right. (Bottom) Extracted *m*/z values used by Waters MassLynx 4.1 software to calculate LexA^δ molecular weight. All extracted values (listed in the table) are within 0.2 Da of the raw LexA δ mass spectrum peaks.
- b) Representative replicate plot of the molecular weight of LexA^δ measured by ESI-MS and calculated by Waters MassLynx 4.1 software. The estimated Lex A^{δ} molecular weight after five replicate runs was 23720 ± 2 Da (calculated 23704 Da). No replicate showed a mass at 23621 Da (denoted by asterisk on plot), which would indicate tyrosine misincorporation at the Q161 position.

Figure S4: Auto-proteolysis activity of LexA^δ compared to wild type LexA

- a) Representative SDS-PAGE gels demonstrating a time course of alkaline-mediated LexA autoproteolysis. Samples of wild type LexA (HIS-LexA) and catalytically-active LexA δ (LexA-Q161 δ -HIS) were mixed with an equal volume of reaction buffer at pH 10.0 and incubated at room temperature for up to 128 min. Arrowheads indicate the full-length LexA proteins, and asterisks represent the Cterminal auto-proteolytic fragment. Time points sampled (from left-to-right on gels) include the following: 0 min, 1 min, 2 min, 4 min, 8 min, 16 min, 32 min, 64 min, 128 min.
- b) Plot of alkaline-mediated LexA auto-proteolysis showing percent of full-length LexA remaining over time. The bands representing full-length and C-terminal auto-proteolytic fragments were quantitated and used to calculate the remaining percentage of each full-length LexA protein. Points represent the average values of three experimental replicates; error bars represent the standard deviation. Data were fit to a single exponential decay curve and best-fit rate constants (*k*) are shown for each sample.
- c) SDS-PAGE gel showing RecA*-mediated auto-proteolysis of wild type LexA and catalytically active LexA^δ proteins. Activated RecA reactions were prepared either with ATPγS or without ATPγS as a negative control, and then these were mixed with either LexA protein and incubated for 90 minutes at room temperature. Arrowheads indicate the full-length LexA proteins, and asterisks represent the Cterminal auto-proteolytic fragment.

Figure S5: One-step reversible binding model for the LexA:RecA* interaction

(Top) The one-step reversible binding reaction model with model parameters indicated in bold type. Fitted parameters include the forward and reverse rate constants, k_1 and k_{-1} , as well as two factors, a and b, used in the conversion of numerically integrated protein concentrations to simulated fluorescence anisotropy signals. (Bottom) Two-dimensional "FitSpace" confidence intervals for these fitted parameters, with higher Chi² ratios indicating better fits to the model. The parameters indicated for each plot are allowed to float around the best-fit value, and fits to the model, as indicated by the Chi² ratio, are plotted.

Figure S6: Dissociation of LexA from RecA*

- a) Experimental setup for measuring the dissociation of LexA from RecA*. Pre-incubated mixtures of LexA δ with non-activated or activated RecA* are rapidly mixed with an excess of 44 mer doublestranded DNA substrates containing the consensus SOS operator sequence or a scrambled operator sequence. The high affinity of LexA for DNA acts as a quench for LexA as it dissociates from RecA*.
- b) Plots of anisotropy changes for LexA^{δ} over time, with the combination of reagents indicated for each curve. A dashed line (black) overlaid on the RecA* + consensus dsDNA curve (purple) indicates a bestfit single exponential decay curve.

Figure S7: LexA auto-proteolytic fragments

- a) Schematic of LexA auto-proteolytic fragments. To generate the fragments, wild type full-length LexA protein is exposed to an alkaline pH, a condition in which it readily undergoes an auto-proteolytic reaction. The N-terminal or C-terminal fragments can be isolated via affinity chromatography depending on the placement of a hexahistidine tag.
- b) Coomassie-stained SDS-PAGE gel of overexpressed and purified LexA variants. From left-to-right, the sample lanes include full-length catalytically inactive LexA (1–202), the N-terminal proteolytic fragment (1–84), and the C-terminal proteolytic fragment (85–202).

Figure S8: LexA N-terminal truncations

- a) Schematic showing the design of the LexA N-terminal domain truncation series. All of these truncations are catalytically-inactive via the S119A amino acid substitution.
- b) Coomassie-stained SDS-PAGE gel of overexpressed and purified LexA truncations.
- c) Plots of percent inhibition as a function of concentration of unlabeled LexA truncations. Data points represent the median value from at least four independent replicates. The solid line represents the best-fit variable-slope sigmoidal dose-response curve for each sample.

Table S1: Fitted values for reversible, one-step binding model parameters

Table S2: Summary of reversible one-step binding model fit results

Model Characteristic	Result		
Data Normalization	Mixed Sigma		
Number of experiments	3		
Number of curves	25		
Data points	10989		
Degrees of Freedom (DoF)	10983		
Sigma w.r.t. fit	0.00244968		
Chi ²	18518		
Chi ² / DoF	1.68606		
p-value	0		
Chi ² Threshold	0.998854		

Table S3: Competition binding assay results for LexA proteolytic fragments

LexA Variant		IC_{50} (μ M)		Hill Slope		
Name	Construct	Best fit	95% CI	Best fit	95% CI	
$"1 - 202"$	$LexA1-202-S119A$	0.34	0.20, 0.55	1.37	1.07, 9.73	
$"1 - 84"$	LexA _{1–84}	6600 *	0, ∞	$0.28*$	$-1.40, 2.32$	
"85-202"	LexA _{85–202}	66	53, 141	1.45	0.70, 2.43	
* parameter estimate p-value > 0.05						

Table S4: Competition bind assay results for LexA N-terminal truncations

Table S5: Competition binding assay results for LexA cleavage loop mutants

- 1. Hostetler, Z. M.; Ferrie, J. J.; Bornstein, M. R.; Sungwienwong, I.; Petersson, E. J.; Kohli, R. M. Systematic evaluation of soluble protein expression using a fluorescent unnatural amino acid reveals no reliable predictors of tolerability. *ACS Chem. Bio.* **2018***, 13*, 2855-2861.
- 2. Sungwienwong, I.; Hostetler, Z. M.; Blizzard, R. J.; Porter, J. J.; Driggers, C. M.; Mbengi, L. Z.; Villegas, J. A.; Speight, L. C.; Saven, J. G.; Perona, J. J.; Kohli, R. M.; Mehl, R. A.; Petersson, E. J. Improving target amino acid selectivity in a permissive aminoacyl tRNA synthetase through counter-selection. *Org. Biomol. Chem.* **2017***, 15*, 3603-3610.
- 3. Studier, F. W. Stable expression clones and auto-induction for protein production in E. coli. In *Structural Genomics*Springer: 2014; pp 17-32.
- 4. Shibata, T.; Osber, L.; Radding, C. M. Purification of recA protein from Escherichia coli. In *Methods in Enzymology*Elsevier: 1983; Vol. 100, pp 197-209.
- 5. Giese, K. C.; Michalowski, C. B.; Little, J. W. RecA-dependent cleavage of LexA dimers. *J. Mol. Biol.* **2008***, 377*, 148-161.
- 6. Johnson, K. A.; Simpson, Z. B.; Blom, T. Global kinetic explorer: a new computer program for dynamic simulation and fitting of kinetic data. *Anal. Biochem.* **2009***, 387*, 20-29.
- 7. Johnson, K. A.; Simpson, Z. B.; Blom, T. FitSpace explorer: an algorithm to evaluate multidimensional parameter space in fitting kinetic data. *Anal. Biochem.* **2009***, 387*, 30-41.
- 8. Culyba, M. J.; Kubiak, J. M.; Mo, C. Y.; Goulian, M.; Kohli, R. M. Non-equilibrium repressor binding kinetics link DNA damage dose to transcriptional timing within the SOS gene network. *PLoS Genet.* **2018***, 14*, e1007405.