Supporting Information

Ratiometric Pulse-chase Amidination Mass Spectrometry as a Probe of

Biomolecular Complex Formation

Feng-Ming James Chang, Matthew A. Lauber, William E. Running, James P. Reilly and David P. Giedroc*

Supplementary Methods, Table S1, Figures S1-S11

Supplementary Methods

Materials and Reagents. Hepes (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid), MES (2-(N-Morpholino)ethanesulfonic acid), Trizma base (2-Amino-2-hydroxymethyl-propane-1,3-diol), endoproteinase Glu-C and α-Chymotrypsin were obtained from Sigma. SMTA and SMTP were synthesized and characterized as outlined previously.[7]

Protein Purification. *Bsu* CsoR, expressed from a pET16b plasmid construct in a BL21 strain of *E. coli*, was purified as previously described[8] and obtained >90% pure. Protein integrity was verified by electrospray ionization/time of flight mass spectrometry at the Indiana University Department of Chemistry Mass Spectrometry Facility on a Waters/Micromass LCT Classic.

LC-ESI MS/MS Analysis of Acetamidinated/Propionamidinated Peptides.
Five μL of each 50 μL sample was injected onto a 15 mm x 100 μm i.d. trapping column packed in house with 5 μm 200 Å pore size Magic C18AQ material and flushed with solvent A (98/2/0.1 water / acetonitrile / formic acid). Peptides were then separated by reversed-phase nano-flow chromatography on a 150 mm x 75 μm i.d. analytical column

pulled to a tip and packed in house with 5 μm 100 Å pore size Magic C18AQ material

(Michrom Biosciences, Auburn, CA) using a 120-min gradient from 5% to 40% solvent B

(0.1% formic acid in acetonitrile) at 250 nL/min (Eksigent nanoLC-2D pump, Eksigent

Technologies, Dublin, CA). Peptides eluting from the column were electrosprayed

directly into the source of an ion-trap mass spectrometer (LCQ Deca XP Plus,

ThermoFinnigan, San Jose, CA) where a continuous cycle of one mass spectra

followed by two tandem mass spectra of the two most intense precursors were acquired.

Collision-induced dissociation (CID) was used to generate fragment ions in tandem

mass spectrometry. Dynamic exclusion was enabled such that each precursor m/z was

selected for tandem mass spectrometry no more than once over a 15 s window.

Ratiometric Quantitation. All MALDI-TOF mass spectra were acquired using

the Flex Control and Flex Analysis software (Bruker Daltonics, Billerica, MA). Integration

of individual species was done by summing the peak areas of the full isotopic

distribution. The peptides containing lysine residues were verified to be fully derivatized

(acetamidinated or propionamidinated) with little detectable unmodified peptide (far less

than 5% of the total).

For pulsed amidination, modified peptides can be resolved as either

acetamidinated (A, mass shifts of 41 Da) or propionamidinated (P, mass shifts of 55 Da).

The mol fraction of acetamidinated (A) peptide, $\Theta(A)$ is defined as

$$\Theta(A) = a(A)/[(a(A) + a(P)] \qquad \text{eq 1}$$

where $a(A)$ and $a(P)$ correspond to the isotope distribution peak area (a) of

acetamidinated and propionamidinated species, respectively. For peptides containing a

single lysine residue, *e.g.,* peptide 60-64 (Table S1), $\Theta(A^{K60}) = \Theta(A)$ and $\Theta(P^{K60}) = 1 -$

Θ(A). Similarly, the mol fractions of doubly, Θ(AA), and triply, Θ(AAA), acetamidinated

peptides in peptides with two lysines, *e.g.,* peptide 18-29, and three lysines, *e.g.,*

peptide 94-101, respectively, are defined as

$$\Theta(AA) = a(AA)/[(a(AA) + a(AP) + a(PP)] \qquad \text{eq 2}$$

$$\Theta(AAA) = a(AAA)/[(a(AAA) + a(AAP) + a(APP) + a(PPP)] \qquad \text{eq 3}$$

We then denote mol fraction of A^{K18}, P^{K26} as χ(A,P) in singly acetamidinated peptide 18-

29 (AP) where the comma in χ(A,P) is indicative of the sequential order in the peptide,

from N-terminus to C-terminus, *i.e.,* $\chi(A^{K18},P^{K26})$ is χ(A,P). χ(A,P) is calculated as the

ratio of the relative abundance (I) of the *b* ion with acetamidinated K18 to the sum of the

relative abundance of these *b* ions with acetamidinated and propionamidinated K18 *b*

ions. χ(A,P) can equivalently be calculated as the ratio of the relative abundance (I) of *y*

ion with propionamidinated K26 to the sum of the relative abundance of these *y* ions

with acetamidinated and propionamidinated K26. Thus, χ(A,P) is calculated as

$$\chi(A, P) = \frac{I(b, A)}{I(b, A) + I(b, P)} = \frac{I(y, P)}{I(y, A) + I(y, P)} \qquad \text{eq 4}$$

where I(*b*, A) and I(*y*, P) are defined as the relative abundance of *b* ion with

acetamidinated K18 and *y* ion with propionamidinated K26, respectively. A single

χ(A,P) value was determined for each unique *b* and *y* ions containing a single

acetamidinated or propionamidinated lysine, the values averaged for all *b* and *y* ions

containing that acetamidinated or proprionamidinated lysine, and the error taken as the

standard deviation of then mean for a χ(A,P) value.

   Similarly, the mol fraction P^{K18} and A^{K26} in singly acetamidinated peptide 18-29 is

defined as χ(P,A) and calculated according to

$$\chi(P,A) = \frac{I(b,P)}{I(b,A) + I(b,P)} = \frac{I(y,A)}{I(y,A) + I(y,P)} \qquad \text{eq 5}$$

The mol fraction of singly acetamidinated peptide 18-29 (acetamidinated at either K18

or K26), $\Theta(AP)$ is calculated from integration of the MALDI-TOF spectra according to

$$\Theta(AP) = a(AP)/[(a(AA) + a(AP) + a(PP)] \qquad \text{eq 6}$$

The mol fraction of A^{K18}, P^{K26} defined as $\Theta(A,P)$ and that of P^{K18} and A^{K26} defined as

$\Theta(P,A)$ for peptide 18-29 is then calculated by eq 7-8

$$\Theta(A,P) = \Theta(AP) \cdot \chi(A,P) \qquad \text{eq 7}$$

$$\Theta(P,A) = Q(AP) \cdot \chi(P,A) \qquad \text{eq 8}$$

with the $\Theta(A^{K18})$ and $\Theta(A^{K26})$ given by eqs 9-10

$$\Theta(A^{K18}) = [\Theta(A,P) + \Theta(AA)]/[\Theta(AA) + \Theta(A,P) + \Theta(P,A) + \Theta(PP)] \qquad \text{eq 9}$$

$$\Theta(A^{K26}) = [\Theta(P,A) + \Theta(AA)]/[\Theta(AA) + \Theta(A,P) + \Theta(P,A) + \Theta(PP)] \qquad \text{eq 10}$$

and $\Theta(P^{K18})$ and $\Theta(P^{K26})$ given by eqs 11-12

$$\Theta(P^{K18}) = [\Theta(P,A) + \Theta(PP)]/[\Theta(AA) + \Theta(A,P) + \Theta(P,A) + \Theta(PP)]=1 - \Theta(A^{K18}) \qquad \text{eq 11}$$

$$\Theta(P^{K26}) = [\Theta(A,P) + \Theta(PP)]/[\Theta(AA) + \Theta(A,P) + \Theta(P,A) + \Theta(PP)]=1 - \Theta(A^{K26}) \qquad \text{eq 12}$$

An exactly analogous scheme was used to calculate the mol fraction of acetamidinated

lysines in peptides containing three lysines, *e.g.,* peptide 94-101, beginning with eq 3 to

ultimately obtain $\Theta(A^{K96})$, $\Theta(A^{K97})$ and $\Theta(A^{K100})$ and corresponding $\Theta(P^{K96})$, $\Theta(P^{K97})$ and

$\Theta(P^{K100})$. To obtain first order rate constant for amidination, k, $\Theta(P^{Ki})$ was plotted as a

function of pulse time, t, and fitted to the $\Theta(P^{Ki}) = A_o \cdot \exp(-kt)$ (Table 1). In some cases,

a fit to a sum of two exponentials was used (Table 1).

Table S1. Calculated and observed mass[a] of acetamidinated (A)/propionamidinated (P) Glu-C[c] and chymotrypsin[d] digested peptides from *Bsu* CsoR

peptide	amino acid sequence	mass (Da)	Lys	Modification[b]	Mass of modified peptides (Da)	
					Calc'd	Obs'd
1-6[c]	**M**E**K**HNE(H)	787.34	α-NH$_2$ K3	AA	869.39	869.5
				AP	883.41	883.5
				PP	897.42	897.5
7-17[c]	(E)H**K**TLNH**K** SS**K**E(K)	1308.70	K8 K13 K17	AAA	1431.78	1431.8
				AAP	1445.80	1445.8
				APP	1459.81	1459.8
				PPP	1473.83	1473.8
18-29[c]	(E)**K**DQITNRL **K**RIE(G)	1513.88	K18 K26	AA	1595.93	1595.9
				AP	1609.95	1610.0
				PP	1623.96	1623.9
60-64[d]	(M)**K**NVAL(H)	544.35	K60	A	585.37	585.3
				P	599.39	599.3
74-92[d,e]	(H)CVADAI**K** SGDGEQAIS ELL(D)	1964.93	K80	A	2005.96	2005.9
				P	2019.97	2019.9
94-101[c]	(D)VF**KK**FT**K** S	984.59	K96 K97 K100	AAA	1107.67	1107.7
				AAP	1121.69	1121.8
				APP	1135.70	1135.8
				PPP	1149.72	1149.7

[a]Detected by MALDI-TOF mass spectrometry. Monoisotopic masses are shown. [b]A, acetamidinated, P, propionamidinated. [c]Glu-C digested peptide. [d]Chymotrypsin digested peptide. [e]Methyl disulfide formed at Cys24[1].

Figure S1. Acetamidinated CsoR binds to the *cop* operator DNA. Analytical G200 gel filtration chromatography of apo CsoR (A) vs. acetamidinated CsoR (B) mixed with 32 bp operator DNA in different monomer mol ratios of CsoR to DNA duplex. 4:1, one tetramer per DNA duplex; 8:1, two tetramers per duplex. CsoR-DNA complex elutes at 14.2 min and free DNA elutes at 15.9 min.
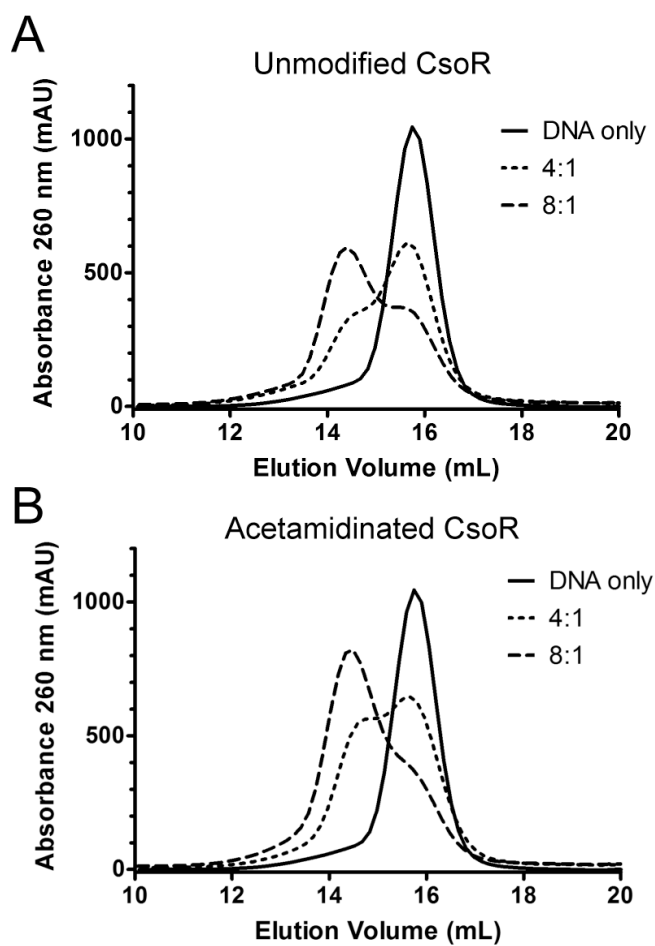
Figure S2. Extent of amidination as a function of reaction time of apo-CsoR, Cu(I)-bound CsoR and DNA-operator bound CsoR. Numbers of modified amines are calculated from the weighted average of *m/z* peaks in mass spectra analyzed by electrospray ionization mass spectrometry (ESI-MS). The smooth lines through the experimental data represent fits to a single exponential function $A=A_{max}(1-e^{-kt})$. A_{max} for apo-CsoR is 10.6 ± 0.2 lysines, with a global second order rate constant of 3.3 ± 0.2 M^{-1} min^{-1}. A_{max} for Cu(I)-bound CsoR is 9.5 ± 0.2 lysines, with second order rate constant of 3.9 ± 0.2 M^{-1} min^{-1}, and for the CsoR-DNA complex, 9.0 ± 0.4 lysines, with second order rate constant of 3.3 ± 0.4 M^{-1} min^{-1},
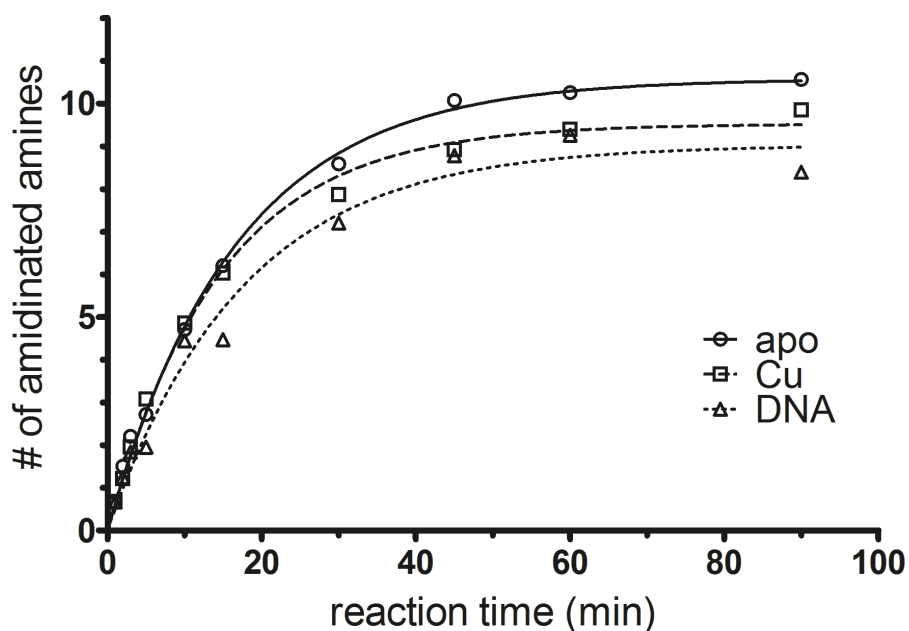
Figure S3. (A) Multiple sequence alignment of known and projected Cu(I)-sensing CsoRs and CsoR orthologs from several bacteria compared to *Bsu* CsoR (BSU_33520). CsoRs from *Staphylococcus aureus* Newman,[2] *Listeria monocytogenes*,[3] *Thermus thermophilus*,[4] and *M. tuberculosis* orthologs RicR[5] and CsoR[6] have been functionally characterized previously. The degree of conservation of all lysine residues in *Bsu* CsoR are represented as darkness of gray circles on the top with α-amino group labeled as an open circle. Two salt bridges formed between the α1 and α1' helices between two protomers within the dimeric four helix bundle indicated by the *red* dashed line (K17 and E31 in *Mtb* CsoR) and *green* dashed line (R20 and D24 in *Mtb* CsoR). R15 and R52 are required for DNA binding by *Mtb* CsoR. Secondary structure elements are labeled based on *Mtb* CsoR structure. (B) Structural illustration for Glu-C digested peptides (peptide 1-6, 7-17, 18-29 and 94-101) and chymotryptic peptides (peptide 60-64 and 74-92) in a homology model of *Bsu* CsoR calculated by threading the sequence of *Bsu* CsoR through the *Mtb* structure (pdb code 2HH7).[6] Predicted N-terminal and C-terminal unstructured regions are shown as dashed lines extending from the folded model that encompasses residues K13-K97. Peptides are highlighted in red and lysine residues are shown as blue sticks. The approximate position of the Cu(I) binding site is indicated between α2 and α2' helices on the periphery of the tetramer.
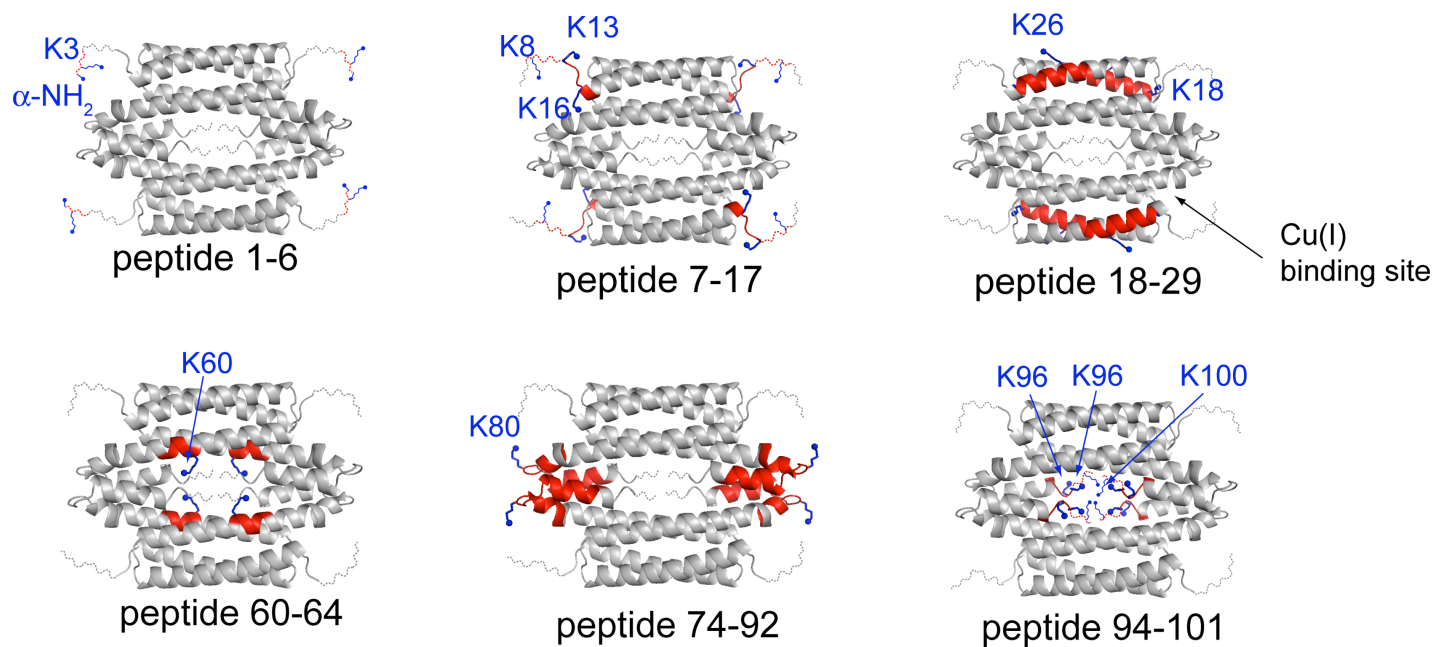
A

B

peptide 1-6

peptide 7-17

peptide 18-29

peptide 60-64

peptide 74-92

peptide 94-101

9
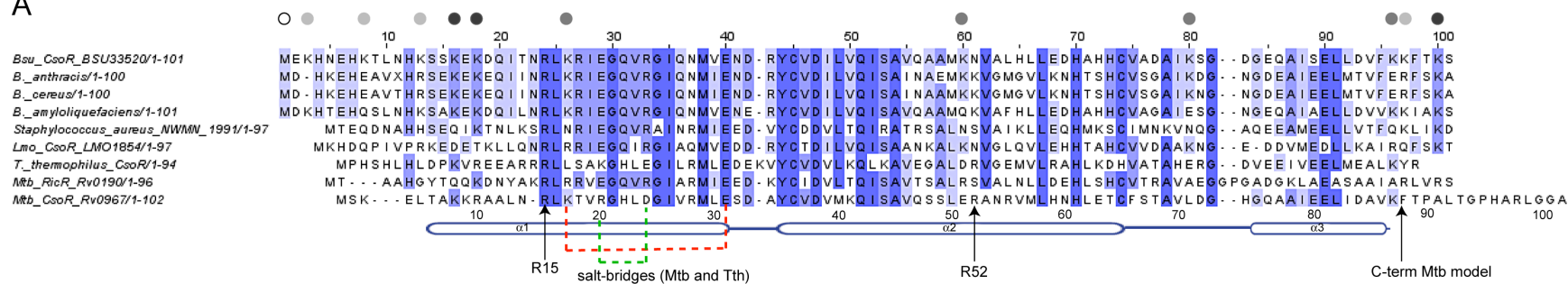
Figure S4. A series of MALDI-TOF mass spectra of chymotryptic peptide 60-64 are shown as a function of pulse amidination time *t* for apo-CsoR (A) and for the apo-CsoR-DNA complex (B). Isotope distribution of two different modified peptides are labeled as A (acetamidinated) and P (propionamidinated) peptides.
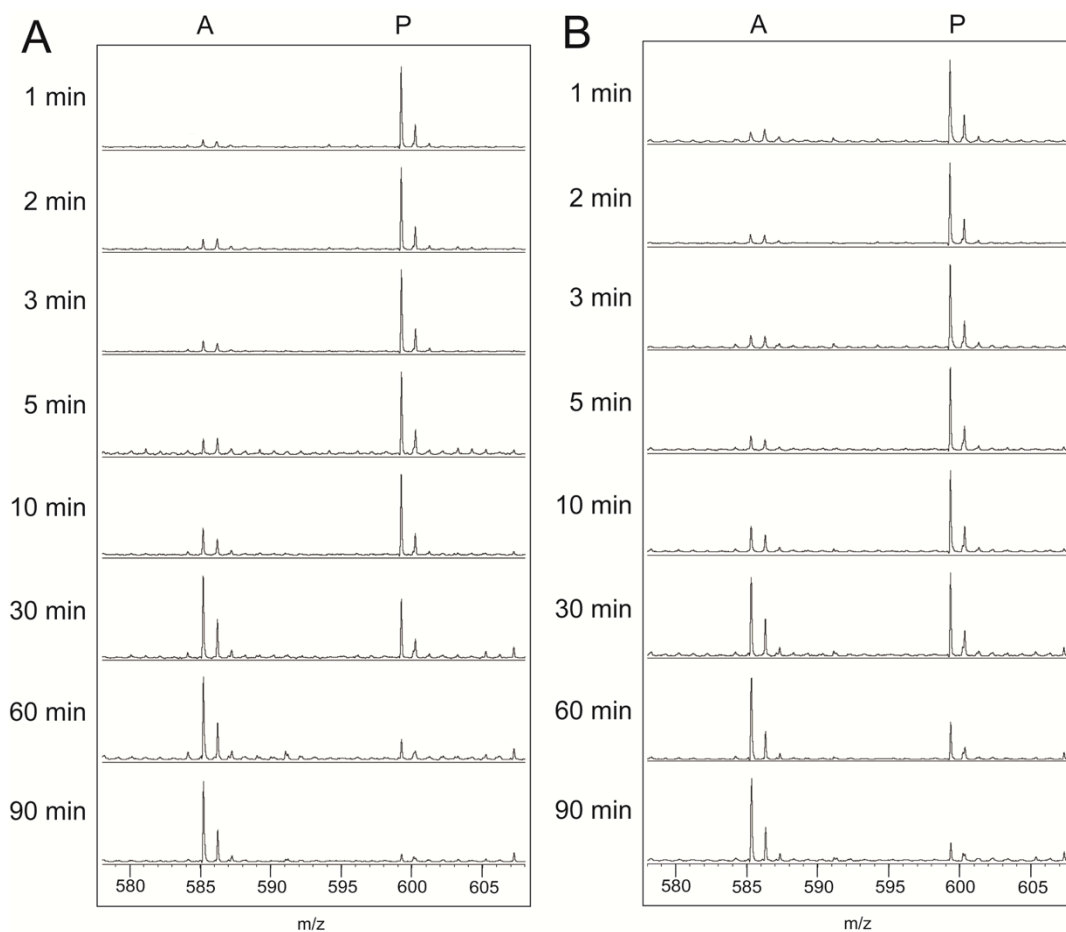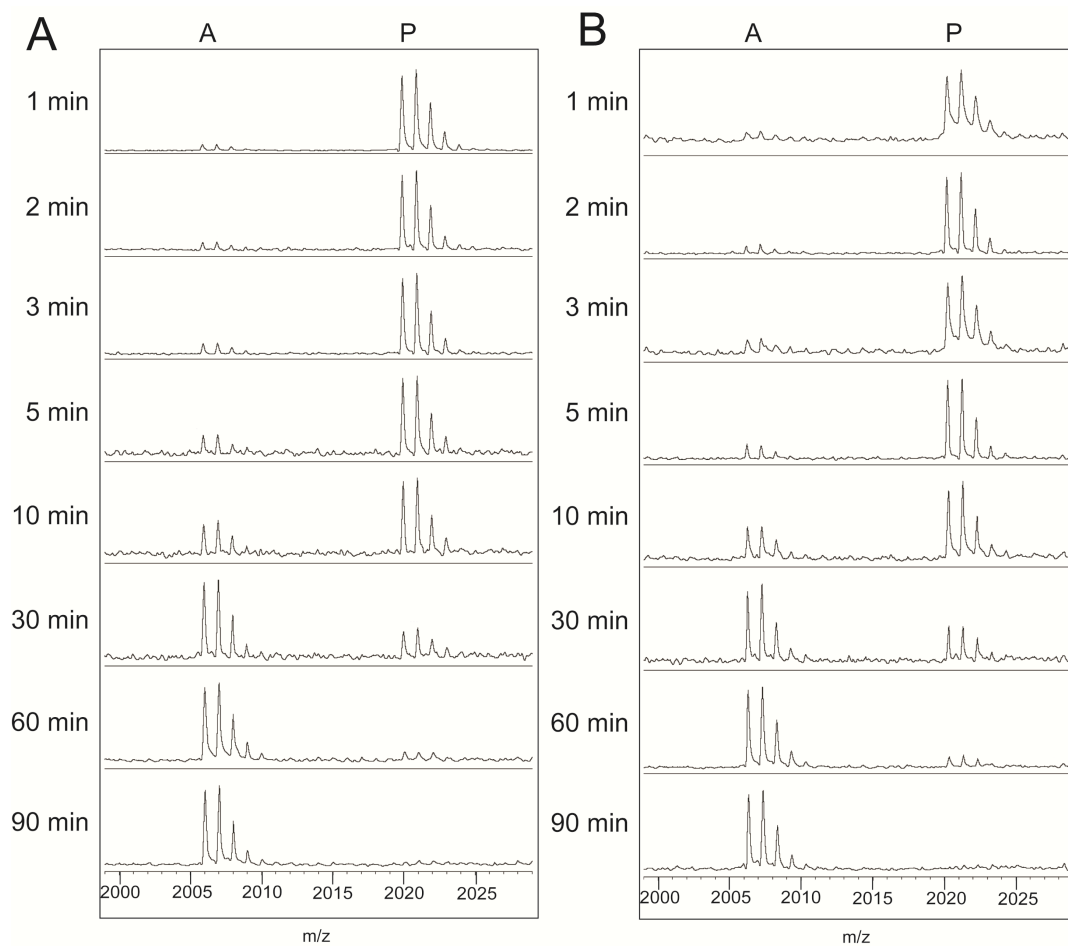
Figure S5. A series of MALDI-TOF spectra of chymotryptic peptide 74-92 are shown as a function of pulse amidination time in apo-CsoR (A) and in the DNA-CsoR complex (B).

Figure S6. Mol fraction of propionamidinated K60, Θ (P^{K60}) and K80, Θ (P^{K80}), as a function of pulse amidination time (t) in apo-CsoR (panels A and C) and in the DNA-CsoR complex (B and D). The continuous curve through the data points represents a fit to a single exponential decay function, Θ (P^i) = $A_o \cdot exp(-kt)$, with A_o and k given in Table 1, main text.
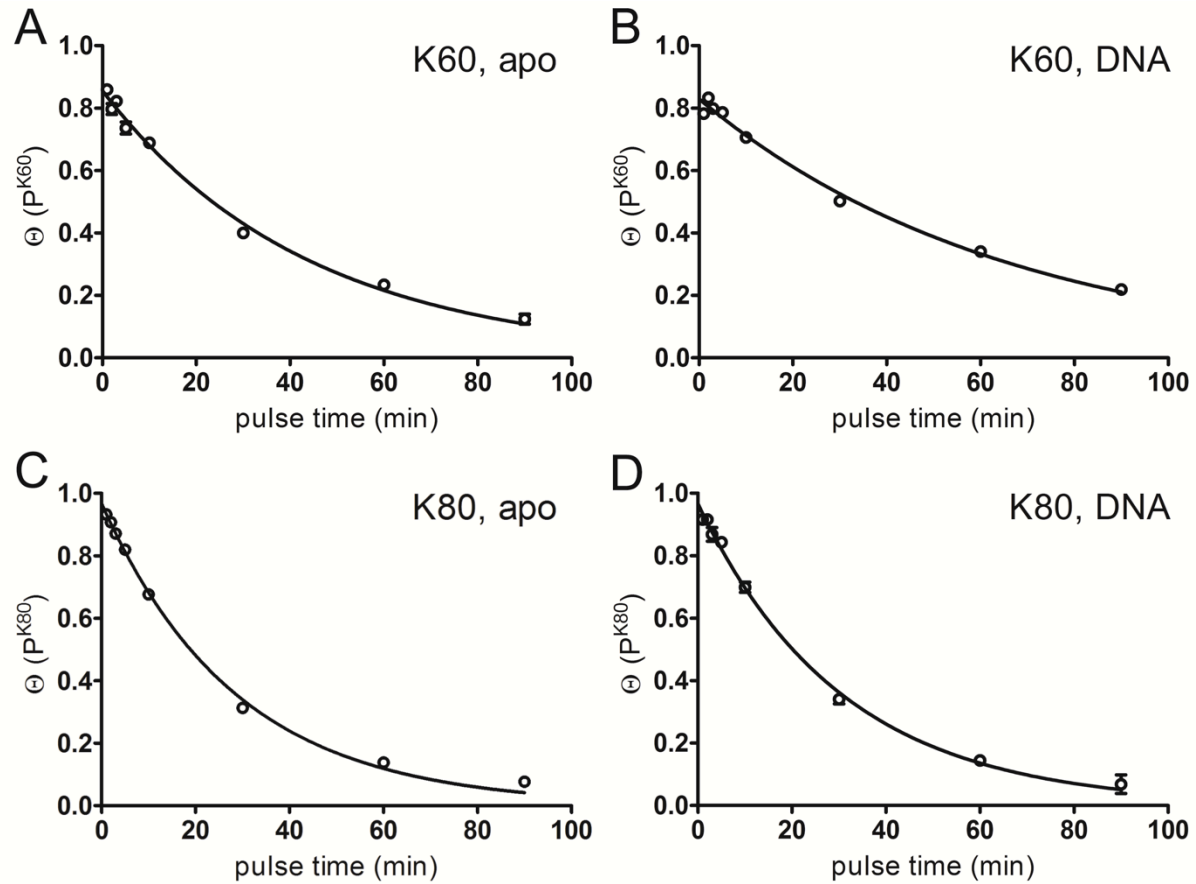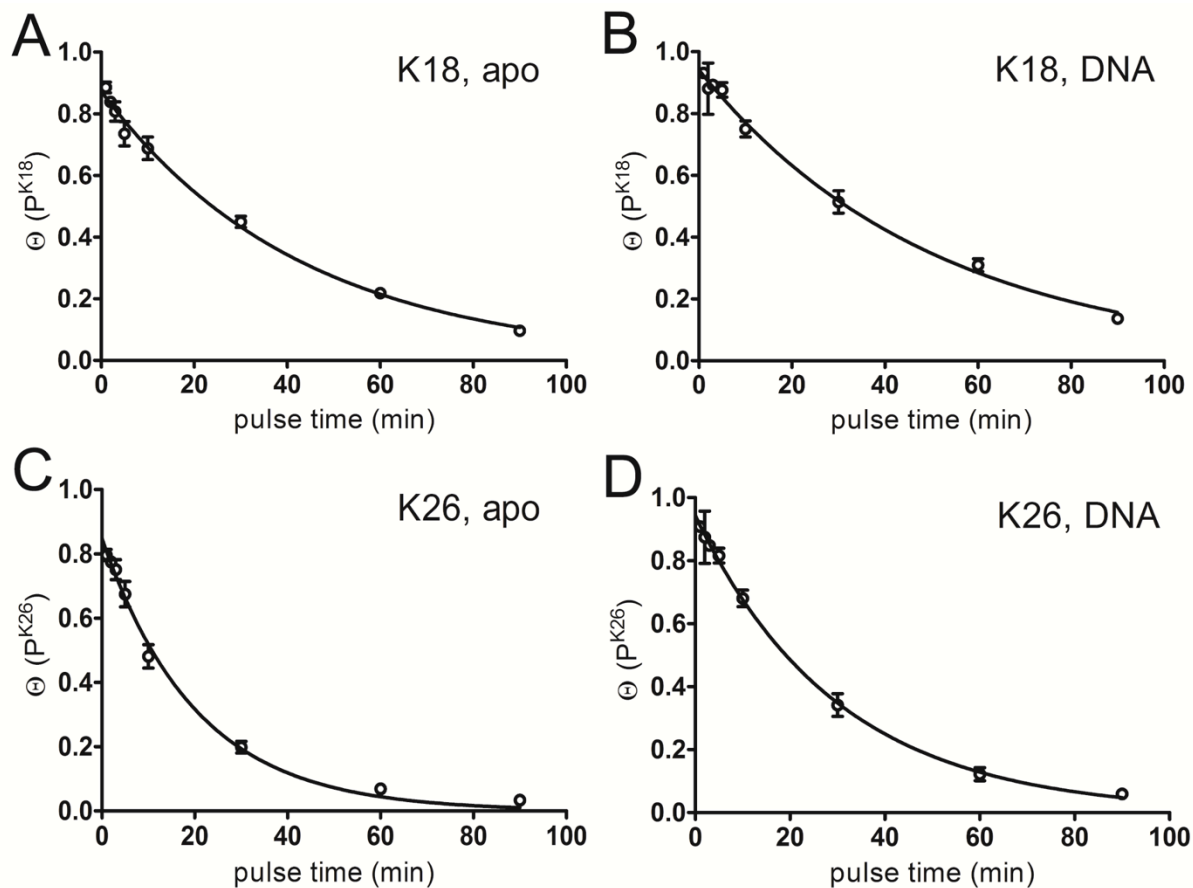
Figure S7. Mol fraction of propionamidinated K18, Θ (P^{K18}) and K26, Θ (P^{K26}), as a function of pulse amidination time (t) in apo-CsoR (panels A and C) and in the DNA-CsoR complex (B and D). The continuous curve through the data points represents a fit to a single exponential decay function, Θ (P^i) = $A_o \cdot \exp(-kt)$, with A_o and k given in Table 1, main text.

Figure S8. Mol fraction of propionamidinated α-amino group, Θ ($P^{N\text{-term}}$) and K3, Θ (P^{K3})) as a function of pulse amidination time in apo CsoR (A and C) and DNA-CsoR complex (B and D). Pseudo first order rate constants are determined by fitting the data with single exponential decay $\Theta = Ae^{-kt}$ for panels A, B, and D; panel C represents a fit to a sum of two exponentials. Fitted parameters are compiled in Table 1, main text.
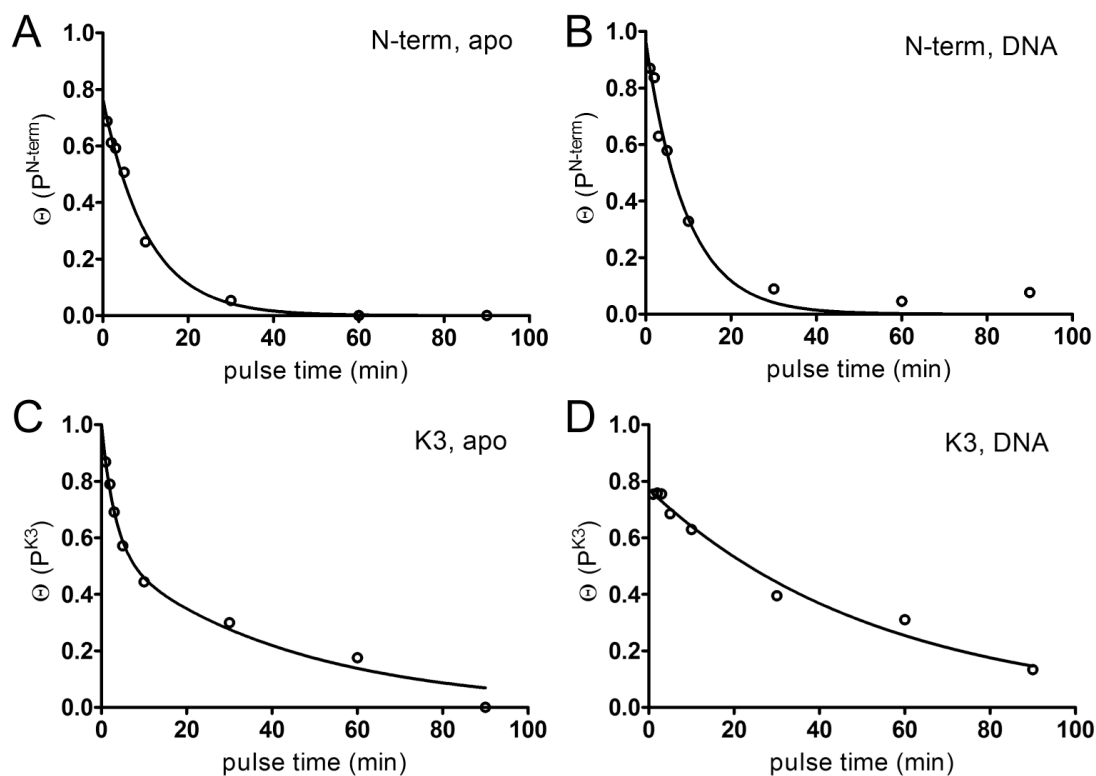
Figure S9. A series of MALDI-TOF spectra of Glu-C digested peptide 94-101 are shown as a function of pulse amidination time in apo-CsoR (A) and in the DNA-CsoR complex (B). The isotope distribution of four different modified peptides are labeled as AAA (triply-acetamidinated), AAP (doubly-acetamidinated), PPA (singly-acetamidinated) and PPP (triply-propionamidinated) with observed *m/z* shown (see Table S1).
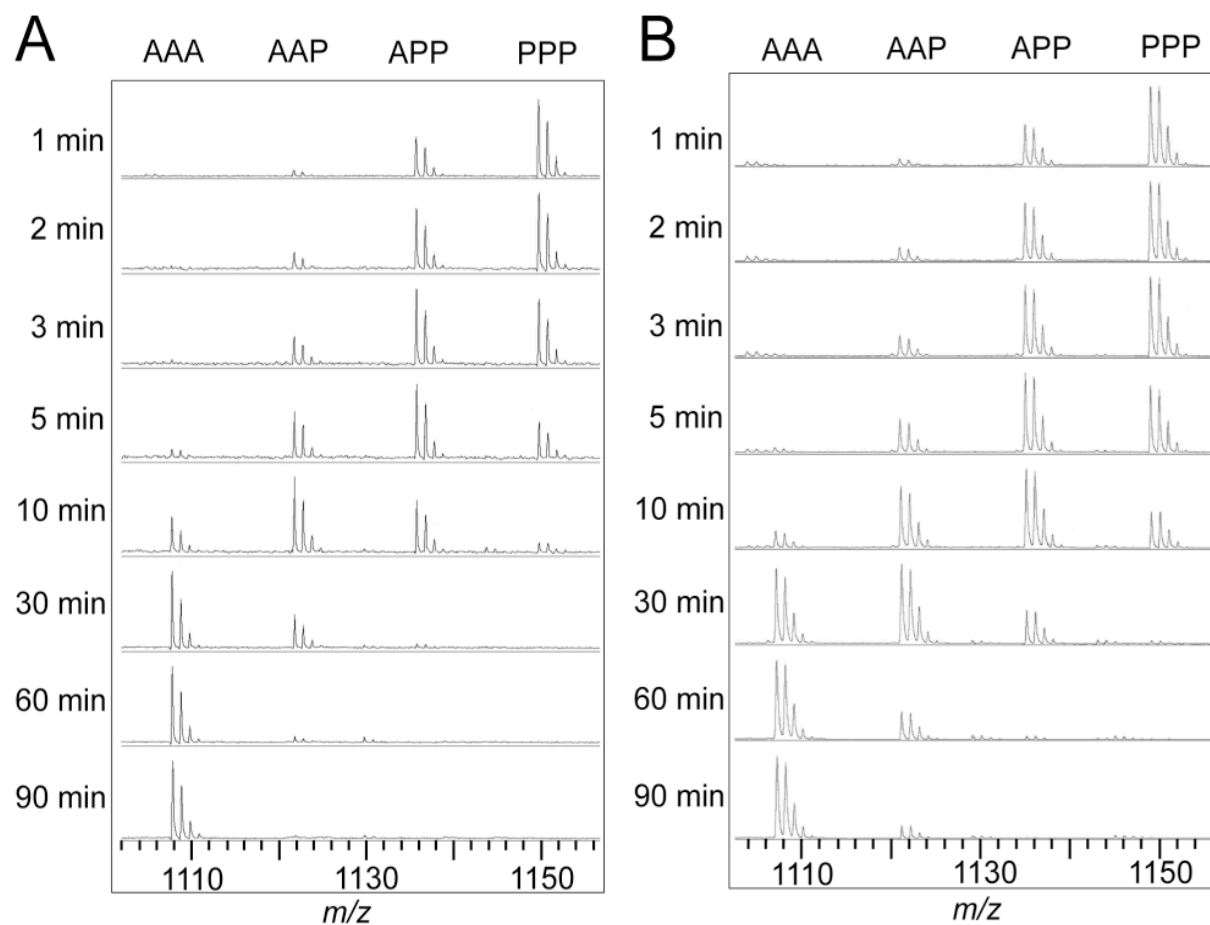
Figure S10. Mol fraction of propionamidinated K96, Θ (P^{K96}), K97, Θ (P^{K97}) and K100, Θ (P^{K100}) plotted as a function of pulse amidination time for apo CsoR (A, C and E) and the DNA-CsoR complex (B, D and F). The solid curves represent a fit to sum of two exponentials for apo-K96, while all other data are fitted to a single exponential function with parameters tabulated in Table 1, main text. The dashed line in panel A is a single exponential that is defined by the fast phase of the reaction.
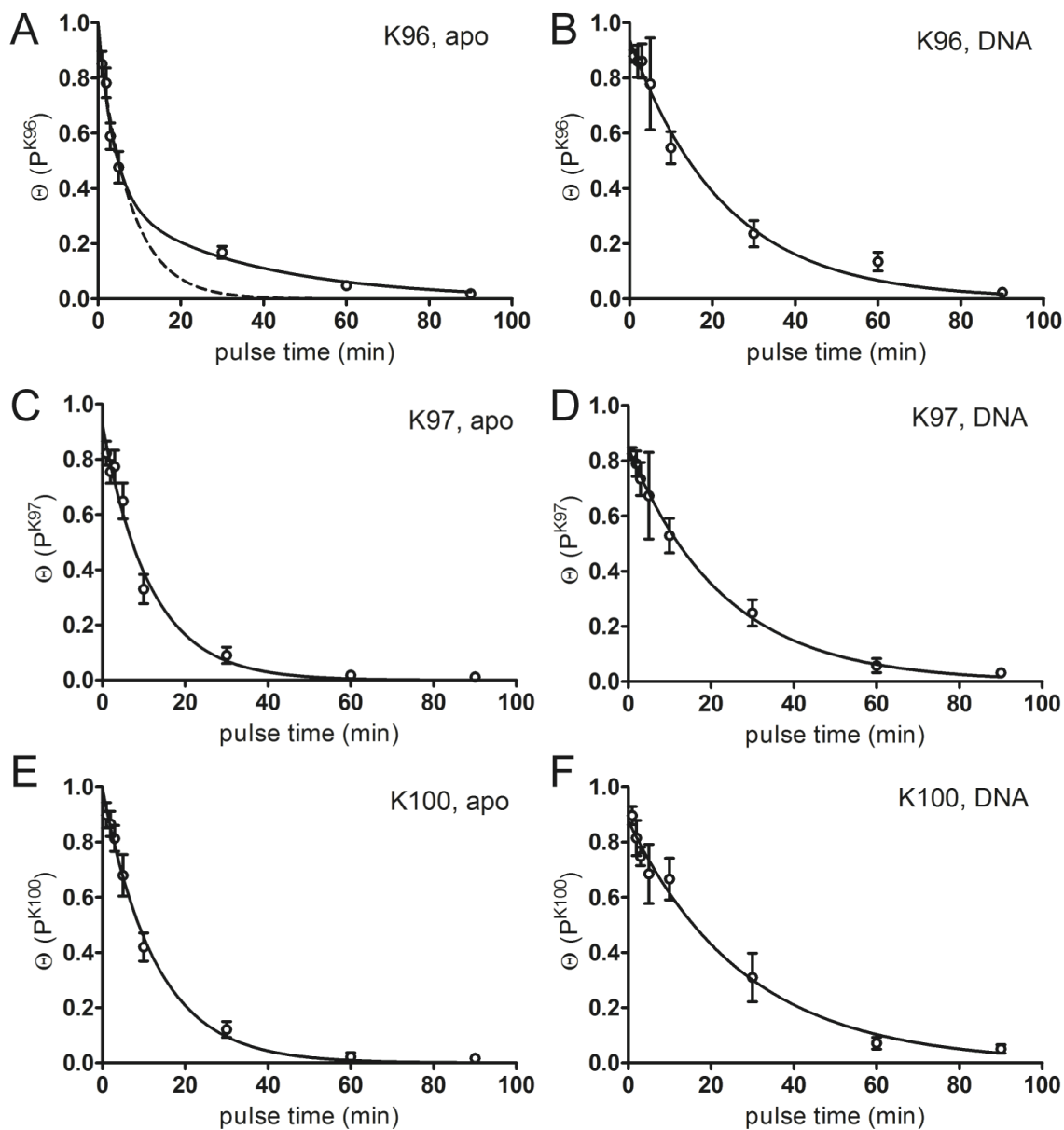
Figure S11. A space filling model of a possible structural relationship between two *Mtb* CsoR tetramers (shaded teal and blue) and a 32 bp DNA operator drawn to scale in a 2:1 complex consistent with previous data.[8] This is not a structure but is simply meant to conceptualize what a 2:1 *Bsu* CsoR tetramer:DNA complex might look like. The kinetics of amidination that we measure here may be reporting on a superposition of "bound" and "unbound, solvent exposed" DNA-binding "faces" of each of the tetramers bound to a single DNA operator in a 2:1 complex.
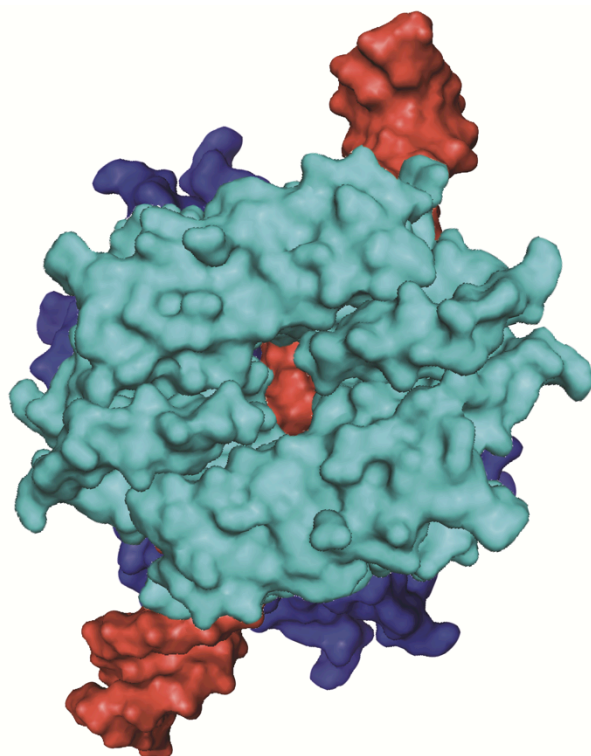
REFERENCES

1. Janecki, D. J., Beardsley, R. L. & Reilly, J. P. *Anal. Chem.* **2005**, 77, 7274-7281.
2. Grossoehme, N., Kehl-Fie, T. E., Ma, Z., Adams, K. W., Cowart, D. M., Scott, R. A., Skaar, E. P. & Giedroc, D. P. *J. Biol. Chem.* **2011**, 286, 13522-13531.
3. Corbett, D., Schuler, S., Glenn, S., Andrew, P. W., Cavet, J. S. & Roberts, I. S. *Mol. Microbiol.* **2011**, *81*, 457-472.
4. Sakamoto, K., Agari, Y., Agari, K., Kuramitsu, S. & Shinkai, *Microbiology* **2010**, 156, 1993-2005.
5. Festa, R. A., Jones, M. B., Butler-Wu, S., Sinsimer, D., Gerads, R., Bishai, W. R., Peterson, S. N. & Darwin, K. H. *Mol. Microbiol.* **2011**, 79, 133-48.
6. Liu, T., Ramesh, A., Ma, Z., Ward, S. K., Zhang, L., George, G. N., Talaat, A. M., Sacchettini, J. C. & Giedroc, D. P. *Nat. Chem. Biol.* **2007**, 3, 60-68.
7. Beardsley, R. L., Reilly, J. P. *J. Proteome Res.* **2003**, *2*, 15-21.
8. Ma, Z., Cowart, D. M., Scott, R. A., Giedroc, D. P. *Biochemistry* **2009**, *48*, 3325-3334.