Biophysical Journal, Volume *118*

Supplemental Information

Differential Local Stability Governs the Metamorphic Fold Switch of Bacterial Virulence Factor RfaH

Pablo Galaz-Davison, José Alejandro Molina, Steve Silletti, Elizabeth A. Komives, Stefan H. Knauer, Irina Artsimovitch, and César A. Ramírez-Sarmiento

Supporting Methods

Initial Structures. Deposited structures of αRfaH (PDB 5OND (1)) and βRfaH (PDB 6C6S (2)) were used for confinement MD simulations. Importantly, the DNA-bound full-length RfaH structure was used instead of the extensively used free RfaH structure (PDB 2OUG (3)). The rationale behind it is that the latter contains a 1-residue misplacement of the last 17 NTD residues (84-100) when compared with all other RfaH NTD structures (PDB 5OND, 6C6S and 6C6T). Since the interdomain linker is expected to be flexible when the protein is isolated in solution, both structures were processed using Rosetta3 (4). For this, fragments were generated for RfaH sequence using Robetta (5) (available at robetta.bakerlab.org), and used along with the LoopModel protocol to generate 500 structures of α- and β-folded full-length RfaH by relaxing this linker. These structures were then minimized by Gradient Descent algorithm, and later deeply minimized (i.e. with changes in rms lower than 10^{-12} Å) using Newton-Raphson Minimization, both implemented in Amber16 through NAB (6).

Free Energy Calculations through a Harmonic Oscillator Approach. Normal Mode Analysis (NMA) was performed for both structures using NAB, employing the AMBER ff14SB force field as we did with all MD/MM procedures indicated herein. The top 3N-6 positive frequencies, with N being the number of particles = 2,609, were used for computing the harmonic oscillator free energy as previously reported (7). Briefly, the harmonic oscillator free energy is (Eq. 1):

$$G_{HO} = -k_B T ln(z_{HO})$$

where z_{HO} is the partition function of the harmonic oscillator, k_B is the Boltzmann constant and T is the absolute temperature. The partition function corresponds to (Eq. 2):

$$z_{HO} = e^{-E/k_B T} \prod_{i=1}^{3N-6} \frac{k_B T}{v_i h}$$

where h is the Planck constant, E is the potential energy at the minimum, and v_i is the *i*-th frequency obtained from NMA, in the appropriate units. Then, free energy for each minimized structure used was calculated as (Eq. 3):

$$G_{HO} = E - k_B T \sum_{i=1}^{3N-6} ln(\frac{k_B T}{h_i v})$$

It should be noted that solving equation 3 for two harmonic oscillators having the same number of particles results in (Eq. 4):

$$\Delta G_{HO} = \Delta E - k_B T \sum_{i=1}^{3N-6} ln(\frac{v_2}{v_1})$$

This shows that only the natural logarithm of the ratio between the frequencies is relevant for the entropic contribution (rightmost summation) of the free-energy difference and implies that as long as v_1 and v_2 are expressed in the same frequency units, the energy difference can be calculated without explicitly evaluating equation 3.

Confinement Simulations and Free Energy Calculations. The aforementioned structures were used as starting configurations for implicit solvent (HCT (8)) confinement MD simulations. In these, a cartesian harmonic constraint is applied on each atom to drive it towards its deeply minimized position. These simulations are carried out for 30 ns at 298 K, using Langevin thermostat alongside SHAKE (9) for hydrogens. No cutoff was used for electrostatics since no PBC was used. In these simulations, the stiffness of the harmonic potential (restraining constant, k) was increased from $k_i = 2.5 \cdot 10^{-5}$ kcal mol^{-1} Å^{-2}, doubling up 25 times until reaching $k_f = 419.2$ kcal mol^{-1} Å^{-2}. For calculating the energy involved in the confinement step for the entire protein as well as for each residue, the squared of the distance of each atom with respect to the minimized structure ($\chi_k = \langle N \cdot \mathrm{RMSD}^2 \rangle_k$, where N is the number of atoms) was averaged throughout each simulation for each structure. As indicated in previous works, these fluctuations decrease exponentially with the increase of the restraining constant ($\chi \approx ak^b$) (Fig. S1) (10). Thus, the free energy was calculated simply as the area below the k, χ curve (Eq. 5):

$$\Delta G_{conf} = \int_{k_i}^{k_f} ak^b dk$$

where k is the restraining constant, a and b are unknown parameters. Since this behavior is not monotonic throughout the confinement steps, trapezoidal numerical

integration for each k_i, k_{i+1} pair is used instead, which can be improved from a linear to an exponential approximation by instead using the primitive of the solution to equation 5 (10, 11) (Eq. 6):

$$\int_{k_i}^{k_{i+1}} ak^b dk = \frac{ak^{b+1}}{b+1}|_{k_i}^{k_{i+1}} = \frac{(ak^b)k}{b+1}|_{k_i}^{k_{i+1}} = \frac{\chi_{i+1}k_{i+1} - \chi_i k_i}{b+1}$$

This shows that only b is required for the numerical integration, which can be isolated from the initial equation by evaluation between two values (7) (Eq. 7):

$$\chi = ak^b;\ \left(\frac{\chi_i}{k_i^b}\right) = \left(\frac{\chi_{i+1}}{k_{i+1}^b}\right);\ b = \frac{ln(\chi_{i+1}) - ln(\chi_i)}{ln(k_{i+1}) - ln(k_i)}$$

Applying the numerical approach to equation 5 results in (Eq. 8):

$$\Delta G_{conf} = \sum_{k_i}^{k_{f-1}} \frac{(\chi_{i+1}k_{i+1} - \chi_i k_i)}{(\frac{ln(\chi_{i+1}) - ln(\chi_i)}{ln(k_{i+1}) - ln(k_i)}) + 1}$$

For a more detailed breakdown of this sum please see (7). This free energy can be broken down into its per-residue contribution just by considering the protein fluctuations to be the results from individual residue contributions (Eq. 9):

$$\chi = \sum_{i=1}^{L} r_i$$

where L is the protein length and r is the squared atom fluctuation for a residue with respect to its position in the minimized structure (12).

Free Energy Difference and Decomposition. Since we cannot decompose the contribution from the normal mode analysis, we used the same approach previously reported, consisting of calculating the change in internal free energy (ΔU) for each residue using Amber16 module *decomp* (without 1,4 long range) (12). The free energy for each structure (and residue) was calculated as (Eq. 10):

$$G = G_{HO} - \Delta G_{conf}$$

therefore, the free energy difference $\Delta\Delta G$ between any pair of structures can be easily calculated as their difference (11). In the case of the per-residue free energy change ($\Delta\Delta G_r$) it is calculated as (Eq. 11):

$$\Delta\Delta G_r = \Delta U_r - \Delta G_{conf(r)}$$

where $\Delta G_{conf(r)}$ is the residue free-energy difference in the confinement step, and the subscript *r* indicates single-residue potential.

Peptide sequences and deuteron incorporation. After pepsin digestion, 27 different peptic peptides were identified for the isolated CTD, 42 for the full-length RfaH protein, and 51 for the full-length NusG (Fig. S3, Tables S1 and S2). To maximize sequence resolution, two considerations were taken: (i) incorporation was calculated for the shortest available peptic peptides; (ii) for two overlapping peptides whose sequence differs only in one overhanging bit (i.e. ACE and ACE**DF**), the deuteron uptake of the overhanging region corresponds to the difference in incorporation between the two peptides. For accuracy, the uncertainty (standard deviation, SD) of each individual peptide was considered and was propagated towards the difference peptide as the sum of their variances. If the resulting SD resulted in more than 20% of the differential uptake along the time intervals, a longer peptide was used instead. For this analysis, only the incorporating amides were considered, therefore the maximum incorporation follows the equation (13) (Eq. 12);

$$N = L_{peptide} - n_{pro} - 1$$

with $L_{peptide}$ being the length of the peptide and n_{pro} the number of proline residues contained in its sequence. The -1 arises from the fast exchange that takes place at the N-terminal of the protein or peptic peptides. However, for most overlapping peptides, the fast exchange of the N-terminal is already taken into account, thus their maximum incorporation was not corrected again for fast exchange.

With the resulting peptic peptides and differential regions calculated (Tables S3-S5), their deuteron uptake was fitted to a single negative exponential as shown below (Eq. 13):

$$\Delta mass_t = \Delta mass_{sat} - \Delta mass_{sat} e^{-kt}$$

where the $\Delta mass_{sat}$ corresponds to a fitting parameter representing maximum deuteron incorporation as obtained from the experiment and k is the global rate of deuteron incorporation (Fig. S3).

The deuteration extent (% deuteration) was calculated simply as the percentage of the maximum saturation reached by $\Delta mass_{sat}$. For a graphical representation, in the differential deuteration extents between the native forms of RfaH and between βRfaH and NusG, the free-amino ends resulting from peptic cleavage were assumed to share the same deuteration as the rest of the peptide.

Supporting Figures

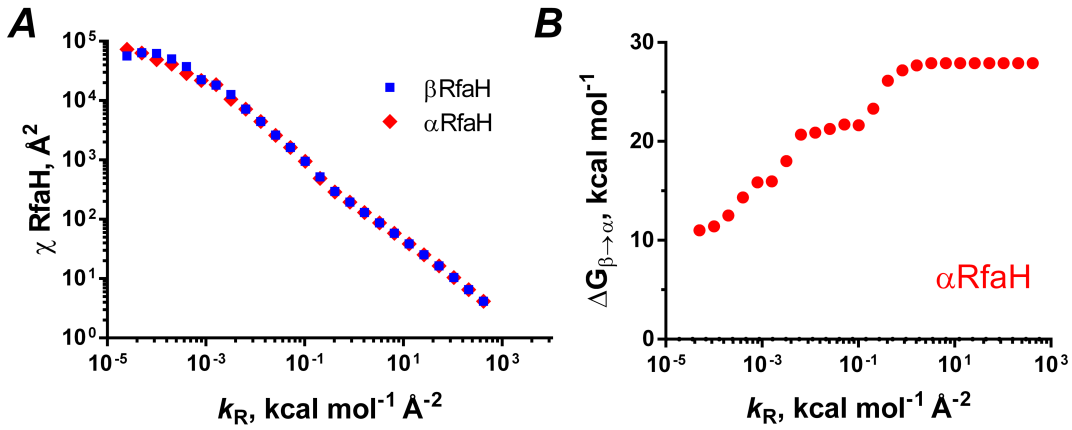


FIGURE S1 Dependence of fluctuations and concomitant free energy on the restraining potential. (*A*) Exponential decrease of the global squared atomic fluctuations χ of αRfaH and βRfaH with the increase of the restraining constant k_R. It can be observed that after a restraining potential of ~1 kcal mol^{-1} Å^2 both systems display the same fluctuation even for different configurations. (*B*) Free-energy difference between both RfaH states. In this, the summation of the contribution of the harmonic oscillator (11 kcal mol^{-1}) is the starting point (unconfined) and the free energy difference during confinement (area below the curve in *A*) is added at each integration step.

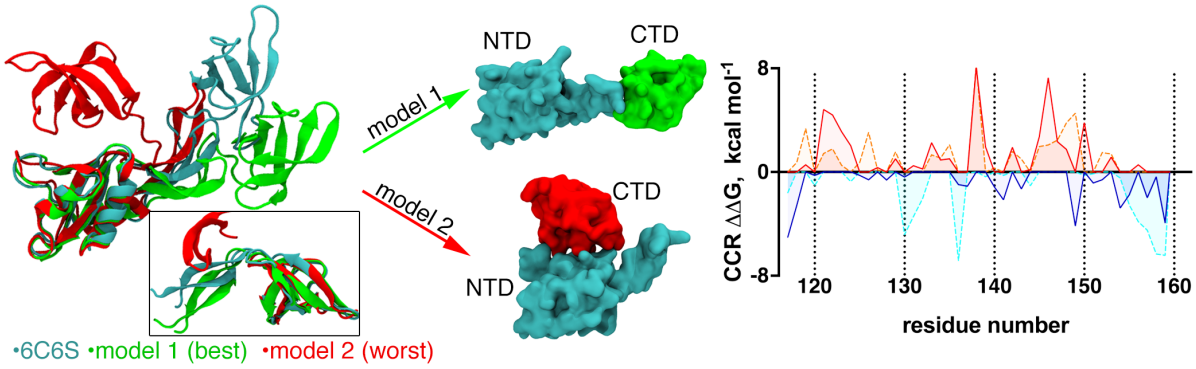
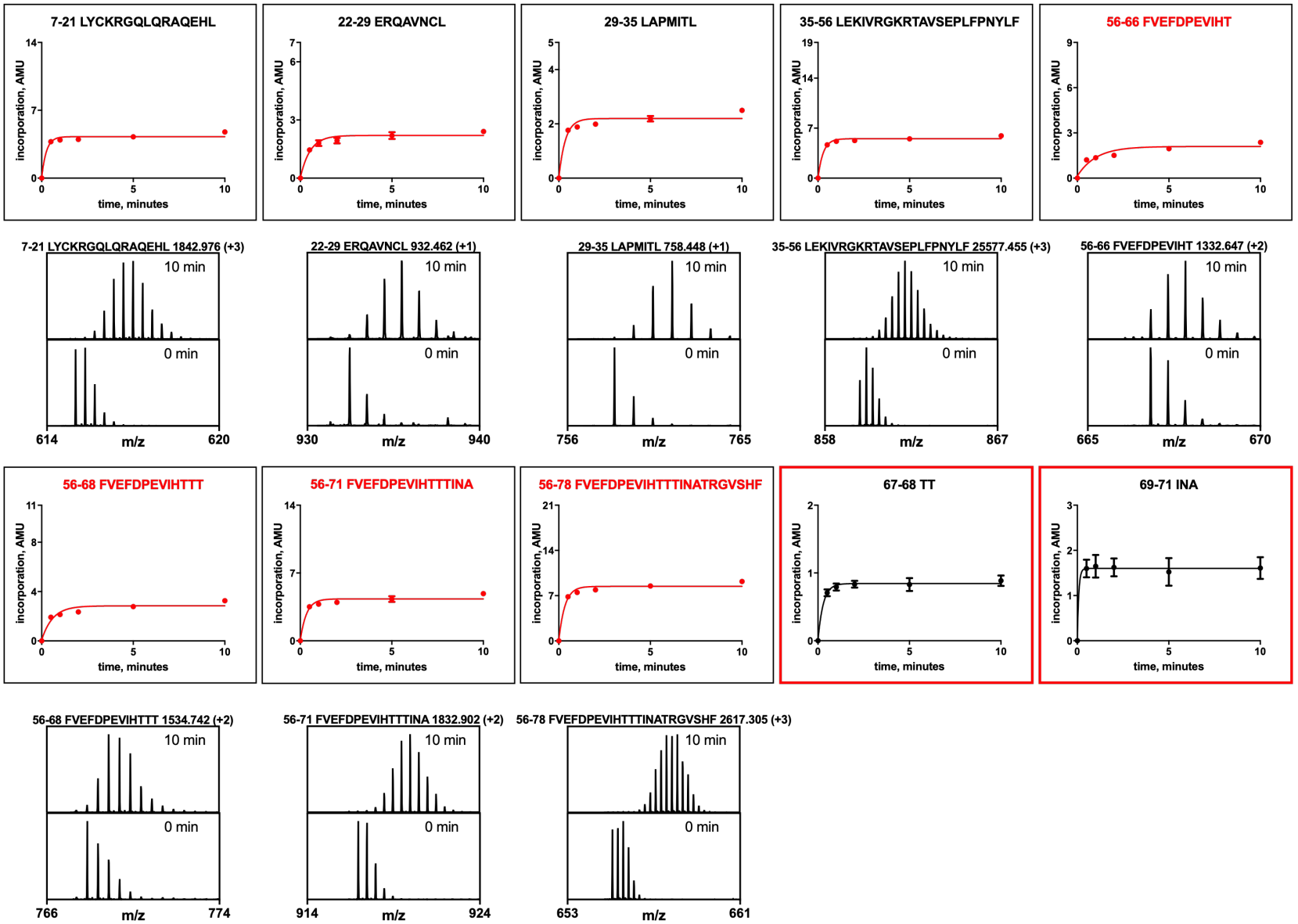
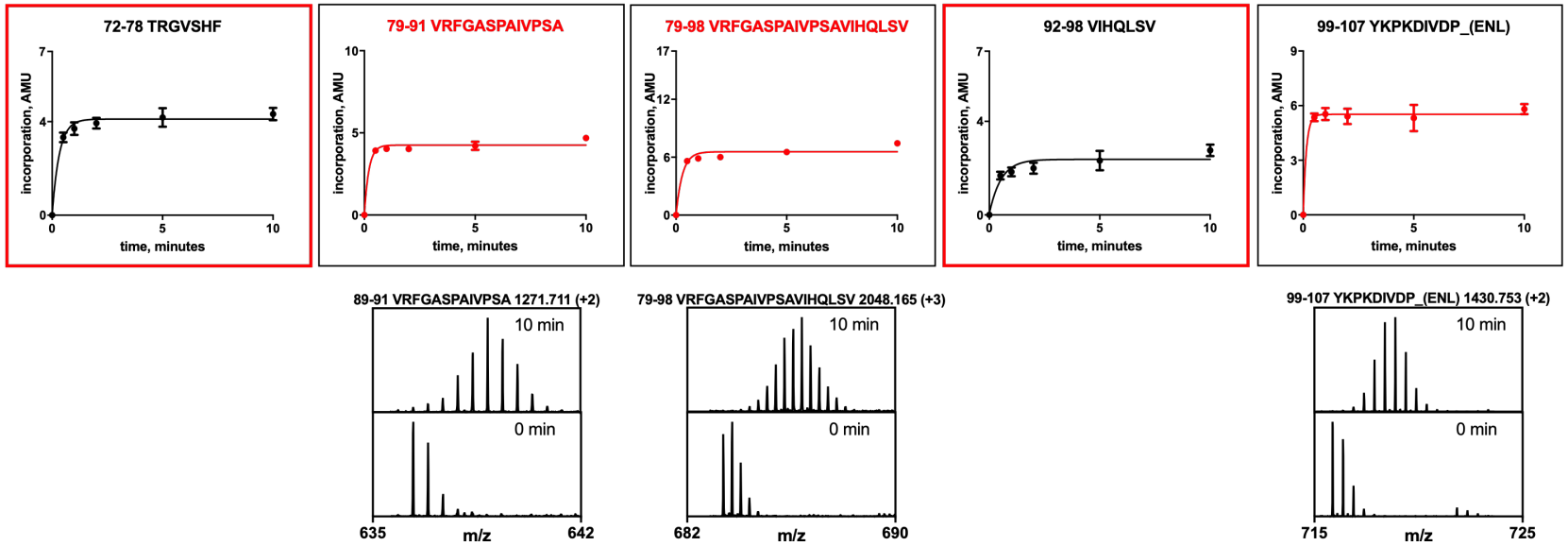


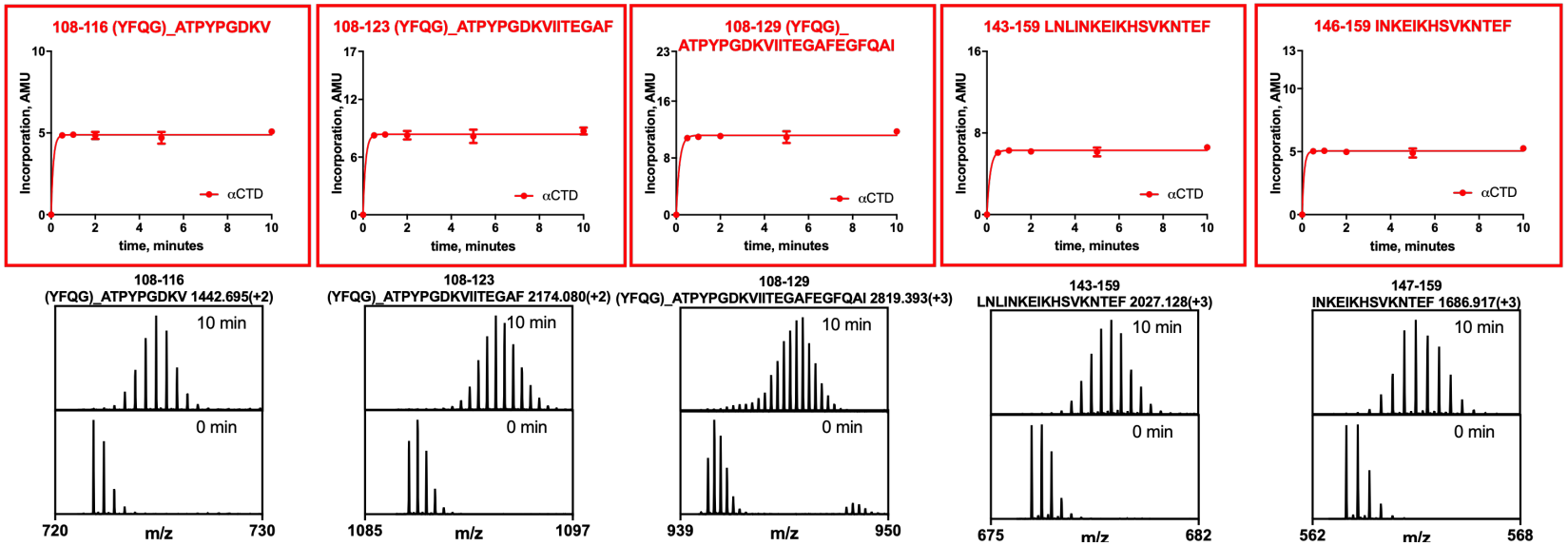
FIGURE S2 Effect of interdomain interactions in the active state of RfaH on the differential stability results from CCR simulations. The NTD and CTD of RfaH in the beta fold were artificially moved apart from each other and connected by an extended loop, and then Rosetta was used to relax the loop regions. In all cases, the modelled structure (green) was highly similar to the experimentally obtained structure in complex with the transcription machinery (cyan). Forcing Rosetta to explore extended loop conformations led to a less-favorable energy structure (red) with a more extensive interaction surface. Regardless, free stability differences obtained by CCR demonstrated that only a few residues (i.e. those involved in forming new interactions in the less favorable structure obtained by Rosetta) have significant changes in energetic stability towards each fold.

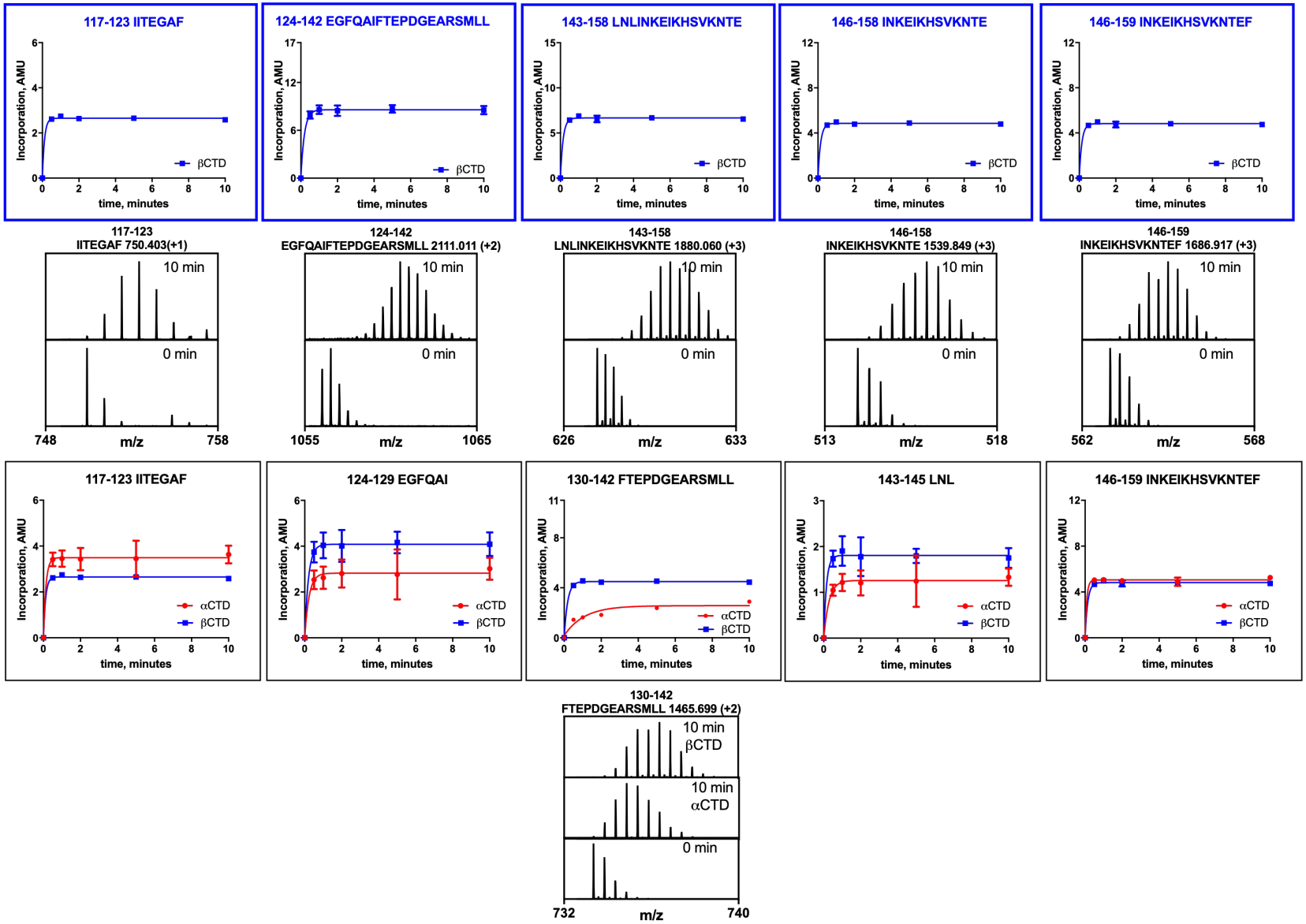
A RfaH-NTD


7-21 LYCKRGQLQRAQEHL
22-29 ERQAVNCL
29-35 LAPMITL
35-56 LEKIVRGKRTAVSEPLFPNYLF
56-66 FVEFDPEVIHT
56-68 FVEFDPEVIHTTT
56-71 FVEFDPEVIHTTTINA
56-78 FVEFDPEVIHTTTINATRGVSHF
67-68 TT
69-71 INA
incorporation, AMU
time, minutes
7-21 LYCKRGQLQRAQEHL 1842.976 (+3)
22-29 ERQAVNCL 932.462 (+1)
29-35 LAPMITL 758.448 (+1)
35-56 LEKIVRGKRTAVSEPLFPNYLF 25577.455 (+3)
56-66 FVEFDPEVIHT 1332.647 (+2)
56-68 FVEFDPEVIHTTT 1534.742 (+2)
56-71 FVEFDPEVIHTTTINA 1832.902 (+2)
56-78 FVEFDPEVIHTTTINATRGVSHF 2617.305 (+3)
10 min
0 min
m/z


72-78 TRGVSHF
79-91 VRFGASPAIVPSA
79-98 VRFGASPAIVPSAVIHQLSV
92-98 VIHQLSV
99-107 YKPKDIVDP_(ENL)
incorporation, AMU
time, minutes
89-91 VRFGASPAIVPSA 1271.711 (+2)
79-98 VRFGASPAIVPSAVIHQLSV 2048.165 (+3)
99-107 YKPKDIVDP_(ENL) 1430.753 (+2)
10 min
0 min
m/z

B RfaH-CTD


108-116 (YFQG)_ATPYPGDKV
108-123 (YFQG)_ATPYPGDKVIITEGAF
108-129 (YFQG)_ATPYPGDKVIITEGAFEGFQAI
143-159 LNLINKEIKHSVKNTEF
146-159 INKEIKHSVKNTEF
Incorporation, AMU
αCTD
time, minutes
108-116 (YFQG)_ATPYPGDKV 1442.695(+2)
108-123 (YFQG)_ATPYPGDKVIITEGAF 2174.080(+2)
108-129 (YFQG)_ATPYPGDKVIITEGAFEGFQAI 2819.393(+3)
143-159 LNLINKEIKHSVKNTEF 2027.128(+3)
147-159 INKEIKHSVKNTEF 1686.917(+3)
10 min
0 min
m/z


117-123 IITEGAF
124-142 EGFQAIFTEPDGEARSMLL
143-158 LNLINKEIKHSVKNTE
146-158 INKEIKHSVKNTE
146-159 INKEIKHSVKNTEF
Incorporation, AMU
time, minutes
βCTD
117-123
IITEGAF 750.403(+1)
124-142
EGFQAIFTEPDGEARSMLL 2111.011 (+2)
143-158
LNLINKEIKHSVKNTE 1880.060 (+3)
146-158
INKEIKHSVKNTE 1539.849 (+3)
146-159
INKEIKHSVKNTEF 1686.917 (+3)
10 min
0 min
m/z
117-123 IITEGAF
124-129 EGFQAI
130-142 FTEPDGEARSMLL
143-145 LNL
146-159 INKEIKHSVKNTEF
αCTD
βCTD
130-142
FTEPDGEARSMLL 1465.699 (+2)
10 min
βCTD
10 min
αCTD
0 min

c NusG-CTD

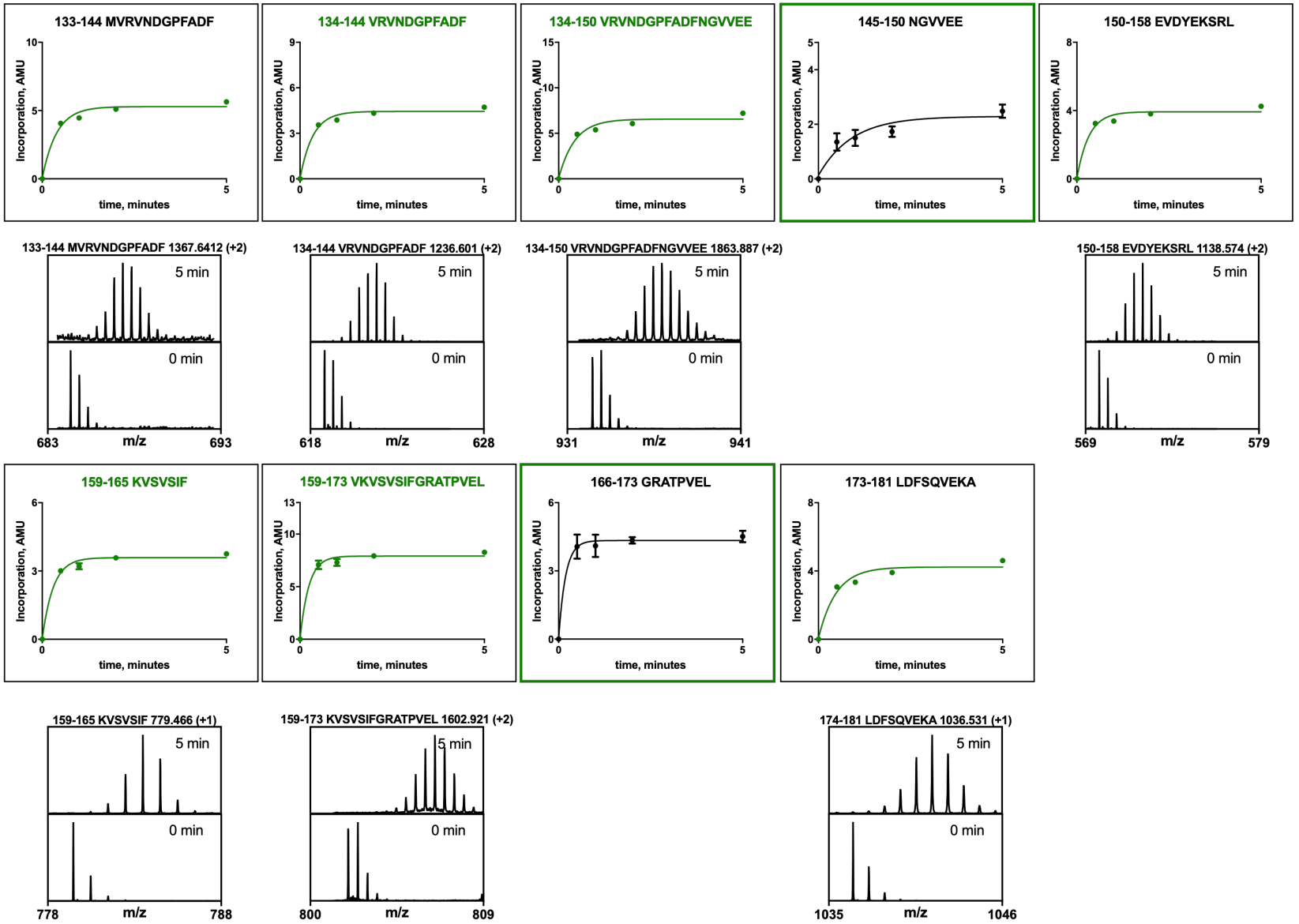

133-144 MVRVNDGPFADF
134-144 VRVNDGPFADF
134-150 VRVNDGPFADFNGVVEE
145-150 NGVVEE
150-158 EVDYEKSRL
Incorporation, AMU
time, minutes
133-144 MVRVNDGPFADF 1367.6412 (+2)
134-144 VRVNDGPFADF 1236.601 (+2)
134-150 VRVNDGPFADFNGVVEE 1863.887 (+2)
150-158 EVDYEKSRL 1138.574 (+2)
5 min
0 min
683
693
618
628
931
941
569
579
m/z
159-165 KVSVSIF
159-173 VKVSVSIFGRATPVEL
166-173 GRATPVEL
173-181 LDFSQVEKA
159-165 KVSVSIF 779.466 (+1)
159-173 KVSVSIFGRATPVEL 1602.921 (+2)
174-181 LDFSQVEKA 1036.531 (+1)
778
788
800
809
1035
1046

FIGURE S3 Deuteron incorporation and mass spectra of different regions identified for full length RfaH (red), its isolated CTD (blue) or NusG CTD (green). Deuterium incorporation was measured between 0-10 min of incubation in deuterated buffer, except for NusG, where the maximum reaction time was 5 min. Data was fitted to a single exponential to determine the maximum extent of deuteron incorporation for each region. Only mass spectra for the minimum (0 min) and maximum (10 min) reaction times are shown. (*A*) Regions identified in the NTD of RfaH. The extent of deuteron exchange for regions indicated in red boxes and lacking mass spectra were determined based on the overlapping of two experimentally observed peptides (indicated by red titles). (*B*) Deuteron incorporation of regions of the CTD of RfaH. Peptides analyzed in the context of full-length RfaH are indicated by red boxes, whereas peptides analyzed in the context of the isolated CTD are shown in blue. These peptides were employed to calculate the extent of exchange of smaller peptide regions of RfaH in both folds by accounting for the overlapping regions between these peptides. Four peptide regions were derived using this approach (residues 117-123, 124-129, 143-145, 146-159), whereas peptide 130-142 was experimentally observed in both full-length RfaH and the isolated CTD, and its mass spectra is consistent with differences in deuterium exchange due to the topology of each native state. (*C*) Deuteron incorporation of regions of the CTD of NusG. The extent of deuteron exchange for regions indicated in green boxes and lacking mass spectra were determined based on the overlapping of two experimentally observed peptides (indicated by green titles).

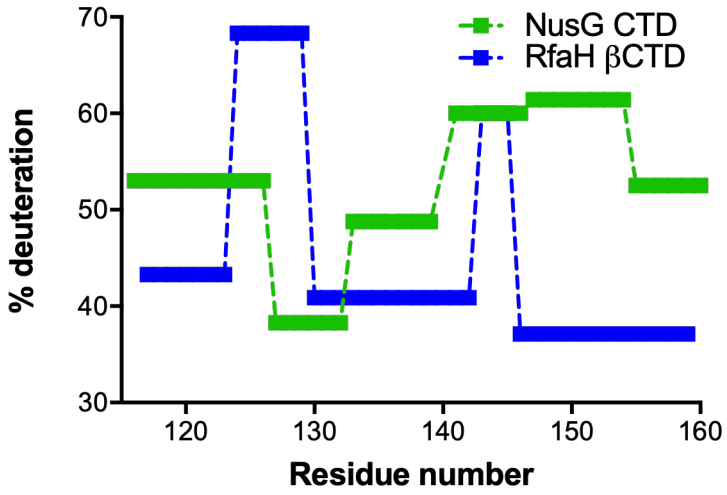


FIGURE S4 Maximum deuteron incorporation in NusG and in the isolated CTD of RfaH in the β-fold. In both cases, the proteins were incubated in deuterated buffer for up to 5 minutes, and its deuteron incorporation fitted to a single exponential function.

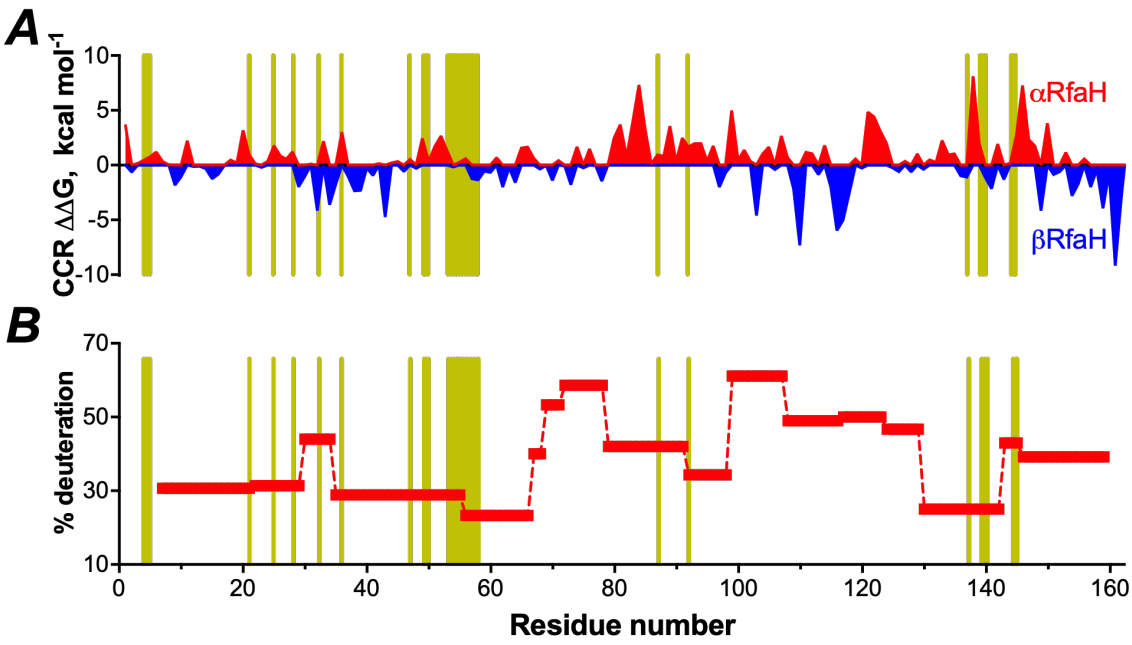


FIGURE S5 Highly stable single amides compared with peptic resolution and computational predictions. Amides with low deuteration due to their stable involvement in H-bonds, based on HDX measured by $^{1}H,^{15}N$-HSQC, are represented as ochre lines perpendicular to the axis, compared to the per-residue preferential stability as assessed from MD simulations (*A*) or for entire peptides or regions resulting from HDXMS on full-length RfaH (*B*).

Supporting Tables

TABLE S1 Deuteron incorporation for RfaH-CTD under different reaction times

Peptic peptide sequence	Deuteron Incorporation, AMU ± SD					
	0.0 min	0.5 min	1.0 min	2.0 min	5.0 min	10.0 min
YFQGATPYPGDKV	0	4.85 ± 0.08	4.9 ± 0.1	4.8 ± 0.2	4.7 ± 0.3	5.1 ± 0.1
YFQGATPYPGDKVIIT	0	6.7 ± 0.1	6.8 ± 0.2	6.8 ± 0.3	6.5 ± 0.5	7.1 ± 0.2
YFQGATPYPGDKVIITE	0	6.9 ± 0.2	7.1 ± 0.3	7.0 ± 0.4	6.8 ± 0.5	7.4 ± 0.3
YFQGATPYPGDKVIITEGAF	0	8.3 ± 0.3	8.3 ± 0.3	8.2 ± 0.4	8.2 ± 0.7	8.7 ± 0.4
YFQGATPYPGDKVIITEGAFEGFQAI	0	10.8 ± 0.3	11.0 ± 0.3	11.1 ± 0.4	10.9 ± 0.8	11.7 ± 0.3
---GATPYPGDKVIITE	0	5.74 ± 0.07	5.8 ± 0.2	5.8 ± 0.2	5.7 ± 0.4	6.1 ± 0.1
---GATPYPGDKVIITEGAF	0	7.05 ± 0.09	7.1 ± 0.2	7.1 ± 0.3	6.9 ± 0.4	7.4 ± 0.1
----------QAIFTEPDGEARSML	0	2.26 ± 0.02	2.77 ± 0.04	3.27 ± 0.03	4.21 ± 0.04	4.92 ± 0.03
----------FTEPDGEARSML	0	1.44 ± 0.02	1.64 ± 0.01	1.83 ± 0.03	2.38 ± 0.02	2.90 ±0.03
----------FTEPDGEARSMLL	0	1.484 ± 0.007	1.72 ± 0.02	1.90 ± 0.01	2.42 ± 0.03	2.95 ± 0.02
----------LNLINKEIKHSVKNTEF	0	6.08 ± 0.09	6.3 ± 0.1	6.2 ± 0.2	6.1 ± 0.4	6.6 ± 0.1
----------INKEIKHSVKNTEF	0	5.03 ± 0.08	5.1 ± 0.1	5.0 ± 0.2	4.9 ± 0.4	5.3 ± 0.1

TABLE S2 Deuteron incorporation for the isolated CTD under different reaction times

Peptic peptide sequence	Deuteron incorporation, AMU ± SD					
	0.0 min	0.5 min	1.0 min	2.0 min	5.0 min	10.0 min
IITEGAF	0	2.61 ± 0.05	2.74 ± 0.08	2.64 ± 0.08	2.65 ± 0.02	2.58 ± 0.03
---EGAFEGF	0	2.27 ± 0.02	2.41± 0.07	2.3 ± 0.1	2.40 ± 0.03	2.33 ± 0.04
-------EGFQAIFTEPDGEARSML	0	7.5 ± 0.3	8.1 ± 0.4	7.9 ± 0.5	8.1 ± 0.3	8.0 ± 0.3
----------QAIFTEPDGEARSML	0	5.57 ± 0.03	6.0 ± 0.2	6.0 ± 0.3	6.17 ± 0.05	6.0 ± 0.1
-----------AIFTEPDGEARSML	0	5.14 ± 0.03	5.5 ± 0.2	5.4 ± 0.3	5.50 ± 0.06	5.4 ± 0.1
----------QAIFTEPDGEARSMLL	0	6.0 ± 0.4	6.5 ± 0.5	6.4 ± 0.6	6.6 ± 0.5	6.5 ± 0.5
-----------AIFTEPDGEARSMLL	0	5.59 ± 0.04	6.0 ± 0.2	5.9 ± 0.3	6.1 ± 0.1	6.0 ± 0.2
------------IFTEPDGEARSML	0	4.29 ± 0.04	4.6 ± 0.1	4.5 ± 0.2	4.59 ± 0.03	4.48 ± 0.09
-------------FTEPDGEARSML	0	3.77 ± 0.09	4.1 ± 0.1	3.9 ± 0.2	4.01 ± 0.09	3.9 ± 0.1
-------------FTEPDGEA	0	2.12 ± 0.02	2.26 ± 0.06	2.2 ± 0.1	2.28 ± 0.03	2.20 ± 0.04
-------------FTEPDGEARS	0	3.22 ± 0.05	3.39 ± 0.08	3.3 ± 0.1	3.33 ± 0.05	3.29 ± 0.08
-------------FTEPDGEARSMLL	0	4.18 ± 0.03	4.5 ± 0.1	4.5 ± 0.2	4.53 ± 0.04	4.4 ± 0.1
-------------------------LNLINKEIKHSVKNTE	0	6.4 ± 0.1	6.9 ± 0.3	6.6 ±0.4	6.7 ± 0.1	6.5 ± 0.2
----------------------------INKEIKHSVKNTE	0	4.69 ± 0.07	5.0 ± 0.2	4.8 ± 0.2	4.88 ± 0.07	4.8 ± 0.1
------------------------LLNLINKEIKHSVKNTE	0	7.5 ± 0.1	7.9 ± 0.2	7.6 ± 0.3	7.8 ± 0.1	7.6 ± 0.2
------------------------LLNLINKEIKHSVKNTEF	0	7.8 ± 0.1	8.3 ± 0.2	7.9 ± 0.4	8.08 ± 0.09	7.9 ± 0.2
------------------------LNLINKEIKHSVKNTEF	0	6.6 ± 0.2	7.1 ± 0.3	6.8 ± 0.4	6.9 ± 0.1	6.8 ± 0.2
------------------------LNLINKEIKHSVKNTEFRKL	0	8.1 ± 0.3	8.7 ± 0.4	8.4 ± 0.5	8.7 ± 0.3	8.7 ± 0.4

TABLE S3 Deuteron incorporation of full-length RfaH

Position	Sequence	*k*, min^{-1}	Δmass, AMU	R^2	%Deut.
7-21	LYCKRGQLQRAQEHL	3.9 ± 1.3	4.3 ± 0.2	0.973	31
22-29	ERQAVNCL	1.8 ± 0.4	2.2 ± 0.1	0.972	31
29-35	LAPMITL	2.8 ± 0.9	2.2 ± 0.1	0.955	44
35-56	LEKIVRGKRTAVSEPLFPNYLF	3.5 ± 0.7	5.5 ± 0.1	0.988	29
56-66[a]	FVEFDPEVIHT	0.9 ± 0.4	2.1 ± 0.2	0.906	23
56-68[a,b]	FVEFDPEVIHTTT	1.7 ± 0.6	2.8 ± 0.2	0.929	25
56-71[b,c]	FVEFDPEVIHTTTINA	3.0 ± 1.0	4.3 ± 0.2	0.964	31
56-78[c]	FVEFDPEVIHTTTINATRGVSHF	3.0 ± 0.8	8.4 ± 0.4	0.979	40
67-68[a]	TT	3.5 ± 0.4	0.8 ± 0.1	0.995	40
69-71[b]	INA	12 ± 41	1.6 ± 0.1	0.996	53
72-78[c]	TRGVSHF	3.0 ± 0.5	4.1 ± 0.1	0.987	59
79-91[d]	VRFGASPAIVPSA	5.0 ± 1.9	4.2 ± 0.2	0.981	42
79-98[d]	VRFGASPAIVPSAVIHQLSV	3.5 ± 1.3	8.4 ± 0.4	0.979	49
92-98[d]	VIHQLSV	1.8 ± 0.7	2.4 ± 0.2	0.925	34
99-107	YKPKDIVDP_(ENL)	7.1 ± 3.0	5.5 ± 0.1	0.995	61
108-116[e]	(YFQG)_ATPYPGDKV	9.9±10.2	4.9 ± 0.1	0.931	49
108-123[e,f]	(YFQG)_ATPYPGDKVIITEGAF	8.6 ± 4.8	8.4 ± 0.1	0.997	49
108-129[f]	(YFQG)_ATPYPGDKVIITEGAFEGFQAI	6.6 ± 2.1	11.2 ± 0.2	0.996	49
117-123[e]	IITEGAF	7.5 ± 2.6	3.5 ± 0.1	0.997	50
124-129[f]	EGFQAI	4.3 ± 1.0	2.8 ± 0.1	0.989	47
130-142	FTEPDGEARSMLL	0.9 ± 0.4	2.6 ± 0.3	0.909	24
143-159[g]	LNLINKEIKHSVKNTEF	6.7 ± 2.0	6.3 ± 0.1	0.996	39
146-159[g]	INKEIKHSVKNTEF	11 ± 19	5.1 ± 0.1	0.997	39
143-145[g]	LNL	3.5 ± 0.5	1.3 ± 0.1	0.994	43

Expressed as average ± std. error of fit

(ENL) and (YFQG) correspond to the TEV cleaving sequence and was not considered in the sequence numbering. a-g: Deuteron incorporations in red were estimated from the difference between two overlapping peptides.

TABLE S4 Deuteron incorporation kinetics of the isolated CTD of RfaH

Position	Sequence	*k*, min^{-1}	Δmass, AMU	R^2	%Deut.
117-123	IITEGAF	8.6 ± 4.0	2.6 ± 0.1	0.997	43
124-142[a]	EGFQAIFTEPDGEARSMLL	5.2 ± 0.3	8.6 ± 0.1	0.999	51
130-142[a]	FTEPDGEARSMLL	5.4 ± 0.4	4.5 ± 0.1	0.999	41
124-129[a]	EGFQAI	4.9 ± 0.4	4.1 ± 0.1	0.999	68
143-158[b]	LNLINKEIKHSVKNTE	6.8 ± 1.6	6.7 ± 0.1	0.998	45
146-158[b]	INKEIKHSVKNTE	6.9 ± 1.3	4.9 ± 0.1	0.999	41
143-145[b]	LNL	6.6 ± 2.4	1.8 ± 0.1	0.995	60
146-159	INKEIKHSVKNTEF	7.0 ± 1.7	4.82 ± 0.06	0.998	37

Expressed as average ± std. error of fit

a-g: Deuteron incorporations in red were estimated from the difference between two overlapping peptides.

TABLE S5 Deuteron incorporation of NusG CTD

Position	Sequence	*k*, min^{-1}	Δmass, AMU	R^2	%Deut.
133-144	MVRVNDGPFADF	2.5 ± 0.6	5.3 ± 0.3	0.983	53
134-144[a]	VRVNDGPFADF	2.9 ± 0.7	4.4 ± 0.2	0.985	49
134-150[a]	VRVNDGPFADFNGVVEE	2.3 ± 0.8	6.6 ± 0.5	0.966	44
145-150[a]	NGVVEE	1.1 ± 0.5	2.3 ± 0.3	0.920	38
150-158	EVDYEKSRL	3.1 ± 1.0	3.9 ± 0. 2	0.976	49
159-165[b]	KVSVSIF	3.3 ± 0.8	3.6 ± 0.1	0.989	60
159-173[b]	KVSVSIFGRATPVEL	4.2 ± 1.1	7.9 ± 0.3	0.992	61
166-173[b]	GRATPVEL	5.3 ± 1.5	4.3 ± 0.1	0.994	61
173-181	LDFSQVEKA	2.1 ± 0.7	4.2 ± 0.3	0.964	52

Expressed as average ± std. error of fit

a-b: Deuteron incorporations in red were estimated from the difference between two overlapping peptides.

TABLE S6 NMR Protection Factors of full-length RfaH

Residue number	Residue type	k_{ex}, s^{-1}	k_{rc}, s^{-1}	PF	ΔG, kcal mol^{-1}
5	TYR	3.55E-04	2.22E+03	6.25E+06	8.95
6	LEU	1.49E-04	1.57E+03	1.05E+07	9.25
22	GLU	3.82E-03	1.02E+03	2.66E+05	7.14
26	VAL	2.17E-03	1.06E+03	4.89E+05	7.49
29	LEU	3.59E-04	4.98E+03	1.39E+07	9.41
33	ILE	1.69E-03	1.28E+03	7.56E+05	7.74
48	GLU	4.43E-03	3.29E+03	7.42E+05	7.73
50	LEU	5.90E-04	8.07E+02	1.37E+06	8.08
51	PHE	2.21E-03	1.89E+03	8.54E+05	7.81
54	TYR	1.15E-03	5.98E+03	5.22E+06	8.85
55	LEU	1.24E-04	1.57E+03	1.27E+07	9.36
56	PHE	1.49E-04	1.89E+03	1.27E+07	9.36
57	VAL	1.67E-04	1.22E+03	7.33E+06	9.04
58	GLU	1.82E-04	1.19E+03	6.56E+06	8.98
59	PHE	1.34E-03	2.17E+03	1.62E+06	8.18
88	VAL	1.99E-03	6.26E+02	3.15E+05	7.24
93	ILE	1.80E-03	7.19E+02	4.00E+05	7.38
138	ARG	1.97E-03	6.41E+03	3.25E+06	8.57
140	MET	6.91E-03	1.04E+04	1.50E+06	8.14
141	LEU	8.23E-04	1.81E+03	2.20E+06	8.35
145	LEU	2.07E-03	2.93E+03	1.42E+06	8.10
146	ILE	1.35E-03	6.12E+02	4.53E+05	7.45

Supporting References

1. Zuber, P.K., I. Artsimovitch, M. NandyMazumdar, Z. Liu, Y. Nedialkov, K. Schweimer, P. Rösch, and S.H. Knauer. 2018. The universally-conserved transcription factor RfaH is recruited to a hairpin structure of the non-template DNA strand. Elife. 7: e36349.
2. Kang, J.Y., R.A. Mooney, Y. Nedialkov, J. Saba, T. V. Mishanina, I. Artsimovitch, R. Landick, and S.A. Darst. 2018. Structural Basis for Transcript Elongation Control by NusG Family Universal Regulators. Cell. 173: 1650-1662.e14.
3. Belogurov, G.A., M.N. Vassylyeva, V. Svetlov, S. Klyuyev, N. V Grishin, D.G. Vassylyev, and I. Artsimovitch. 2007. Structural Basis for Converting a General Transcription Factor into an Operon-Specific Virulence Regulator. Mol. Cell. 26: 117–129.
4. Leaver-fay, A., M. Tyka, S.M. Lewis, O.F. Lange, R. Jacak, K. Kaufman, P.D. Renfrew, C. a Smith, I.W. Davis, S. Cooper, A. Treuille, D.J. Mandell, Y.A. Ban, S.J. Fleishman, J.E. Corn, and D.E. Kim. 2011. Rosetta 3: An Object-Oriented Software Suite for the Simulation and Design of Macromolecules. Methods Enzymol. 487: 545–574.
5. Kim, D.E., D. Chivian, and D. Baker. 2004. Protein structure prediction and analysis using the Robetta server. Nucleic Acids Res. 32: W526–W531.
6. Case, D.A., R.M. Betz, D.S. Cerutti, T.E. Cheatham III, T.A. Darden, R.E. Duke, T.J. Giese, H. Gohlke, A.W. Goetz, N. Homeyer, S. Izadi, P. Janowski, J. Kaus, A. Kovalenko, T.S. Lee, S. LeGrand, P. Li, C. Lin, T. Luchko, R. Luo, B. Madej, D. Mermelstein, K.M. Merz, G. Monard, H. Nguyen, H.T. Nguyen, I. Omelyan, A. Onufriev, D.R. Roe, A. Roitberg, C. Sagui, C.L. Simmerling, W.M. Botello-Smith, J. Swails, R.C. Walker, J. Wang, R.M. Wolf, X. Wu, L. Xiao, and P.A. Kollman. 2016. AMBER 2016.
7. Cecchini, M., S. V. Krivov, M. Spichty, and M. Karplus. 2009. Calculation of free-energy differences by confinement simulations. Application to peptide conformers. J. Phys. Chem. B. 113: 9728–9740.
8. Hawkins, G.D., C.J. Cramer, and D.G. Truhlar. 1996. Parametrized models of aqueous free energies of solvation based on pairwise descreening of solute atomic charges from a dielectric medium. J. Phys. Chem. 100: 19824–19839.
9. Ryckaert, J.P., G. Ciccotti, and H.J.C. Berendsen. 1977. Numerical integration of the cartesian equations of motion of a system with constraints: molecular dynamics of n-alkanes. J. Comput. Phys. 23: 327–341.
10. Tyka, M.D., A.R. Clarke, and R.B. Sessions. 2006. An efficient, path-independent method for free-energy calculations. J. Phys. Chem. B. 110: 17212–17220.

11. Ovchinnikov, V., M. Cecchini, and M. Karplus. 2013. A simplified confinement method for calculating absolute free energies and free energy and entropy differences. J. Phys. Chem. B. 117: 750–762.
12. Roy, A., A. Perez, K.A. Dill, and J.L. MacCallum. 2014. Computing the relative stabilities and the per-residue components in protein conformational changes. Structure. 22: 168–175.
13. Zhang, Z., and D.L. Smith. 2008. Determination of amide hydrogen exchange by mass spectrometry: A new tool for protein structure elucidation. Protein Sci. 2: 522–531.