Supplementary Information:

4 Additional Figures

Estimation of the rate constant of the second faster process

In order to estimate the rate of this second process for BeF_3^--activated $NtrC^R$, we fit the dispersion data along with the $R_{20}\beta$ to a fast-form of the Carver-Richards equation where the inactive to active interconversion and the second faster process involving the tyrosine rotation were both modeled as two-state and occurring independently from each other according to:

$$R_2^{eff}(\nu_{CPMG})=R_{20}\beta(\nu_{CPMG\rightarrow\infty})+(p_A p_I \Delta\omega_{IA}^2/k_{ex})[1-(4\nu_{CPMG}/k_{ex})\tanh(k_{ex}/4\nu_{CPMG})]$$

$$+ (p_1 p_2 \Delta\omega_{12}^2/k_{ex\_tyr})[1-(4\nu_{CPMG}/k_{ex\_tyr})\tanh(k_{ex\_tyr}/4\nu_{CPMG})]$$

The populations (p_I and p_A), change in chemical shift ($\Delta\omega_{IA}$) for each residue, and the global rate of exchange (k_{ex}) for the inactive to active conformational exchange were determined from the global fit of the ^{15}N CPMG relaxation dispersion data alone. These values were then fixed in the two-process Carver-Richards equation. The $R_{20}\beta$ values for each residue was set to the quantity calculated from the independent experiments, and the population was fixed to that of the major state as estimated by the 71μs MD simulation for the BeF_3^--activated protein (p_1=0.65). A global rate of exchange for the second process was determined for the residues in activated $NtrC^R$ that showed a significant difference (more than 1.5σ) between $R_{20}\beta$ and $R_2\beta_{1000Hz}$ yielding a rate of exchange for the tyrosine dynamics (k_{ex_tyr}) of 30,000 ± 12,000 s^{-1}. Varying the relative populations of the second process between 65-95% did not affect the overall rate of k_{ex_tyr} but did affect

the fitted delta omegas. The rate constant determined is more than an order of magnitude faster than the inactive to active interconversion rate of 1,950 ± 290 s^{-1}. [2]

This method of determining the rate of the tyrosine conformational exchange assumes that both exchange processes are two-state. According to the MD simulation of BeF_3^--activated $NtrC^R$, however, Y101 is interconverting between all three $\chi 1$ rotameric states, consequently a strict two-site assumption is not valid as discussed in more detail below.

Parameters for the second process are underdetermined

We used the software Sparta+[37] to calculate the $\delta^{15}N$ chemical shift differences ($\Delta\delta^{15}N_{comp}$) in the structures associated with the three rotameric states. Since the timescale observed for the *t* to *g*+ interconversion in our simulation is much faster than the process detected in the NMR experiments, we assume that the fast second process might correspond to a transition involving exchange between *g*- and one of the other rotameric states (or the average of them). Although in the simulation we only sample a couple of transitions to *g*- (Fig3A), the timescale is in qualitative agreement with the estimated k_{ex_tyr} of 30 KHz from the fit of NMR data.

In Fig S2 we show, color coded on the structure of the active state of $NtrC^R$, the predicted $\Delta\delta^{15}N_{comp}$ for three different transitions: *t* to *g*- ; *g*+ to *g*- ; average between *t* and *g*+ , and *g*-. We also plot the same data in Fig S3. In this case, the $\Delta\delta^{15}N_{NMR}$ are calculated from the R_{ex} contribution of the second process assuming different populations and interconversion rate constants. The plot shows that a rate for the second process on the order of 20 to 30 KHz, and populations on the order of 0.65 to 0.85. Although the comparison with the $\Delta\delta^{15}N$ estimated from the R_{ex} show a qualitative agreement with the

values predicted by the simulations, a precise correspondence could not be established. These comparisons in fact indicate the intrinsic limitations of such an approach.

Besides the issues discussed in the main text, we want to point out a few more problems that undermine the possibility of a more quantitative analysis. The faster process can only be well detected for $NtrC^R$ forms in which the inactive to active transition is at least an order of magnitude slower than the Y101 dynamics meaning that the NMR data shown in Fig. S2 and S3 are from the BeF_3^- activated protein. The simulations were however performed for apo $NtrC^R$ in its active state. Both the chemical shifts in the NMR experiment as well as the relative population of the rotamer states can be affected by the BeF_3^- activation. The calculations of ^{15}N chemical shifts using empirical approaches in Sparta+ are clearly approximations, with typical standard deviations for ^{15}N of comparable magnitude to the deviations we are hoping to detect. Moreover, additional rearrangements besides the ones captured in the 71 μs simulation could happen on longer time regimes (note that the approximate time constant of the process detected by NMR is about 30 μs), for which a much longer simulation would be needed.

We tried to get more experimental data using ^{13}C R1rho experiments [78, 79], which allows one to monitor R^{eff} suppression at higher field strengths. However the scatter of the data prohibited a much better estimation of the rate constant of the process than obtained by the combination of CPMG and the independently determined $R_{20}\beta$. Even with this method, the dilemma of not being able to separate populations from chemical shift differences, in addition to the restriction to fit to a 2-site exchange, did not warrant further pursuit.

Supplementary Figures Legends:

Figure S1. Time Scale for Y101 motion and correlation to T82

A) Histogram of the length of the time intervals spent in the Y101 rotameric *t* state (see Fig 3A). In counting the transitions, events with lifetimes shorter than 50 ns have not been considered. To obtain an estimate of the typical timescale associated with the aromatic ring rearrangement we have fit the data to a single exponential curve. The result is slightly different if we include the first data bin in the fit, blue fit (which might be affected by our extremely simplified 1 dimensional definition of the transition between states). In either case the associated timescale is on the order of ~1μs.

B) The faster motion of Y101 relative to the active/inactive transition does not significantly affect significantly the behavior of T82, which sits in a single side-chain rotamaric state for almost 100% of the time regardless of whether Y101 is in the *t* or *g*+ states (left, middle panels). However, the situation is quite different for the portion of the simulation in which Y101 is in the rotameric *g*- state and thus far from the active site region. With Y101 in rotameric *g*- state (right panel), the side chain of T82 has considerably more freedom to rotate and it populates all three of its rotameric states.

Figure S2: Comparison of chemical shift differences due to the Y101 rotation predicted from the MD simulations using Sparta+ with R_{ex} values for the second faster process estimated from the the CPMG dispersion at 1000Hz and $R_{20}\beta$.

The NMR R_{ex} values are determined from the difference in $R_2\beta_{1000Hz}$ and the $R_{20}\beta$. Such an approach is warranted for all residues since the R_{ex} suppression from 0 to 1000 Hz is small relative to these values as seen from residues that only sense this second faster process. Without

assuming any specific populations or rate of exchange, we plotted in arbitrary units the square root of R_{ex} and display in grey amino acids for which no experimental information was available (A). In panels B-D the $\Delta\delta^{15}N$ were computed with Sparta+ for the three possible exchange processes : $t \Leftrightarrow g-$; $g+ \Leftrightarrow g-$; and average t ,$g+ \Leftrightarrow g-$. For each of the single states (t , $g+$ and $g-$) the chemical shifts have been averaged over 5000 frames extracted from the simulation. The figure has a reasonable qualitative correspondence between the NMR results and the MD predictions, showing that sizable chemical shifts differences are expected in a broad area around Y101.

Figure S3: Attempt to estimate populations and rate constants for the fast process using the chemical shift differences calculated from MD with estimated values from the NMR relaxation with varying populations and rate constants.

The $\Delta\delta^{15}N$ computed with Sparta+ on the frames extracted from the simulation are plotted together with corresponding $\Delta\delta^{15}N$ estimated from the NMR relaxation data for population of the major state equal to 0.65 (A) or 0.85 (B), for two different values of k_{ex_tyr}. Regions of the sequence for which no experimental data are available are colored in grey. Although a qualitative agreement can be observed, at closer inspection it is clear that no precise one-to-one connection between predicted and measured data can be established. Moreover, it is clear that due to limited data available and limited accuracy of the predicted chemical shifts, a broad range of values for populations and k_{ex_tyr} are in equally good qualitative agreement with the predicted data and a reliable fit of the data to extract rate of exchange and populations is not possible.

Figure S4. Conservation of the aromatic residue in position 101 in prokaryotic two component system families.

A collection of protein sequences belonging to various two-component system response

regulators families, as classified by the two components systems prokaryotic database [43], has been compared to a reference sequence (CheY from E. coli; NCBI Reference Sequence: NP_288319.1) to determine the conservation of the aromatic residue corresponding to $NtrC^R$ Y101 for several receiver domain families. Each sequence has been aligned to the reference sequence using the program MUSCLE.[80] The frequency of occurrence for the 20 amino acids is shown for (F) cumulated statistics on all analyzed sequences belonging to response regulators and hybrid histidine kinases, and for only sequences classified as belonging to (A) NtrC family, (B) CheY family, (C) CheB family, (D) CheB family and (E) OmpR family.

Figure S1

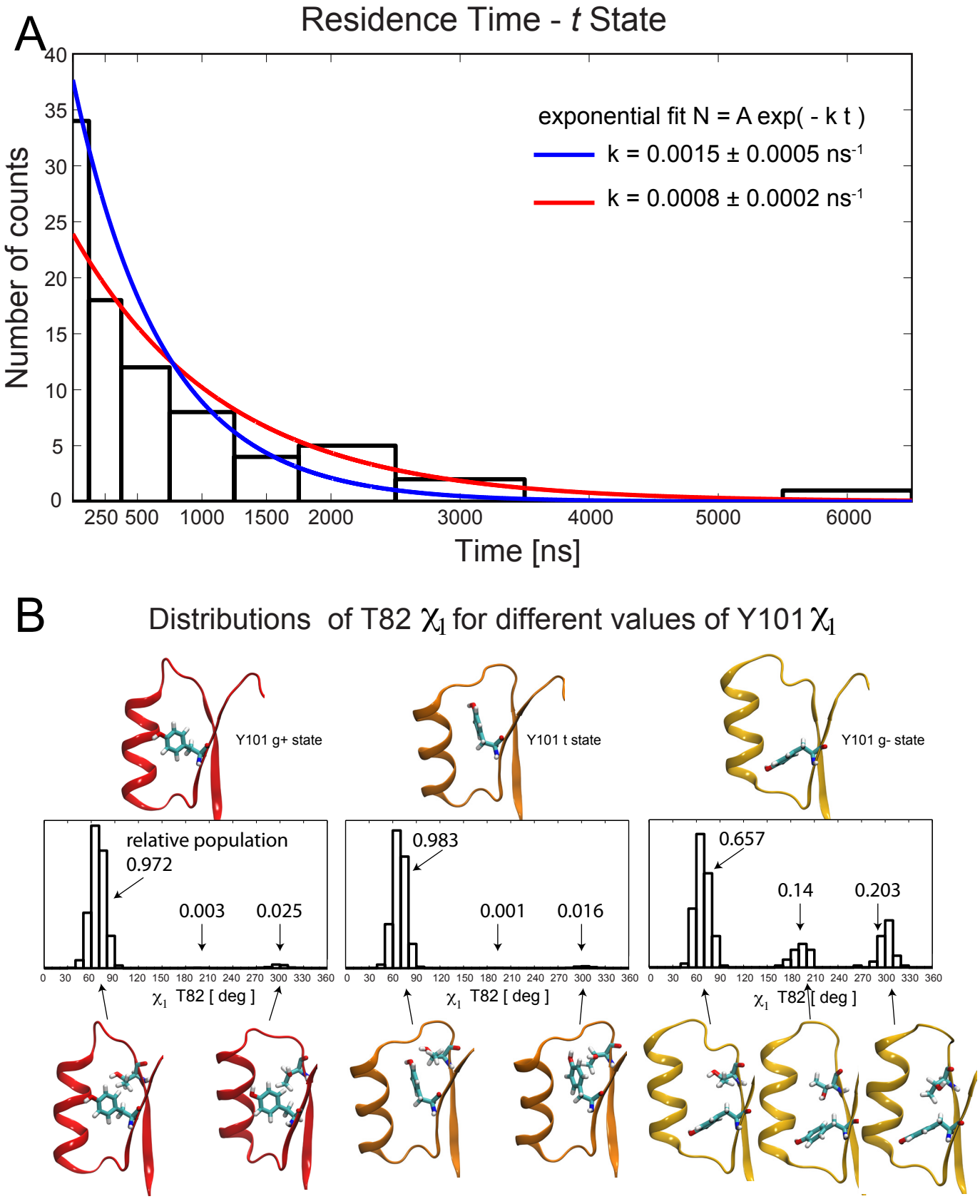

A
Residence Time - t State
Number of counts
Time [ns]
exponential fit N = A exp(- k t)
k = 0.0015 ± 0.0005 ns-1
k = 0.0008 ± 0.0002 ns-1
B
Distributions of T82 χ1 for different values of Y101 χ1
Y101 g+ state
Y101 t state
Y101 g- state
relative population
0.972
0.003
0.025
0.983
0.001
0.016
0.657
0.14
0.203
χ1 T82 [deg]

Figure S2

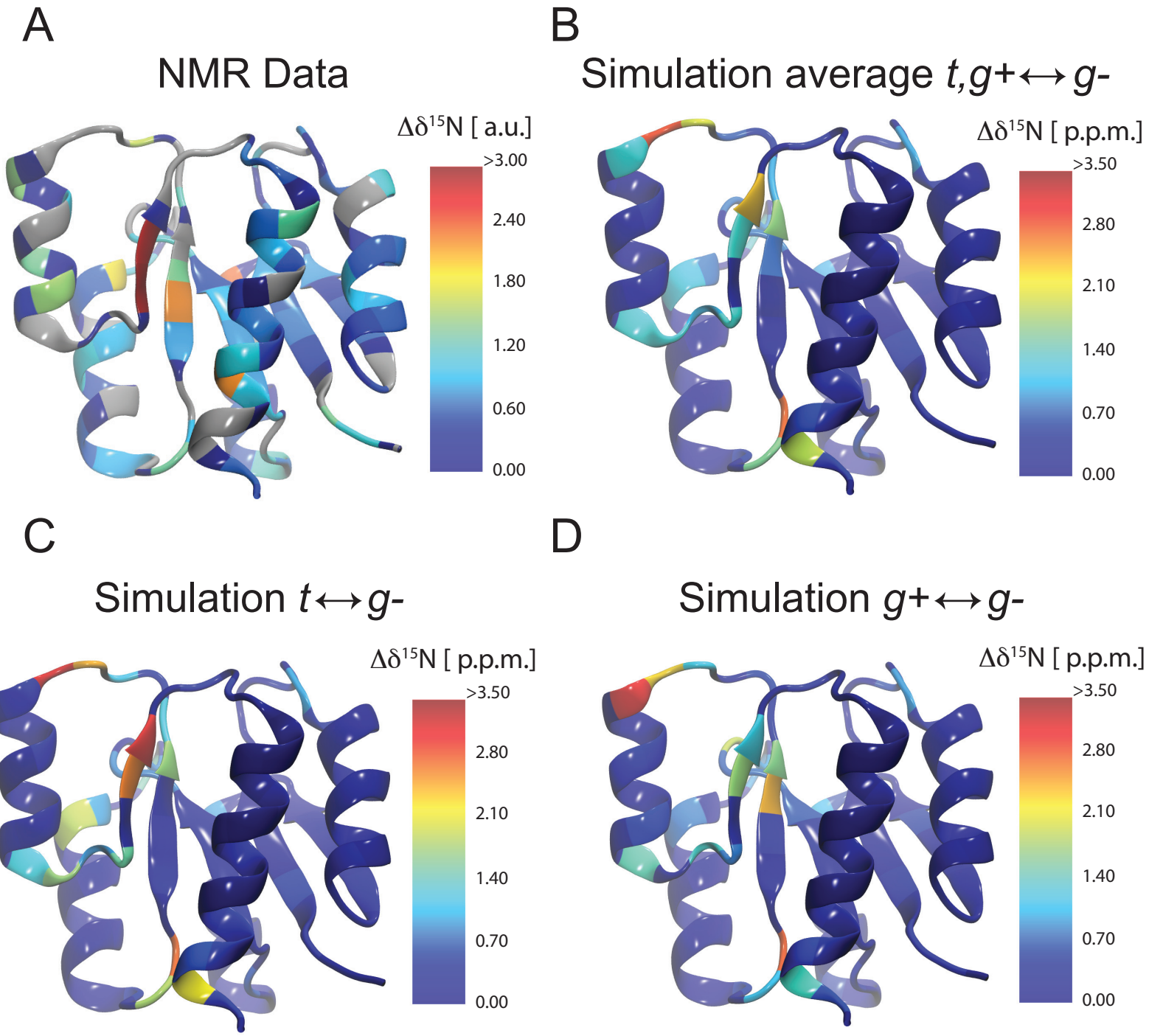

A
NMR Data
Δδ15N [a.u.]
>3.00
2.40
1.80
1.20
0.60
0.00
B
Simulation average t,g+↔g-
Δδ15N [p.p.m.]
>3.50
2.80
2.10
1.40
0.70
0.00
C
Simulation t↔g-
Δδ15N [p.p.m.]
>3.50
2.80
2.10
1.40
0.70
0.00
D
Simulation g+↔g-
Δδ15N [p.p.m.]
>3.50
2.80
2.10
1.40
0.70
0.00

Figure S3

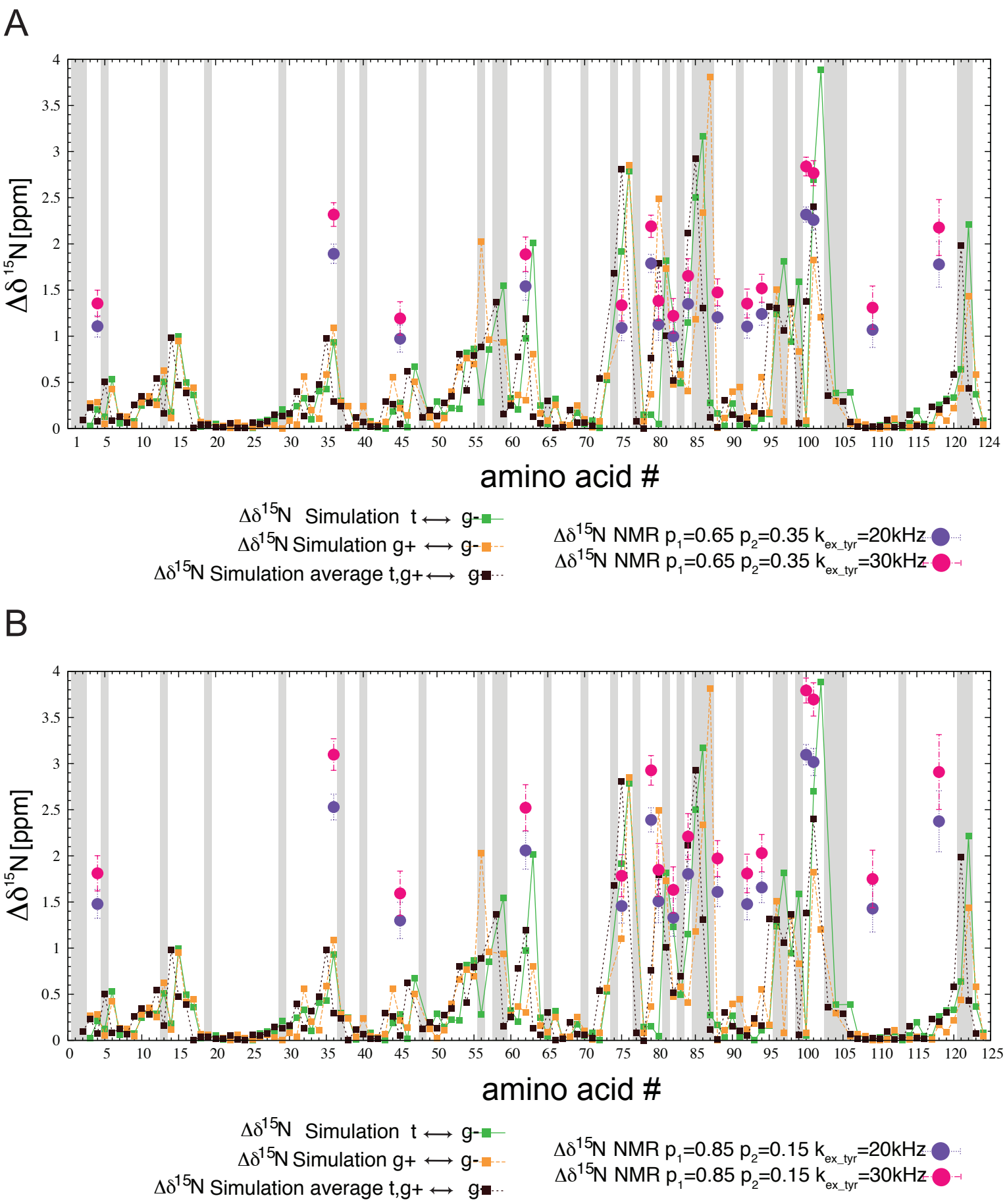
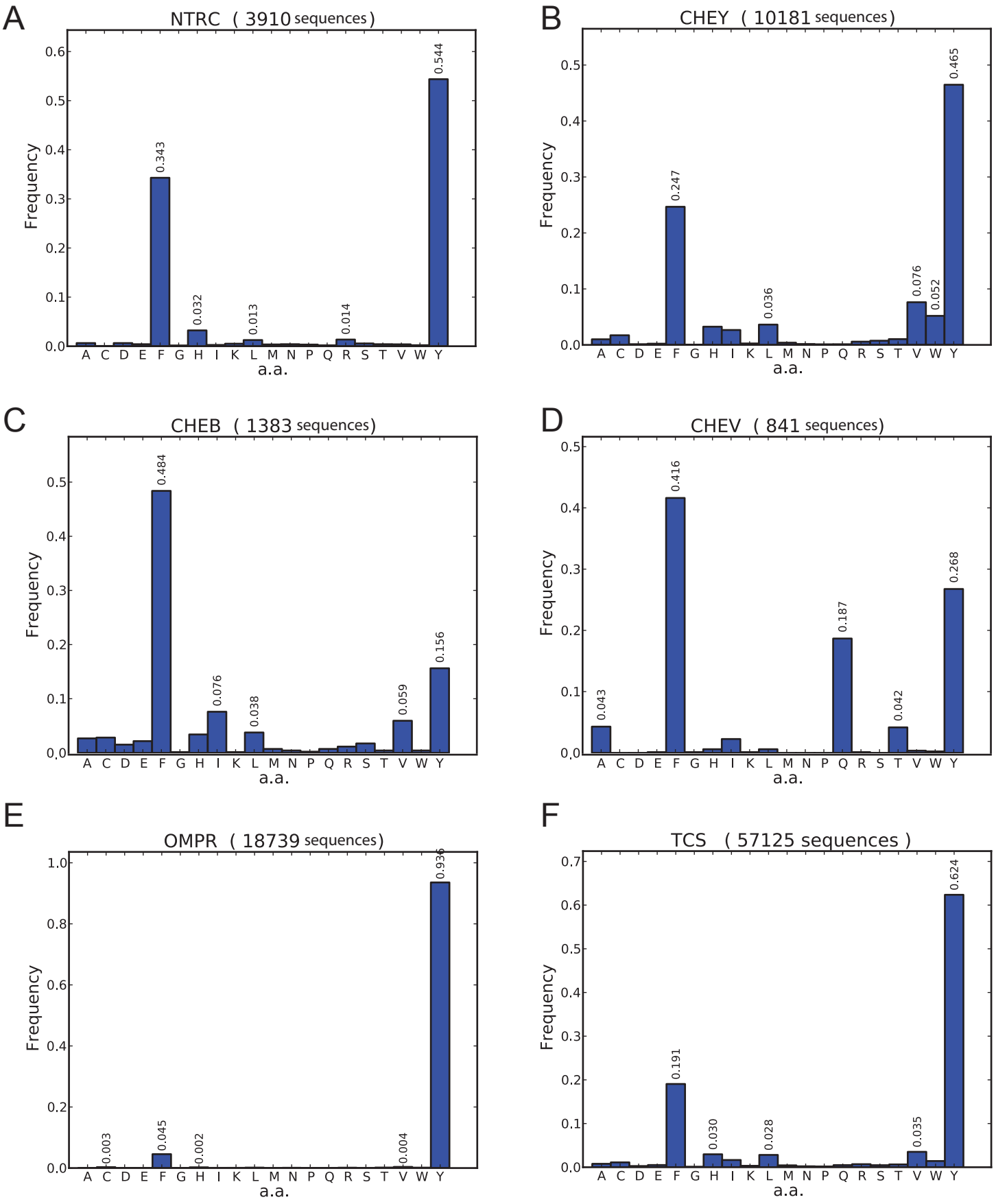

A
Δδ15N [ppm]
amino acid #
Δδ15N Simulation t ⟷ g-
Δδ15N Simulation g+ ⟷ g-
Δδ15N Simulation average t,g+ ⟷ g-
Δδ15N NMR p1=0.65 p2=0.35 kex_tyr=20kHz
Δδ15N NMR p1=0.65 p2=0.35 kex_tyr=30kHz
B
Δδ15N [ppm]
amino acid #
Δδ15N Simulation t ⟷ g-
Δδ15N Simulation g+ ⟷ g-
Δδ15N Simulation average t,g+ ⟷ g-
Δδ15N NMR p1=0.85 p2=0.15 kex_tyr=20kHz
Δδ15N NMR p1=0.85 p2=0.15 kex_tyr=30kHz

Figure S4


A
NTRC (3910 sequences)
Frequency
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.343
0.032
0.013
0.014
0.544
A C D E F G H I K L M N P Q R S T V W Y
a.a.
B
CHEY (10181 sequences)
Frequency
0.0
0.1
0.2
0.3
0.4
0.5
0.247
0.036
0.076
0.052
0.465
A C D E F G H I K L M N P Q R S T V W Y
a.a.
C
CHEB (1383 sequences)
Frequency
0.0
0.1
0.2
0.3
0.4
0.5
0.484
0.076
0.038
0.059
0.156
A C D E F G H I K L M N P Q R S T V W Y
a.a.
D
CHEV (841 sequences)
Frequency
0.0
0.1
0.2
0.3
0.4
0.5
0.043
0.416
0.187
0.042
0.268
A C D E F G H I K L M N P Q R S T V W Y
a.a.
E
OMPR (18739 sequences)
Frequency
0.0
0.2
0.4
0.6
0.8
1.0
0.003
0.045
0.002
0.004
0.936
A C D E F G H I K L M N P Q R S T V W Y
a.a.
F
TCS (57125 sequences)
Frequency
0.0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.191
0.030
0.028
0.035
0.624
A C D E F G H I K L M N P Q R S T V W Y
a.a.