Biophysical Journal, Volume 98

Supporting Material

Folding intermediate and folding nucleus for I-N and U-I-N transitions in apomyoglobin: contributions by conserved and non-conserved residues

Ekaterina N. Samatova, Bogdan S. Melnik, Vitaly A. Balobanov, Natalya S. Katina, Dmitry A. Dolgikh, Gennady V. Semisotnov, Alexei V. Finkelstein, and Valentina E. Bychkova

SUPPORTING MATERIALS

The complete set of experimental data for equilibrium urea-induced denaturing transitions of apomyoglobin and its mutants is presented in Fig. 1S.

The complete set of experimental data for kinetics of de- and renaturation transitions of apomyoglobin mutants is presented in Fig. 2S.

Direct fitting the chevron plots.

Chevron plot parameters obtained by the commonly used (traditional, "established") procedures for three-state chevron plots [1-3] and by our analytical approach [4] coincide with experimental chevron plots having a more or less complete refolding limb. In the opposite case, the commonly used procedures (based solely on chevron plots) are not applicable, while our approach allows describing the folding process. The advantage of the latter approach is that we use experimentally determined parameter f_I(M) (urea concentration-dependent fraction of the intermediate state), obtained from the burst phase fluorescence amplitude. That is, we use an additional source of experimental information. As we reported in 2005 [4], this approach allowed calculating energetic parameters of the folding/unfolding reaction even from incomplete plots for the I-N transition.

As an example, we give below some calculations based solely on chevron plots and neglecting an information that can be extracted from the burst phase fluorescence amplitude. These calculations are presented for the WT and I28A apoMb, where the experimentally obtained ln(k_{obs}([M])) plots show refolding limbs, as well as for the mutants F33A and L61A, showing chevron plots with incomplete refolding limbs. For the calculations, we used the following equation:

$$\ln(k_{obs}([M])) = \ln\left[\frac{k_{IN}^{water}\cdot\exp(m_{IN}\cdot[M])}{1+\exp((C_{1/2}-[M])/m_{UI})} + k_{IN}^{water}\cdot\exp(m_{NI}\cdot[M])\right] \quad (1S)$$

This equation is analogous to those used in the papers [1-3], and it is based on the equation $k_{obs} = k_{NI} + f_I \cdot k_{IN}$ presented in [5, 6]. Here, k_{NI}, k_{IN}, and f_I are determined as

$$k_{NI}([M]) = k_{NI}^{water}\cdot\exp(m_{NI}\cdot[M]), \quad (2S)$$

$$k_{IN}([M]) = k_{IN}^{water}\cdot\exp(m_{IN}\cdot[M]), \quad (3S)$$

$$f_I([M]) = \frac{1}{1 + \exp((C_{1/2} - [M])/m_{UI})}. \quad (4S)$$

Automatically calculated parameters values and their errors are shown in the following Table:

WT		I28A	
$\ln k_{IN}^{water}$	=2.73 ± 0.08	$\ln k_{IN}^{water}$	=1.94 ± 0.07
m_{NI}	=-0.38 ± 0.09	m_{NI}	=-0.2 ± 0.1
$\ln k_{NI}^{water}$	=-1.58 ± 0.09	$\ln k_{NI}^{water}$	=-1.4 ± 0.1
m_{IN}	=0.68 ± 0.02	m_{IN}	=0.80 ± 0.03
$C_{1/2}$	=2.21 ± 0.09	$C_{1/2}$	=1.9 ± 0.1
m_{UI}	=-0.33 ± 0.03	m_{UI}	=-0.36 ± 0.04
F33A		L61A	
$\ln k_{IN}^{water}$	=2.4 ± 0.3	$\ln k_{IN}^{water}$	=1.5 ± 0.4
m_{NI}	=0.0 ± 0.7	m_{NI}	=0 ± 1
$\ln k_{NI}^{water}$	=0 ± 9	$\ln k_{NI}^{water}$	=0 ± 3
m_{IN}	=1 ± 3	m_{IN}	=1 ± 1
$C_{1/2}$	=3 ± 3	$C_{1/2}$	=2 ± 1
m_{UI}	=-0.5 ± 0.5	m_{UI}	=-0.3 ± 0.5

As seen, the errors (specifically, for $\ln k_{NI}^{water}$ and $C_{1/2}$) are within the admissible limits for WT and I28A apoMb, but for F33A and L61A the errors are too high, while the chevron plots are well described by fitting of the used equations (the best fit of which is shown as continuous lines in all panels of Fig. 3S). Yet, the approach that uses additional information extracted from the burst phase fluorescence amplitude [4] gives much smaller errors, see Table 1 in the main text. This is why we used the approach [4] that has been developed in our laboratory and published in 2005 to describe a three-state kinetic process.

Supplementary REFERENCES

1. Matouschek A., L. Serrano, and A.R. Fersht. 1992. The folding of an enzyme. IV. Structure of an intermediate in the refolding of barnase analysed by a protein engineering procedure. *J. Mol. Biol.* 224:819-835.

2. Parker M.J., M. Lorch, R.B. Sessions, and A.R. Clarke. 1998. Thermodynamic properties of transient intermediates and transition states in the folding of two contrasting protein structure. *Biochemistry*. 37:2538-2545.
3. Ferguson N., A.P. Capaldi, R. james, C.Kleanthous, and S.E. Radford. 1999. Rapid folding with and without populated intermediates in the homologous four-helix proteins Im7 and Im9.
4. Baryshnikova, E.N., B.S. Melnik, A.V. Finkelstein, G.V. Semisotnov, and V.E. Bychkova. 2005. Three-state protein folding: experimental determination of free-energy profile. *Protein Sci*. 14:2658-2667.
5. Parker, M.J., J. Spencer, and A.R. Clarke. 1995. An integrated kinetic analysis of intermediates and transition states in protein folding reactions. *J. Mol. Biol.* 253:771-786.
6. Baldwin, R.L. 1996. On-pathway versus off-pathway folding intermediates. *Fold. Des.* 1:R1-R8.

Supplementary FIGURES LEGENDS

Supplementary FIGURE 1S. Equilibrium urea-induced unfolding of WT apomyoglobin and some of its mutant forms at pH 6.2: (a, c, e, d), as detected by molar ellipticity at 222 nm ($[\theta]_{222}$); (b, d, f, h), as detected by changing intensity of tryptophan fluorescence at 335 nm (I_{335}). Continuous lines in panels c, e, d are the same best-fitted 3-state WT line shown in panel a, and continuous lines in panels d, f, h are the same best-fitted 3-state WT line shown in panel b.

Supplementary FIGURE 2S. Upper panels: observed folding/unfolding rate constant (k_{obs}) *versus* the final urea concentration M for apomyoglobin mutants. All rate constants are measured in sec^{-1} and given in a logarithmic scale. The left part of each chevron corresponds to folding experiments (non-filled symbols), the right part ─ to unfolding experiments (dark blue filled symbols). Grey symbols show the folding rate constants $k_{IN}([M])$ for the I→N transition; these are calculated from Eq. 2 using the corresponding $k_{obs}([M])$, extrapolations of unfolding rates $k_{NI}([M]) = k_{NU}([M])$, and $f_I([M])$ plots shown for each protein in the lower panel. The $k_{NI}([M])$ values were calculated from the folding data in the region where the $f_I([M]) > 0.1$. Linear approximations of $\ln(k_{NI}([M]))$ are marked by dark blue dashed line, and those of $\ln(k_{IN}([M]))$ by dark blue dash-dot lines. Light blue lines represent calculated

chevron plots using an equation $k_{obs} = k_{NI} + f_I \bullet k_{IN}$. Continuous black lines in the upper panels represent chevron plots for the WT protein (see Fig. 3, WT panel). Light blue lines in the lower panels correspond to two-state fitting of f_I *vs* urea concentration. Continuous black lines in the lower panels represent f_I plots for WT protein (see Fig. 3, WT panel).

Supplementary FIGURE 3S. Best-fitted approximation of experimental data by equations (1) - (4). Comparison with the corresponding panels of Fig. 3 and Fig. 3S shows that this "best fit" (red lines) is, actually, not better than that obtained by our analytical approach (see Fig. 3 and Fig. 2S), though the latter leads to smaller errors in the chevron plot parameters.

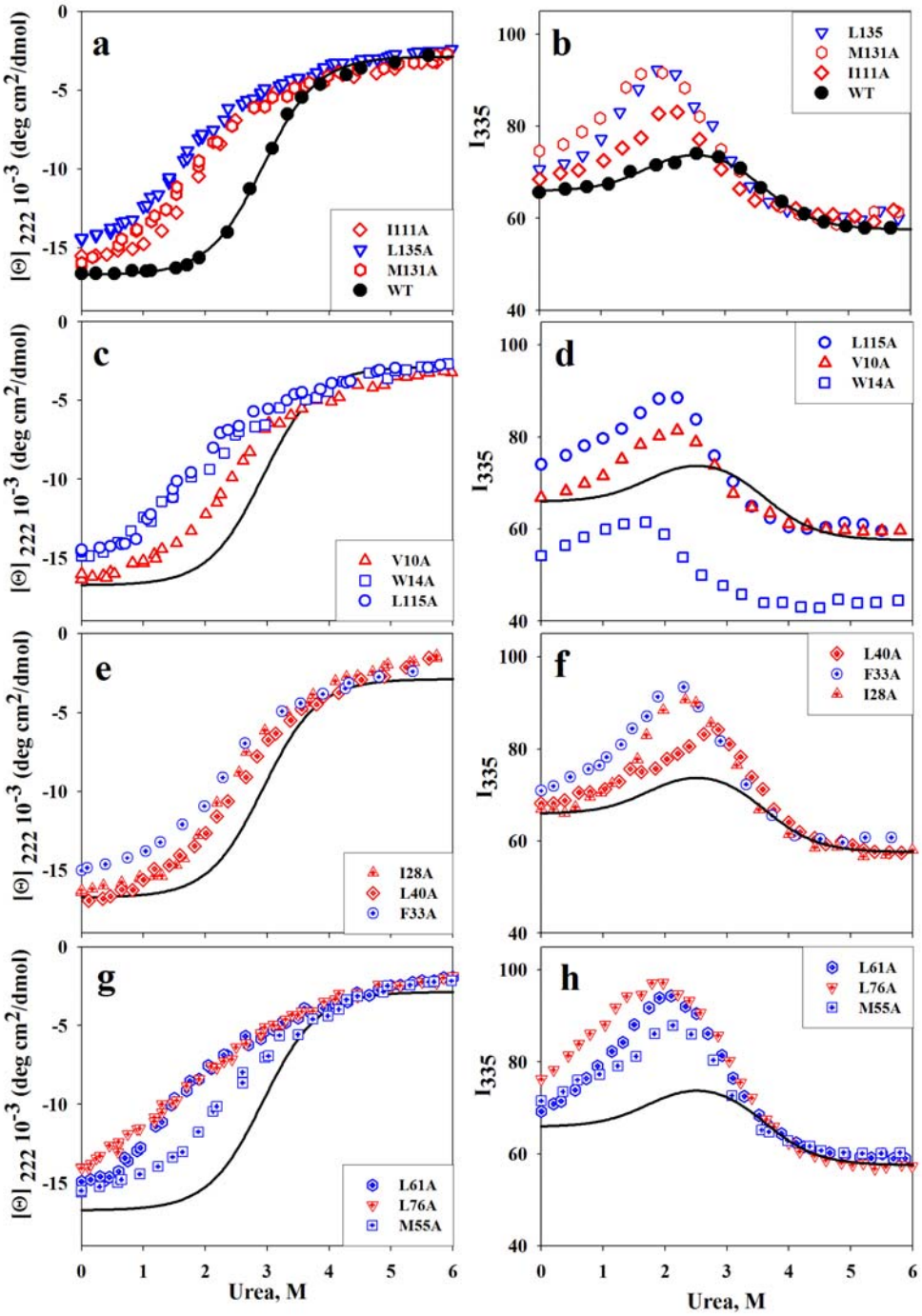
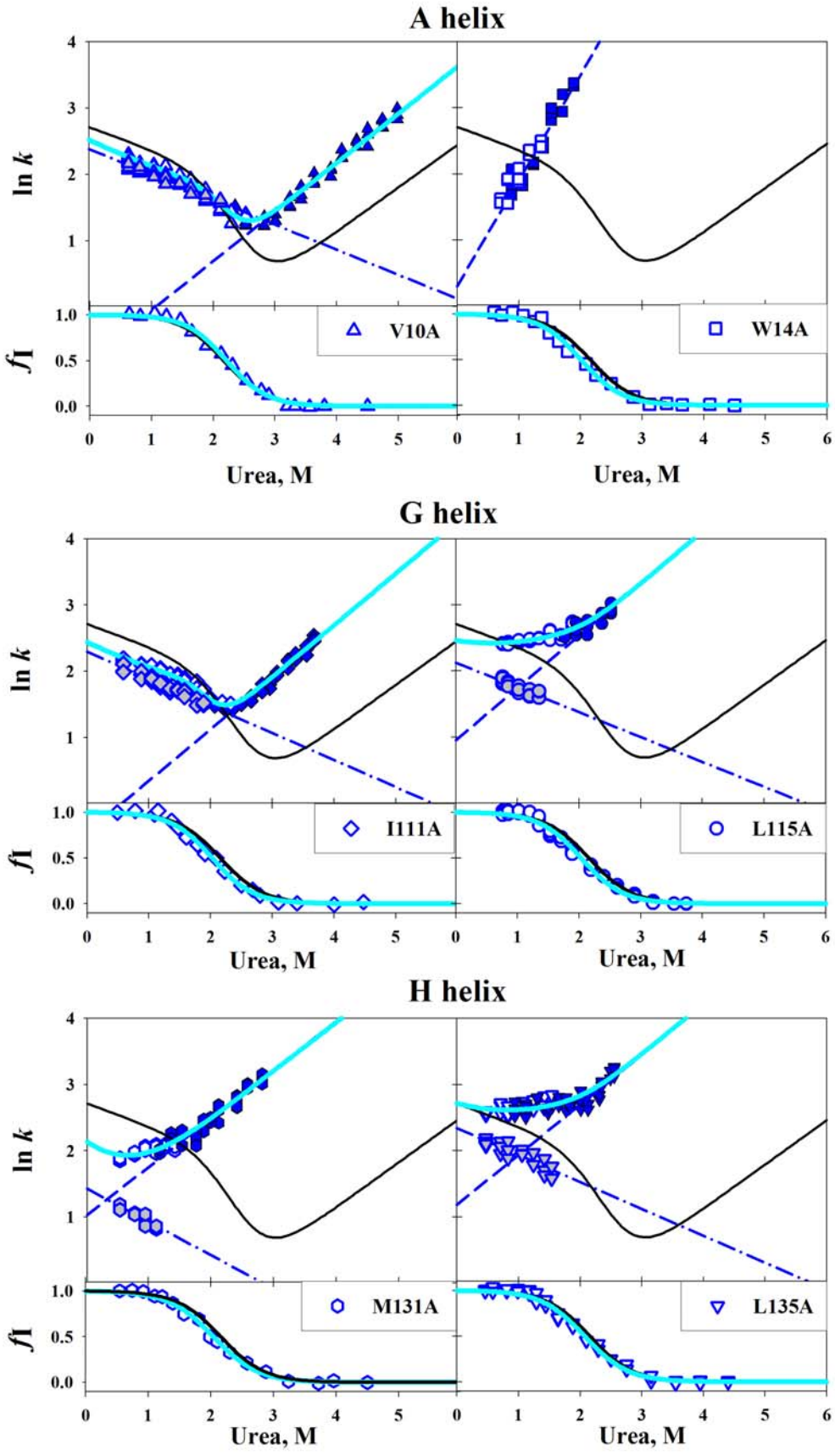

a
I111A
L135A
M131A
WT
b
L135
M131A
I111A
WT
c
V10A
W14A
L115A
d
L115A
V10A
W14A
e
I28A
L40A
F33A
f
L40A
F33A
I28A
g
L61A
L76A
M55A
h
L61A
L76A
M55A
[Θ] 222 10-3 (deg cm2/dmol)
I335
Urea, M

FIGURE 1S


A helix
ln k
fI
V10A
W14A
Urea, M
Urea, M
G helix
ln k
fI
I111A
L115A
Urea, M
Urea, M
H helix
ln k
fI
M131A
L135A
Urea, M
Urea, M

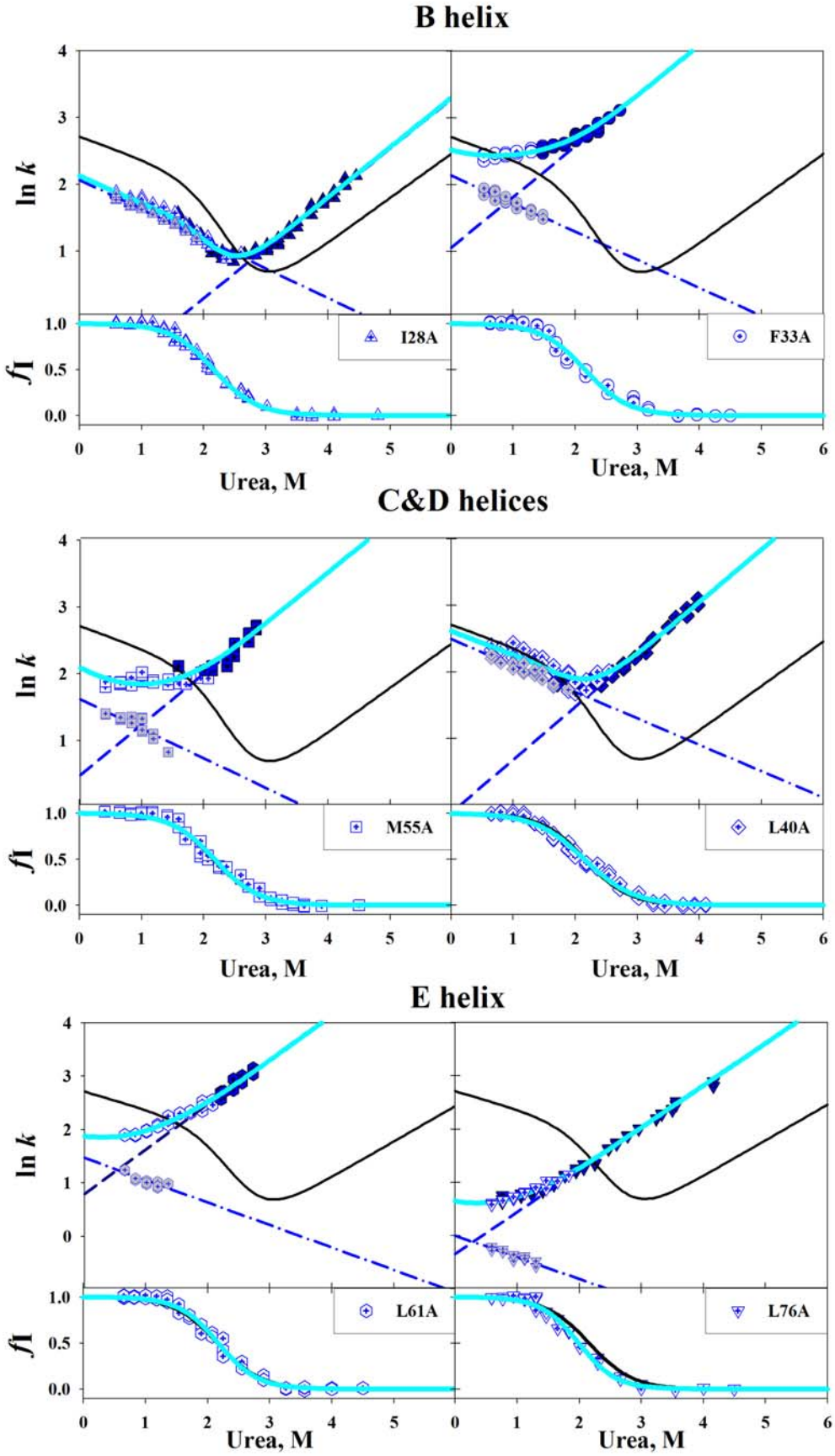

B helix
ln k
fI
I28A
F33A
Urea, M
C&D helices
M55A
L40A
E helix
L61A
L76A

FIGURE 2S

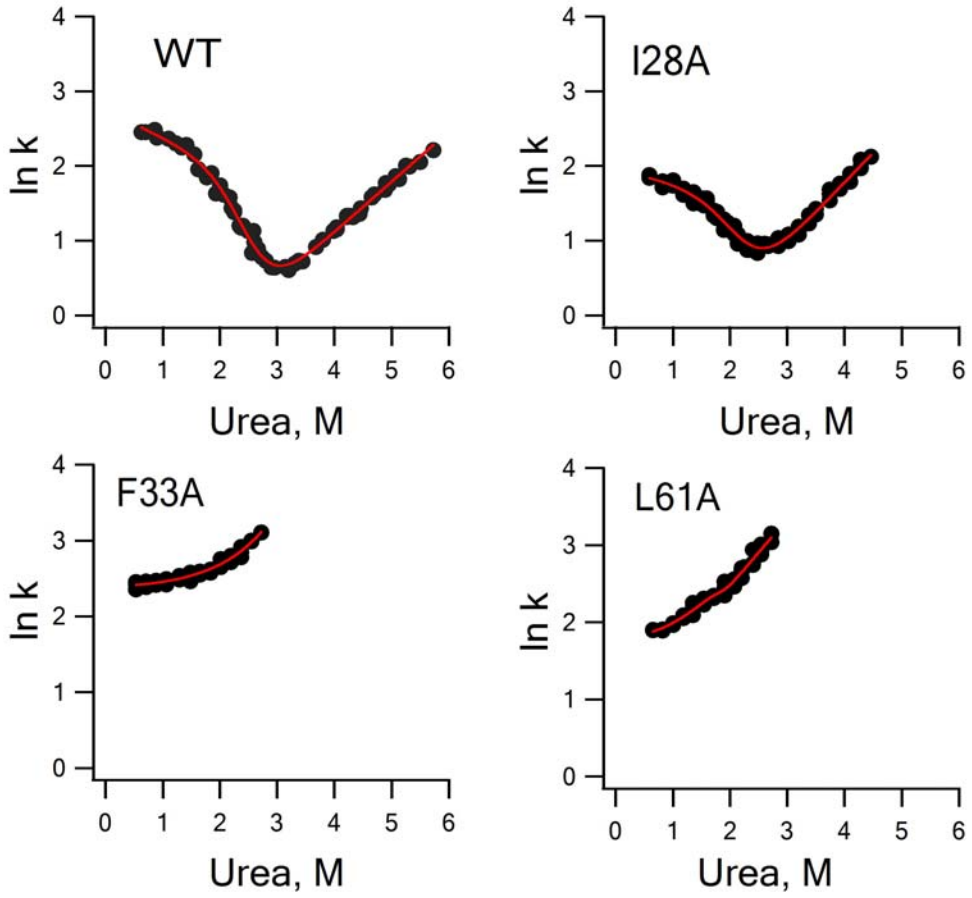


FIGURE 3S