Supporting Information

Dawson et al. 10.1073/pnas.0809138106

SI Text

Mutagenesis. H87Y mutant was obtained by modifying a WT TrAvrPto construct in the original vector (1). Primers 5′-(GGCGGACATGCAG*TAT*AGGTACATGACGGGAGCG)-3′ and 3′-(CGCTCCCGTCATGTACCT*ATA*CTGCATGTCC-GCC)-5′ with italicized mutation site were used as a sense and antisense primers in a PCR using pQE9 WT TrAvrPto as a template. The reaction was performed with the QuikChange II Site-Directed Mutagenesis kit (Stratagene) following the procedure provided by the manufacturer.

Protein Expression and Purification. *E. coli* M15 (pREP4) cells (Qiagen) containing the WT or H87Y plasmid were grown in a 10-mL starter culture in enriched LB media (containing 0.5% wt/vol D-glucose, 2 mM $MgSO_4$, 0.1 mM $CaCl_2$, 0.05 mM $FeSO_4$) with 100 μg/mL ampicillin (sodium salt) and 40 μg/mL kanamycin monosulfate at 37 °C for ≈10 h. The cells were pelleted and inoculated into 50 mL of M9 media, enriched in a similar manner with additional 19 mM $^{15}NH_4Cl$ and 1 mM thiamine HCl, for an overnight growth. The resulting cell culture was gently pelleted, resuspended, and added to a final 1-L M9 minimal media prepared in the same way and grown to an OD at 600 nm of 0.8−1.0, at which point the media was transferred to a 15 °C shaker, and protein overexpression was induced with 0.9 mM isopropyl β-D-1-thiogalactopyranoside for 24 h. Natural abundance protein samples were prepared in a similar fashion by growing a 10-mL overnight enriched LB culture, followed by an inoculation into a 1-L enriched LB media. Cells were harvested and frozen at −80 °C.

The cells were resuspended in 50 mM NaH_2PO_4, 300 mM NaCl, and 0.2 mM phenylmethanesulphonylfluoride and lysed using a Cell Disruptor model W-375 (Ultrasonics, Inc.). Upon ultracentrifugation, the protein was purified using a 3-mL Ni-NTA (Qiagen) column, eluted upon N-terminal $(His)_6$ tag cleavage with protease 3C from a picornavirus (3Cpro) and dialyzed against the final pH 7 McIlvaine's citric acid-phosphate buffer (2) diluted to a 230 μM ionic strength (with 5 mM sodium azide). Concentrated protein was stored at 4 °C. Buffer ionic strength was kept constant throughout the experiments by buffer-exchanging into adjusted amounts of 230 μM ionic strength 38.3 mM citric acid and 76.7 mM disodium phosphate solutions.

CD Spectroscopy. The CD experiments were performed using an Aviv Biomedical Circular Dichroism Spectrometer, model 202-01. The thermal denaturation of 80 μM TrAvrPto was quantified by monitoring CD signal at 220 nm as a function of temperature (Fig. S2*A*) over the range 1–80 °C from pH 4.04 to 8.00 for a total of 12 pH samples. CD 260–190 nm scan spectra (Fig. S2*B*) were recorded at 25 °C with the exception of a pH 4.04 sample scan which was additionally recorded at 1 °C. 0.1-cm path-length quartz cuvette was used. Scan data are shown in the 260- to 202-nm range due to the interference of the solvent signal in the 200- to 190-nm range. Mean residue CD extinction coefficient $\Delta\varepsilon_{MRW}$ was calculated as:

$$\Delta\varepsilon_{MRW} = \frac{S \times MRW}{32{,}980 \times C_{mg/mL} \times L}, \qquad [S1]$$

where S is the observed CD signal in millidegrees, MRW is the mean residue weight (calculated by using 12,029.3 Da, and 107 residues for WT TrAvrPto), L is the path length in cm, and $C_{mg/ml}$ is the protein concentration in mg/mL (3).

Determination of p_F from temperature denaturation curves was done using a standard fitting analysis as described (4). Briefly, p_F was obtained at a given temperature point by:

$$p_F = \frac{y_{obs} - y_F}{y_U - y_F}, \qquad [S2]$$

where y is $\Delta\varepsilon_{MRW}$ for the CD experiment, and $\Delta\varepsilon_{MRW,F}$ and $\Delta\varepsilon_{MRW,U}$ were extrapolated from the linear fits of the regions where the data gives either the fully folded or unfolded protein signal.

Errors in thermal denaturation measurements were determined to be dominated by the temperature uncertainty. Uncertainty in temperature includes variation in the temperature calibration of the different instruments (approximately ±0.1 °C), incomplete equilibration (approximated as ±0.05 °C), and temperature measurement error (±0.05 °C), and was estimated to be a total of ±0.2 °C. This temperature error was propagated via the linear relationship between temperature and p_F to yield the uncertainty in p_F.

Tryptophan Fluorescence Spectroscopy. The Trp fluorescence temperature denaturation measurements were performed using a Cary Eclipse Fluorescence Spectrophotometer (Varian Instruments). The thermal denaturation of 80 μM TrAvrPto was quantified from pH 4.00 to 7.09 for a total of 7 pH values over the temperature range of 7–80 °C (Fig. S3). Fluorescence was measured at 343 nm upon excitation of the buried Trp (Fig. 1*B*) at 295 nm. Fluorescence scans in the range of 300–400 nm were taken to confirm the emission peak maximum ≈343 nm for most samples. The folded population values were obtained from the fluorescence signal using Eq. **S2** where y denotes the measured intensity. Errors in p_F were estimated from the temperature uncertainty as described for CD.

The acid denaturation by Trp fluorescence was measured using a Synergy HT 96-well plate reader (BioTek). Data were acquired in a time-resolved fluorescence mode using an excitation filter of 286 ± 5 nm and an emission filter at 360 ± 20 nm. Scans were measured in triplicate to obtain the average values and their uncertainties. The pF was calculated from Eq. **S2**, with y_{obs} denoting the observed intensity, the folded (y_F) and unfolded (y_U) limits of fluorescence signal taken as the average of the 8 highest pH points (y_F) and as the lowest fluorescence signal from the unfolded WT data (y_U).

NMR Spectroscopy. NMR data were collected by using a Varian Inova 600-MHz spectrometer (Varian Instruments). Temperature was kept constant at a 25 °C setting, which was later measured to correspond to 26.6 °C via a methanol standard calibration. The fHSQC spectra (5) were acquired with a recycle delay of 5 s of 0.3–0.6 mM samples at pH 2.88 to 7.14 for a total of 13 pH values. To correct for the differences in relaxation rates of the U and F states during forward and reverse INEPT steps, the fHSQC experiments were also performed with a recycle delay of 1 s and J_{NH} set to 95 and 32 Hz to obtain peak volumes at INEPT delays of 1/(4J) and 3/(4J), respectively. The average transverse relaxation of the in-phase and anti-phase ^1H magnetizations during INEPT transfer steps, R_{2H}^{avg}, was obtained for each residue from the fHSQC data with varied INEPT delays as:

$$R_2^j = \frac{\ln(I_j^{\tau_1} - I_j^{\tau_2})}{4\tau_2 - 4\tau_1}, \quad [S3]$$

using peak intensities I from the 2 measurements at $\tau_1 = 1/(4J)$ and $\tau_2 = 3/(4J)$ to calculate R_2 for each residue j. The same peak pairs were monitored at pH 5.05–6.02 range in a total of 6 pH steps.

Reported p_F values of each pH sample were measured as:

$$p_F = \frac{V_F/A_F}{V_F/A_F + V_U/A_U}, \quad [S4]$$

where the values for F and U peak volumes V were obtained from the fHSQC with a long recycle delay to ensure full relaxation of magnetization (6). Applied correction factors A for each individual residue accounted for the different relaxation rates of the U and F states during the INEPT transfer. The correction factors were then obtained by calculating the resulting effect of R_2 relaxation for the total INEPT delay times ($4\tau = 4 \times (1/4J) = 1/J$):

$$A^j = \exp(R_2^j \times 1/J), \quad [S5]$$

where $J = 95$ Hz for the fHSQC experiment with a recycle delay of 5 s. Histidine side-chain pK_a values were quantified via a series of HMQC NMR experiments at pH 4.62–7.83 in a total of 13 pH steps (pH 4.85–7.83 shown in Fig. S4). A J-coupling of 20 Hz was found empirically to yield the optimal spectra, compensating for spin relaxation during the experiment. The desired resonances were recorded by using a 10,000-Hz (8,507.9-Hz) sweep width centered at 4.75 ppm (230 ppm) in the ^{1}H (^{15}N) dimension. A recycle delay of 1 s was used for all spectra. The chemical shift data δ(pH) were fitted to:

$$\delta(\text{pH}) = \frac{\delta_{\text{His}^+} + \delta_{\text{His}^0} \times 10^{\text{pH-pKa}}}{1 + 10^{\text{pH-pKa}}}, \quad [S6]$$

where δ_{His^+} and δ_{His^0} are the chemical shifts for the His$^+$ and His^0, respectively, and pK_a is the pH at which half of the His are protonated, His$^+$. The ^{15}N and ^{1}H chemical shifts were fitted separately, with the average values and the corresponding standard deviation reported as the final pK_a values for H41, H54, H75, H87, H103, H125, and H130. Fitting of the data was done with Excel Solver 12.1.1 (Frontline Systems).

All NMR data were processed using NMRPipe and NMRDraw (7). Spectra were visualized, and peak volumes were determined using Sparky (8). Integration settings assumed a Gaussian line shape with allowed peak motions and adjustment of line widths.

Data Fitting and Estimation of Errors. The experimental data (CD, Trp fluorescence, and NMR) were fitted to the model (Fig. 4*A* and Eq. **2**.) using nonlinear least-squares method. The Levenberg–Marquardt (9) algorithm was used to iteratively fit the simulated curve to the dataset and determine the fitted parameters (e.g., $\Delta G^{\text{neutral}}$, pK_a^F and pK_a^U) for the best-fit curve. Normally distributed random errors were weighted by the experimental errors and used to generate 1,000 synthetic datasets. Each synthetic dataset was again iteratively fit to the model and the fitted parameters were determined as above. The standard deviation of these fitted parameters gave their error estimation. All of the fitting routines were implemented in MATLAB 2007a (MathWorks, Inc.) and the final plots were generated in Microsoft Excel 2007.

Alignment of AvrPto and AvrPtoB Homologues. The TrAvrPto sequence (GI 75346155, residues 29–133) was used in a PSI-BLAST search (BLASTP 2.2.19+) (10) to identify homologous sequences, and yield a single ortholog with an E-value below the 0.005 threshold: AvrPto1 from *P. syringae* pv. *syringae* B728a (GI 66048143). Similarly, a PSI-BLAST search on AvrPtoB (residues 121–200 of the sequence GI 81763648, *P. syringae* pv. *tomato*), selected based on previous structure prediction (11), yielded a total of 11 homologous sequences, 7 of which display distinct sequence variations in the region of interest. TrAvrPto alignment with its ortholog and the AvrPtoB alignment with the 7 unique orthologs were done separately using CLUSTALW2 (12). Final alignment in Fig. S3*C* was arranged based on predicted secondary structure alignment between AvrPto and AvrPtoB sequences from *P. syringae* pv. *tomato* (11). The alignment figure was done using ESPript 2.2 (13).

1. Dawson JE, Nicholson LK (2008) Folding kinetics and thermodynamics of *Pseudomonas syringae* effector protein AvrPto provide insight into translocation via the type III secretion system. *Protein Sci* 17:1109–1119.
2. McIlvaine TC (1921) A buffer solution for colorimetric comparison. *J Biol Chem* 49:183–186.
3. Martin SR, Schilstra MJ (2008) Circular dichroism and its application to the study of biomolecules. *Methods Cell Biol* 84:263–293.
4. Greenfield NJ (2006) Determination of the folding of proteins as a function of denaturants, osmolytes or ligands using circular dichroism. *Nat Protocols* 1:2733–2741.
5. Mulder FAA, Spronk CAEM, Slijper M, Kaptein R, Boelens R (1996) Improved HSQC experiments for the observation of exchange broadened signals. *J Biomol NMR* 8:223–228.
6. Tollinger M, Skrynnikov NR, Mulder FA, Forman-Kay JD, Kay LE (2001) Slow dynamics in folded and unfolded states of an SH3 domain. *J Am Chem Soc* 123:11341–11352.
7. Delaglio F, et al. (1995) NMRPipe: A multidimensional spectral processing system based on UNIX pipes. *J Biomol NMR* 6:277–293.
8. Goddard TD, Kneller DG (2008) Sparky 3 (University of California, San Francisko), ver. 3.115.
9. Marquardt DW (1963) An algorithm for least-squares estimation of nonlinear parameters. *J Soc Ind Appl Math* 11:431–441.
10. Altschul SF, et al. (1997) Gapped BLAST and PSI-BLAST: A new generation of protein database search programs. *Nucleic Acids Res* 25:3389–3402.
11. Xiao F, et al. (2007) The N-terminal region of *Pseudomonas* type III effector AvrPtoB elicits Pto-dependent immunity and has two distinct virulence determinants. *Plant J* 52:595–614.
12. Thompson JD, Higgins DG, Gibson TJ (1994) CLUSTAL W: Improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. *Nucleic Acids Res* 22:4673–4680.
13. Gouet P, Courcelle E, Stuart DI, Metoz F (1999) ESPript: Analysis of multiple sequence alignments in PostScript. *Bioinformatics* 15:305–308.

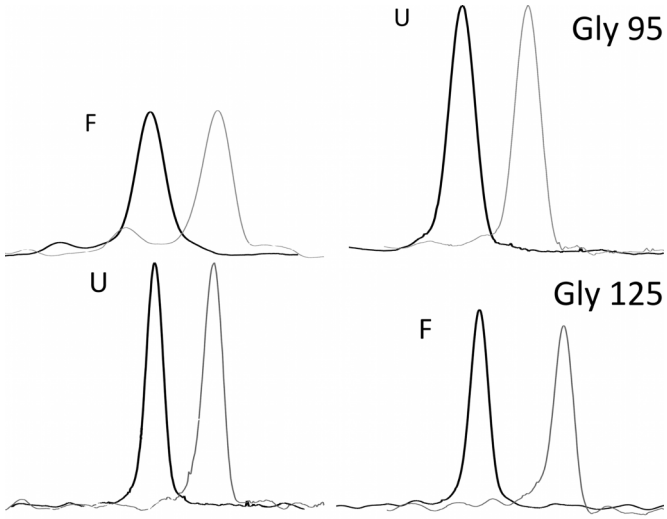

Fig. S1. Overlay of 1-dimensional slices through peaks in ^{15}N–^{1}H fHSQC spectrum of WT TrAvrPto corresponding to the folded and unfolded states of residues G95 and G128 in samples at pH 5.0 and protein concentrations of 80 μM (gray lines) and 420 μM (black lines).

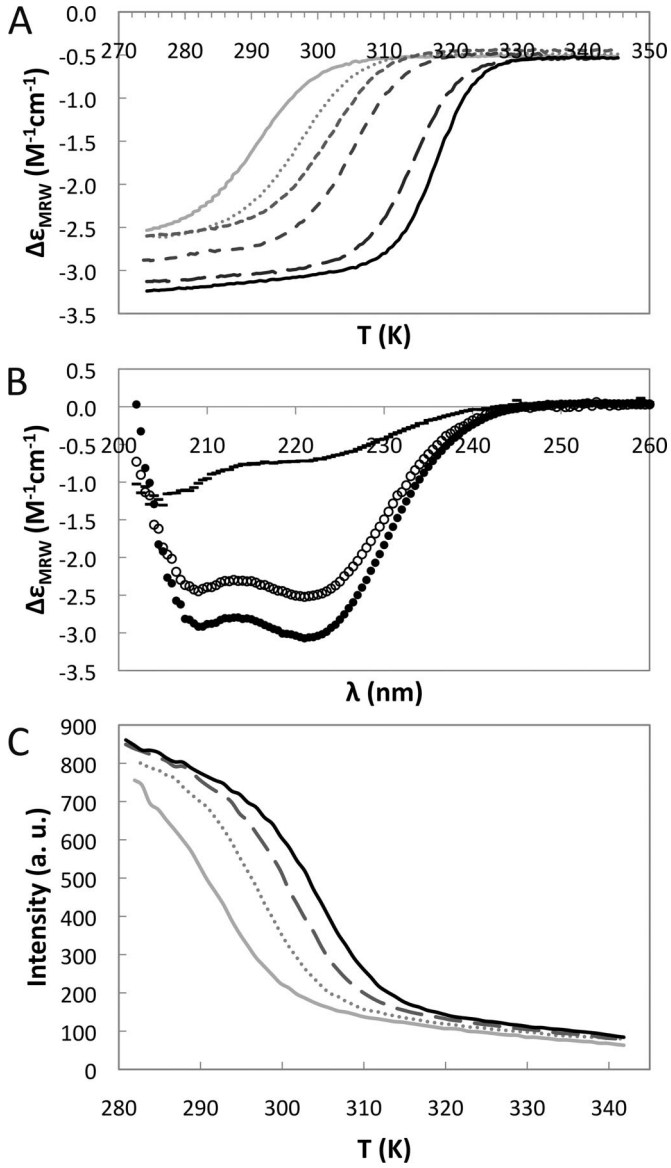

Fig. S2. WT TrAvrPto CD and fluorescence data. (*A*) Mean residue CD extinction coefficient $\Delta\varepsilon_{MFW}$ shown for the temperature denaturation curves at pH 4.04 (solid light-gray line), pH 4.46 (dotted light-gray line), pH 5.12, 5.47, pH 6.49 (dashed gray lines in increasing dash length), pH 8.00 (solid black line). (*B*) Scans of WT TrAvrPto showing mean residue CD extinction coefficients over 202–260 nm at pH 4.04 at 25 °C (dashes), pH 4.04 at 1 °C (open circles), and pH 8.00 at 25 °C (filled circles). (*C*) Tryptophan fluorescence temperature denaturation curves for WT AvrPto at pH 4.00 (the lightest solid gray line), 4.66 (dotted gray line), 5.10 (dashed gray line), and 5.41 (black solid line).

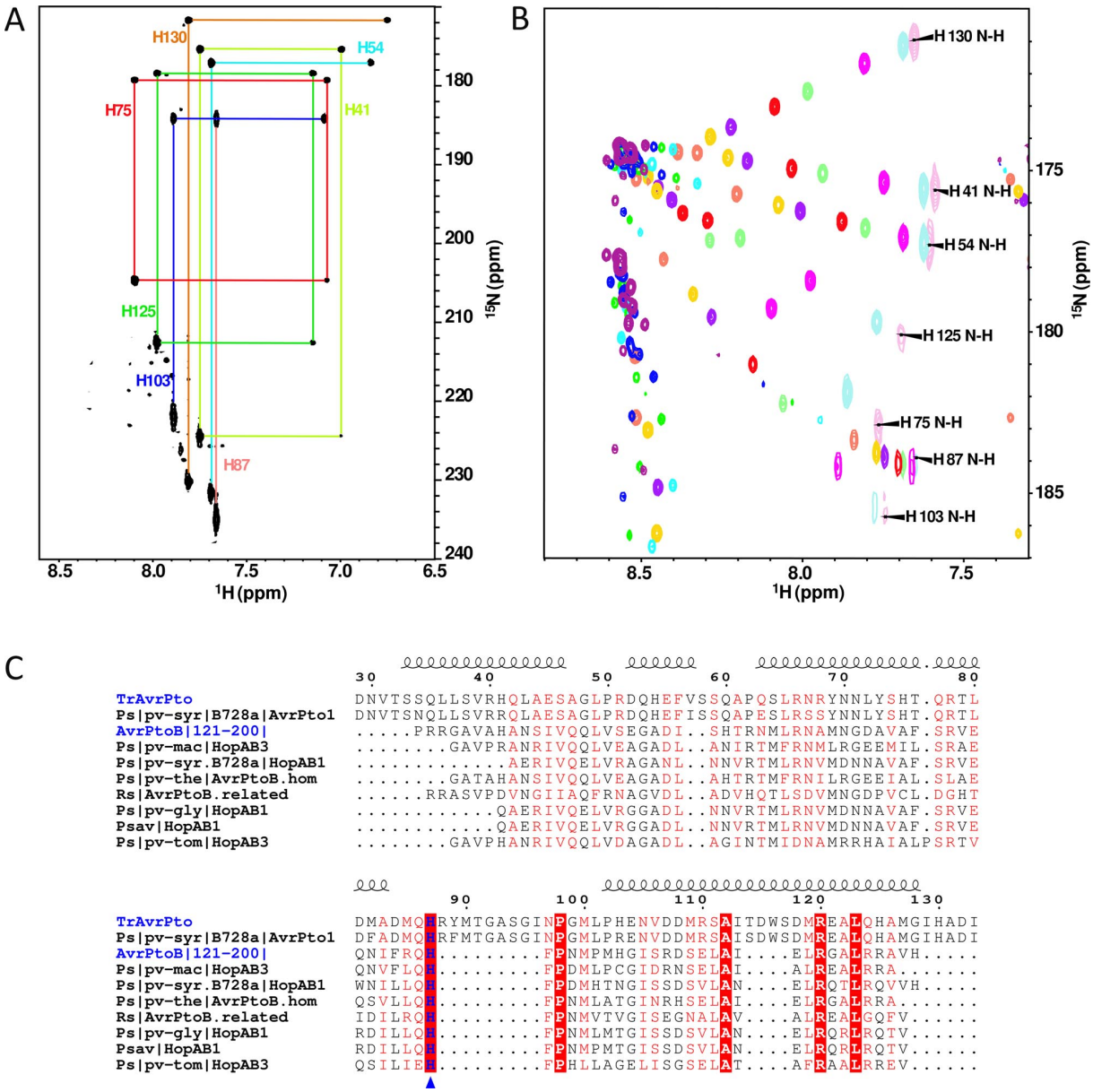

Fig. S3. Histidine pK_a values and sequence conservation indicate the functional significance of H87. (*A*) ^{15}N–^{1}H HMQC spectrum of TrAvrPto at pH 7.0. The peak patterns indicate that each side chain is predominantly in the Nε2-H tautomeric state. (*B*) Expanded region (corresponding to the ^{15}Nε2, ^{1}Hε1 peak region) of ^{15}N–^{1}H HMQC spectra of WT TrAvrPto at pH 4.85 (maroon), 5.05 (blue), 5.20 (green), 5.39 (cyan), 5.63 (coral), 5.90 (gold), 6.02 (purple), 6.32 (red), 6.58 (light green), 6.93 (magenta), 7.51 (light blue), and 7.83 (pink). In the absence of tautomer exchange, as one ^{15}N is protonated, its peak should shift to higher ^{15}N frequency (downfield), whereas the other ^{15}N peak moves to lower frequency (upfield). This is precisely what is observed for H130, with the protonating ^{15}Nε2 shifting to higher ^{15}N frequency. Conversely, fast and intermediate tautomer exchange typically causes the protonated ^{15}N peak to move downfield with increase in pH. The remaining 6 histidines of TrAvrPto display this behavior. Additionally, the line widths of H87 and H103 broaden in the ^{15}N dimension as pH is increased, consistent with the intermediate exchange with neutral tautomers. The side chains H41, H54, H75, and H125 display little to no line broadening, indicating that the tautomerization rates of these side chains are faster than those of H87 and H103. (*C*) Alignment of the sequences of AvrPto and AvrPtoB homologs, found by PSI-BLAST search of TrAvrPto and AvrPtoB$_{121–200}$ (names in blue). AvrPto1 ortholog from *P. syringae* pv. *syringae* B728a (residues 29–133, denoted as Ps|pv-syr|B728a|AvrPto1) is 90% identical to TrAvrPto. Homologous sequences for AvrPtoB$_{121–200}$ were found to share from 41% (*Ralstonia solanacearum* AvrPtoB-related protein, residues 131–208, denoted as Rs|AvrPtoB.related) to 66% (*P. syringae* pv. *theae* avirulence protein AvrPtoB homolog, residues 130–204, denoted as Ps|pv-the341AvrPtoB.hom) identity in the sequence region shown. The remaining sequences are *P. syringae* pv. *maculicola* HopAB3 (60% identity, residues 141–215, denoted as Ps|pv-mac|HopAB3), *P. syringae* pv. *syringae* strain B728a HopAB1 (59% identity, residues 98–169, denoted as Ps|pv-syr.B728a|HopAB1), *P. syringae* pv. *glycinea* HopAB1 (61% identity, residues 94–165, denoted as Ps|pv-gly|HopAB1), *P. savastanoi* HopAB1 (62% identity, residues 94–165, denoted as Psav|HopAB1), and *P. syringae* pv. *tomato* HopAB1 (49% identity, residues 138–214, denoted as Ps|pv-tom|HopAB3). The only absolutely conserved histidine is H87 (highlighted by a blue triangle). Secondary structure designations derived from the deposited NMR structure (PDB ID code 1R5E).

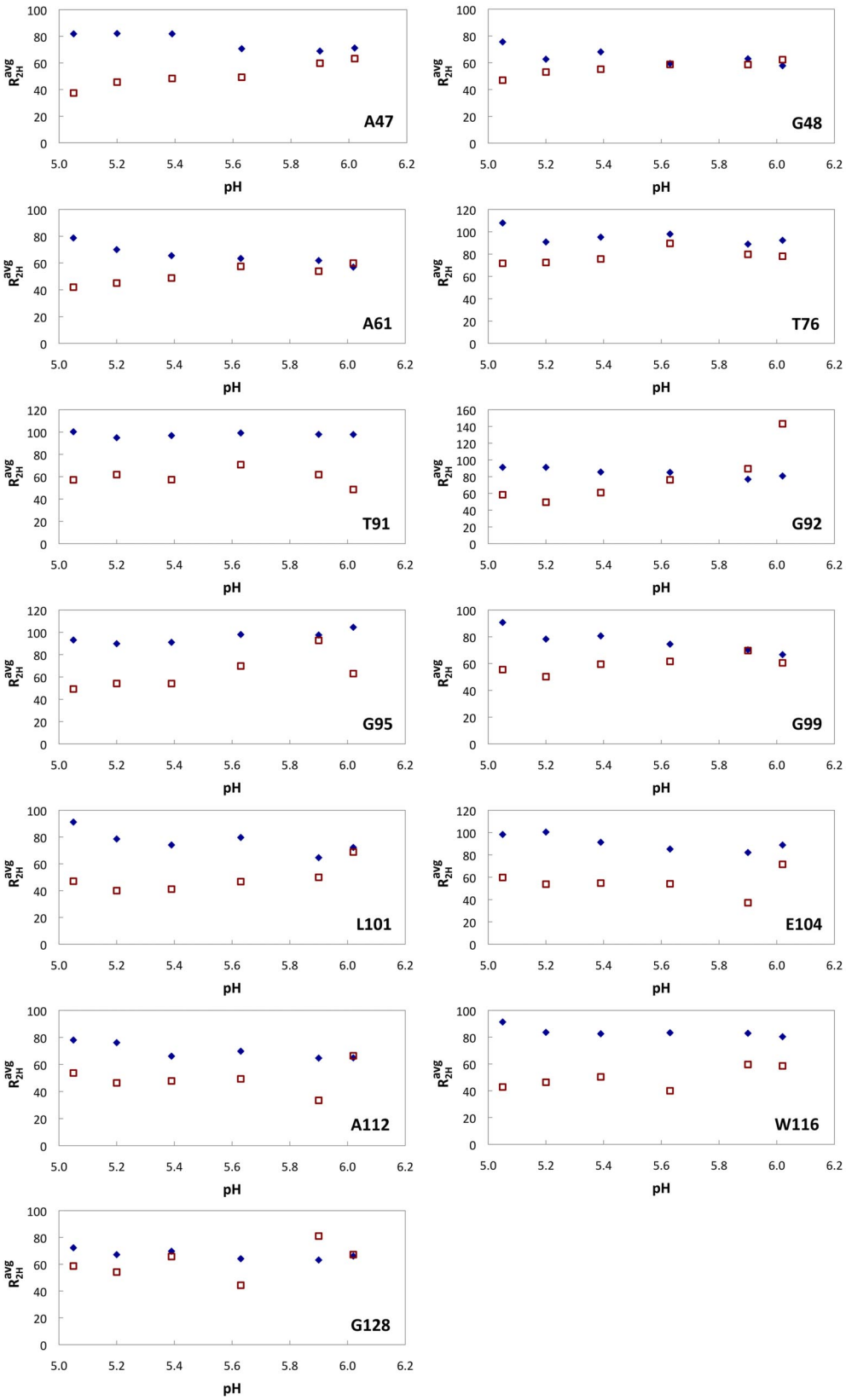

Fig. S4. Average proton transfer relaxation rates, R_{2H}^{avg} (s^{-1}), for the folded (filled diamonds) and unfolded (open squares) peaks of the analyzed residues at different pH values.

Table S1. Concentration independence of WT TrAvrPto shown by p_F values of individual residues at 420 μM and 80 μM measured at pH 5.0 by using ^{15}N–^{1}H fHSQC

	A47	G48	A61	T76	T91	G92	G95	G99	L101	E104	A112	W116	G128
$p_F^{420\mu M}$	0.523	0.514	0.555	0.510	0.516	0.524	0.517	0.542	0.539	0.509	0.493	0.525	0.500
$p_F^{80\mu M}$	0.515	0.515	0.543	0.505	0.508	0.506	0.524	0.543	0.559	0.508	0.486	0.525	0.482