Supplementary Material

Figure S1: Result of combinatorial scan on the titration curves of Glu78 and Glu172 in BCX WT. A total of 640,000 parameter combinations were tested to investigate if the fitter found an optimal and unique solution. The fitted solution is marked by a green cross. Insert: all 640,000 combination, main plot: zoomed in around the best solution. Top: Mean unsigned error plotted as function of Euclidean distance in parameter space (intrinsic pKa values and interaction energies) to optimal solution. Bottom: Mean unsigned error plotted as function of interaction energy.

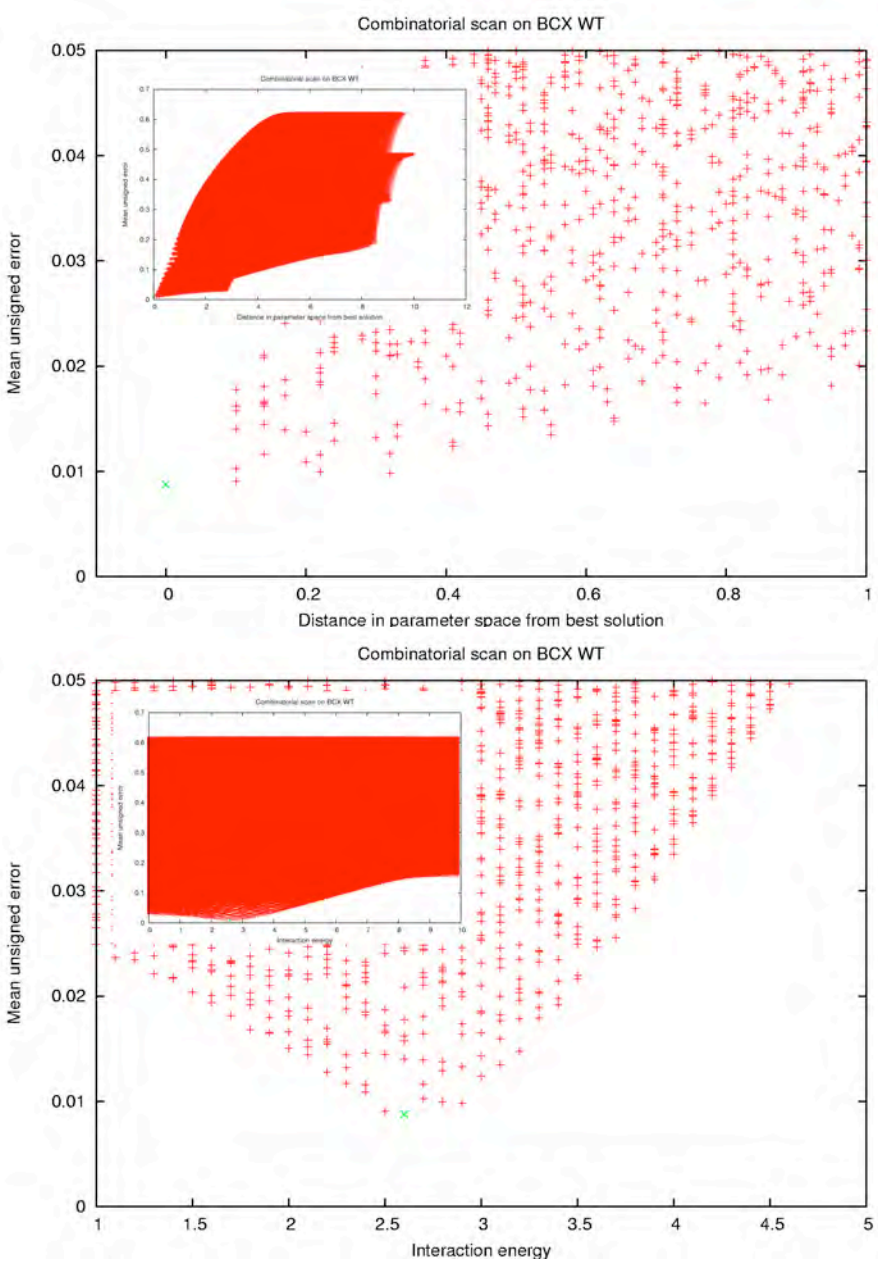
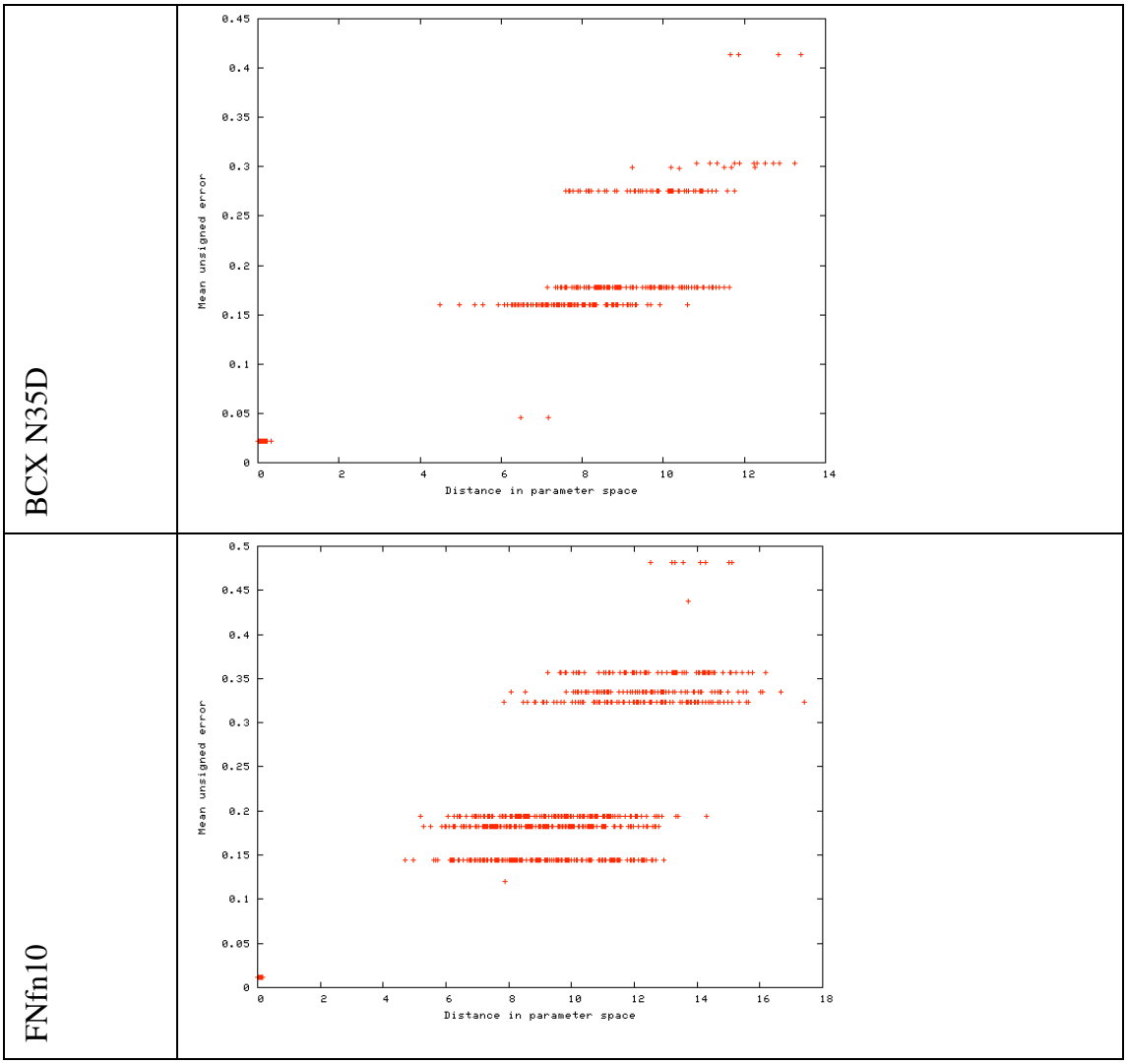


Figure S2: Uniqueness tests of BCX N35D, FNfn10, RNase A, and OHT. In order to investigate, if the fitter will converge to a unique solution independently on the starting conditions, 1000 fits were done with random starting conditions for each system. The graphs present the error versus distance in parameter space of each of the 1000 solutions. Generally, fits either converge to an optimal and unique solution, or the fit does not converge, resulting in an error that is significantly higher than that of the optimal solution. Two runs for BCX N35D resulted in fits far away in parameter space from the optimal solution and with an error that is only twice that of the optimal solutions. Theses two cases occurred due to premature invocation of the convergence criterion and manual continuation of the fits lead to the optimal solution.



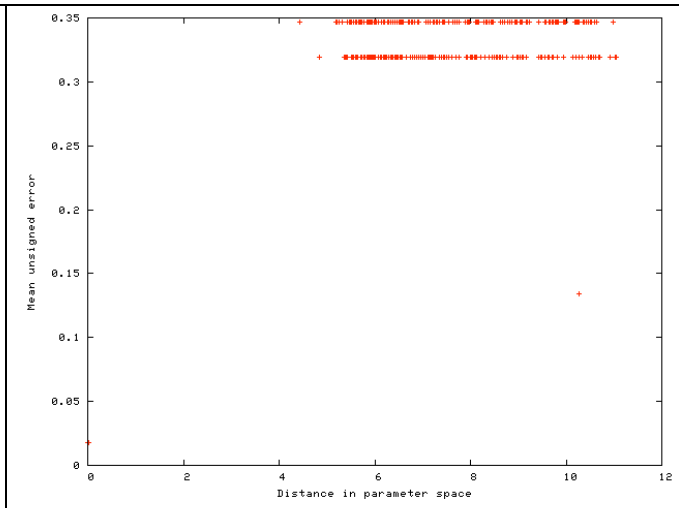
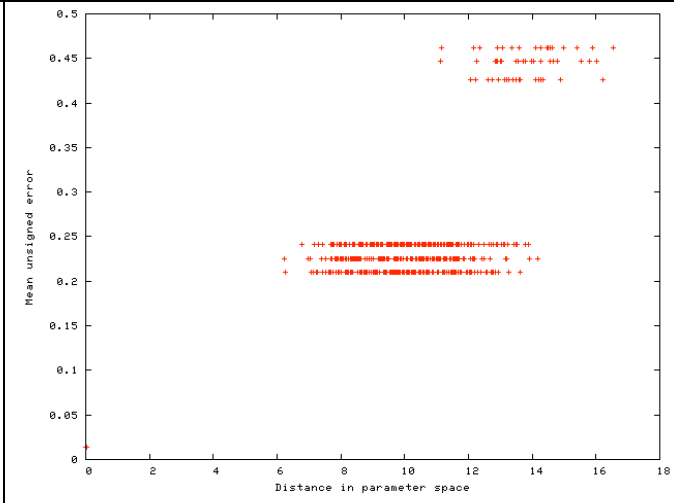
RNase A	 Mean unsigned error Distance in parameter space 0 0.05 0.1 0.15 0.2 0.25 0.3 0.35 0 2 4 6 8 10 12
HOT	 Mean unsigned error Distance in parameter space 0 0.05 0.1 0.15 0.2 0.25 0.3 0.35 0.4 0.45 0.5 0 2 4 6 8 10 12 14 16 18

Figure S3: Global fit of NMR titration curves for FNfn10: Asp7: black curve, Glu9: blue curve, Asp23: red curve. NMR titration data is represented by dots.

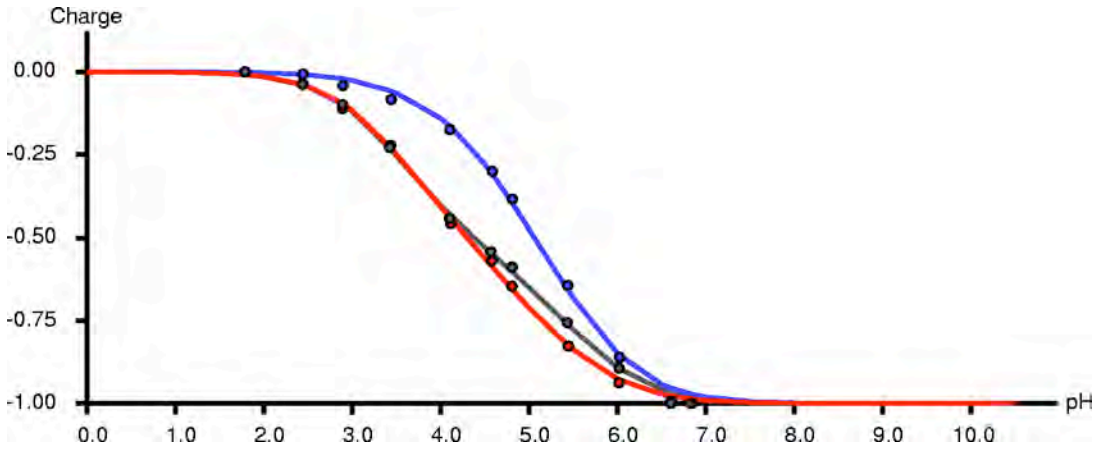
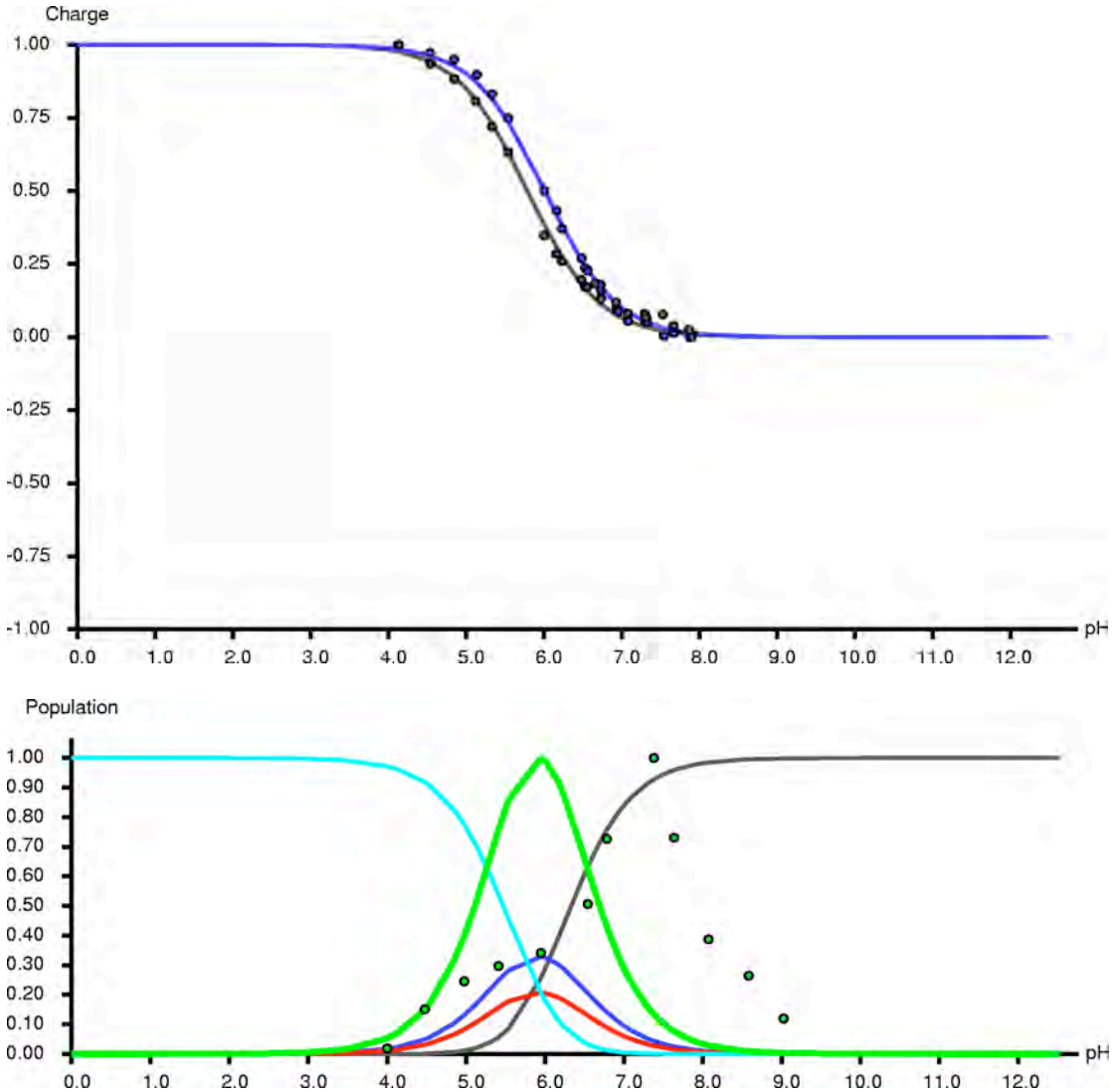
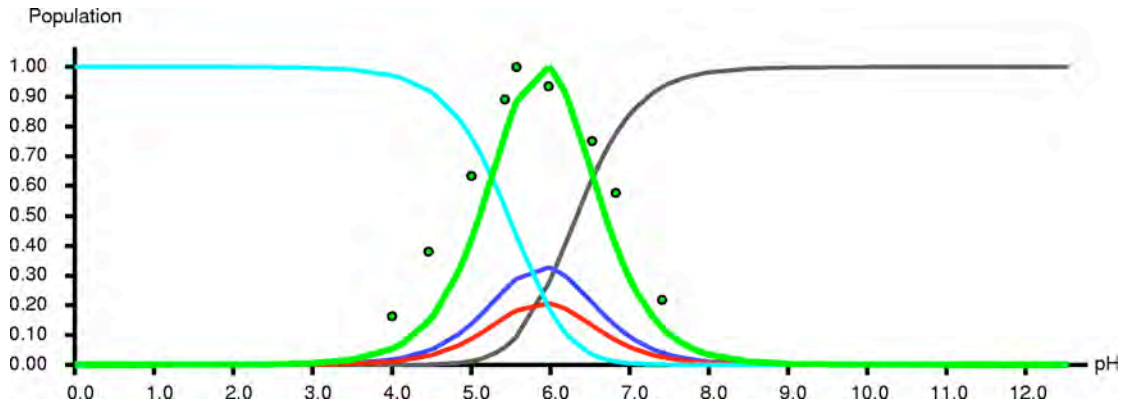


Figure S4: Global fit of His12 (black curve) and His119 (blue curve) of bovine pancreatic ribonuclease (1st graph). Fitted and experimental pH activity profile of bovine pancreatic ribonuclease measured in water (2nd graph), 50% (v/v) dioxane (3rd graph), and 50% (v/v) formamide (4th graph). The protonation state populations are colored as follows: cyan: [+1,+1], blue: [0,+1], red: [+1,0], and black: [0,0], where '0' and '+1' signifies neutral and charged forms of [His12, His119], respectively. The CCPS [0,1] is scaled to unity (green line)




Population
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
11.0
12.0
pH

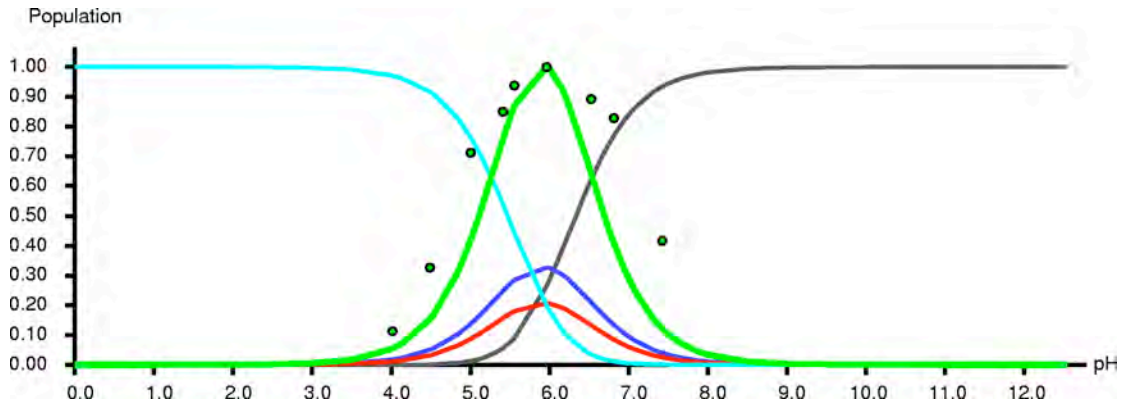

Population
1.00
0.90
0.80
0.70
0.60
0.50
0.40
0.30
0.20
0.10
0.00
0.0
1.0
2.0
3.0
4.0
5.0
6.0
7.0
8.0
9.0
10.0
11.0
12.0
pH

Figure S5: Global fit of RNase A including the titration curves of His12 (black curve), His119 (blue curve), an unknown base (red curve), and the pH-activity profile measured in water. Including an extra base yields a reasonable simultaneous fit to the titration curves and the pH-activity profile. However, the maximal population of the CCPS (red line; scaled to unity: green line) would in this case only be 0.9 %. Other protonation states include: black: [0,0,0], blue: [0,0,+1], purple: [+1,0,0], cyan: [0,+1,+1], yellow: [+1,0,+1], orange: [+1,+1,0], gray: [+1,+1,+1], where '0' and '+1' signifies neutral and charged forms of the residues [His12, His119, Unknown], respectively.

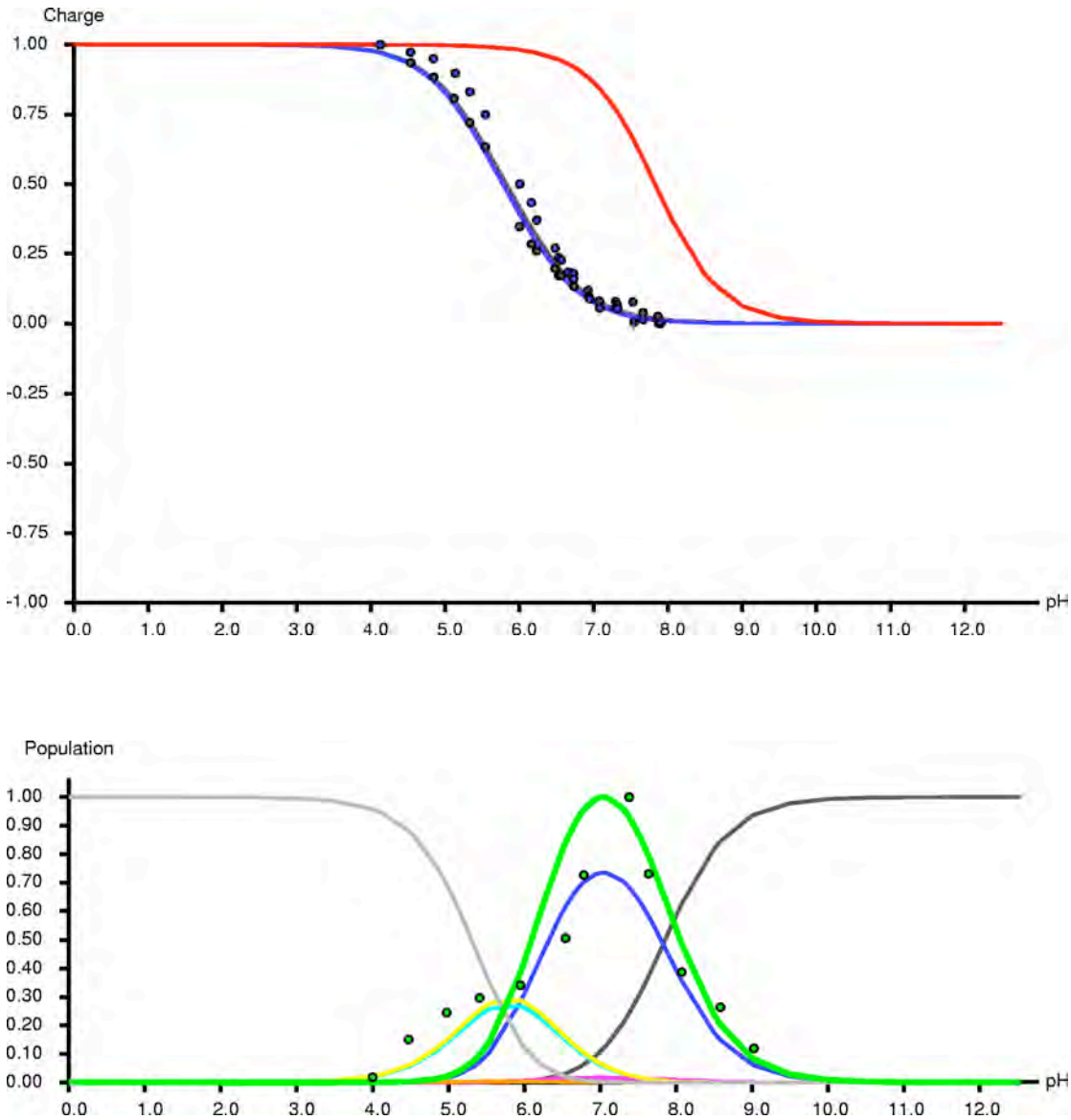


Figure S6: Global fit of Asp58 (black), Asp60 (red), and Asp61 (blue) of human oxidized thioredoxin. See Tables 1 and 2 in main article for fitted parameters.

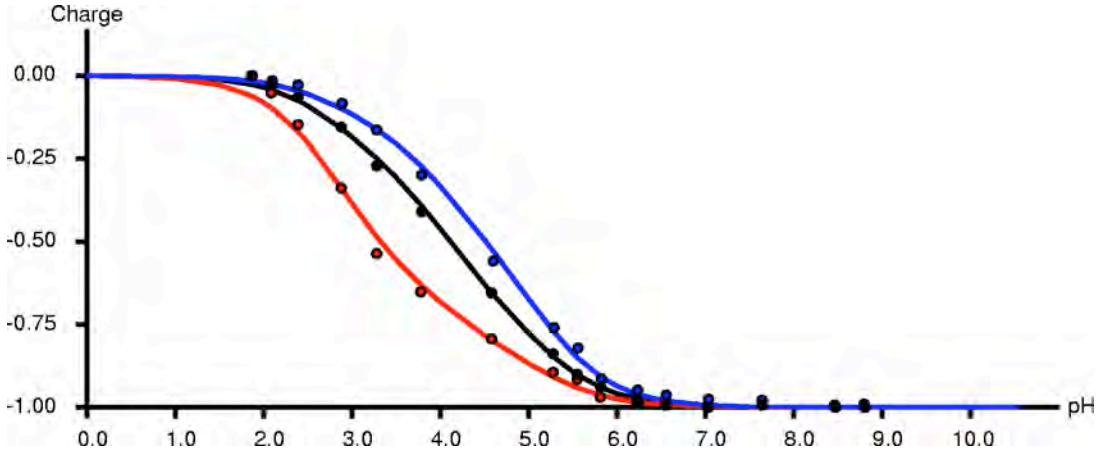


Figure S7: Theoretical FT-IR titration curve (red line) calculated from GloFTE fit of BCX wild type titration curves and pH-activity profile, compared with FT-IR data (dots) from Davoodi *et al*.[1] The theoretical FT-IT is calculated by summing the protonation fractions of the theoretical titration curves as described in Material and Methods.

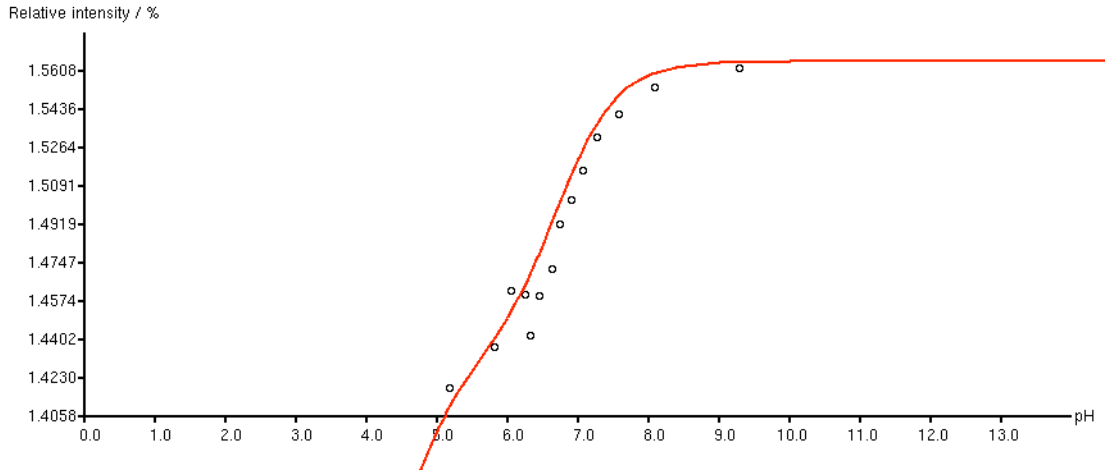


Table S1: Identification of Glu78 and Glu172 in BCX WT using the titration curves of both Glu78 and Glu172. The correct residues are ranked as number 7/8 (There is no way to distinguish the order of residues when only two residues included in the fit).

Glu78[a]	Glu172[a]	Length of difference matrix[b]
Arg136	Arg132	0.10309
Arg132	Arg136	0.10309
Asp101	Asp83	0.26176
Asp83	Asp101	0.26176
His149	Arg136	0.46364
Arg136	His149	0.46364
Glu172	Glu78	1.70385
Glu78	Glu172	1.70385
His149	Arg132	1.99750
Arg132	His149	1.99750

[a] Residue assigned to NMR titration curve

[b] 'Length of difference matrix' is calculated using the Frobenius inner product, as described in Materials and Methods.

Table S2: Identification of Glu78 or Glu172 when one residue is known. One residue (Glu78 or Glu172) is assumed to be known and the other residue is predicted using the fitted interaction energy between the residues. Values in the 'Energy' column are the interaction energies of the predicted residues with the known residue (Glu78 or Glu172). 'Diff' is the absolute difference between the fitted interaction energy between Glu78 and Glu172 and calculated energy between Glu78 or Glu172 and the predicted residue. Glu78 and Glu172 are ranked as number two and four, respectively. However, taking the high expected intrinsic pKa value of tyrosine residues into account, the top ranking tyrosine residues can be discarded in both cases, leaving the correct residues correctly predicted.

Identification of Glu78			Identification of Glu172		
Residue	Energy	Diff	Residue	Energy	Diff
Tyr65	3.51	0.86	Tyr166	1.86	0.79
Glu78	3.85	1.20	Tyr79	1.70	0.95
Tyr166	1.24	1.41	Tyr65	1.49	1.16
Tyr174	1.18	1.47	Glu172	3.85	1.20
Tyr69	4.12	1.48	Tyr174	0.75	1.90
Tyr79	1.01	1.64	Asp11	0.66	1.98
Tyr26	0.91	1.73	Tyr128	0.65	2.00
Tyr88	0.85	1.80	Tyr88	0.57	2.08
Asp11	0.77	1.87	Tyr26	0.52	2.12
Asp83	0.63	2.02	Tyr5	0.48	2.16

Table S3: Identification of the number of interacting titratable groups for BCX N35D when including the titration curves E78 and E172 and the pH-activity profile. The top row shows the minimal setup (two theoretical acidic groups). The following rows shows the fits with an increasing number of theoretical groups and different combinations of acid-base assignments of the theoretical groups. The total error is the average of the error on the titration curves and the error on the pH-activity profile. Improvements in the last three columns are relative to the minimal setup in the first row. Optimal setups are chosen when adding an extra theoretical group does not significantly (< 5 percentage points) improve the overall fit. Optimal setups are **in bold**. The optimal setups are all seen to include three theoretical groups and in each case the CCPS set to one deprotonated and two protonated groups). The conclusion of this analysis is that the experimental NMR titration curves of E78 and E172 are optimal reproduced using three theoretical groups. The specific assignment of acids and bases does change the overall quality of the fit.

Loaded residues		No. groups	No. acids	No. bases	No. deprotonated	Error on titration curves	Error on activity data	Total error	Improve-ment on titration curves / %	Improve-ment on activity data / %	Total improve-ment / %
E78(acid)	E172(acid)	2	2	0	0	0.0688	0.197	0.133	0	0	0
E78(acid)	E172(acid)	2	2	0	1	0.122	0.346	0.234	-76.9	-75.7	-76
E78(base)	E172(acid)	2	1	1	0	0.0688	0.197	0.133	-0.01	0	0
E78(acid)	E172(base)	2	1	1	0	0.0688	0.197	0.133	0	0	0
E78(base)	E172(acid)	2	1	1	1	0.122	0.346	0.234	-76.9	-75.7	-76
E78(acid)	E172(base)	2	1	1	1	0.122	0.346	0.234	-76.9	-75.7	-76
E78(base)	E172(base)	2	0	2	0	0.0688	0.197	0.133	0	0	0
E78(base)	E172(base)	2	0	2	1	0.122	0.346	0.234	-76.9	-75.7	-76
E78(acid)	E172(acid)	3	3	0	0	0.0679	0.198	0.133	1.32	-0.3	0.12
E78(acid)	**E172(acid)**	**3**	**3**	**0**	**1**	**0.0276**	**0.0626**	**0.0451**	**59.9**	**68.3**	**66.1**
E78(acid)	E172(acid)	3	3	0	2	0.122	0.346	0.234	-76.9	-75.7	-76
E78(acid)	E172(base)	3	2	1	0	0.0684	0.197	0.133	0.67	-0.03	0.15
E78(base)	E172(acid)	3	2	1	0	0.0678	0.198	0.133	1.54	-0.27	0.2
E78(acid)	E172(acid)	3	2	1	0	0.0663	0.203	0.134	3.66	-2.75	-1.09
E78(base)	**E172(acid)**	**3**	**2**	**1**	**1**	**0.0276**	**0.0626**	**0.0451**	**59.9**	**68.3**	**66.1**
E78(acid)	**E172(acid)**	**3**	**2**	**1**	**1**	**0.0275**	**0.0623**	**0.0449**	**60**	**68.4**	**66.2**

E78(acid)	**E172(base)**	**3**	**2**	**1**	**1**	**0.0276**	**0.0626**	**0.0451**	**59.9**	**68.3**	**66.1**
E78(base)	E172(acid)	3	2	1	2	0.122	0.347	0.234	-76.9	-75.7	-76
E78(acid)	E172(acid)	3	2	1	2	0.122	0.346	0.234	-76.9	-75.7	-76
E78(acid)	E172(base)	3	2	1	2	0.122	0.346	0.234	-76.9	-75.7	-76
E78(base)	E172(base)	3	1	2	0	0.0678	0.198	0.133	1.53	-0.38	0.12
E78(acid)	E172(base)	3	1	2	0	0.068	0.197	0.133	1.15	-0.1	0.22
E78(base)	E172(acid)	3	1	2	0	0.0661	0.203	0.134	3.91	-2.83	-1.09
E78(base)	**E172(acid)**	**3**	**1**	**2**	**1**	**0.0276**	**0.0626**	**0.0451**	**59.9**	**68.3**	**66.1**
E78(base)	**E172(base)**	**3**	**1**	**2**	**1**	**0.0276**	**0.0626**	**0.0451**	**59.9**	**68.3**	**66.1**
E78(acid)	**E172(base)**	**3**	**1**	**2**	**1**	**0.031**	**0.0614**	**0.0462**	**54.9**	**68.9**	**65.3**
E78(base)	E172(base)	3	1	2	2	0.122	0.346	0.234	-76.9	-75.7	-76
E78(base)	E172(acid)	3	1	2	2	0.122	0.346	0.234	-76.9	-75.7	-76
E78(acid)	E172(base)	3	1	2	2	0.122	0.346	0.234	-76.9	-75.7	-76
E78(base)	E172(base)	3	0	3	0	0.0663	0.202	0.134	3.62	-2.56	-0.96
E78(base)	**E172(base)**	**3**	**0**	**3**	**1**	**0.0276**	**0.0626**	**0.0451**	**59.9**	**68.3**	**66.1**
E78(base)	E172(base)	3	0	3	2	0.122	0.346	0.234	-76.9	-75.7	-76
E78(acid)	E172(acid)	4	4	0	0	0.0692	0.197	0.133	-0.6	0.03	-0.13
E78(acid)	E172(acid)	4	4	0	1	0.0276	0.0611	0.0444	59.8	69	66.6
E78(acid)	E172(acid)	4	4	0	2	0.0278	0.0627	0.0452	59.7	68.2	66
E78(acid)	E172(acid)	4	4	0	3	0.122	0.347	0.234	-77	-75.8	-76.1
E78(acid)	E172(acid)	4	3	1	0	0.0681	0.198	0.133	1.02	-0.42	-0.05
E78(acid)	E172(base)	4	3	1	0	0.0689	0.197	0.133	-0.09	0.04	0.01
E78(base)	E172(acid)	4	3	1	0	0.0693	0.197	0.133	-0.7	0.07	-0.13
E78(acid)	E172(base)	4	3	1	1	0.0491	0.0596	0.0544	28.7	69.8	59.1
E78(acid)	E172(acid)	4	3	1	1	0.0296	0.0623	0.0459	57	68.4	65.5
E78(base)	E172(acid)	4	3	1	1	0.0289	0.0614	0.0451	58	68.9	66.1
E78(acid)	E172(acid)	4	3	1	2	0.0277	0.0627	0.0452	59.7	68.2	66
E78(acid)	E172(base)	4	3	1	2	0.0278	0.0626	0.0452	59.6	68.2	66
E78(base)	E172(acid)	4	3	1	2	0.0279	0.0626	0.0453	59.4	68.3	66
E78(acid)	E172(base)	4	3	1	3	0.122	0.347	0.234	-76.9	-75.8	-76.1
E78(acid)	E172(acid)	4	3	1	3	0.122	0.347	0.234	-77	-75.9	-76.2
E78(base)	E172(acid)	4	3	1	3	0.122	0.347	0.234	-77	-75.8	-76.1
E78(acid)	E172(base)	4	2	2	0	0.0682	0.197	0.133	0.86	-0.16	0.1
E78(base)	E172(base)	4	2	2	0	0.0692	0.197	0.133	-0.58	0.22	0.01
E78(acid)	E172(acid)	4	2	2	0	0.0673	0.201	0.134	2.17	-2.16	-1.04
E78(base)	E172(acid)	4	2	2	0	0.068	0.198	0.133	1.24	-0.48	-0.04
E78(acid)	E172(base)	4	2	2	1	0.0306	0.0628	0.0467	55.5	68.2	64.9
E78(acid)	E172(acid)	4	2	2	1	0.0293	0.0623	0.0458	57.5	68.4	65.6
E78(base)	E172(base)	4	2	2	1	0.0266	0.0591	0.0428	61.4	70	67.8
E78(base)	E172(acid)	4	2	2	1	0.0308	0.0625	0.0466	55.2	68.3	64.9
E78(acid)	E172(base)	4	2	2	2	0.0278	0.0626	0.0452	59.6	68.3	66
E78(base)	E172(acid)	4	2	2	2	0.0277	0.0627	0.0452	59.7	68.2	66
E78(base)	E172(base)	4	2	2	2	0.028	0.0625	0.0453	59.3	68.3	66
E78(acid)	E172(acid)	4	2	2	2	0.0277	0.0627	0.0452	59.8	68.2	66
E78(acid)	E172(base)	4	2	2	3	0.122	0.347	0.234	-76.9	-75.8	-76.1
E78(acid)	E172(acid)	4	2	2	3	0.122	0.347	0.234	-76.9	-75.8	-76.1
E78(base)	E172(base)	4	2	2	3	0.122	0.347	0.234	-76.9	-75.8	-76.1
E78(base)	E172(acid)	4	2	2	3	0.122	0.347	0.234	-77	-75.8	-76.1
E78(base)	E172(acid)	4	1	3	0	0.0671	0.202	0.135	2.47	-2.44	-1.17
E78(base)	E172(base)	4	1	3	0	0.0678	0.198	0.133	1.47	-0.55	-0.03
E78(acid)	E172(base)	4	1	3	0	0.0685	0.197	0.133	0.47	-0.15	0.01

E78(base)	E172(acid)	4	1	3	1	0.0301	0.0624	0.0462	56.3	68.4	65.2
E78(base)	E172(base)	4	1	3	1	0.036	0.0706	0.0533	47.8	64.2	59.9
E78(acid)	E172(base)	4	1	3	1	0.0271	0.0613	0.0442	60.6	68.9	66.8
E78(acid)	E172(base)	4	1	3	2	0.0277	0.0626	0.0451	59.8	68.2	66.1
E78(base)	E172(acid)	4	1	3	2	0.0276	0.0627	0.0451	59.9	68.2	66.1
E78(base)	E172(base)	4	1	3	2	0.0275	0.0622	0.0448	60.1	68.5	66.3
E78(acid)	E172(base)	4	1	3	3	0.122	0.347	0.234	-76.9	-75.8	-76.1
E78(base)	E172(acid)	4	1	3	3	0.122	0.347	0.234	-77	-75.8	-76.1
E78(base)	E172(base)	4	1	3	3	0.122	0.347	0.234	-76.9	-75.8	-76.1
E78(base)	E172(base)	4	0	4	0	0.068	0.199	0.133	1.18	-0.81	-0.29
E78(base)	E172(base)	4	0	4	1	0.0302	0.0619	0.0461	56.1	68.6	65.4
E78(base)	E172(base)	4	0	4	2	0.0509	0.0619	0.0564	26.1	68.6	57.6
E78(base)	E172(base)	4	0	4	3	0.122	0.347	0.234	-76.9	-75.8	-76.1

Table S4: Identification of titrating residues in *E. coli* reduced thioredoxin. Both interpretations (I1 and I2 – see main article) of the NMR titration data of Cys32 and Cys35 are fitted using the GloFTE method and the three interacting residues Asp26, Cys32, and Cys35 are attempted identified. Identification is unsuccessful for both interpretations.

Data used	Rank of correct identification (groups correctly identified)
Asp26, Cys32, pH-activity profile (I1)	>100 (Asp26, Cys32, Cys35)
Asp26, Cys32, Cys35, pH-activity profile (I2)	>100 (Asp26, Cys32, Cys35)

Figure S8: RET Interpretation #1:Simultaneous fit to the titration curve of Cys32 (black curve) as measured by [2] and the titration curve of Asp26 (blue curve) as measured by [3] together with the pH-activity profile[4]. Three theoretical groups are employed in the fit, and it is seen that the titration curve of the third theoretical curve (red curve) bears a striking resemblance to the titration curve of Cys35 employed in RET Interpretation #2.

The protonation states are colored as follows: black: [0,0,0], blue: [0,0,-1], red: [0,-1,0], cyan: [0,-1,-1], purple [-1,0,0], yellow: [-1,0,-1], orange: [-1,-1,0], gray: [-1,-1,-1], where '0' and '-1' signifies neutral and charged forms of the residues [Asp26, Cys32, Unknown], respectively.

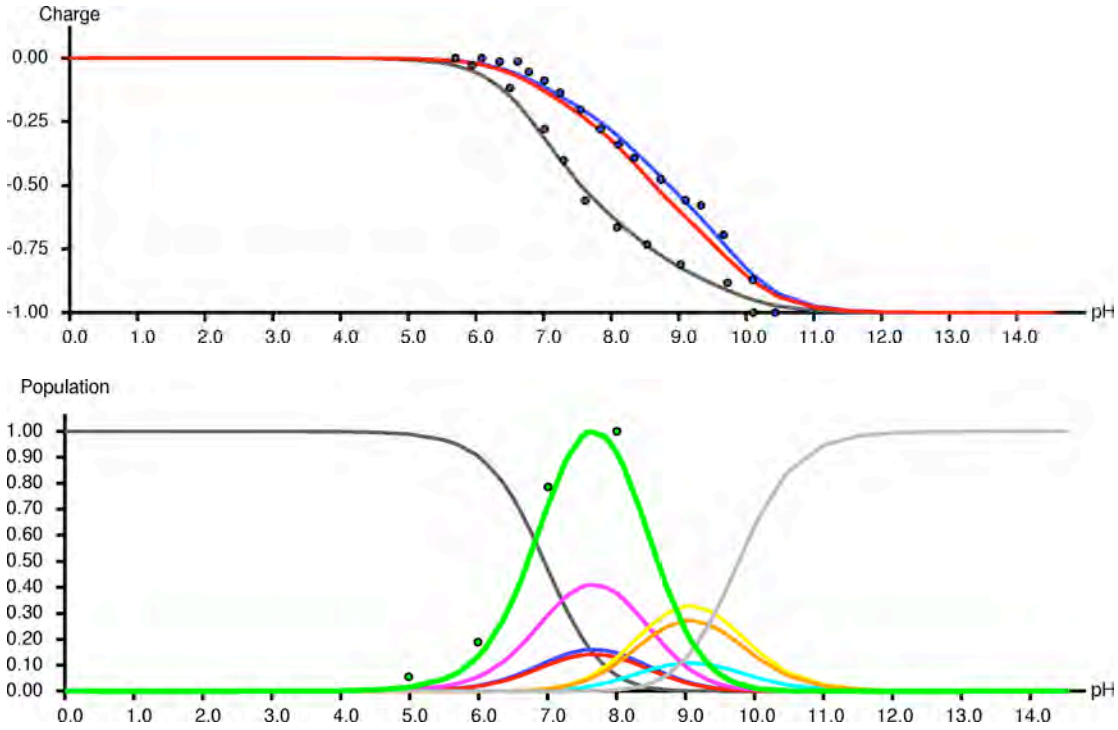
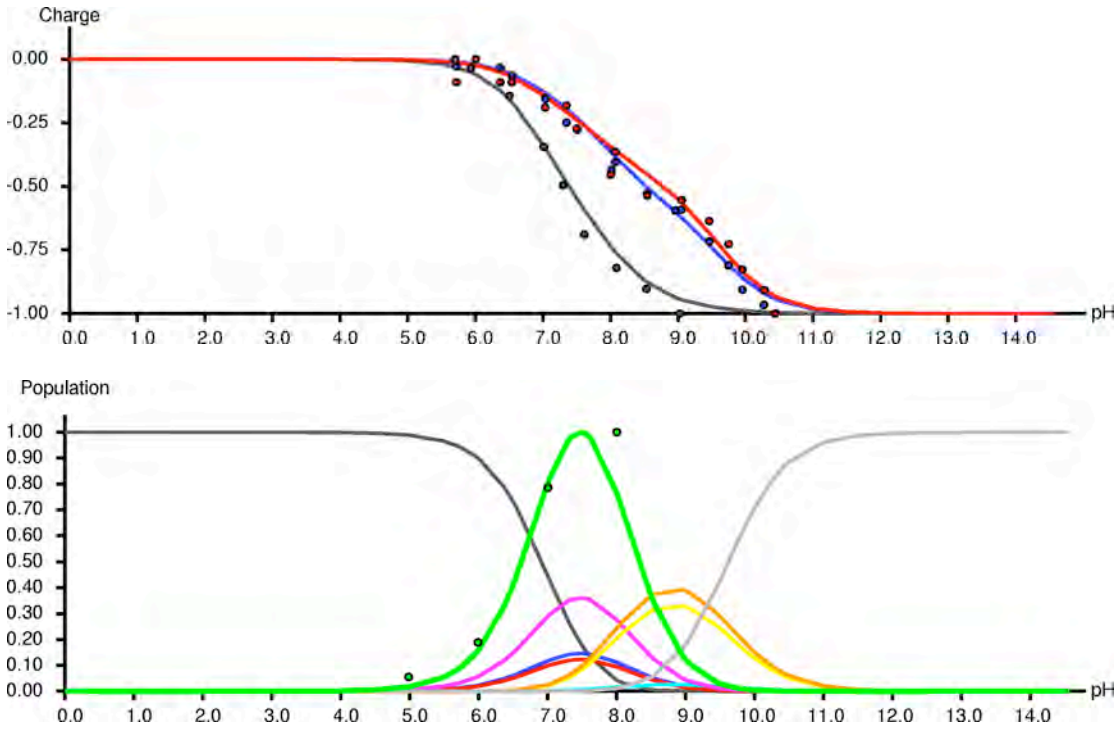


Figure S9: RET Interpretation #2: Global fit of the titration curves of Cys32 (black curve), Cys35 (red curve) as measured in [5], and Asp26 (blue curve) a measured in [3] and pH-activity profile[4] of RET. The protonation states are colored as follows: black: [0,0,0], blue: [0,0,-1], red: [0,-1,0], cyan: [0,-1,-1], purple [-1,0,0], yellow: [-1,0,-1], orange: [-1,-1,0], gray: [-1,-1,-1], where '0' and '-1' signifies neutral and charged forms of the residues [Asp26, Cys32, Cys35], respectively.



References

1. Davoodi, J., Wakarchuk, W.W., Campbell, R.L., Carey, P.R. & Surewicz, W.K. Abnormally high pKa of an active-site glutamic acid residue in Bacillus circulans xylanase. The role of electrostatic interactions. *Eur J Biochem* **232**, 839-43 (1995).
2. Chivers, P.T. et al. Microscopic pKa values of Escherichia coli thioredoxin. *Biochemistry* **36**, 14985-91 (1997).
3. Jeng, M.F. & Dyson, H.J. Direct measurement of the aspartic acid 26 pKa for reduced Escherichia coli thioredoxin by 13C NMR. *Biochemistry* **35**, 1-6 (1996).
4. Dyson, H.J. et al. Effects of buried charged groups on cysteine thiol ionization and reactivity in Escherichia coli thioredoxin: structural and functional characterization of mutants of Asp 26 and Lys 57. *Biochemistry* **36**, 2622-36 (1997).
5. Jeng, M.F., Holmgren, A. & Dyson, H.J. Proton sharing between cysteine thiols in Escherichia coli thioredoxin: implications for the mechanism of protein disulfide reduction. *Biochemistry* **34**, 10101-5 (1995).