

Supplementary Information Section

Hydrogen exchange reveals Hsp104 architecture,
structural dynamics, and energetics in physiological solution

Xiang Ye^{1,2}, Jiabei Lin², Leland Mayne^{1,2}, James Shorter², and S. Walter Englander^{1,2}

¹Johnson Research Foundation, ²Department of Biochemistry and Biophysics, Perelman School of Medicine, University of Pennsylvania, Philadelphia, PA, USA.

Correspondence should be addressed to XY, xiangye@penmedicine.upenn.edu;

or SWE, engl@penmedicine.upenn.edu.

Methods

The HX MS experiment. In this work, we expose Hsp104 preparations to hydrogen to deuterium (H to D) exchange for increasing times in D₂O buffer at physiological pH and analyze timed samples for carried D. To achieve structural resolution, we utilize a fragment separation analysis (1-4).

In a typical HX experiment, 20 μ M Hsp104 (in monomer) was incubated under the desired buffer conditions at 30°C for 5 min before H to D labeling. Labeling was initiated by 10-fold dilution into the corresponding D₂O buffer. After varying HX labeling times, the HX reaction was quenched by addition of ice-cold quench buffer of equal volume to reach a final pH of 2.5 and 1.5 M GdmCl, 0°C. The quenched samples were proteolyzed into small fragments by brief passage through an immobilized pepsin column, the fragments were separated roughly by fast HPLC, and the HPLC eluant was electrosprayed continuously into a mass spectrometer which repeatedly scans the incoming peptide ions at the rate of \sim 2 scans/sec. Analysis of the extensive MS data used the ExMS program (5) to search through the \sim 1000 high resolution scans taken during the HPLC elution. ExMS finds and verifies the isotopic mass envelopes of hundreds of variably deuterated peptides at each HX time point and determines the accumulated deuterons on each fragment in terms of the peptide Δ mass. HX time points shorter than 10 s were collected using a Biologic stopped-flow SFM4 device as described in (6, 7).

Initially, Hsp104 was buffer exchanged into H₂O-buffer containing 20 mM HEPES, 10 mM MgCl₂, 2 mM DTT, pH 7.45 with [KCl] with 750 mM KCl to produce the monomer state, 10 mM KCl for the apohexamer state, and 150 mM KCl for hexameric states. D₂O buffer used for labeling at pD 7.84 contains the same components as the H₂O-buffer. For H to D labeling at pD 8.9, HEPES was replaced with 20 mM bicine and the buffer pD was adjusted so that the pD reached 8.9 after 9 to 1 mixing with H₂O-buffer. For HX in the presence of nucleotides, both H₂O-buffer and D₂O-buffer were supplemented with 5 mM ADP, AMPPNP, or ATP γ S. Because the ATP γ S state forms slowly, the protein was always preincubated for 1 hr before measuring HX. Under ATP turnover conditions, an ATP regenerating system was used (10 mM ATP, 10 mM phosphoenolpyruvate, 20 units/ml of pyruvate kinase).

The analysis described here was based on 352 unique peptides (Fig. 1C) that passed the ExMS2 autochecks and were consistently found by ExMS under all of the conditions studied with good signal/noise, no spurious m/z peaks, and consistency with overlapping peptides. The comparison of overlapping fragments provides many internal consistency checks and allows subpeptide resolution. To achieve even higher resolution we used the HDsite program (8) which formally compares HX data from many overlapping peptides. HDsite was operated in the centroid mode to increase fitting reliability. Calculated HX rates are plotted against the Hsp104 sequence in Fig. 3. HDsite analysis can in principle resolve HX rates to individual residues but limited peptide overlap often limits attainable resolution to short sets of neighboring “switchable” residues.

HX MS data analysis. We used an upgraded version (ExMS2, to be published) of the ExMS software (5) to identify deuterated peptides in the MS scans. Thanks to the high mass accuracy and scan stability of the ThermoFisher Exactive Plus orbitrap mass spectrometer, we were able to set a very narrow mass tolerance range of 2 ppm in searching for the individual m/z peaks of each peptide. This is critical to minimize false peptide identification for the large Hsp104 protein. Peptides that passed the multiple ExMS2 autocheck criteria were subject to further manual inspection.

HX MS data interpretation. HX behavior can be understood in terms of the view originally suggested by Linderstrøm-Lang in his pioneering studies of protein hydrogen exchange which first revealed the dynamic character of protein structure (9). Lang visualized the HX process as in Eqs. 1-3 (10-12).



$$k_{ex} = k_{op} k_{ch} / (k_{op} + k_{cl} + k_{ch}) \quad \text{Eq. 2}$$

$$\Delta G_{op} = -RT \ln K_{op} = -RT \ln (k_{op}/k_{cl}) = -RT \ln (1/(Pf - 1)) \quad \text{Eq. 3}$$

In this view, amide NH hydrogens protected by H-bonding (“closed”) can only exchange when the protecting H-bond is separated in a dynamic “opening” reaction, during which time they exchange at the free peptide characteristic rate (k_{ch}) (13, 14). We now know that HX opening can occur either by way of cooperative multi-site unfolding or by more local site by site fluctuations (15).

Cooperative unfolding reactions, characteristic for low-stability elements, expose a whole set of amides to concurrent exchange. If refolding is relatively slow ($k_{cl} \ll k_{ch}$), that set of amides will all exchange at the same single exponential rate, equal to the segment unfolding rate (k_{op}), termed EX1 or monomolecular exchange. The mass spectrum will appear bimodal if the MS envelopes for the exchanged (heavier, deuterated) and the not-yet-exchanged (lighter) populations are sufficiently separate, as in Fig. 2C. The exchanged population fraction grows in time at the exact expense of the not-yet-exchanged population fraction.

Alternatively, refolding may be relatively fast ($k_{cl} \gg k_{ch}$). In this case exchange occurs in the EX2 limit (bimolecular exchange with rate dependent on pH). The HX rate is then given by k_{ch} multiplied down by the equilibrium fraction of time open (Eq. 2). Many sites exchange with the same protection factor ($Pf = k_{ch}/k_{ex}$), given by comparing the measured HX rate curve (k_{ex}) with the reference rate curve known for equivalent unstructured sites (13, 14). The mass centroid plot is parallel to the reference curve, and the offset on the time axis represents the common equilibrium protection factor. In this case, Pf approximates the fraction of time open, therefore the equilibrium constant for unfolding and leads to the free energy that stabilizes against the unfolding reaction (Eq. 3).

A different situation arises when the HX-determining opening reaction involves small residue-level distortions, called local fluctuations, rather than a cooperative multi-residue unfolding. Near neighbor sites can then exchange at very different rates. The mass centroid plot can spread over a wide HX time range reflecting the range of HX rates among the various amides in that peptide (e.g., Fig. 2D). This mode tends to dominate when the segment is flexible or when HX due to any competing segmental unfolding reaction is slower. Local fluctuations reclose very rapidly (16) and so always show EX2 character. In this case the unimodal mass envelope simply slides to higher mass in time as the varied sites exchange, as in Fig. 2A. Finally, any given segment may display both bimodal EX1 and sliding EX2 behaviors if its various amide sites exchange through these two different opening modes. The time-dependent mass spectra and mass centroid plots usually make these alternative behaviors and the dynamic information they provide easily recognizable.

The connection with structural energetics. Eq. 3 connects measured HX rates to detailed structural energetics. HX rates relate to the stability of local structure against the opening reactions that support exchange. Local destabilization will increase unfolding and produce faster HX, and the reverse. Thus HX is able to qualitatively distinguish energetically uphill and downhill changes. The quantitative validity of Eq. 3 is less clear. It seems likely to be more correct for unfolding reactions but less so for local fluctuations where the open state is locally constrained and may be slower than the unstructured reference k_{ch} value, by perhaps 10-fold (17). For example, the current results find several cases with apparent Pf in the low range of 10 or so even though no protecting H-bonds appear to be present. Further studies like the current one can contribute to this important question.

References

1. Englander JJ, Rogero JR, & Englander SW (1985) Protein hydrogen exchange studied by the fragment separation method. *Anal Biochem* 147(1):234-244.
2. Mayne L, Kan ZY, Chetty PS, Ricciuti A, Walters BT, & Englander SW (2011) Many overlapping peptides for protein hydrogen exchange experiments by the fragment separation-mass spectrometry method. *J Am Soc Mass Spectrom* 22(11):1898-1905.
3. Mayne L (2016) Hydrogen Exchange Mass Spectrometry. *Methods Enzymol* 566:335-356.
4. Gallagher ES & Hudgens J (2016) Mapping protein-ligand interactions with proteolytic-fragmentation, hydrogen/deuterium exchange-mass spectrometry. *Methods Enzymol* 566:357-404.

5. Kan ZY, Mayne L, Chetty PS, & Englander SW (2011) ExMS: data analysis for HX-MS experiments. *J Am Soc Mass Spectrom* 22(11):1906-1915.
6. Hu W, Walters BT, Kan ZY, Mayne L, Rosen LE, Marqusee S, & Englander SW (2013) Stepwise protein folding at near amino acid resolution by hydrogen exchange and mass spectrometry. *Proc Natl Acad Sci U S A* 110(19):7684-7689.
7. Ye X, Mayne L, Kan ZY, & Englander SW (2018) Folding of maltose binding protein outside of and in GroEL. *Proc Natl Acad Sci U S A* 115(3):519-524.
8. Kan ZY, Walters BT, Mayne L, & Englander SW (2013) Protein hydrogen exchange at residue resolution by proteolytic fragmentation mass spectrometry analysis. *Proc Natl Acad Sci U S A* 110(41):16438-16443.
9. Englander SW, Mayne L, Bai Y, & Sosnick TR (1997) Hydrogen exchange: the modern legacy of Linderstrom-Lang. *Protein Sci* 6(5):1101-1109.
10. Linderstrøm-Lang KU & Schellman JA (1959) Protein structure and enzyme activity. *The Enzymes*, eds Boyer PD, Lardy H, & Myrback K (Academic Press, NY), 2nd Ed, pp 443-510.
11. Hvidt A & Nielsen SO (1966) Hydrogen exchange in proteins. *Adv Protein Chem* 21:287-386.
12. Englander SW & Kallenbach NR (1983) Hydrogen exchange and structural dynamics of proteins and nucleic-acids. *Q Revs Biophys* 16:521-655.
13. Bai Y, Milne JS, Mayne L, & Englander SW (1993) Primary structure effects on peptide group hydrogen exchange. *Proteins* 17(1):75-86.
14. Connelly GP, Bai Y, Jeng MF, & Englander SW (1993) Isotope effects in peptide group hydrogen exchange. *Proteins* 17(1):87-92.
15. Skinner JJ, Lim WK, Bedard S, Black BE, & Englander SW (2012) Protein dynamics viewed by hydrogen exchange. *Protein Sci* 21(7):996-1005.
16. Anderson JS, Hernandez G, & Lemaster DM (2008) A billion-fold range in acidity for the solvent-exposed amides of *Pyrococcus furiosus* rubredoxin. *Biochemistry* 47(23):6178-6188.
17. Maity H, Lim WK, Rumbley JN, & Englander SW (2003) Protein hydrogen exchange mechanism: local fluctuations. *Protein Sci* 12(1):153-160.

Fig. SI 1: Mass centroid plots for the 352 peptides used in this study are at the end of this document.

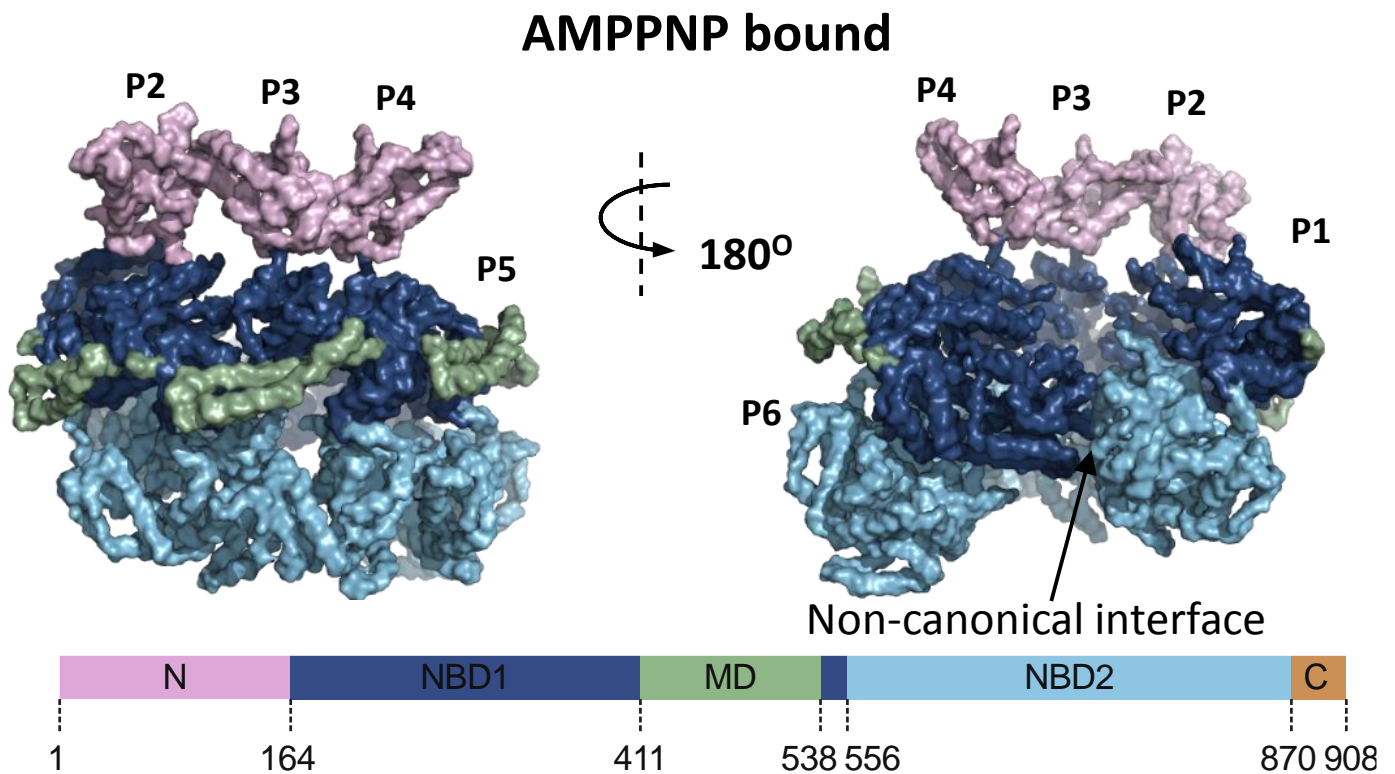
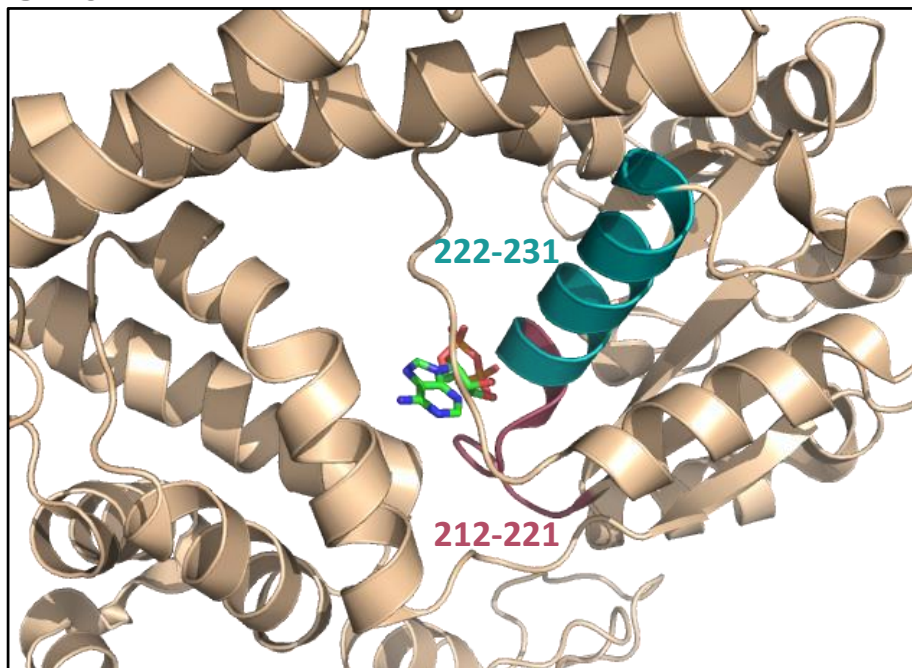
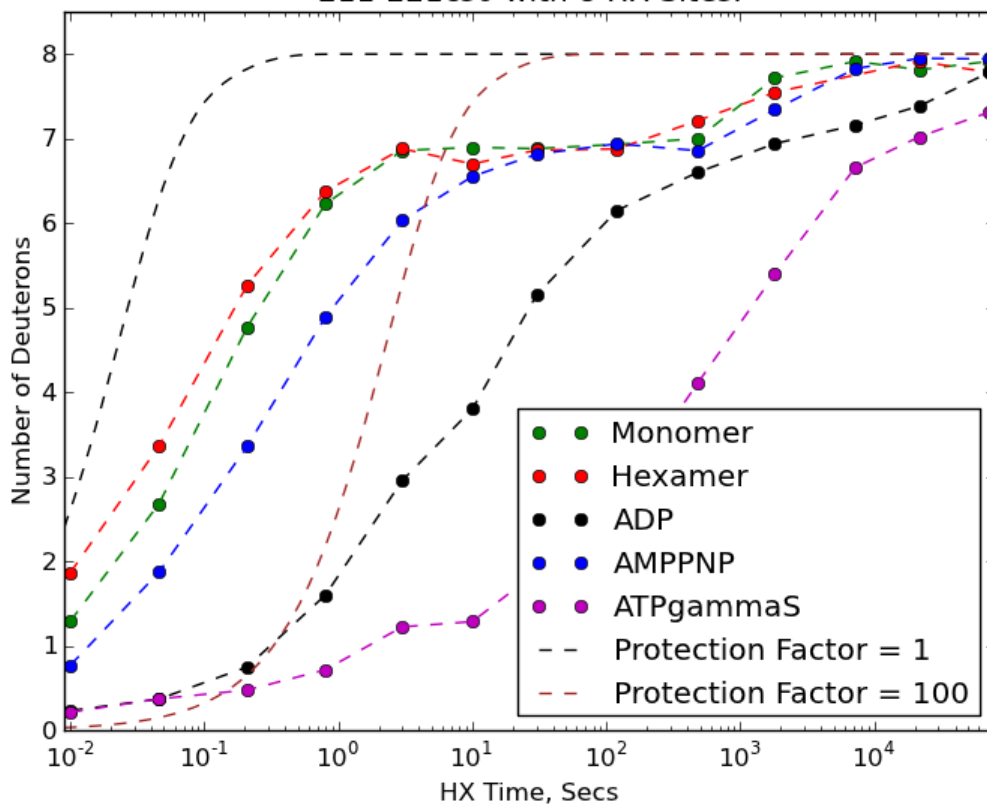


Fig. SI 2. Cryo-EM structures of Hsp104 with AMPPNP (5KNE) bound. The AMPPNP structure is color coded as in the bar below to indicate the different protomer domains. The hexamer structure is asymmetric with two kinds of interfaces. The canonical interfaces (left) all have NBD1 to NBD1 and NBD2 to NBD2 contacts from P1 around to P6. The non-canonical interface (right) has offset NBD1 to NBD2 contacts from P6 to P1 with the other two faces exposed to solvent. Peptides that occupy these hetero-positions in the AMPPNP state (Fig. 4, main text) and in the ADP state (Fig. SI 6) exchange at different rates and display bimodal HX MS profiles.

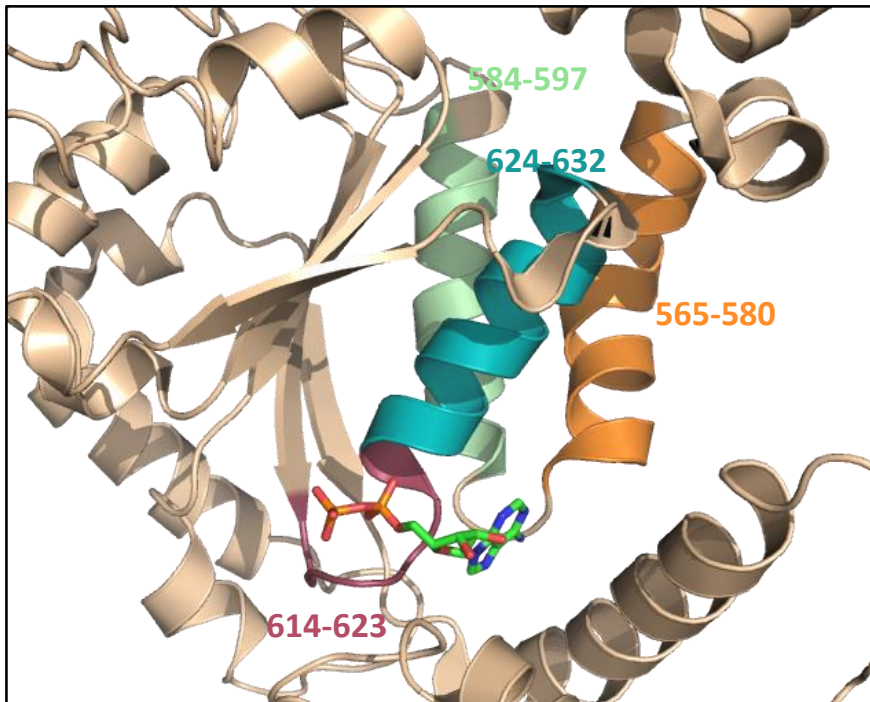
SI 3A



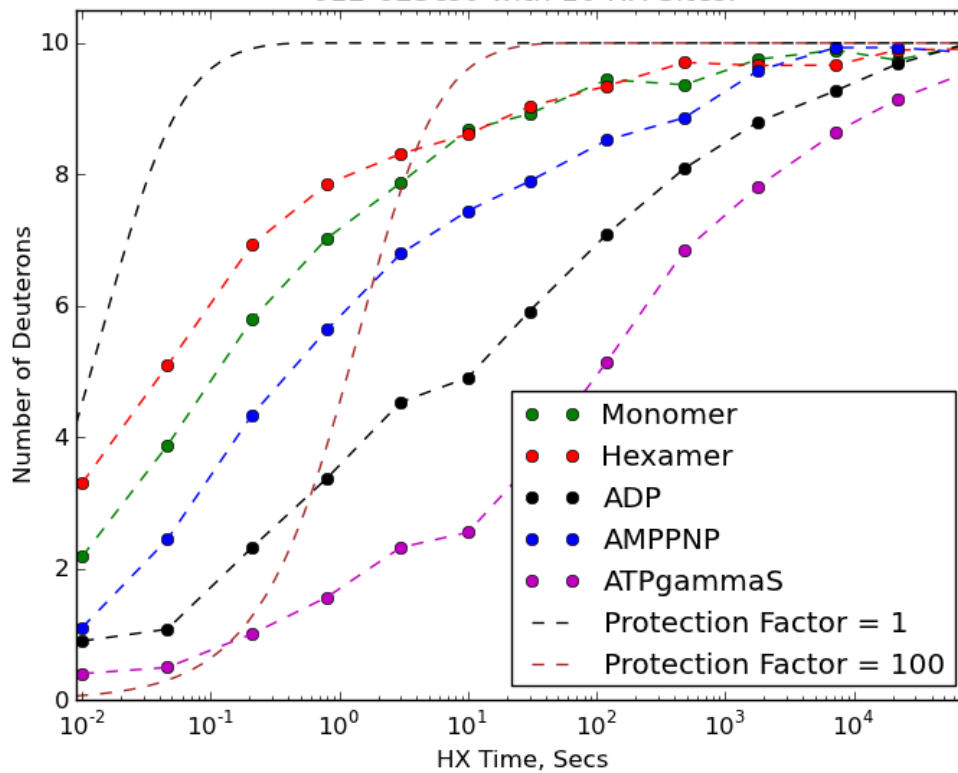
211-221cs0 with 8 HX Sites.



SI 3B



612-623cs0 with 10 HX Sites.



SI 3C

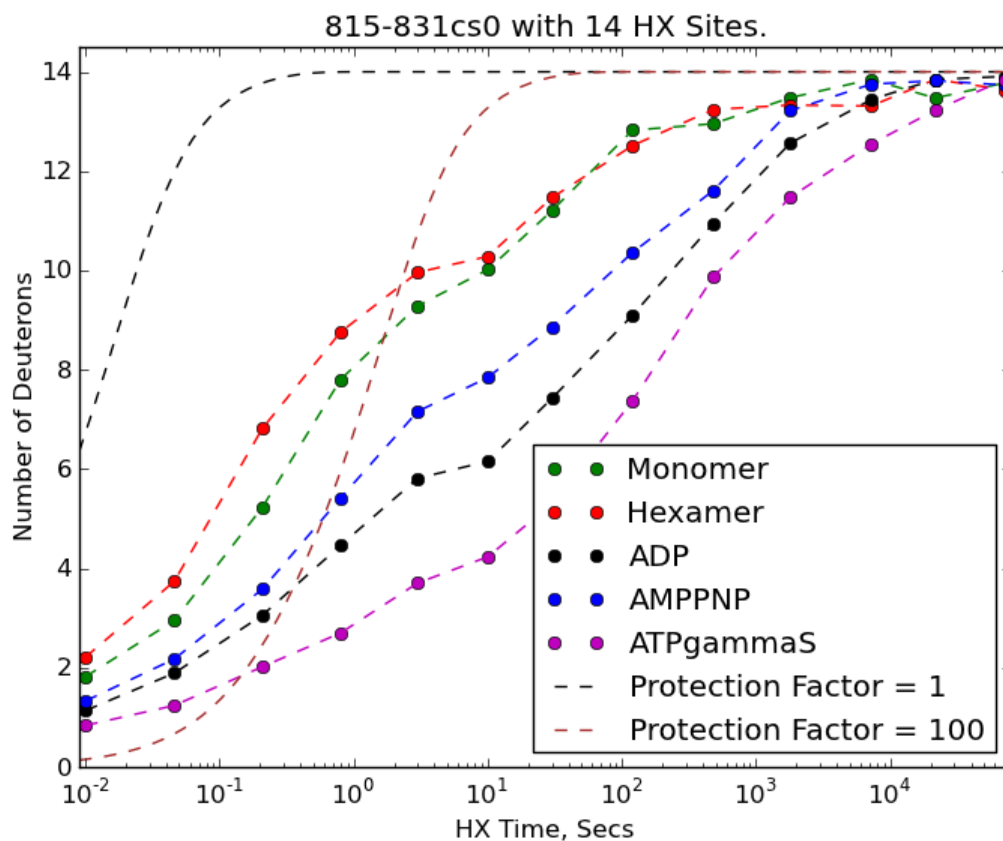
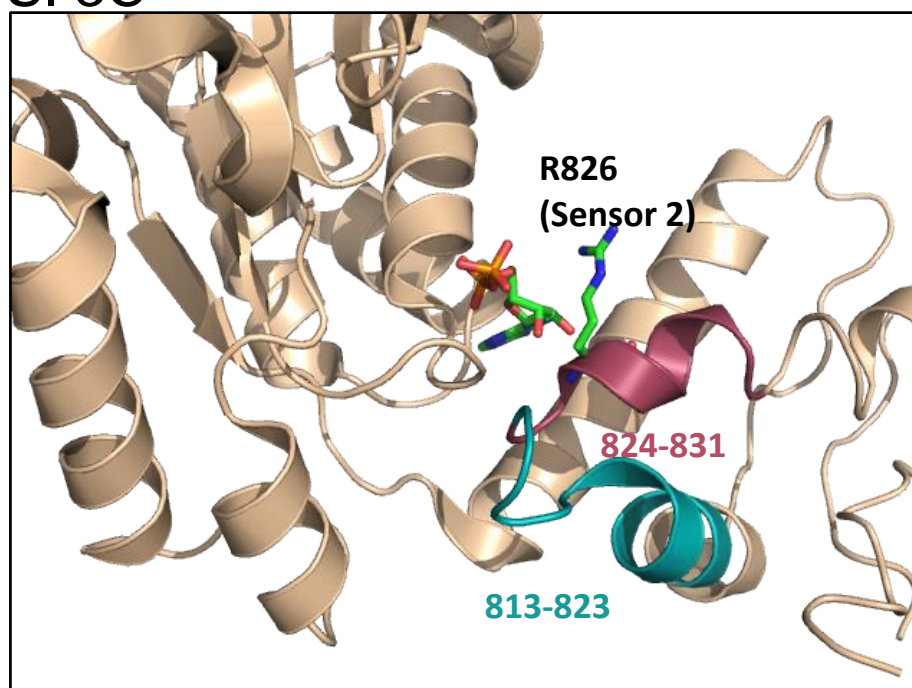
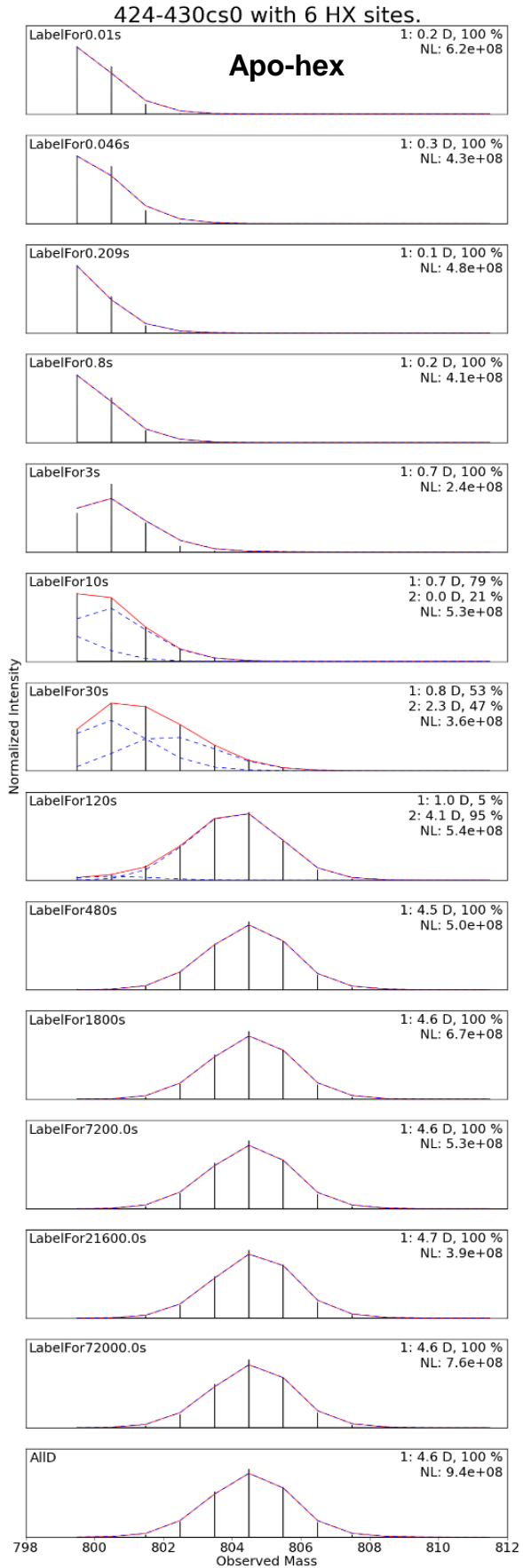
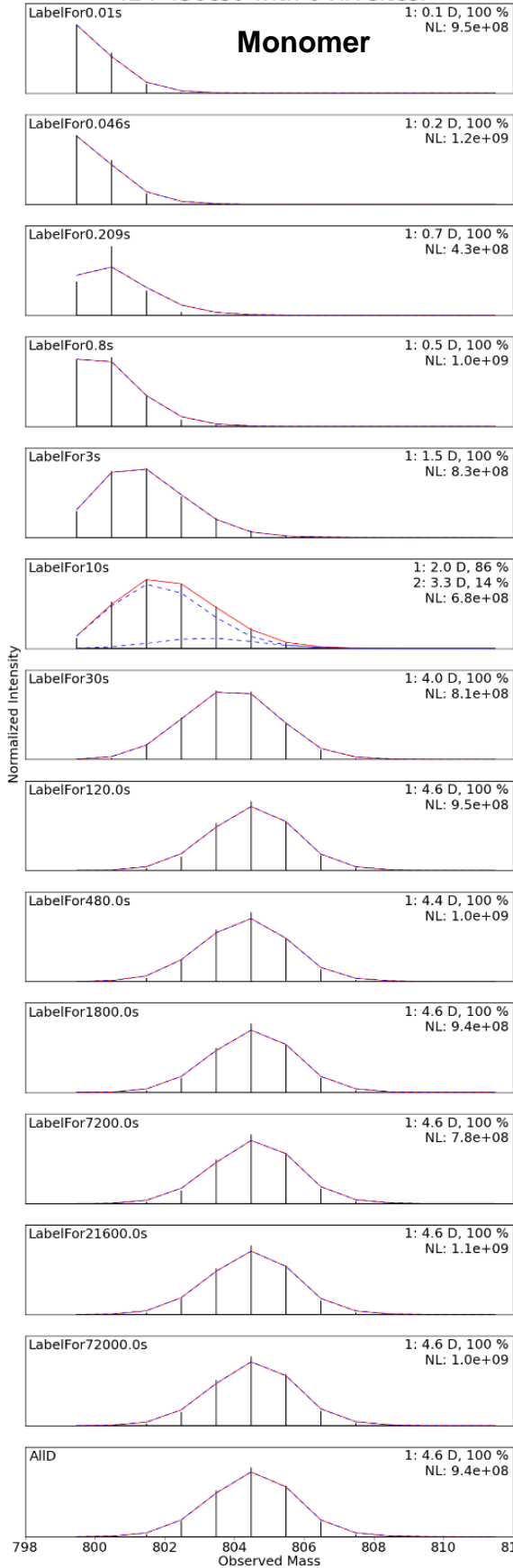
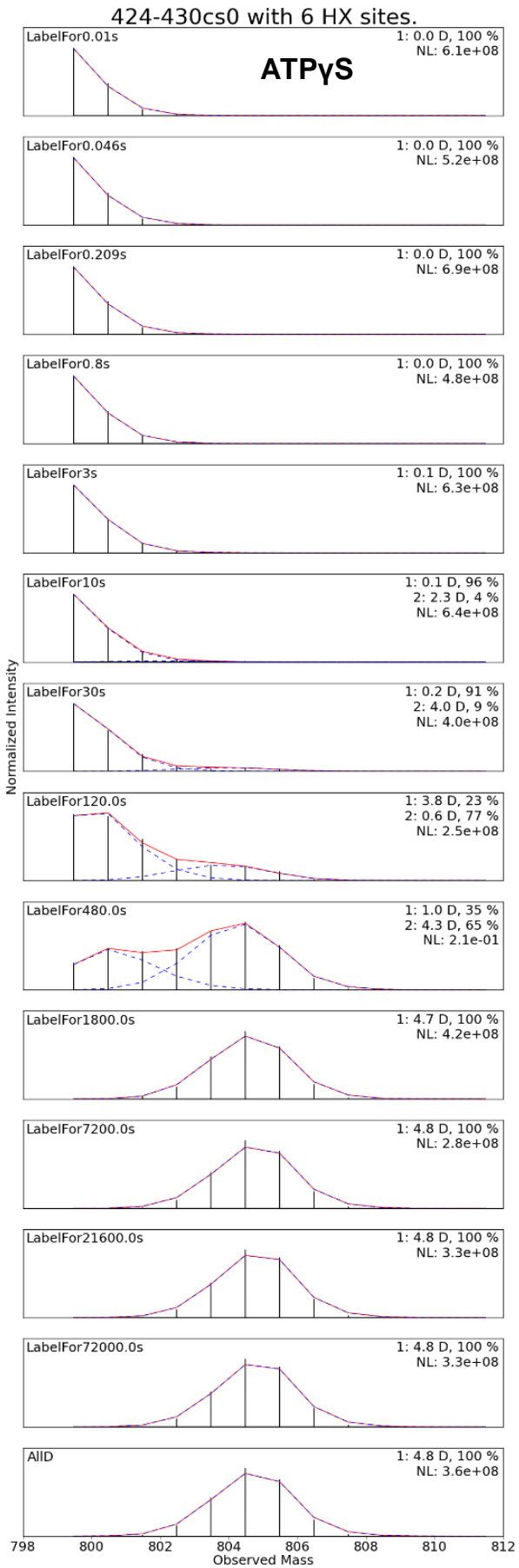
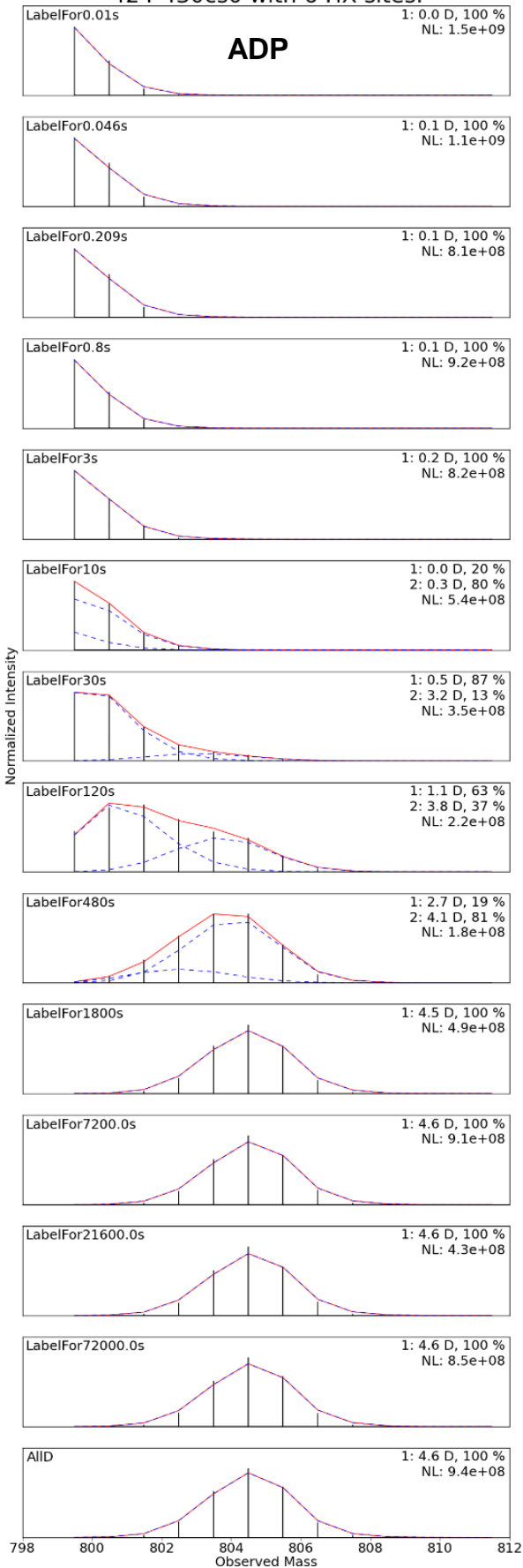


Fig. SI 3. Local and non-local effects. Relative to the monomer, hexamer formation slightly destabilizes (faster HX) Walker A in NBD1 (A) and in NBD2 (B) and sensor 2 in NBD2 (C), even though they are not at the interprotomer surface. The binding of different nucleotides increasingly stabilizes them in the order ADP<AMPPNP<ATP γ S. ADP is represented in stick.

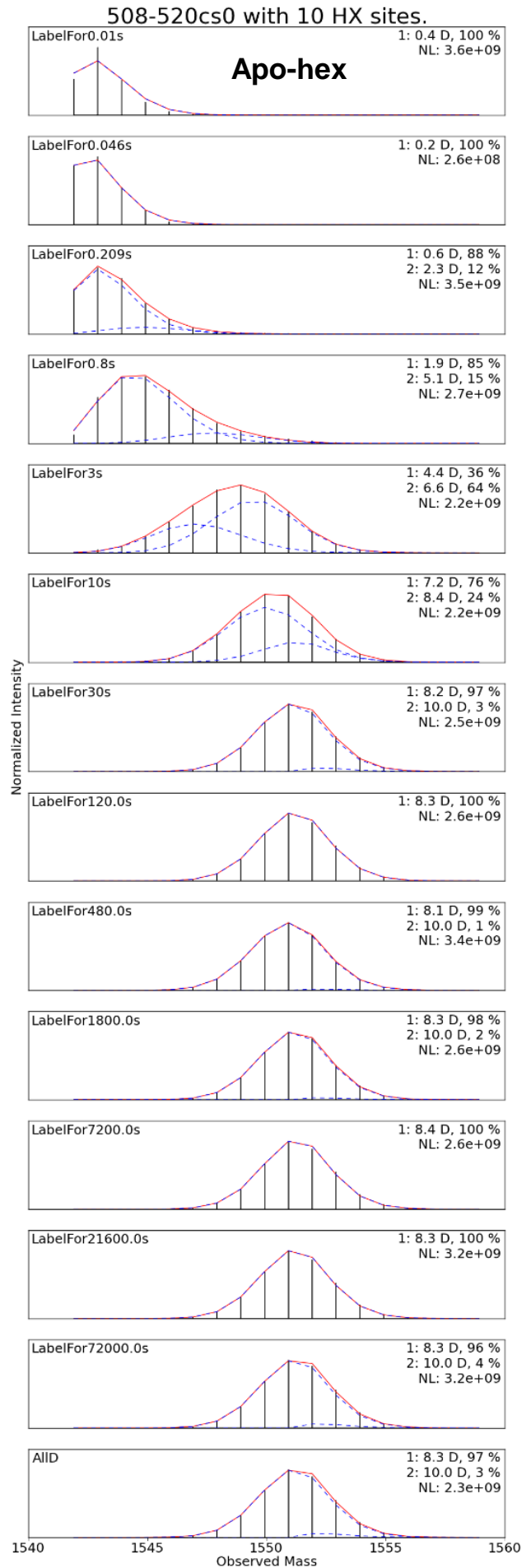
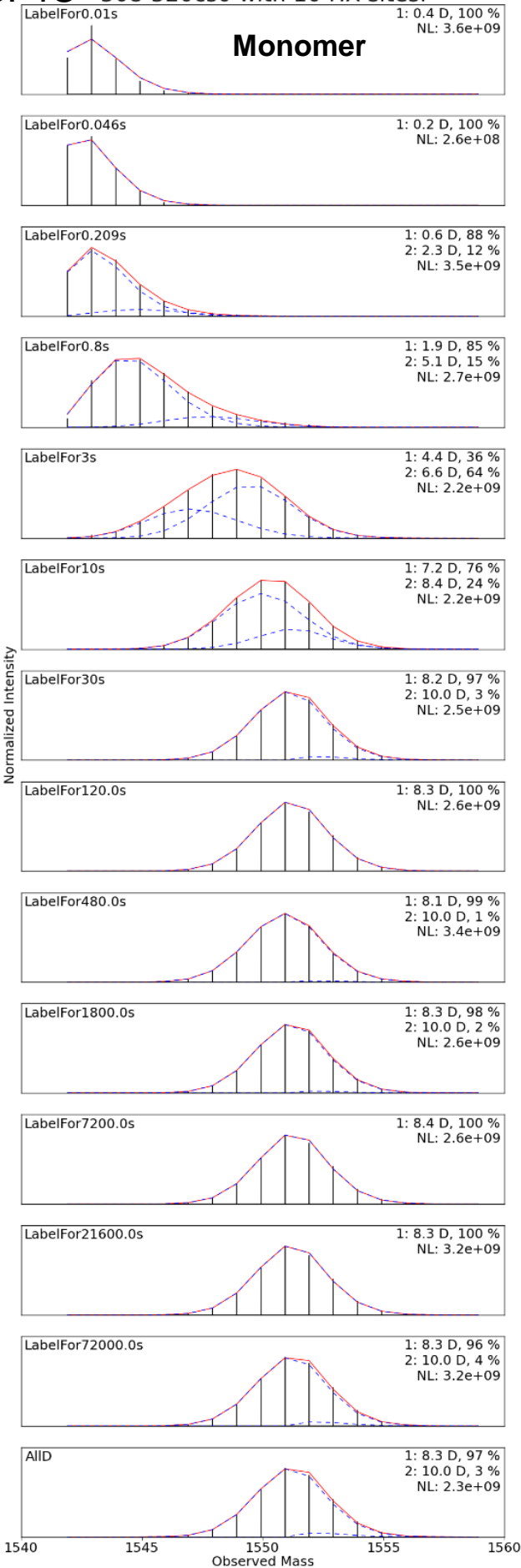
SI 4A 424-430cs0 with 6 HX sites.



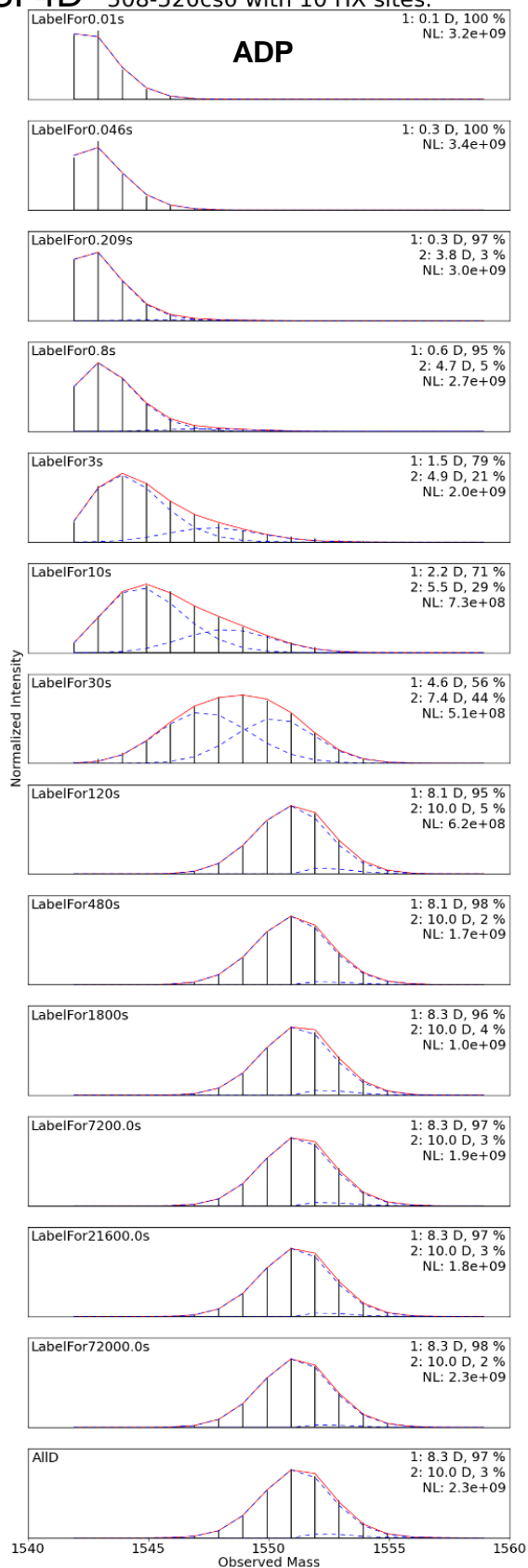
SI 4B 424-430cs0 with 6 HX sites.



SI 4C 508-520cs0 with 10 HX sites.



SI 4D 508-520cs0 with 10 HX sites.



508-520cs0 with 11 HX sites.

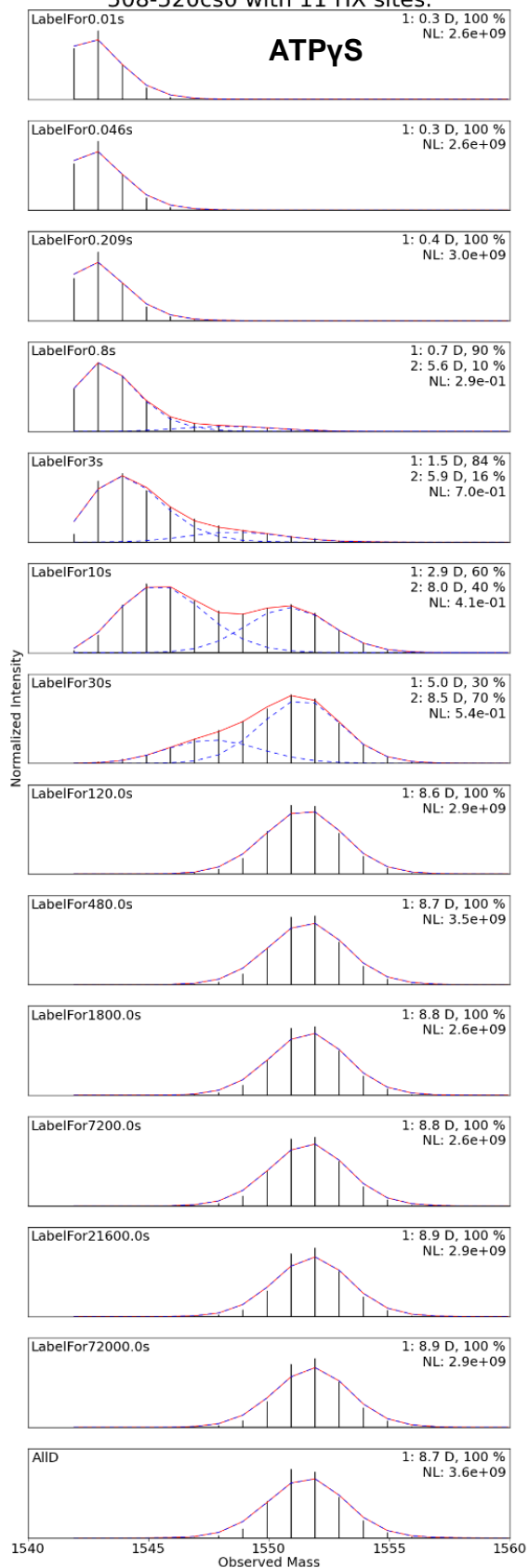
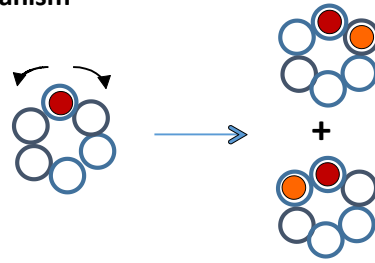
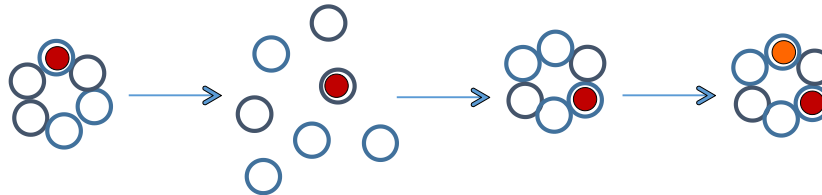


Fig. SI 4. EX1 peptides in the MD. MS plots for two M domain EX1 peptides, 424-430 in (A&B) and 508-520 in (C&D). For both, HX becomes progressively slower and bimodality more distinct in the progression from monomer to apohexamer to ADP to ATPyS.

Processive mechanism



**Random mixing mechanism
(hexamer-monomer equilibrium)**



**Random mixing mechanism
(two hexamer colliding)**

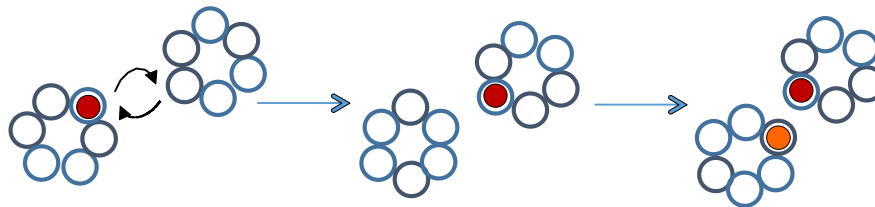


Fig. SI 5. The processive and random mixing models. These models were considered in Monte Carlo simulations to generate the slow protomer interchange process shown in main text Fig. 4D (AMPPNP). In the processive model, protomers exposed in the non-canonical interface interchange systematically with those from the canonical interface one position at a time. In the mixing models, the hexamer either fully dissociates and randomly reassembles or two hexamers collide and trade individual monomers. The HX data show that the protomers interchange on a long time scale and spend time in both the canonical and non-canonical interfaces but the detailed mechanisms pictured cannot be distinguished. With ATP cycling, a fast apparently functional 1 sec time scale interchange mechanism is observed.

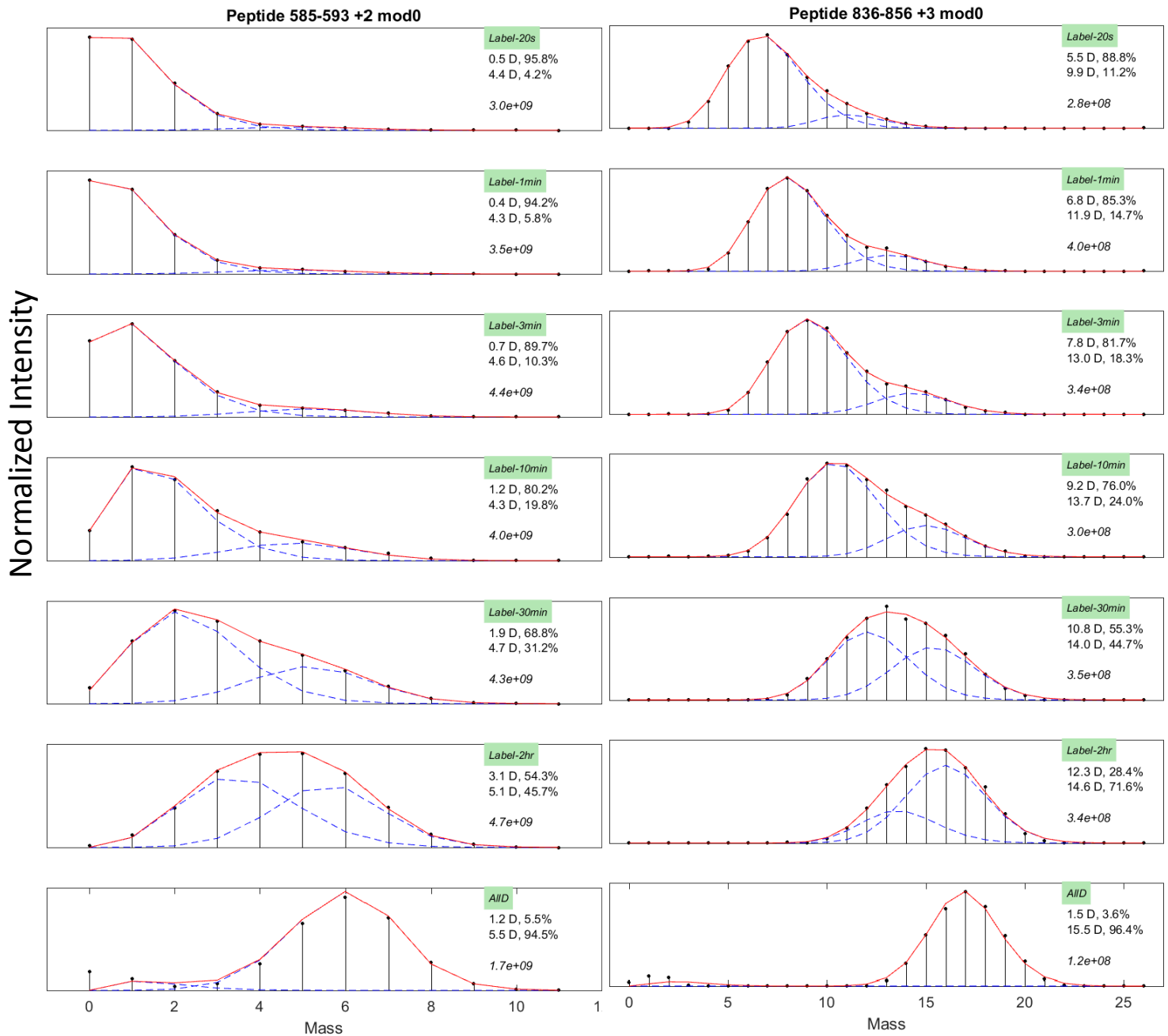


Fig. S1 6. Heterogeneous HX of non-canonical interface peptides in the ADP-bound state. Bimodal HX MS reveals faster and slower sub-populations in a 1:5 ratio as expected for the ADP-bound asymmetric spiraled hexamer, as for the AMPNP state in main text Fig. 4. We used pD 8.9 (25°C, 150 mM KCl) to make HX faster in order to compete better with the slow non-functional interchange process that follows on a much longer time scale.

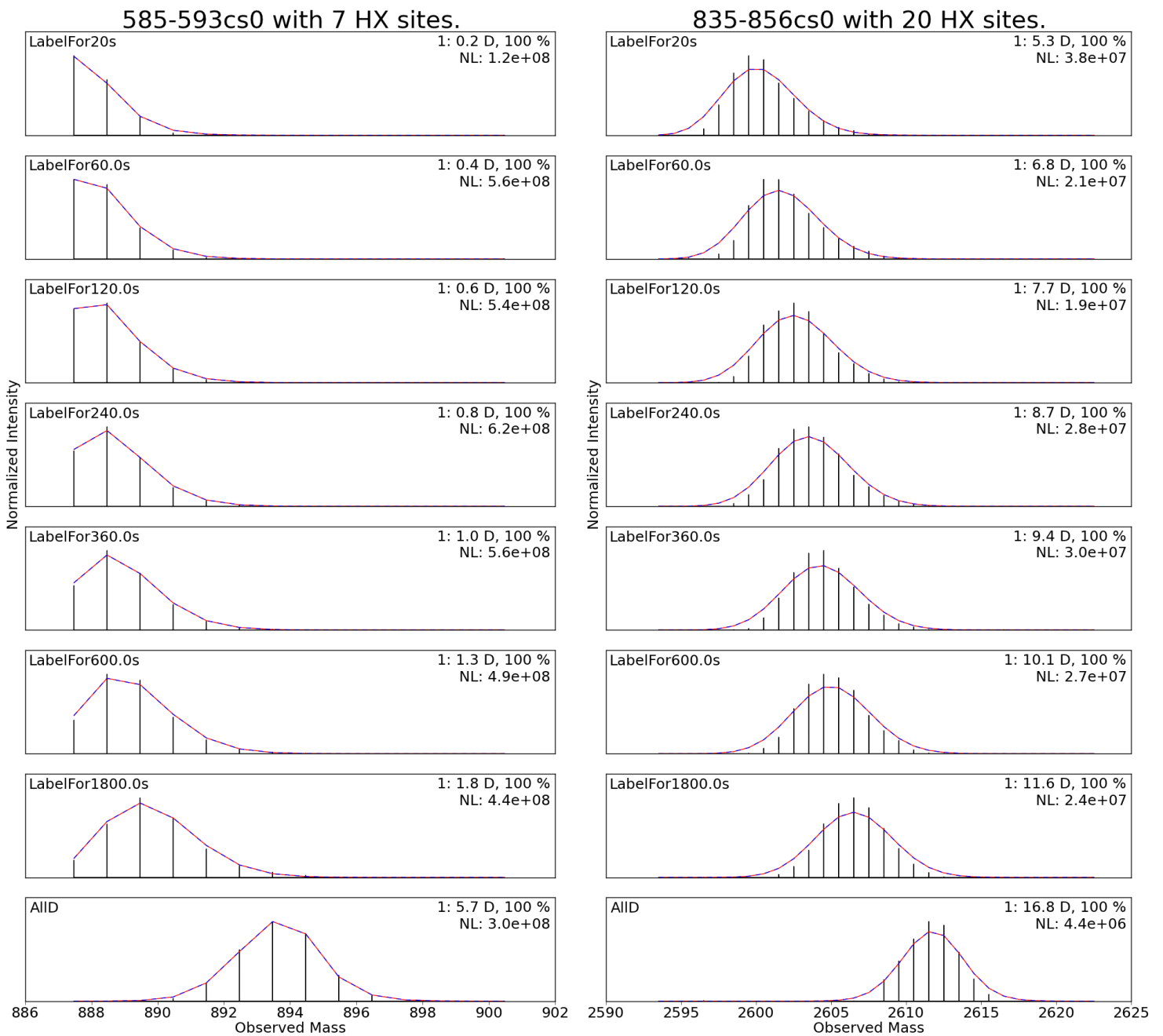


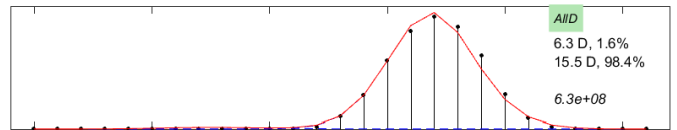
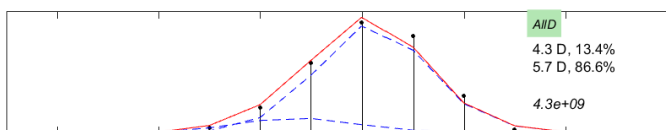
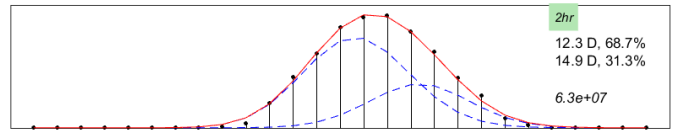
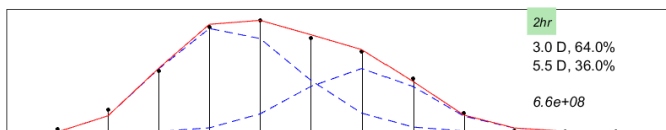
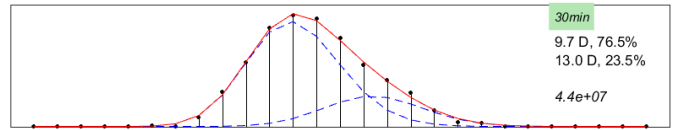
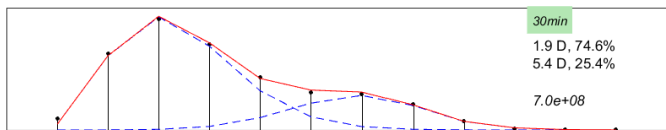
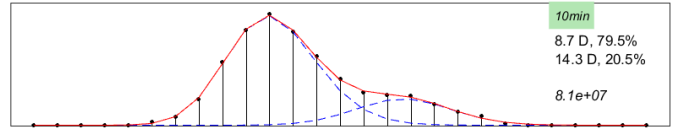
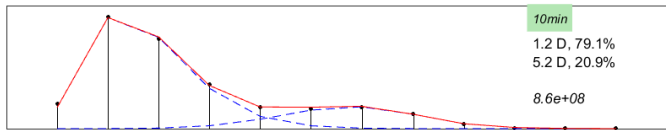
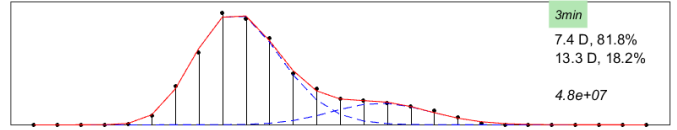
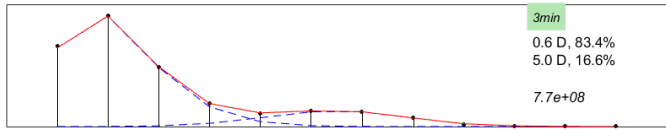
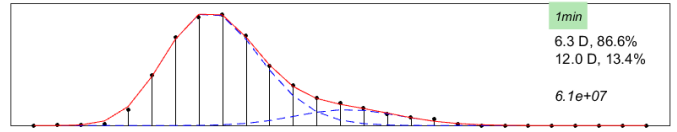
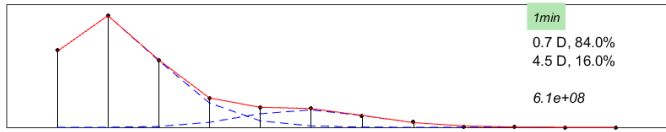
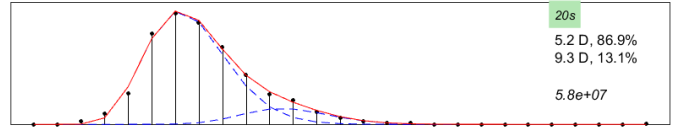
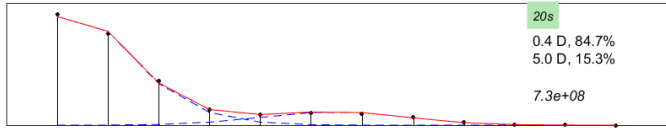
Fig. S1 7. Loss of bimodality due to loss of the non-canonical interface in the presence of ATP γ S. The non-canonical interface is lost in the flat closed ATP γ S state. Conditions are the same as for Fig. 4 except 5 mM ATP γ S was used.

SI 8A

Peptide 585-593 +1 mod0

Peptide 836-856 +3 mod0

Normalized Intensity



0 2 4 6 8 10 12 Mass

0 5 10 15 20 25 Mass

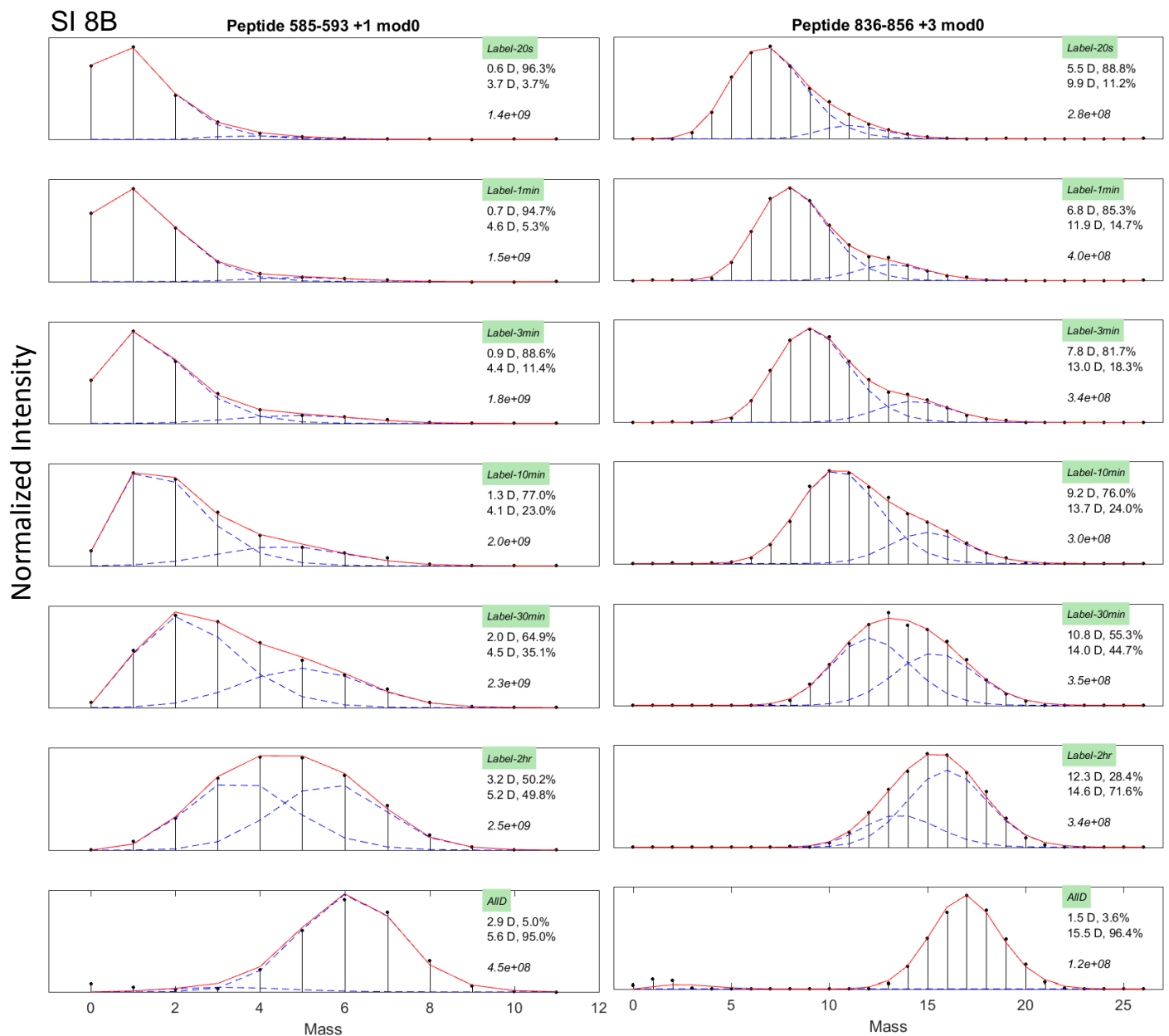


Fig. SI 8. Slow protomer interchange in the presence of low salt ([KCl] = 10 mM). MS spectra of the two heterogeneous exchanging peptides in the presence of AMPPNP (A) or ADP (B). The slow upswep of the less protected fraction of these two heterogeneously exchanging peptides in the non-canonical interface is shown in main text Fig. 4. Low salt stabilizes protomer interaction and slows the later upswep, indicating that the upswep represents a slow non-functional protomer dissociation process.

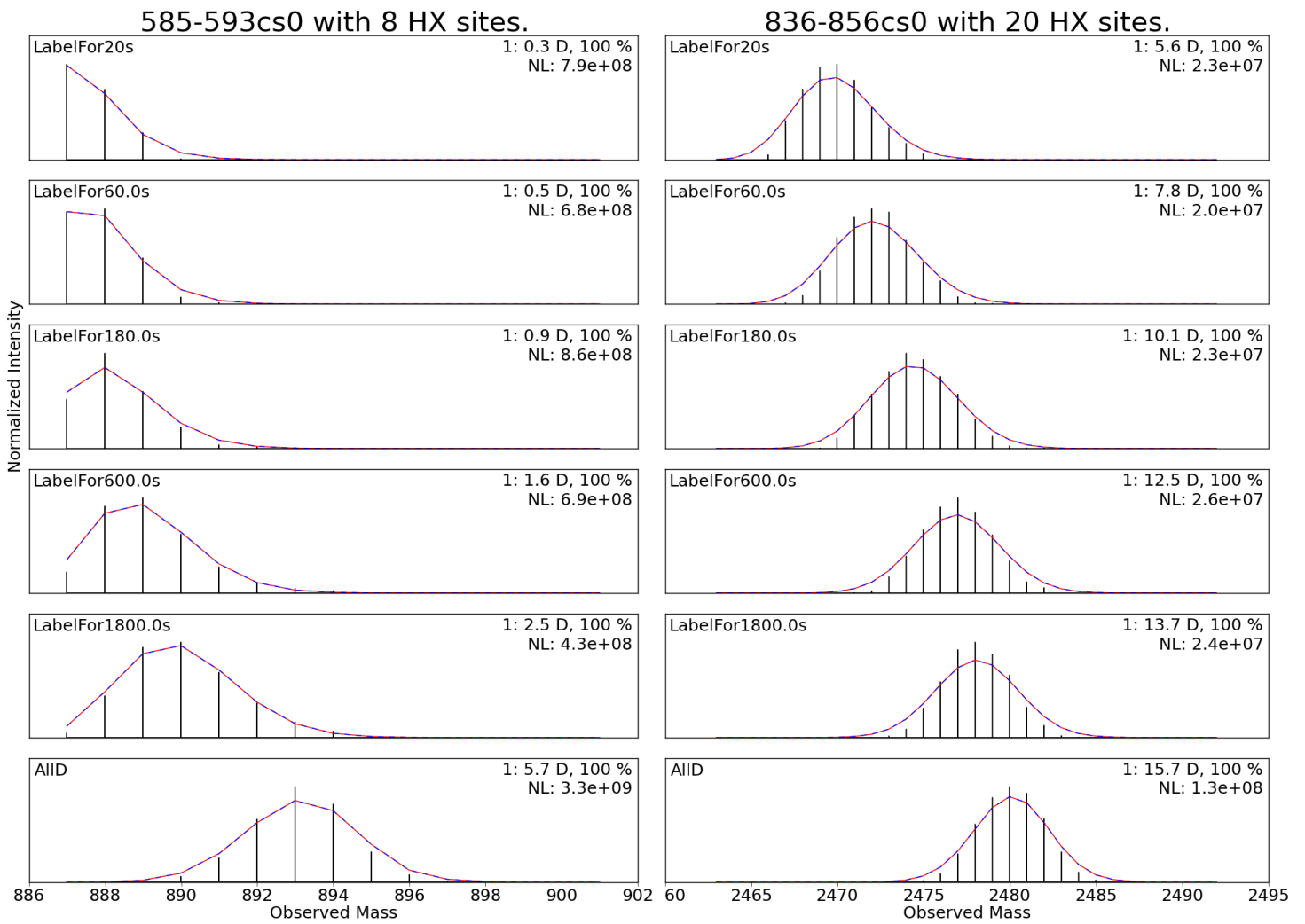
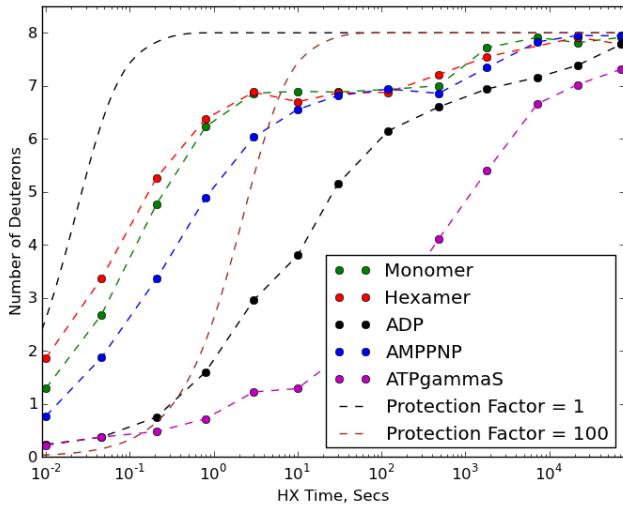


Fig. SI 9. Protomer interchange accelerated by ATP turnover. Active turnover in the presence of 10 mM ATP (with an ATP regenerating system) results in fast conformational averaging between the canonical and non-canonical interfaces, switching bimodal HX behavior to unimodal.

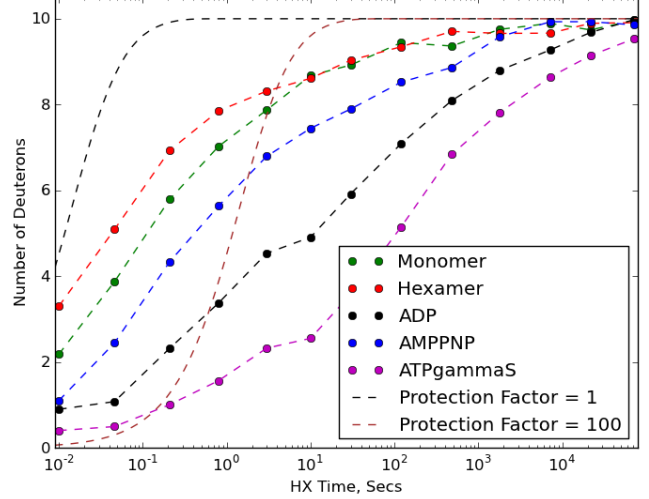
Walker A NBD1

211-221cs0 with 8 HX Sites.



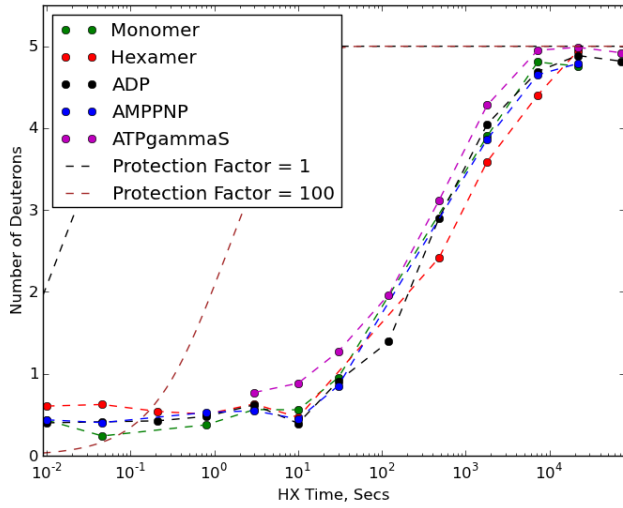
Walker A NBD2

612-623cs0 with 10 HX Sites.



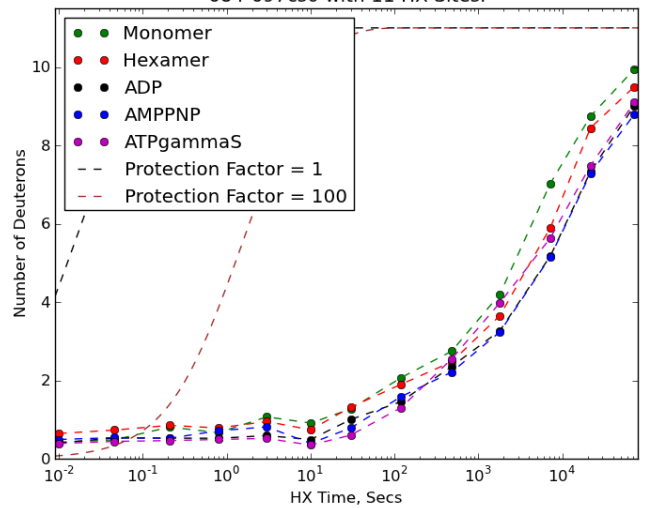
Walker B NBD1

282-288cs0 with 5 HX Sites.



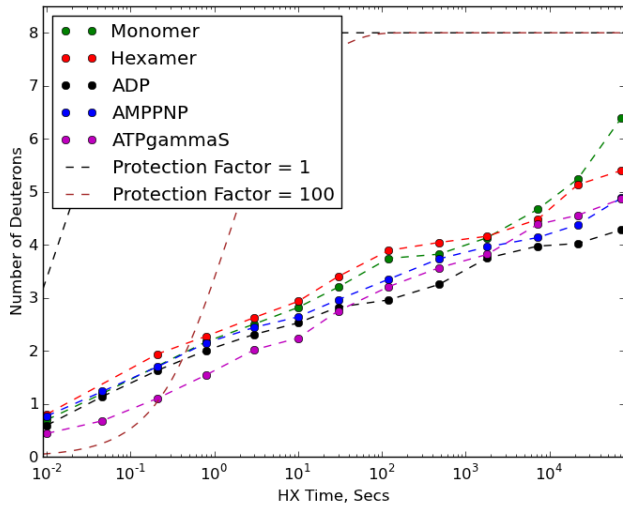
Walker B NBD2

684-697cs0 with 11 HX Sites.



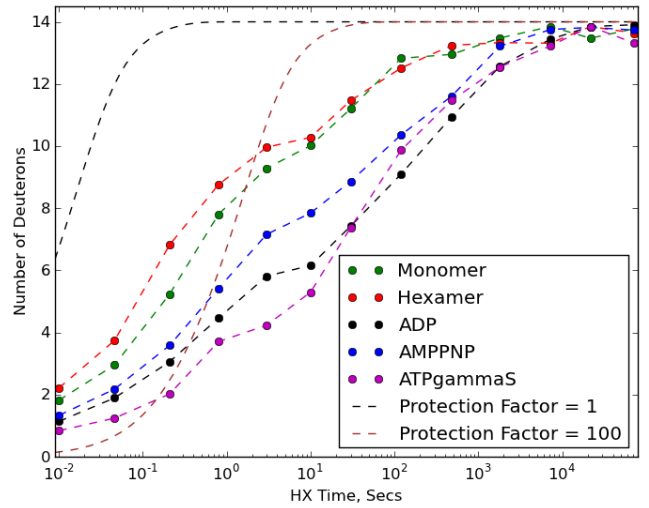
Sensor1 NBD1

311-320cs0 with 8 HX Sites.



Sensor2 NBD2

815-831cs0 with 14 HX Sites.



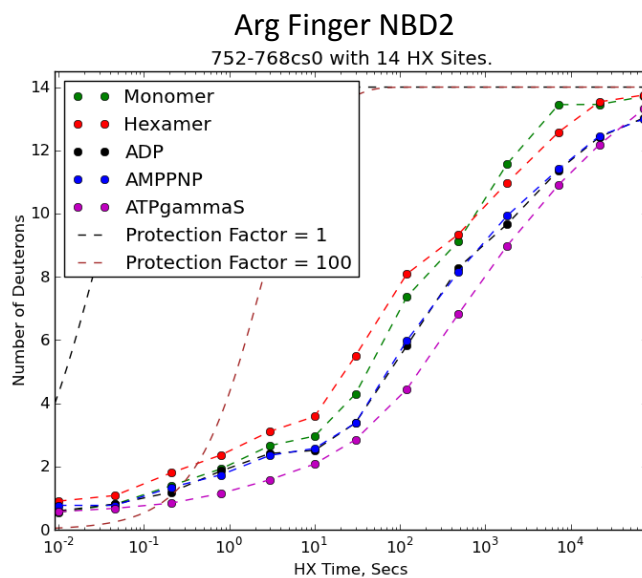
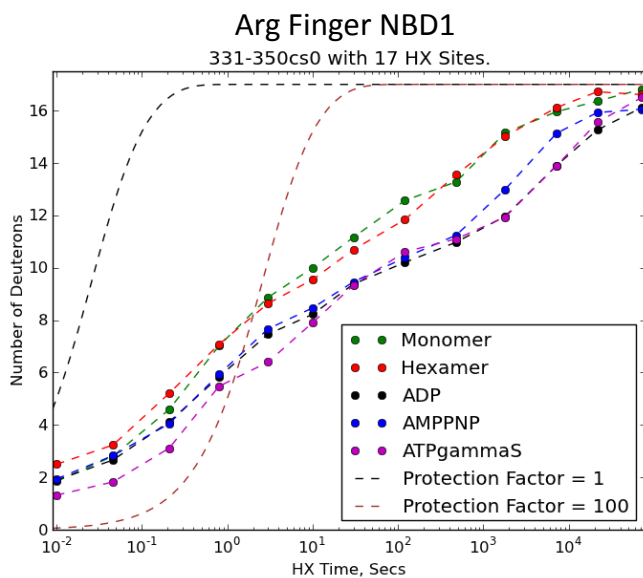


Fig. SI 10. Centroid plots of structural elements involved in nucleotide binding. The Walker A segment of both NBDs is by far the most affected by ATP γ S binding making it the most likely candidate for inducing the formation of the ATP-characteristic closed conformation.

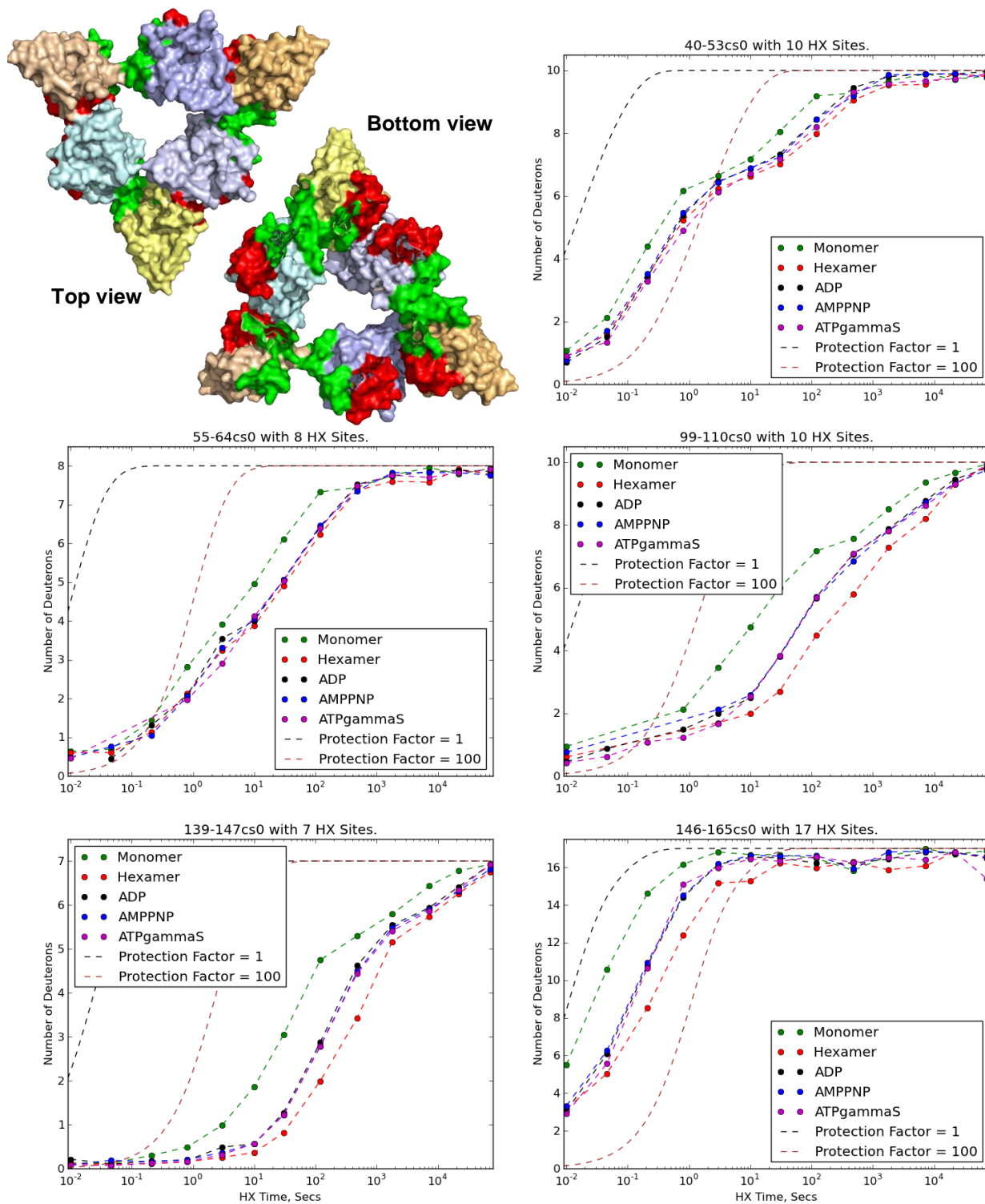
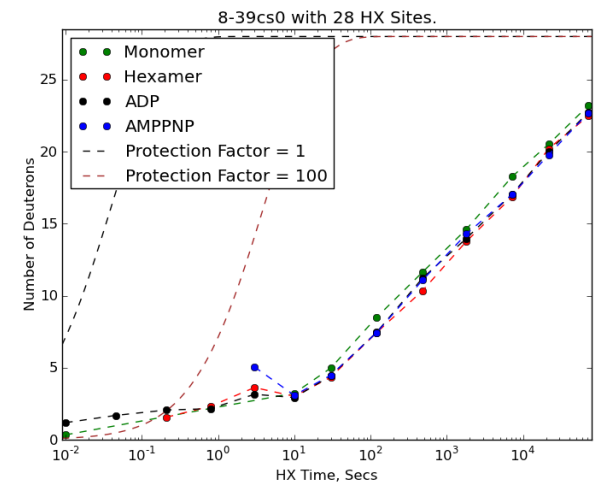
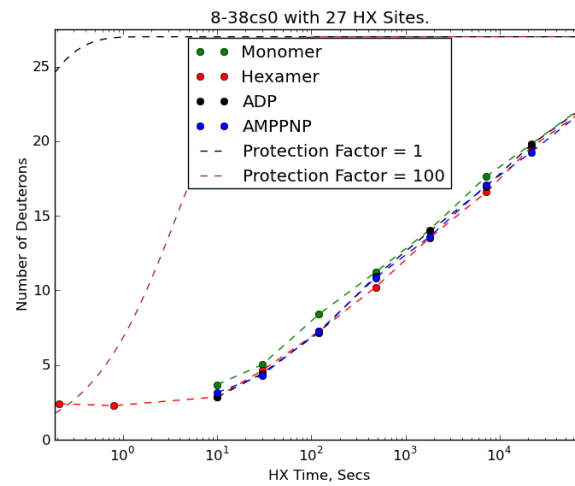
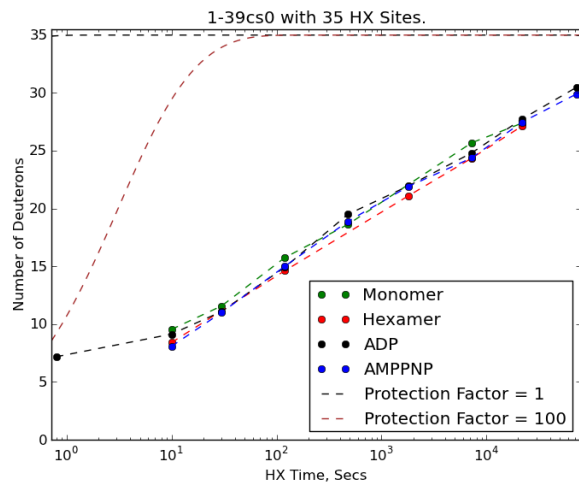
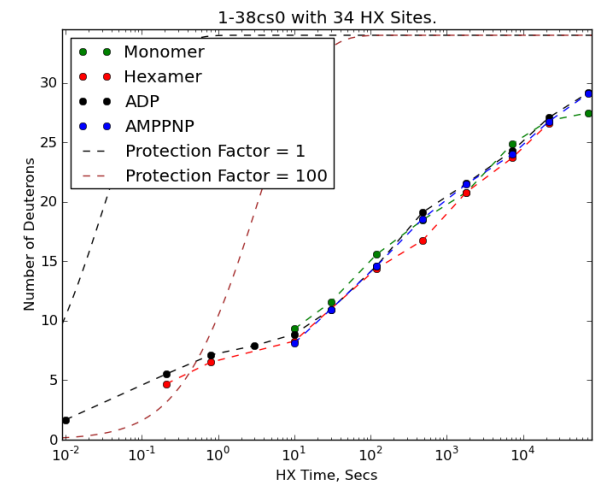
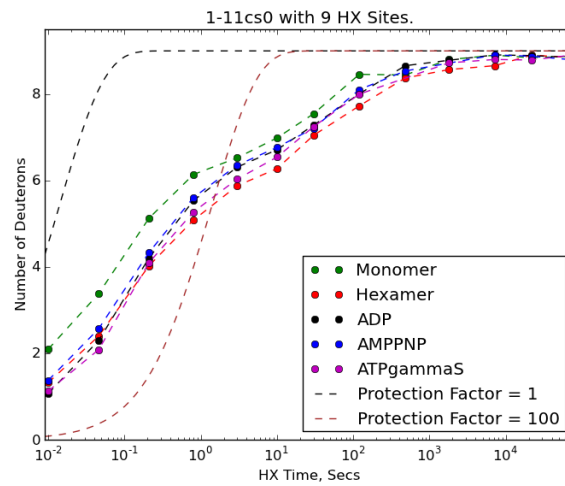
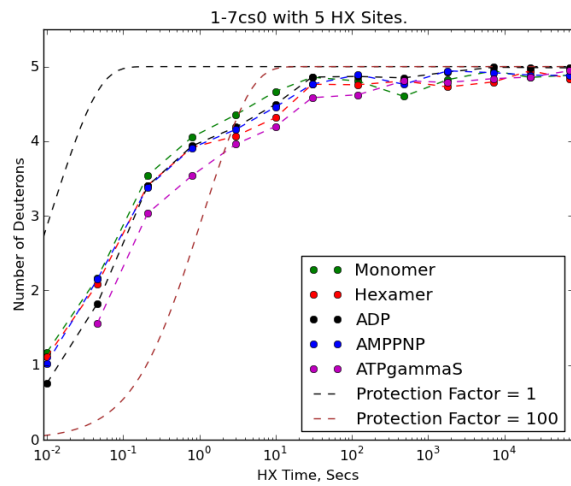
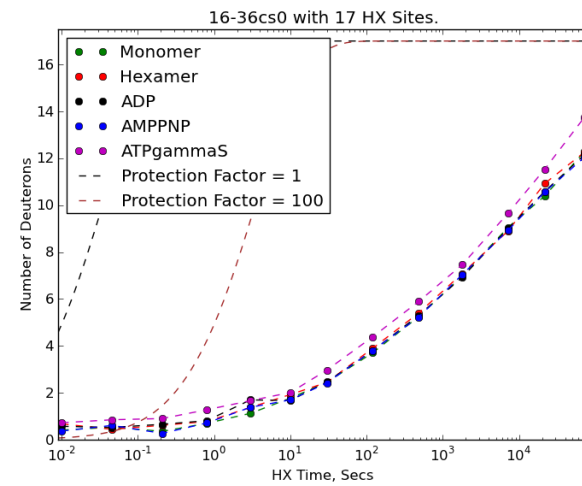
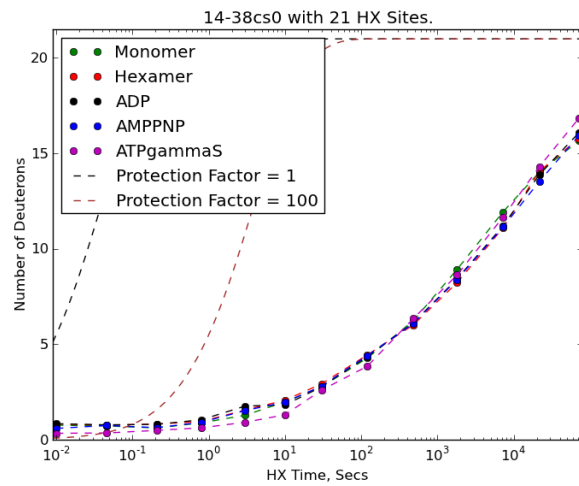
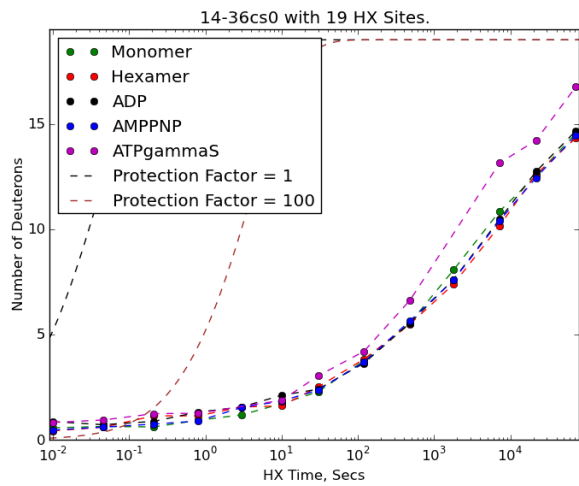
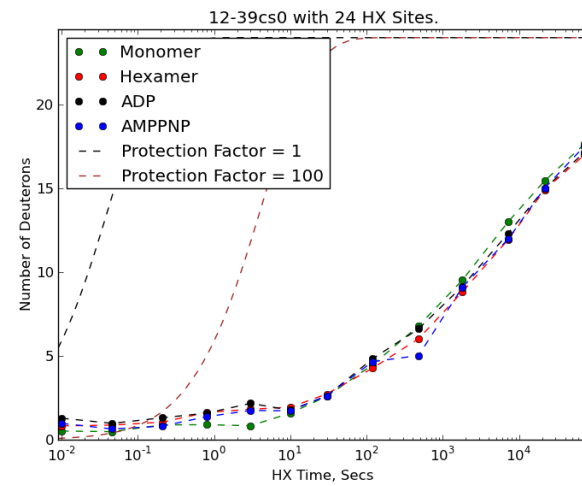
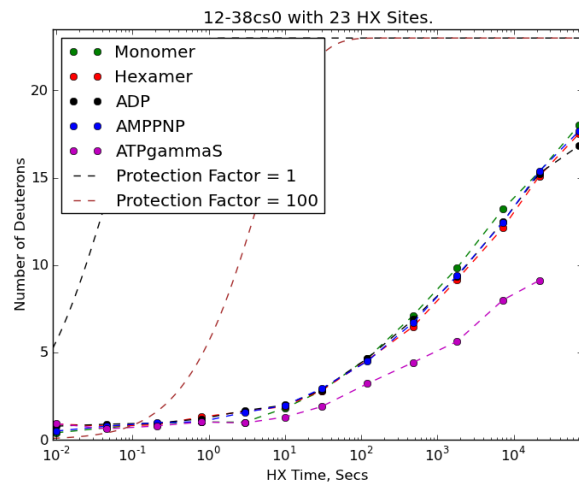
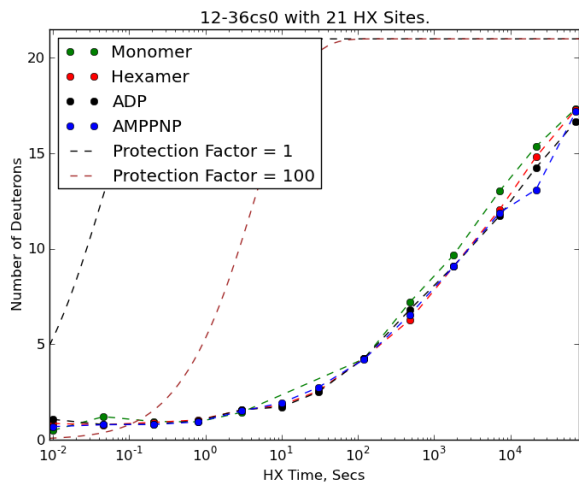
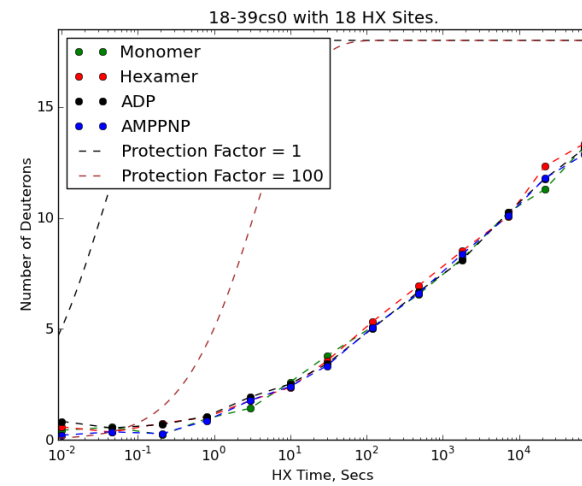
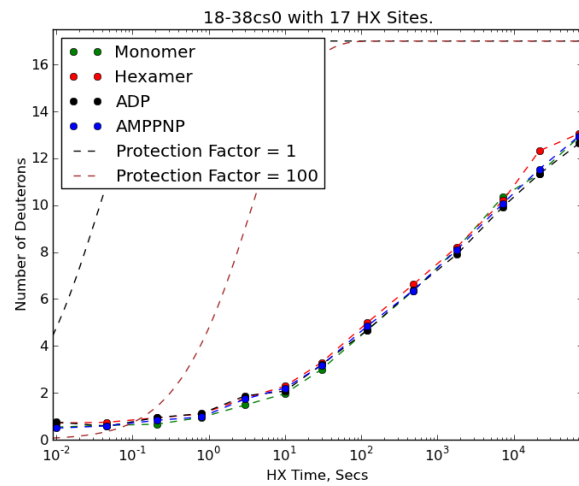
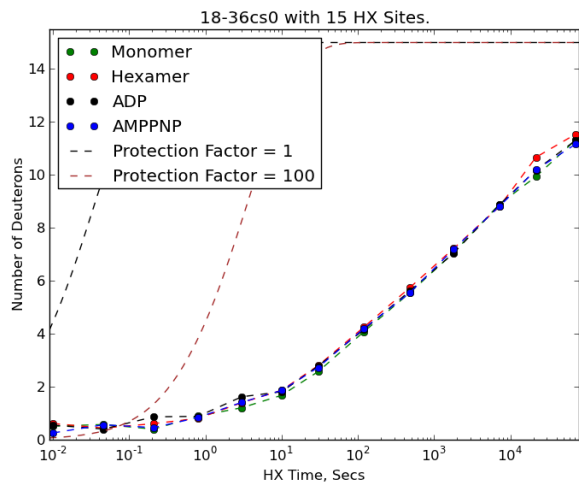
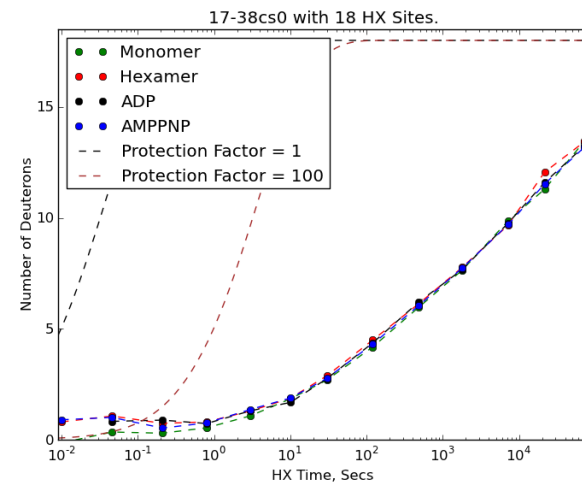
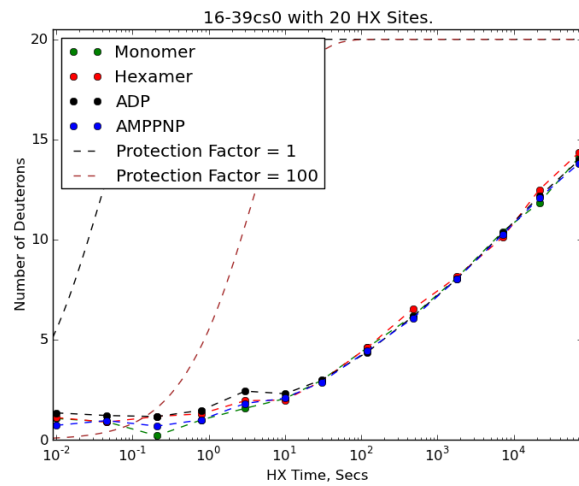
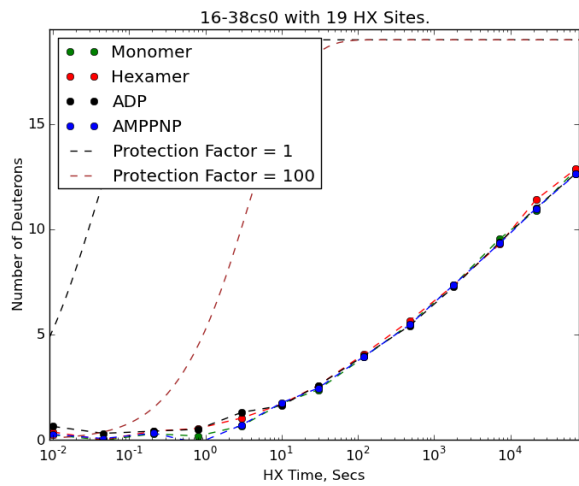
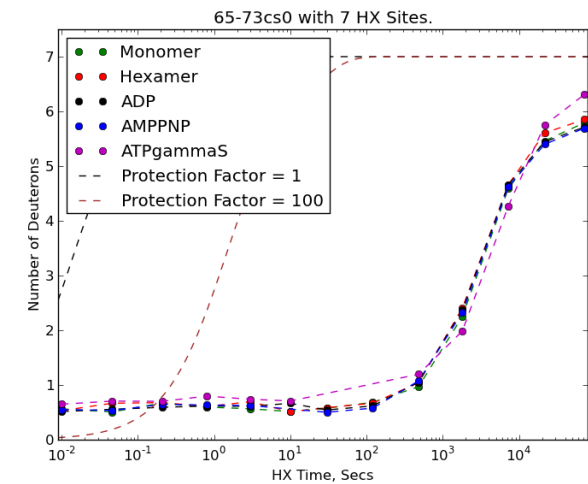
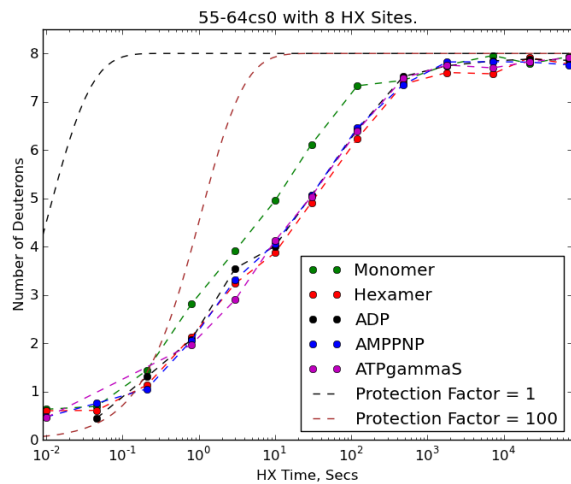
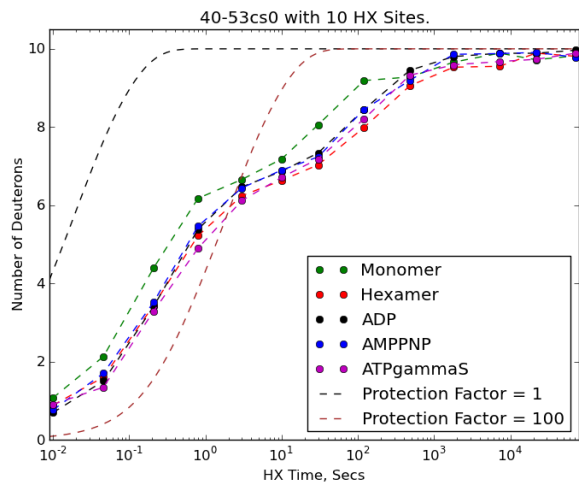
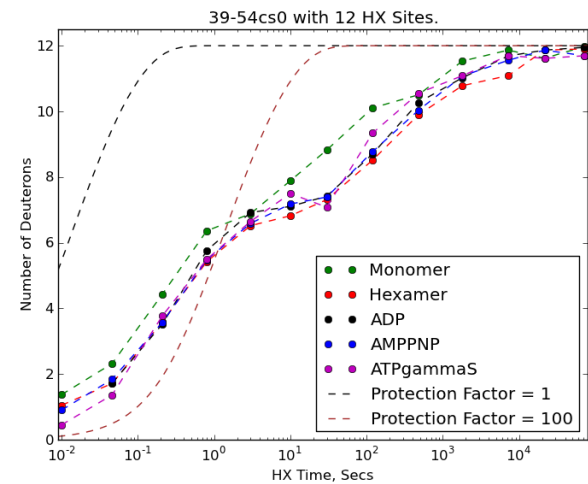
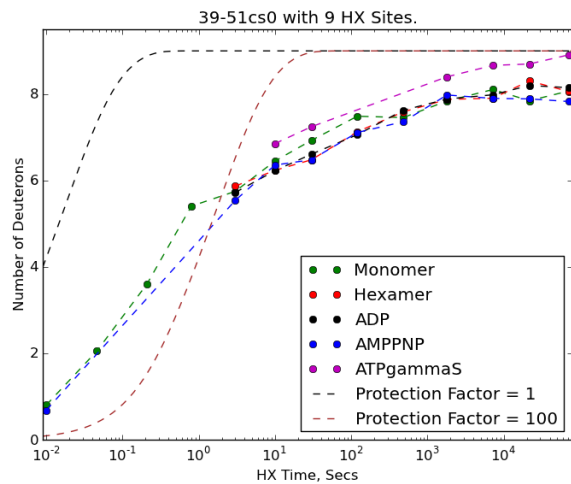
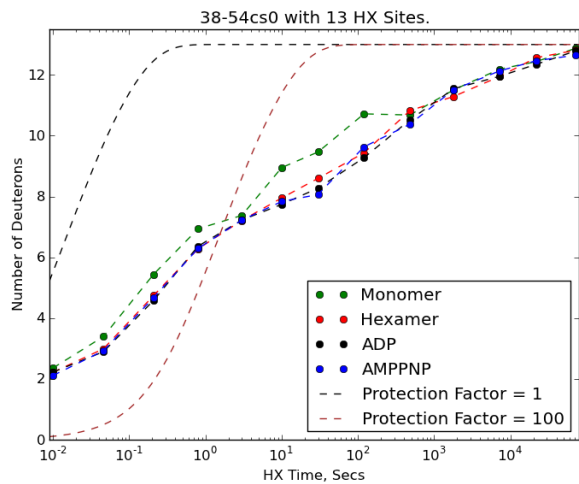


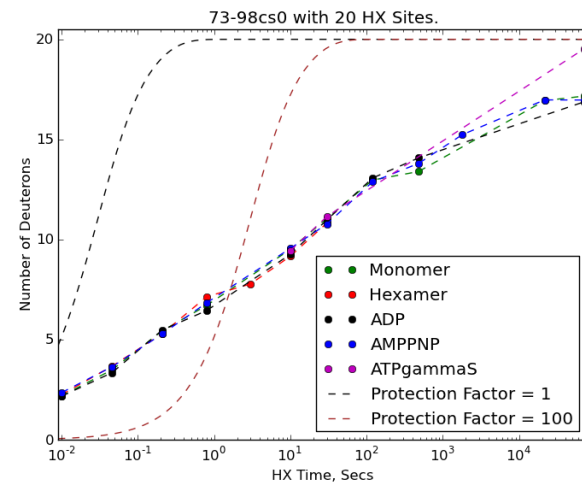
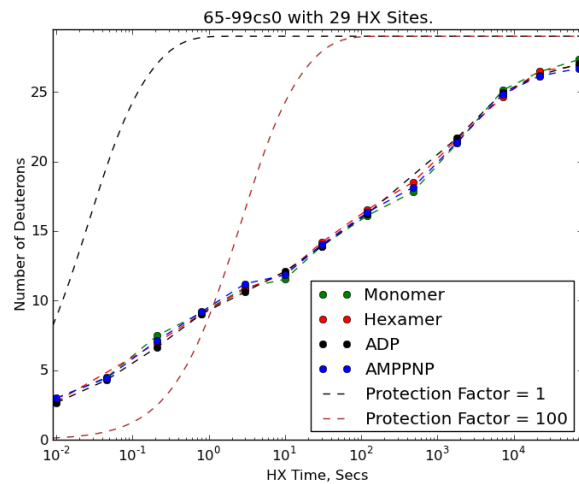
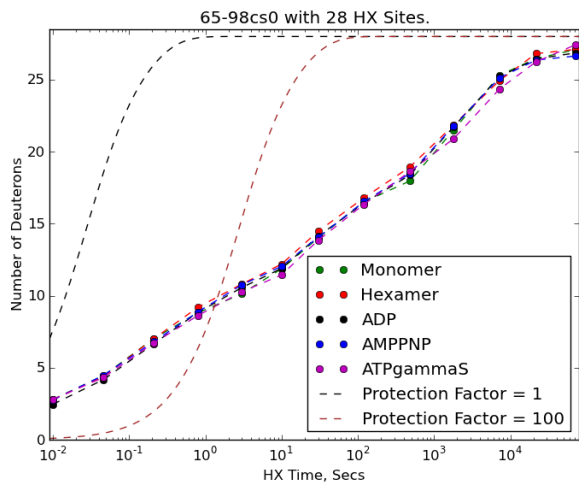
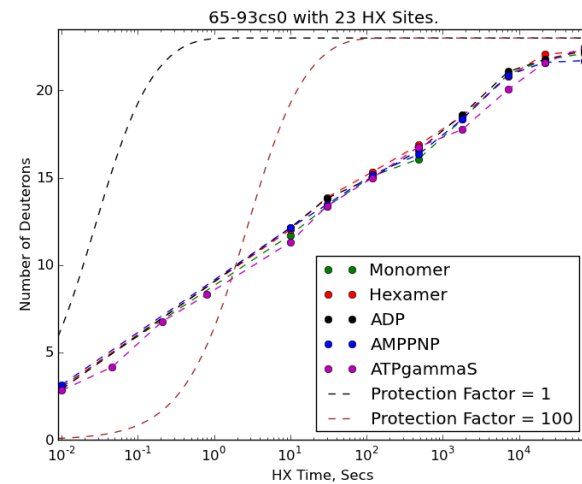
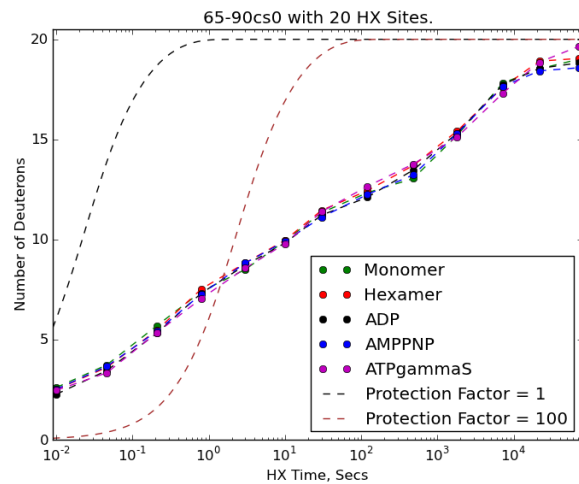
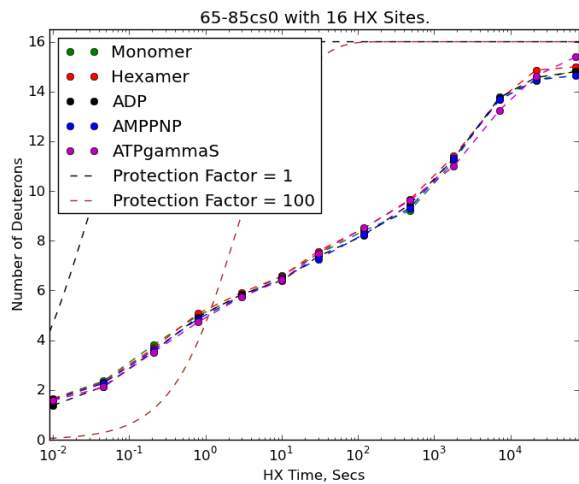
Fig. SI 11. Differential HX in the NTD. The top left panel shows the arrangement of the six NTDs in one class of the closed conformation resolved in (12). The regions slowed by hexamer formation are concentrated at the side of the NTDs adjacent to NBD1, indicating direct association. Regions that are stabilized by the four hexameric conditions to a similar extent are colored red; they include 40-53 and 55-64 as shown in the first two mass centroid plots. Regions that are stabilized more by the bare apohexamer condition than by the three nucleotide-bound conditions are green. They are represented by the other three peptides in this figure.

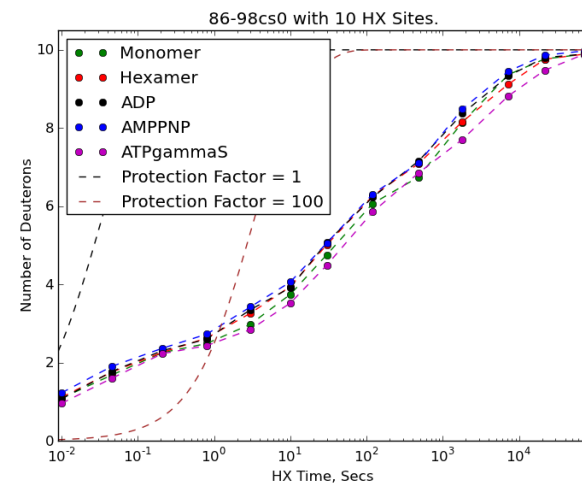
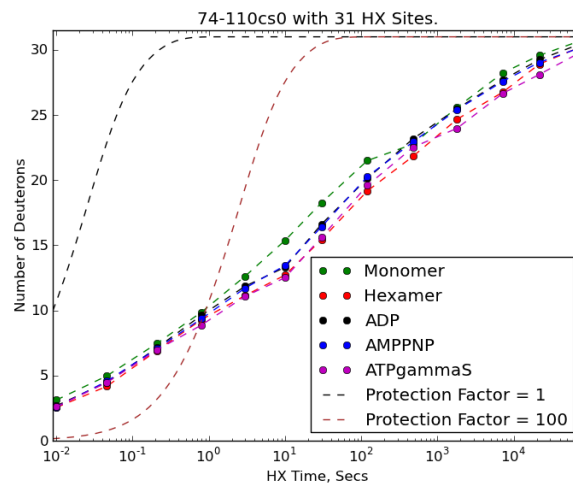
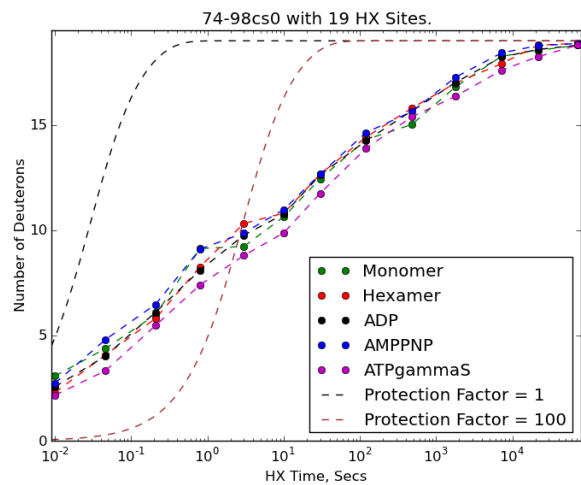
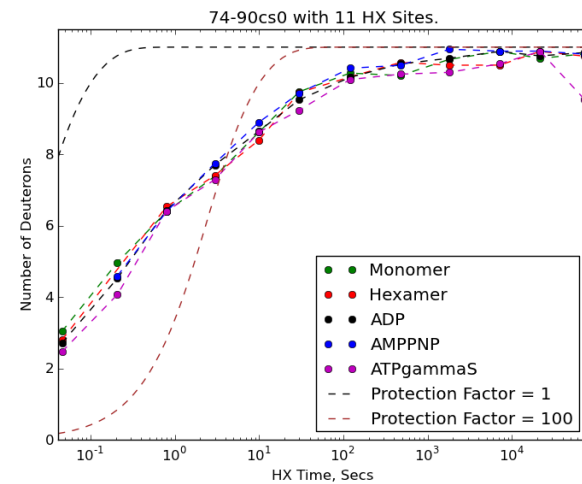
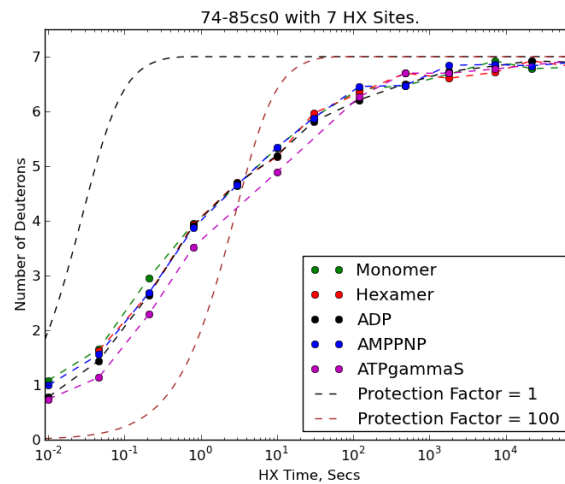
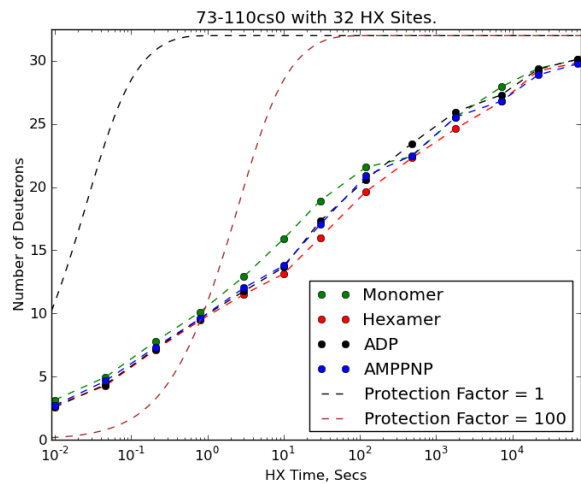


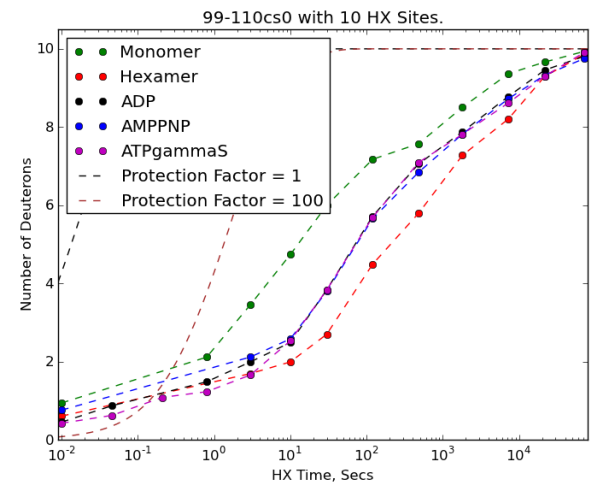
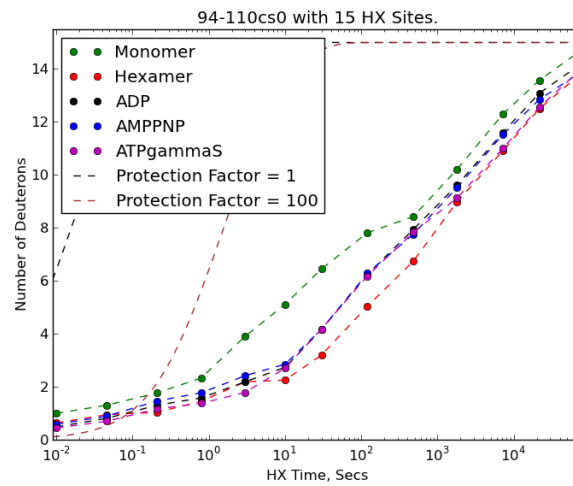
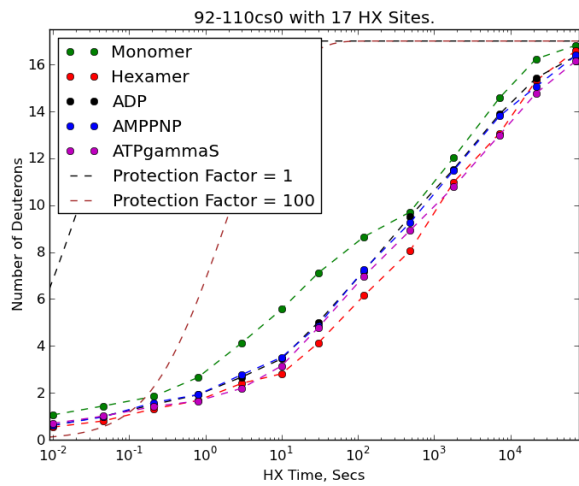
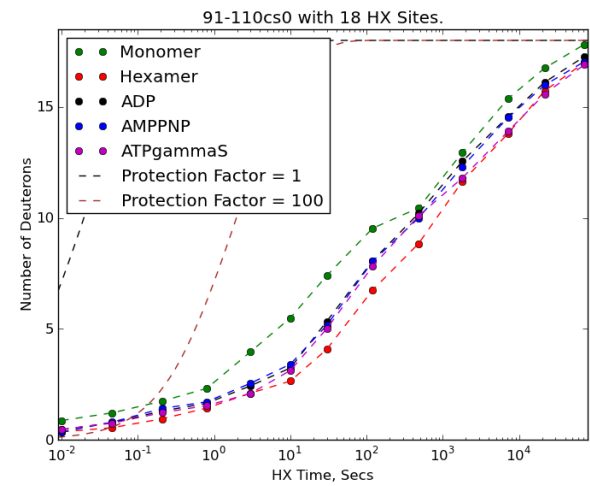
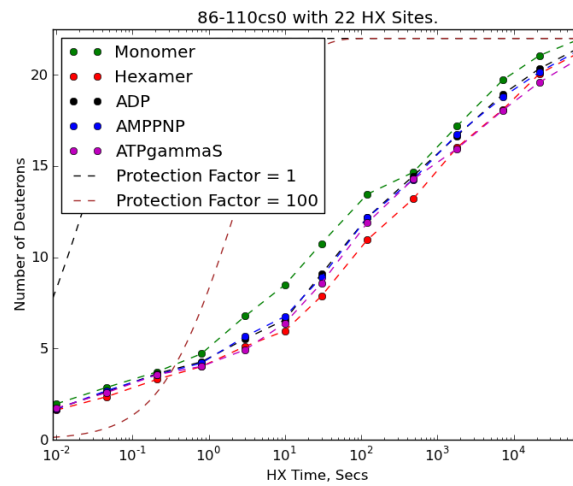
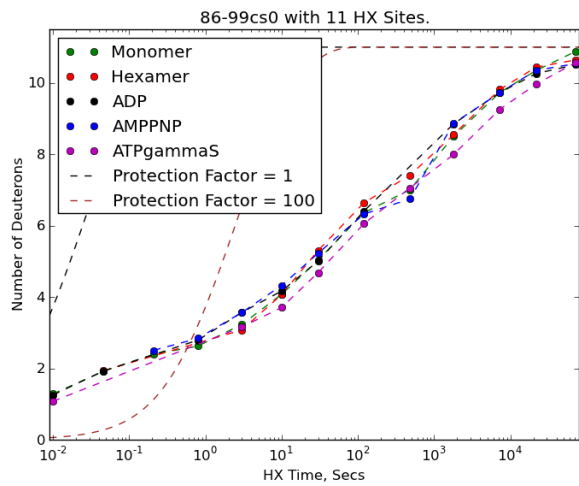


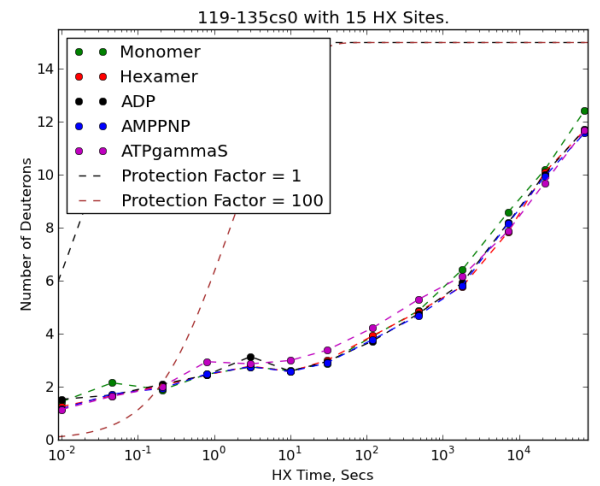
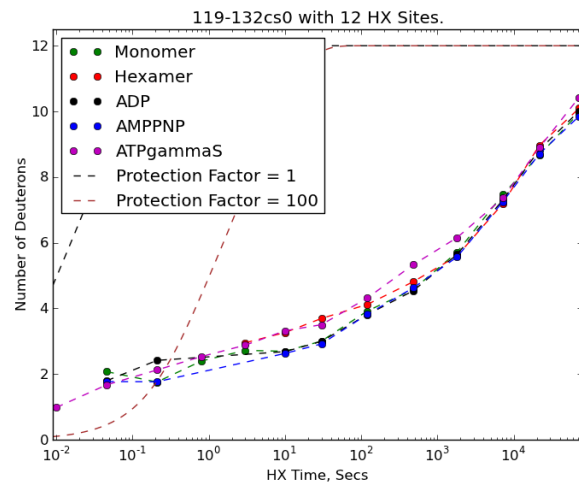
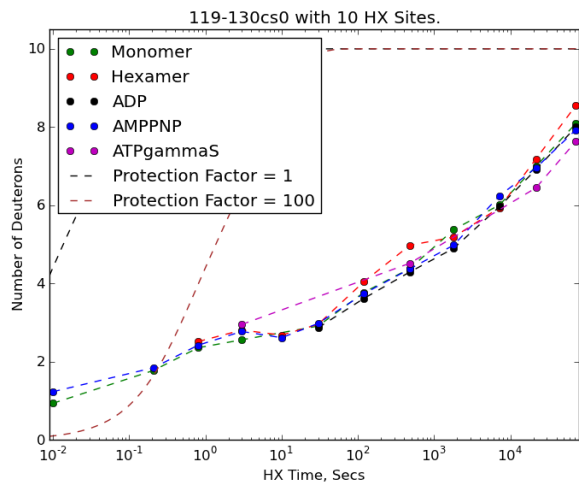
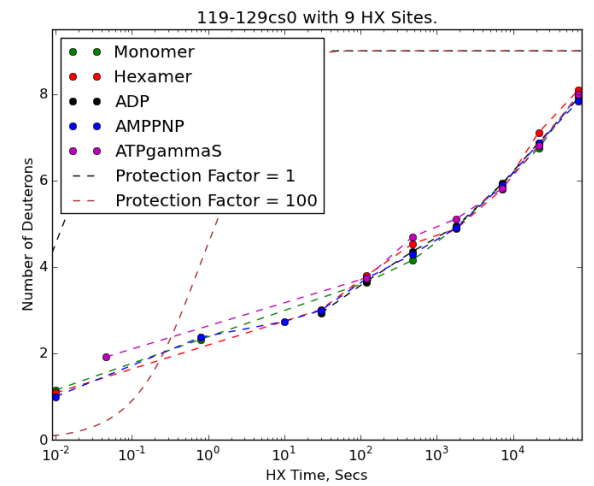
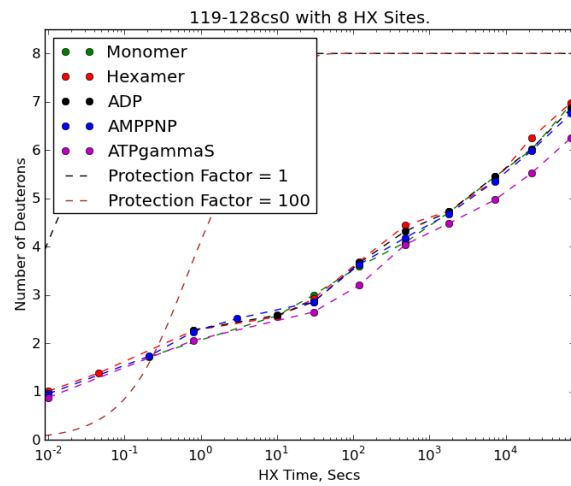
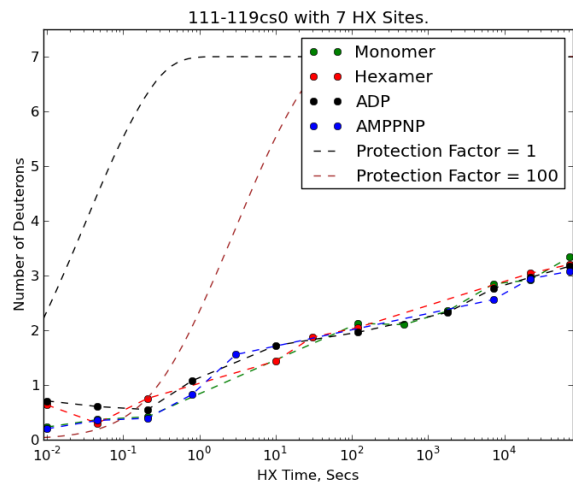


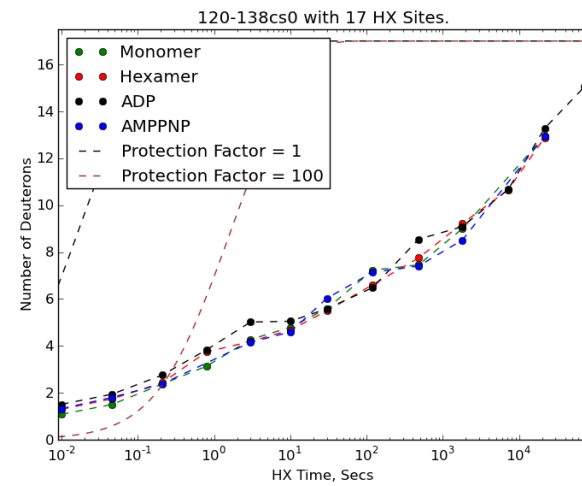
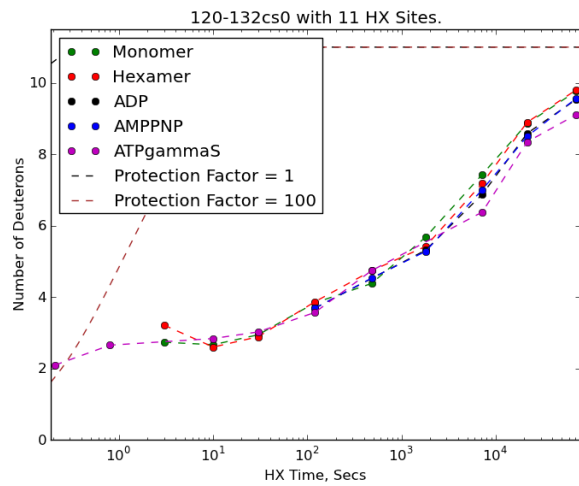
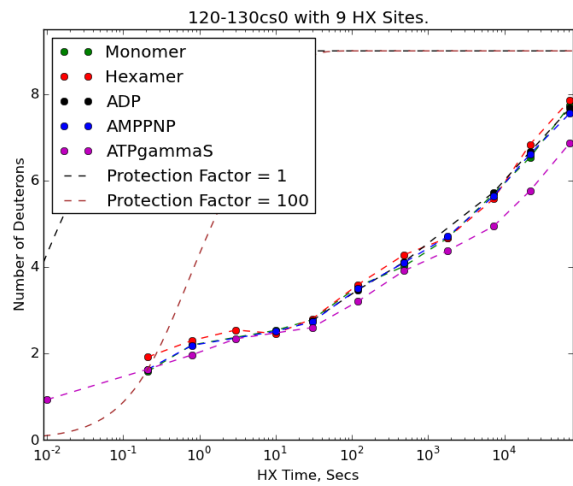
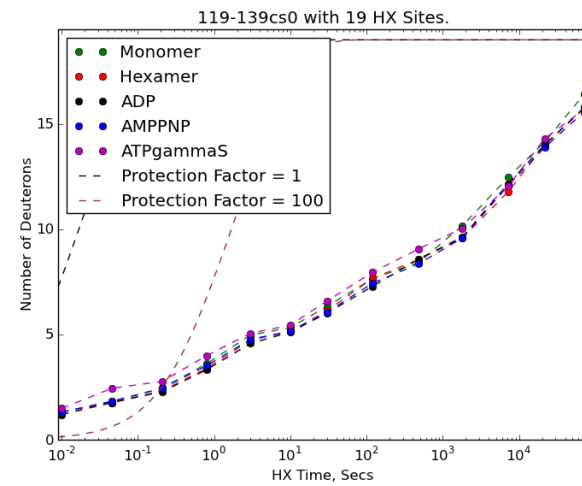
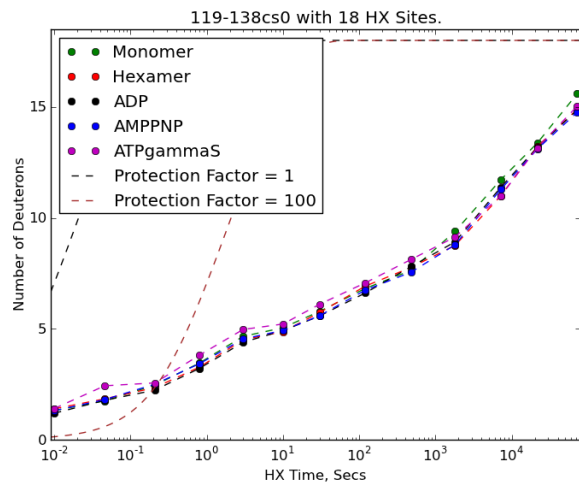
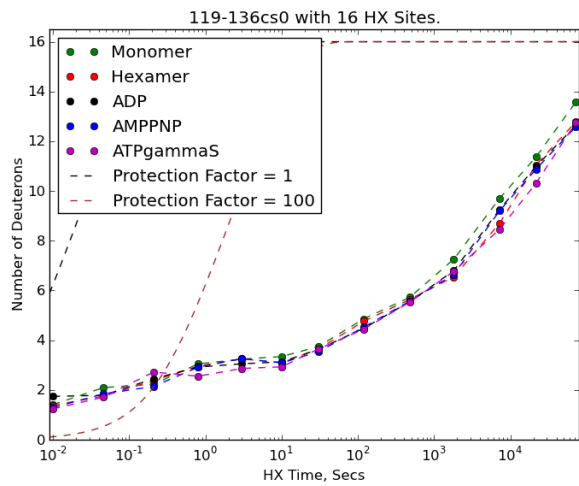


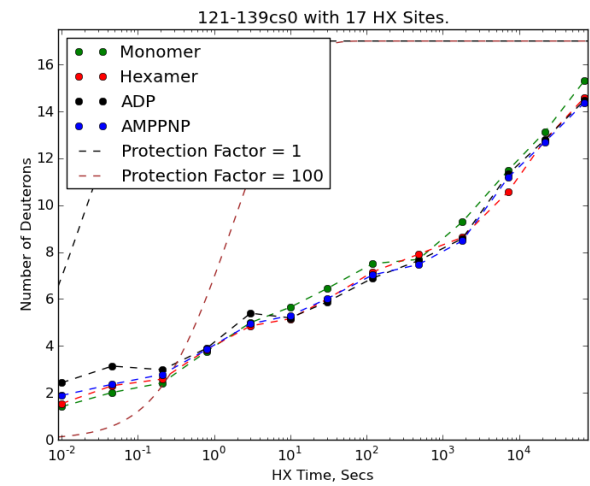
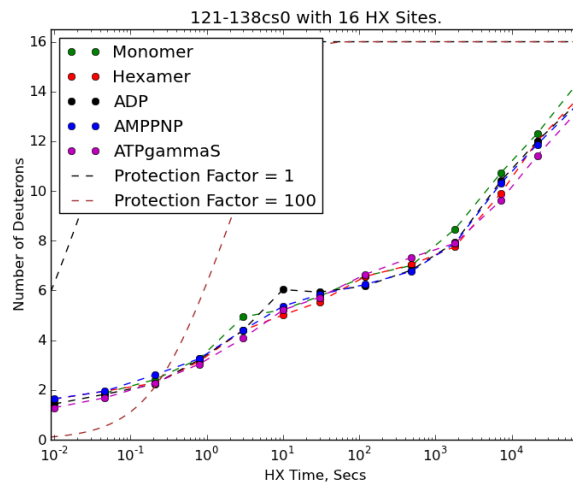
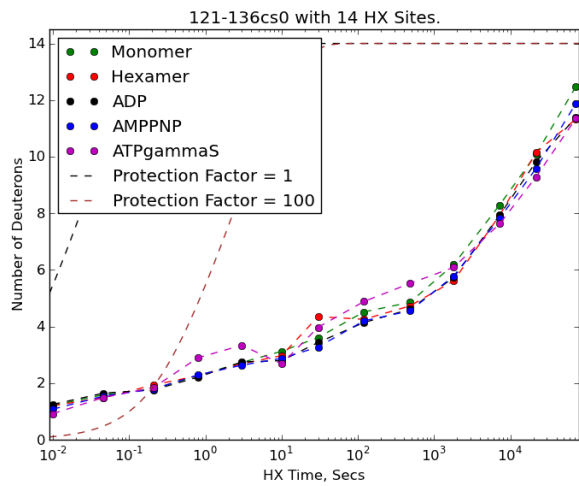
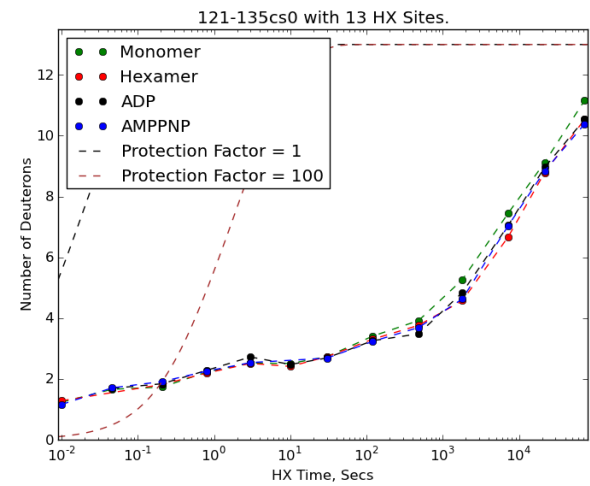
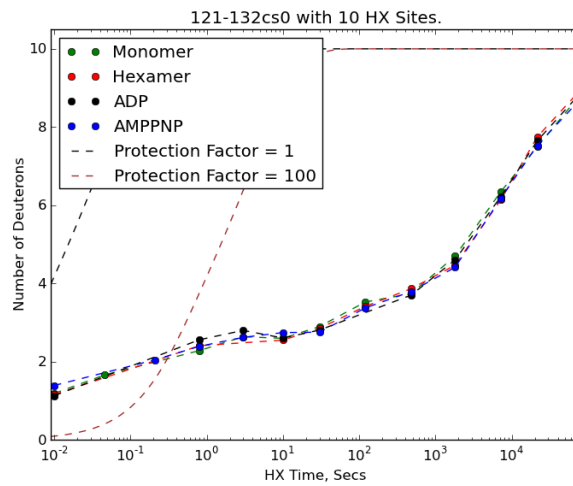
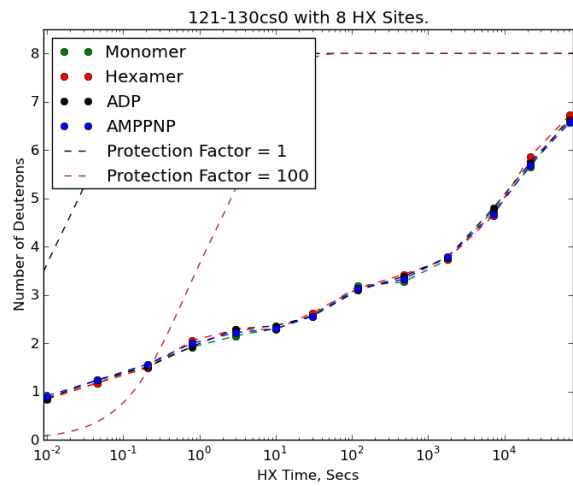


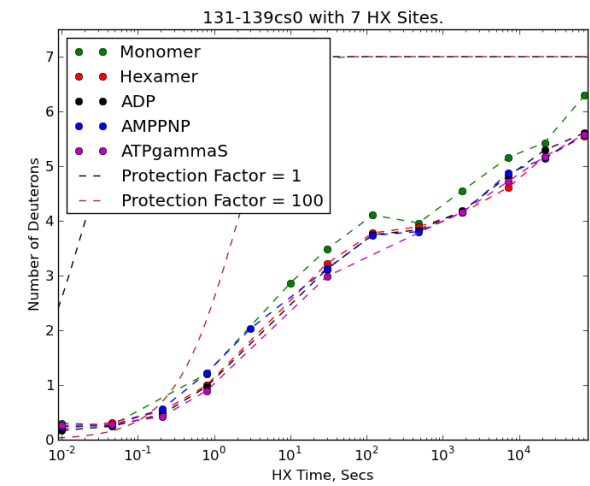
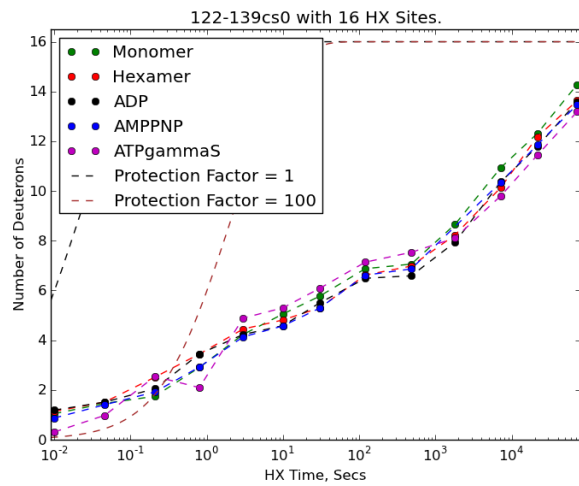
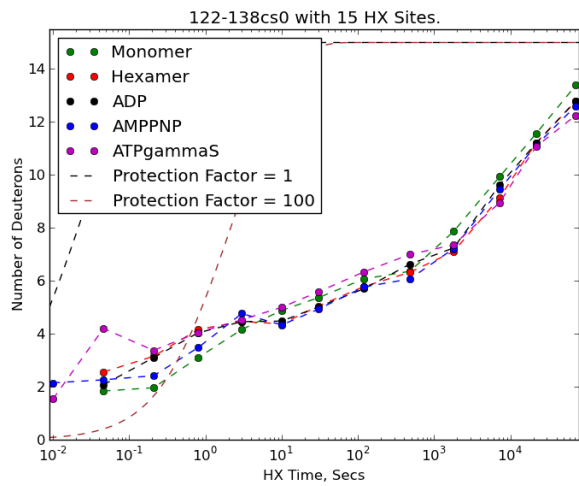
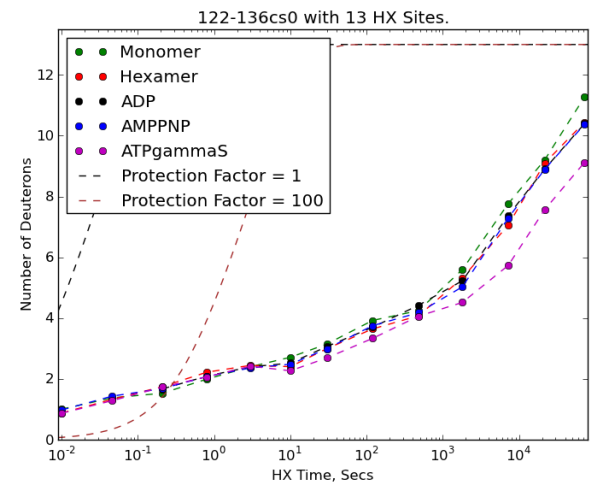
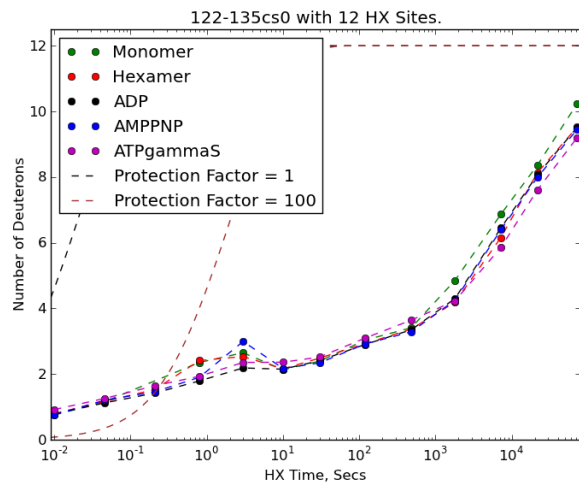
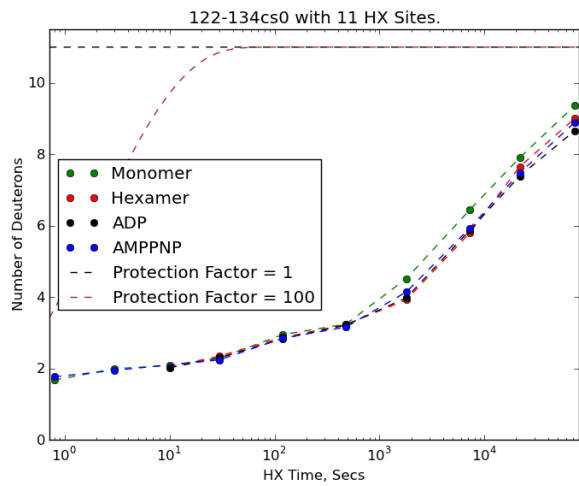


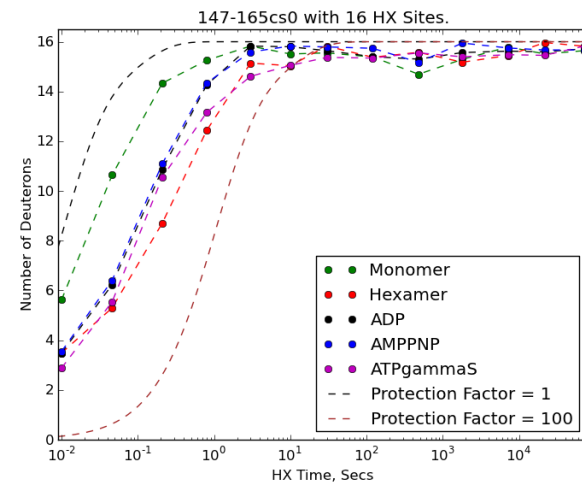
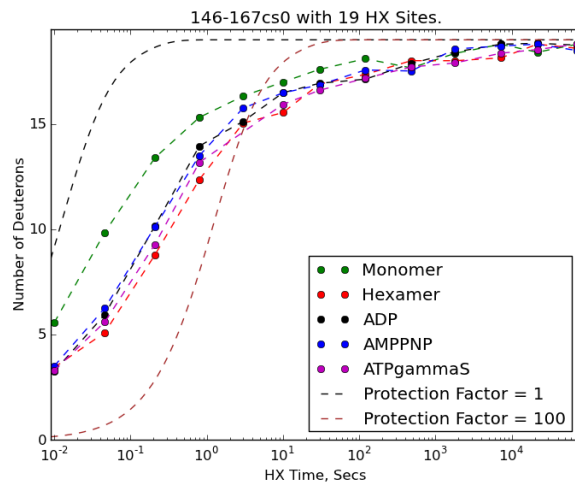
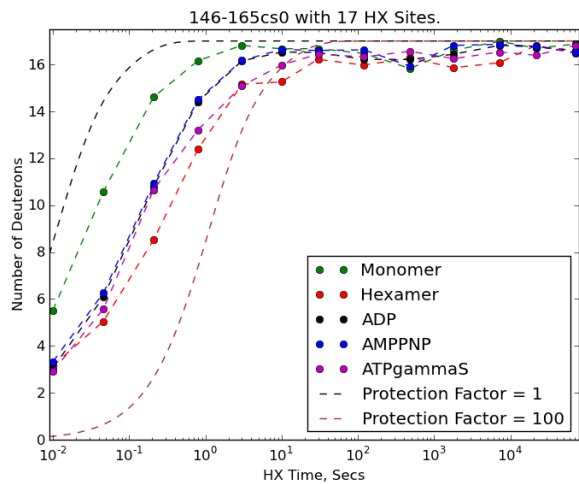
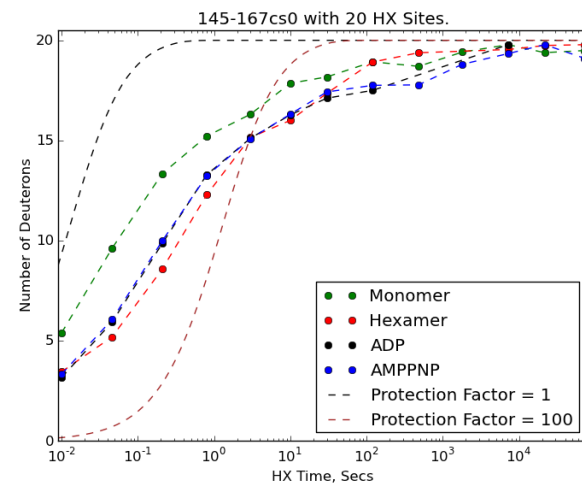
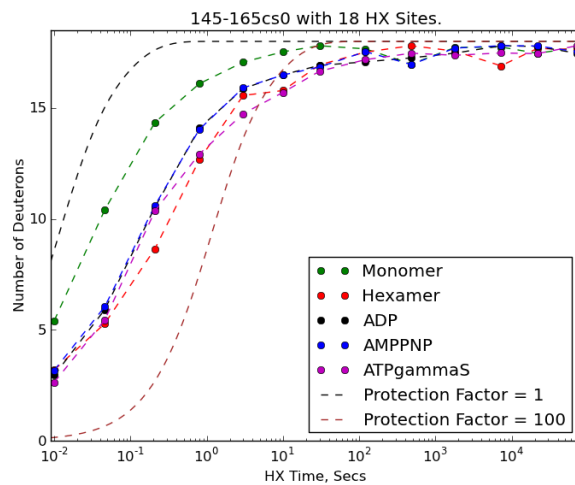
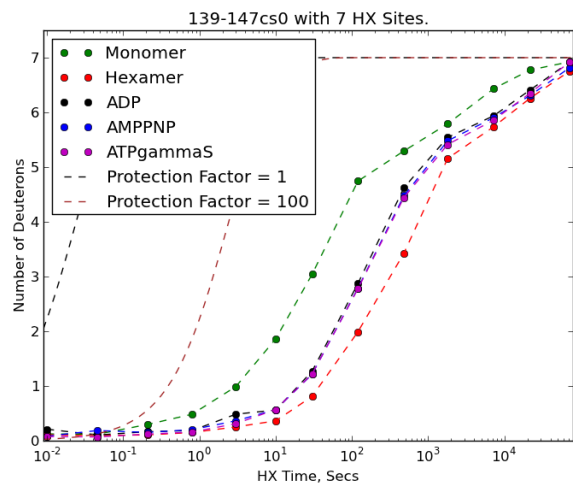


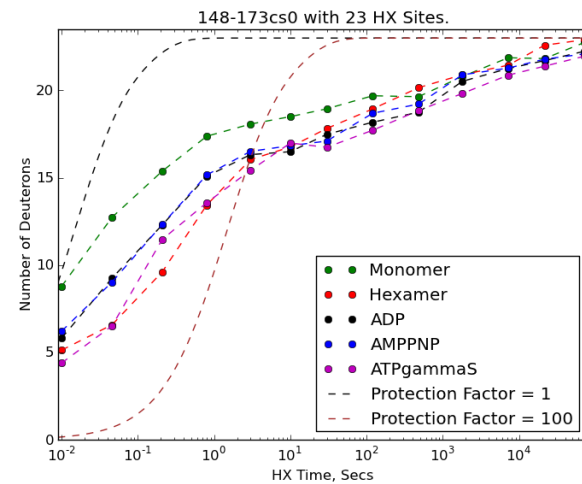
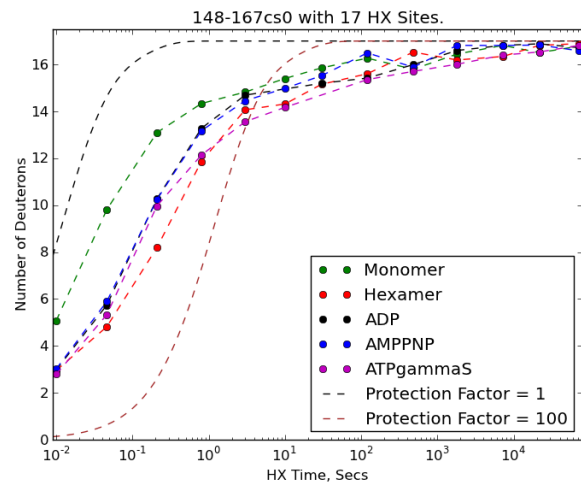
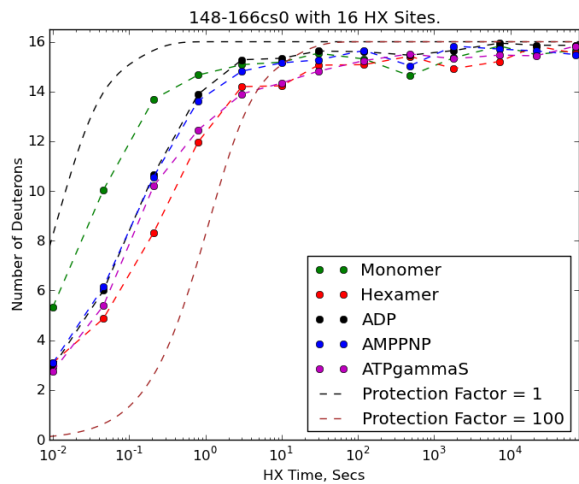
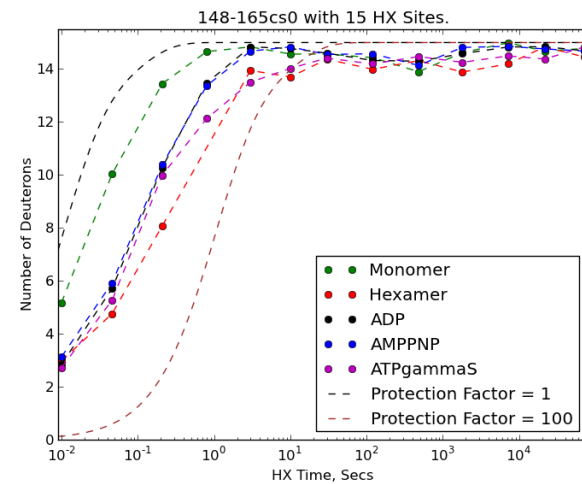
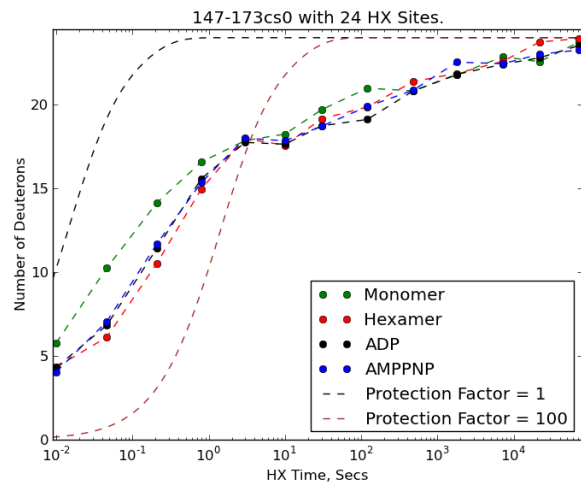
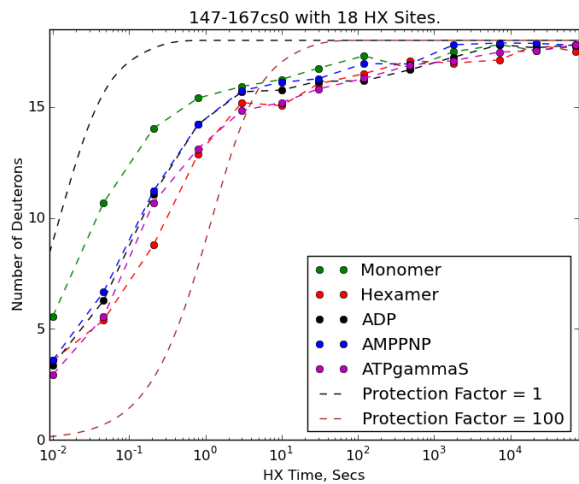


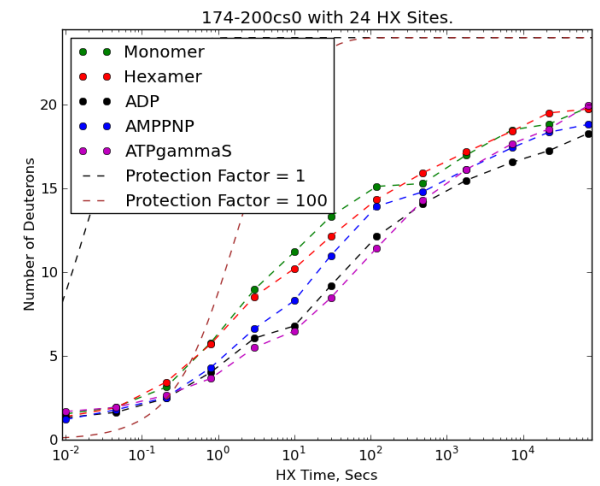
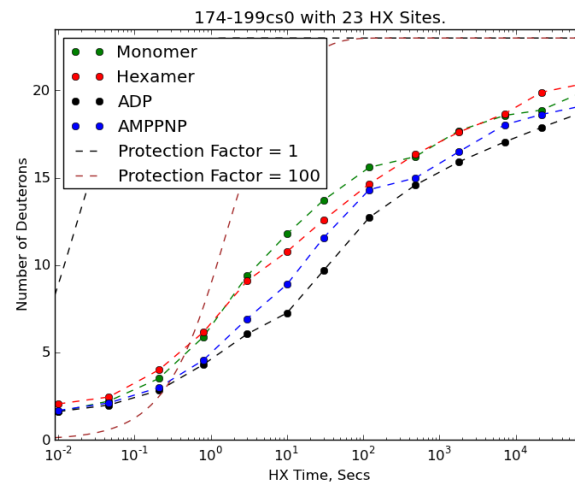
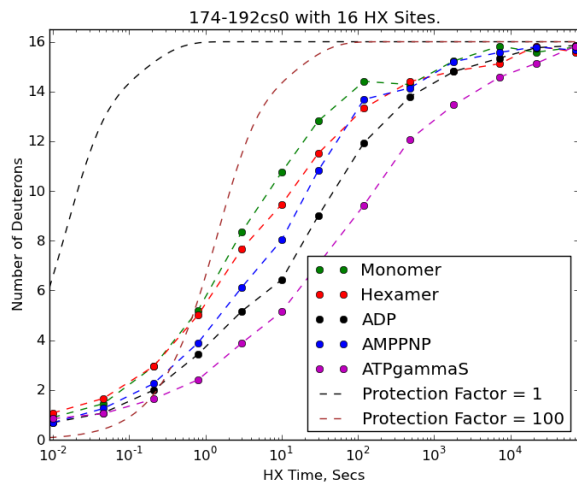
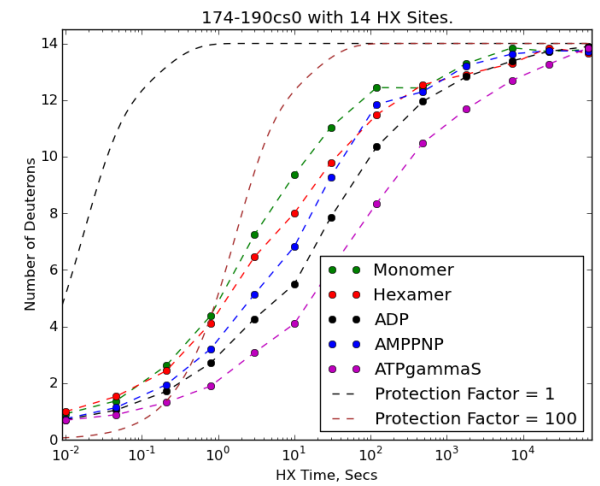
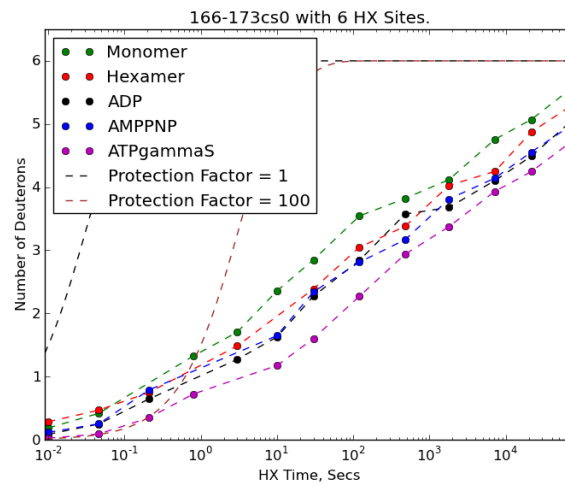
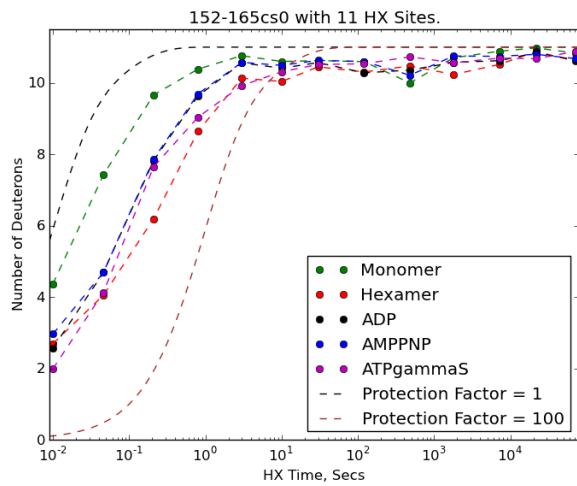


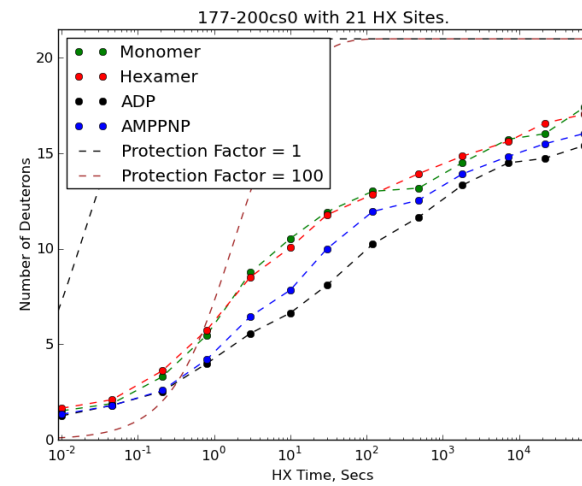
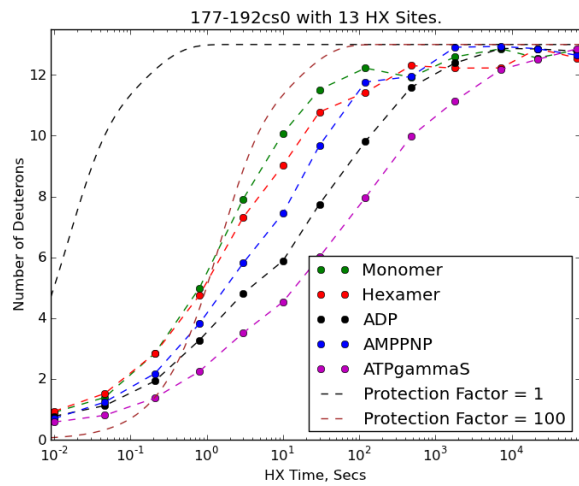
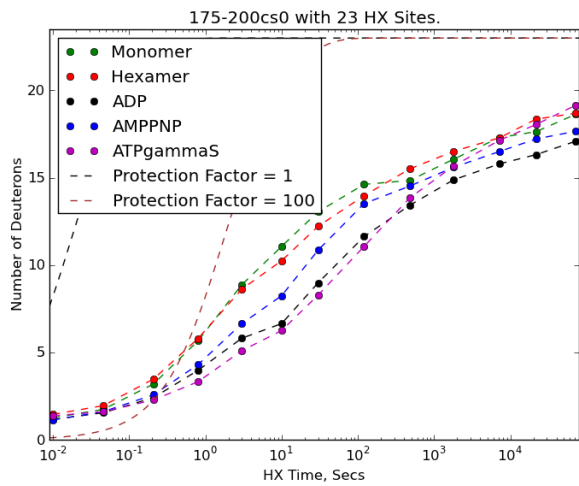
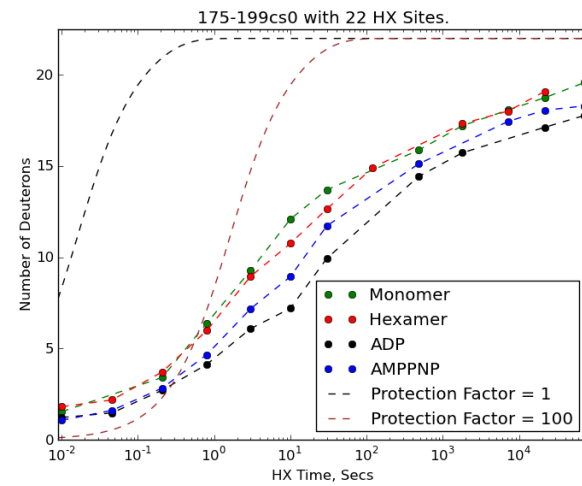
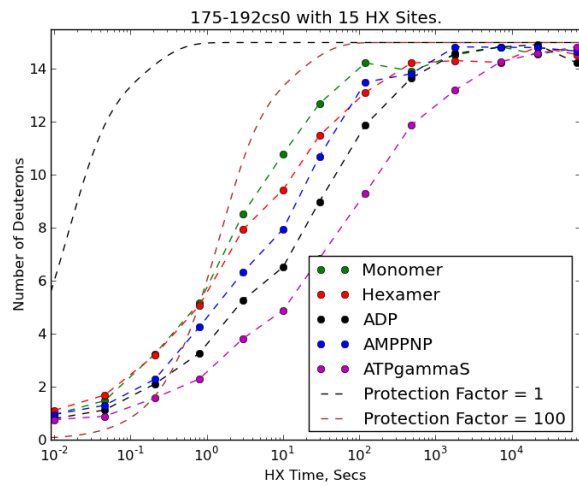
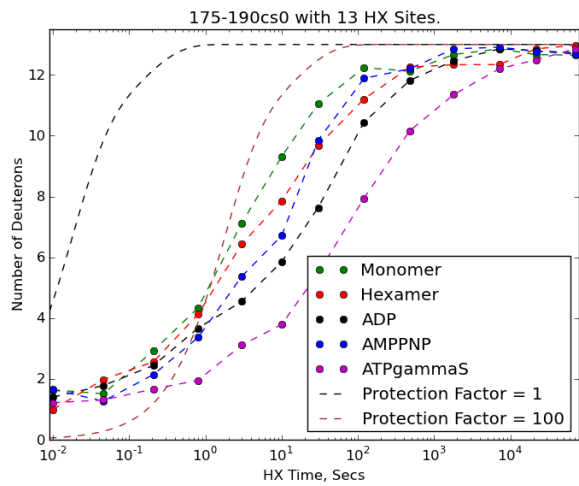


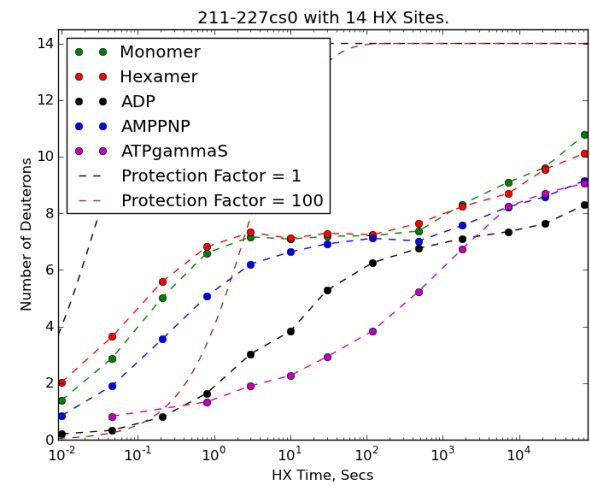
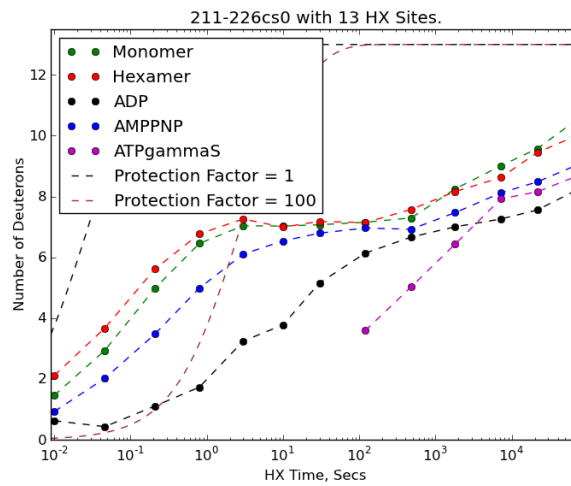
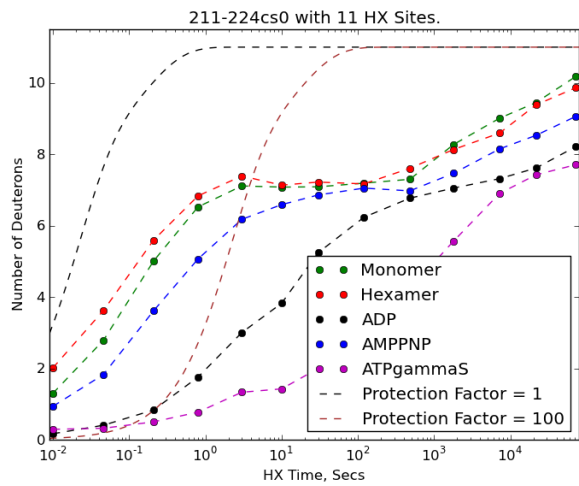
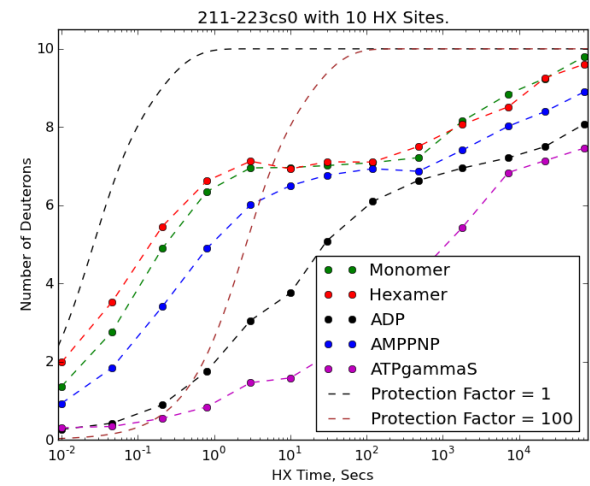
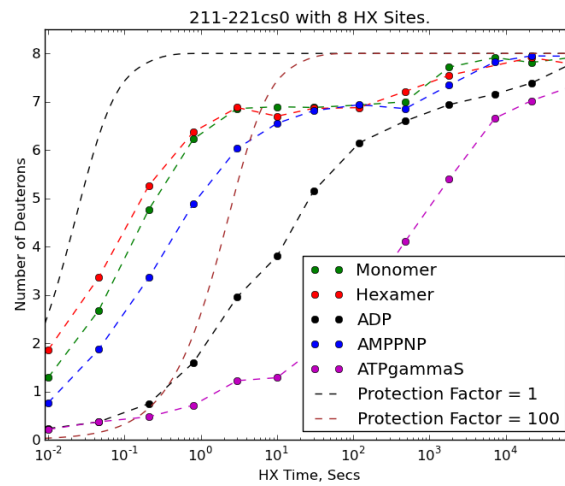
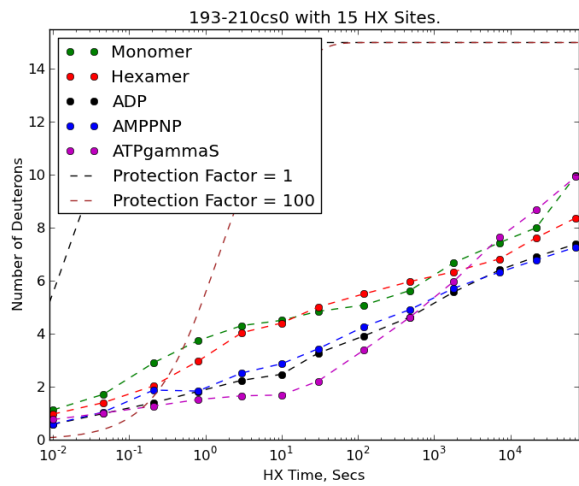


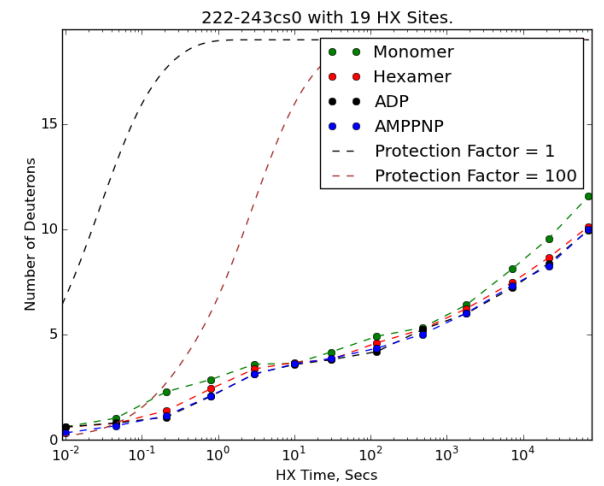
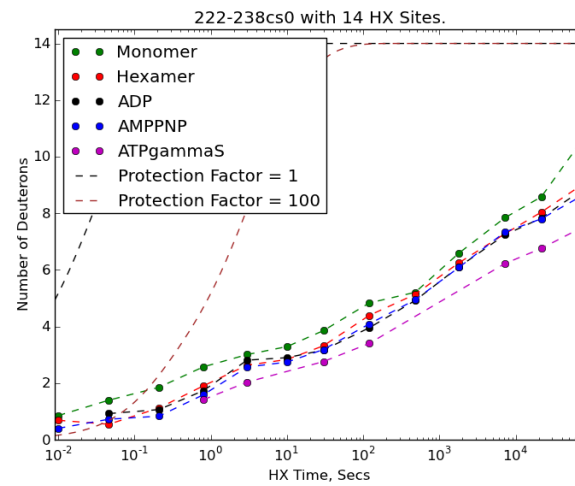
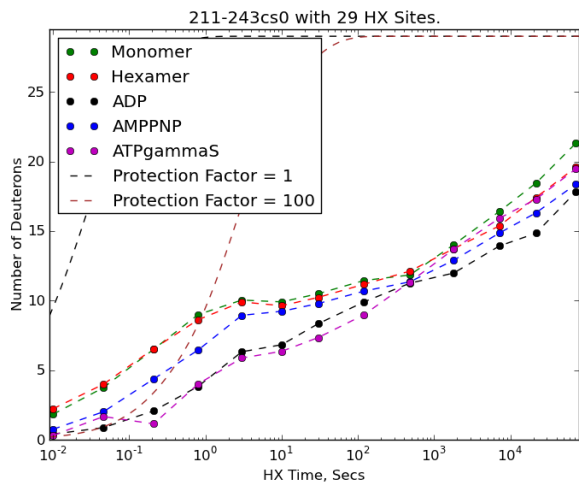
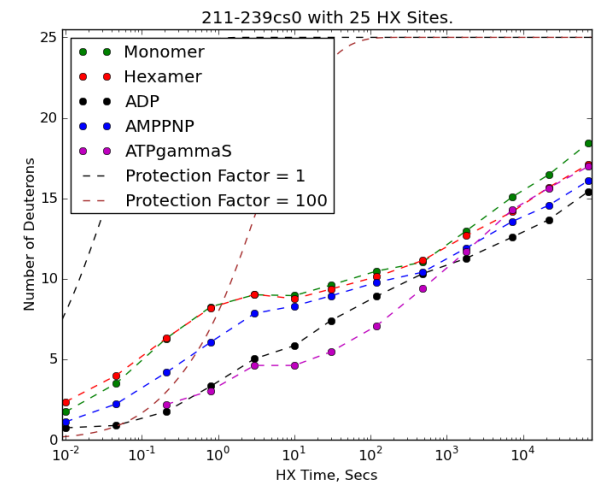
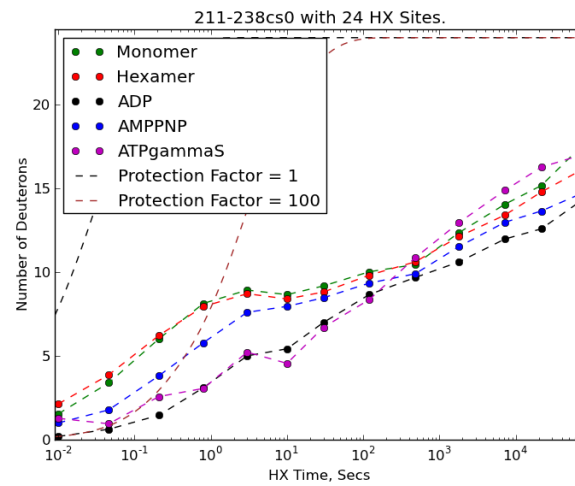
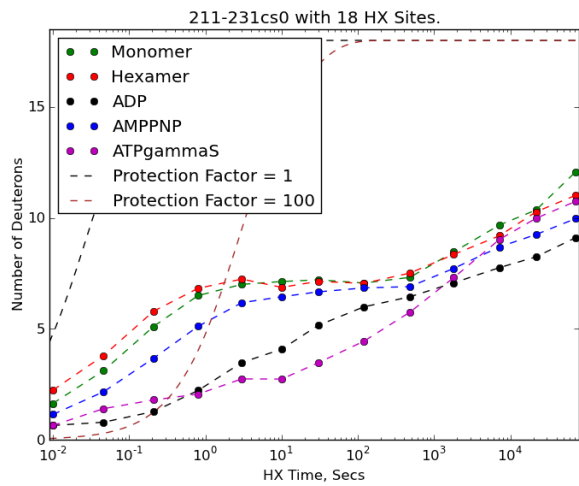


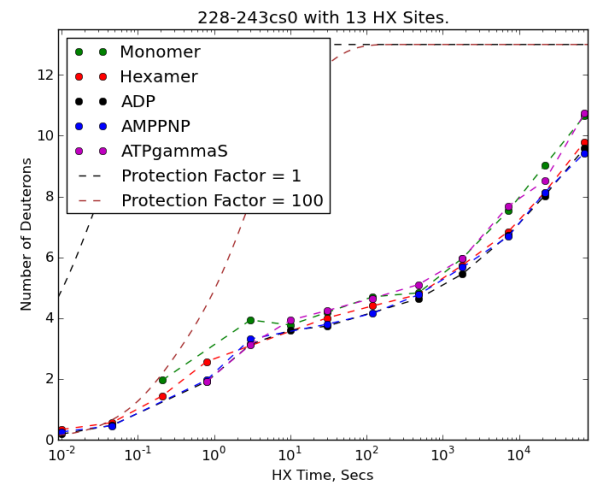
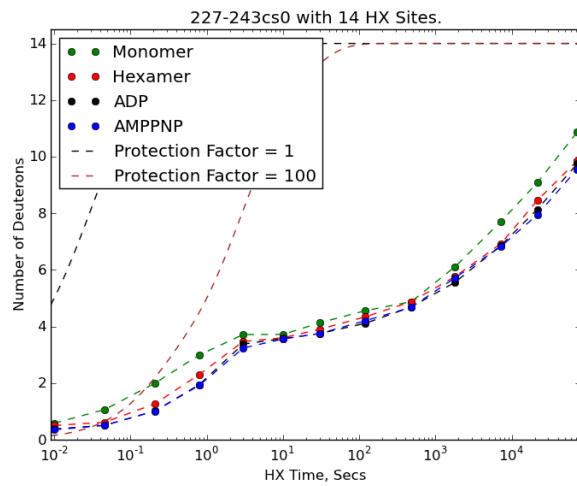
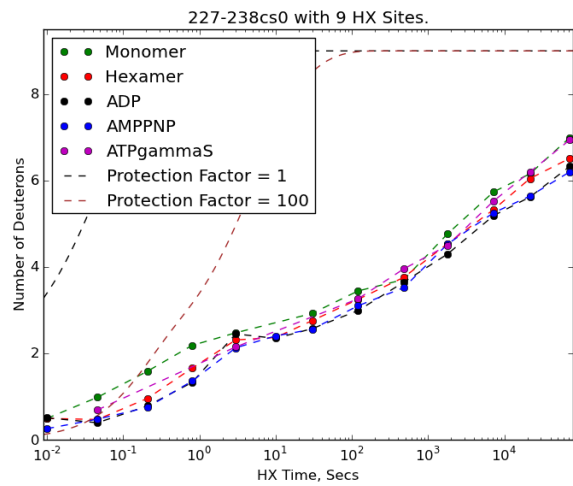
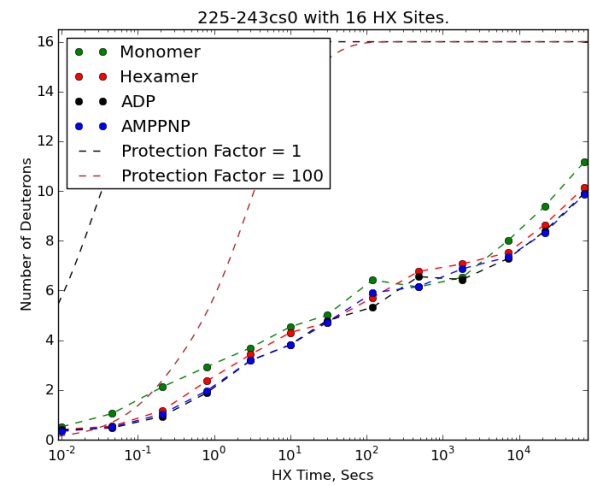
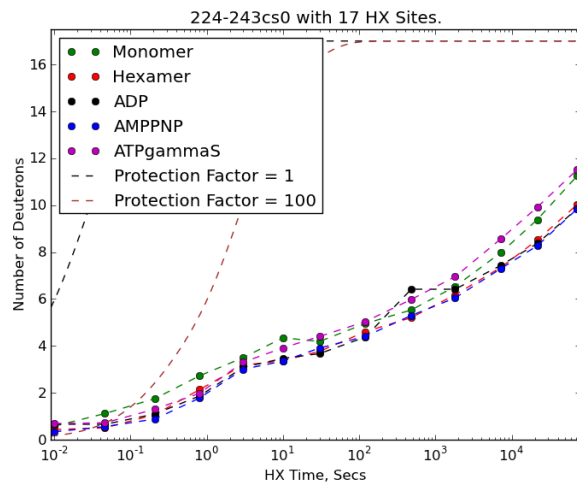
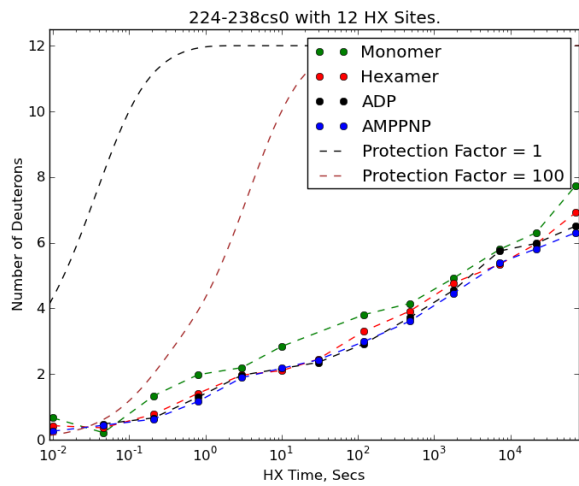


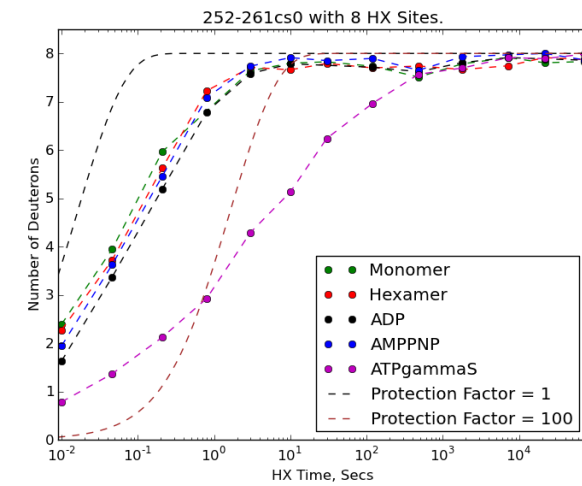
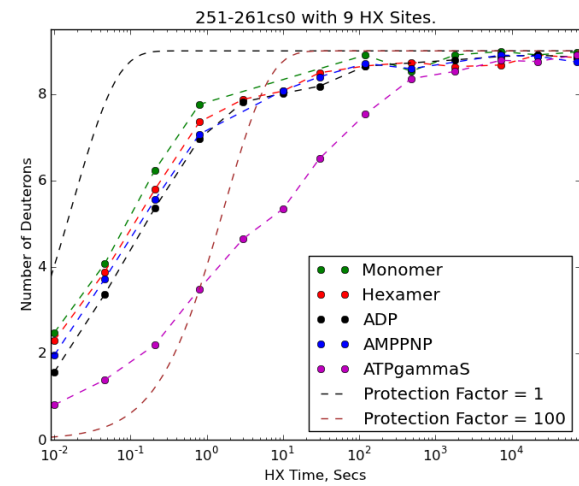
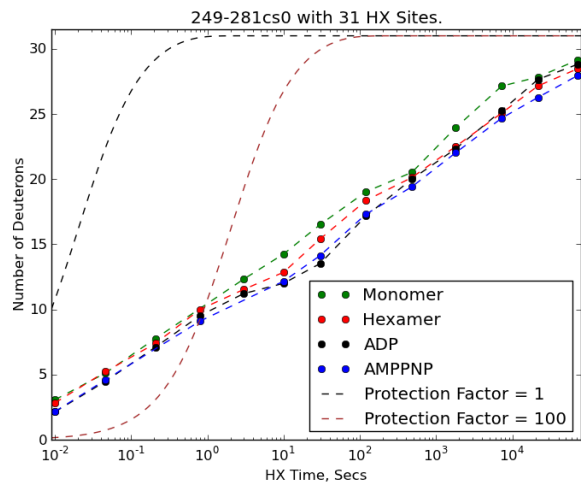
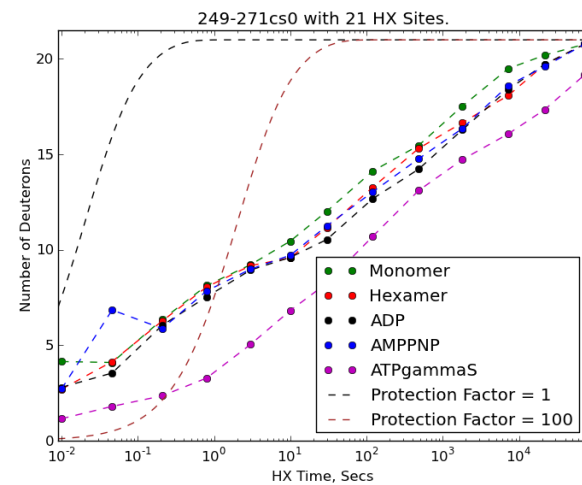
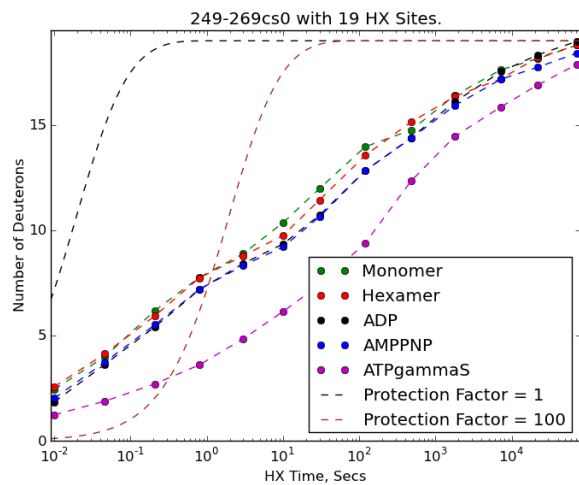
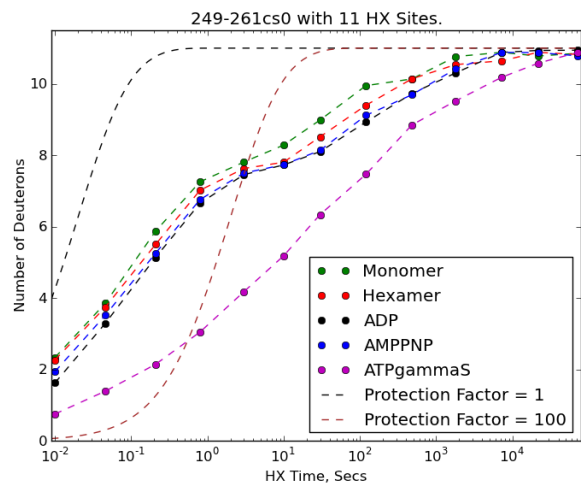


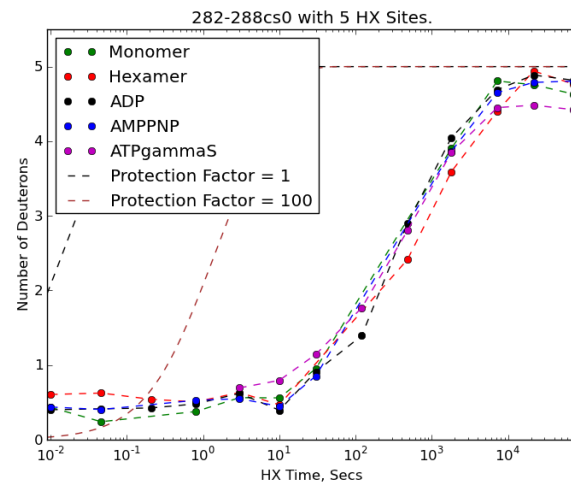
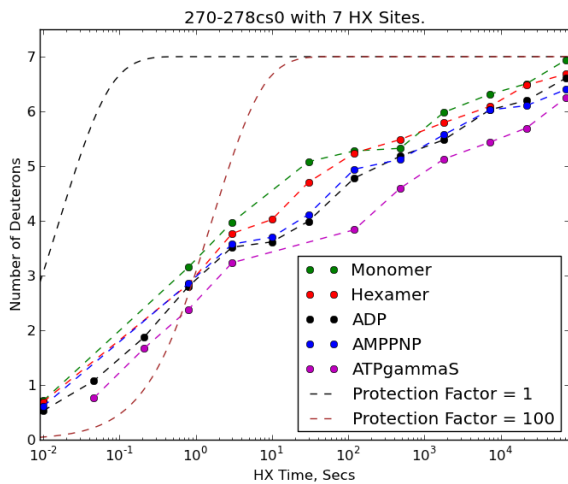
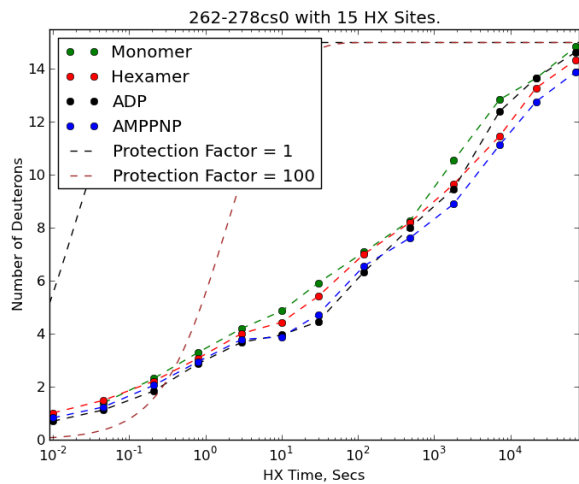
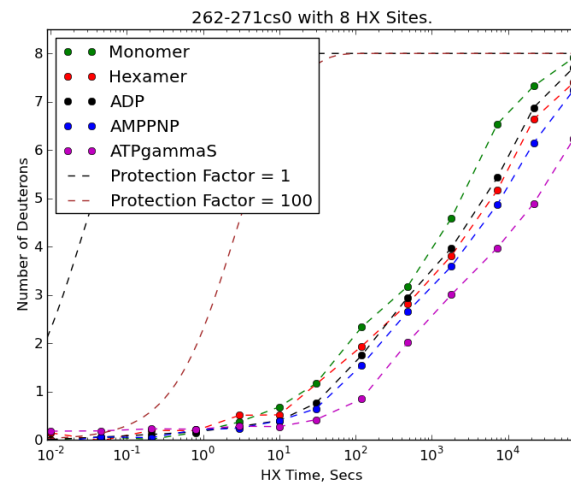
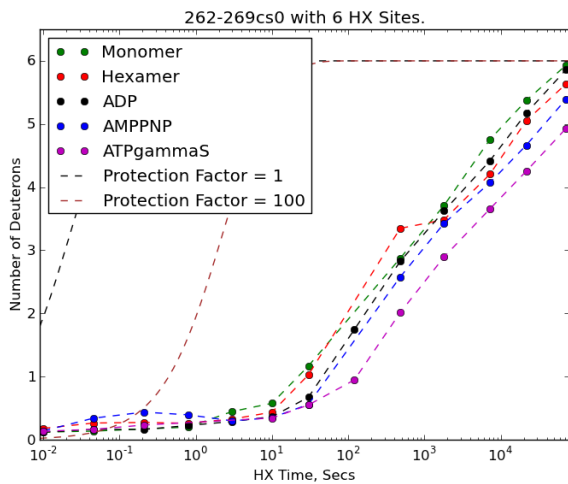
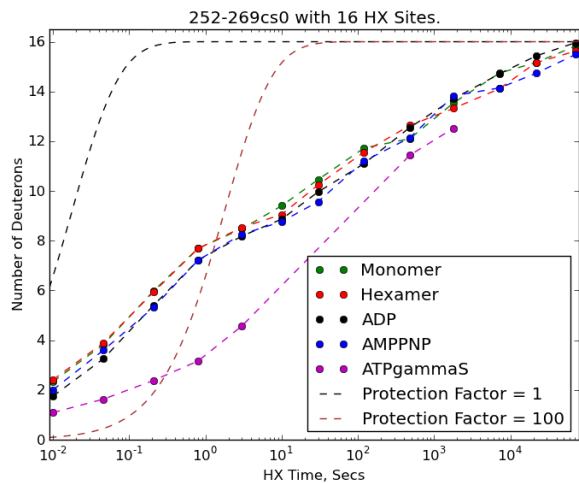


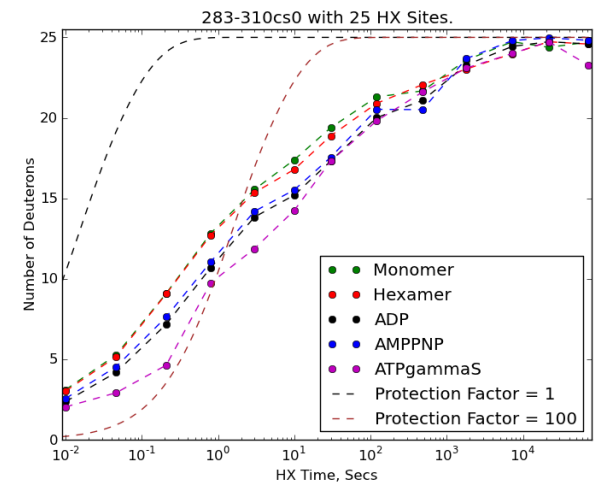
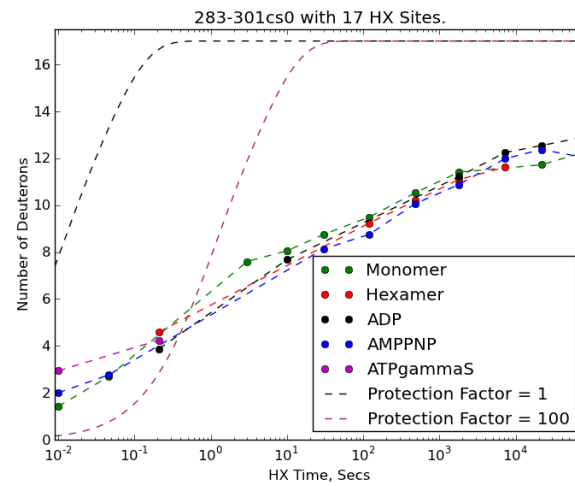
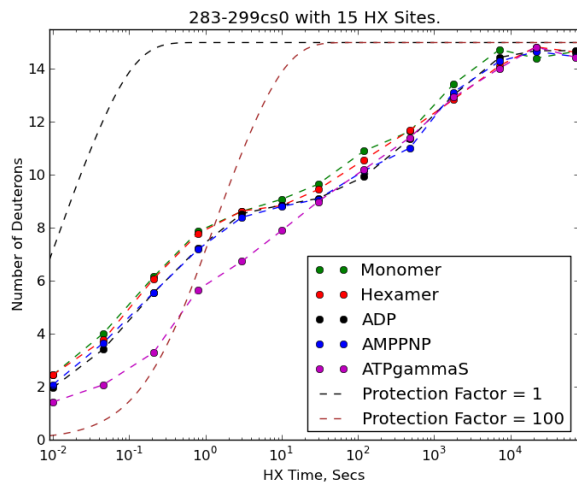
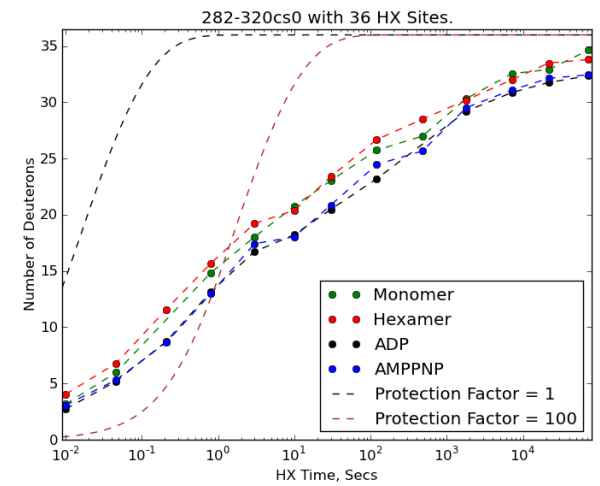
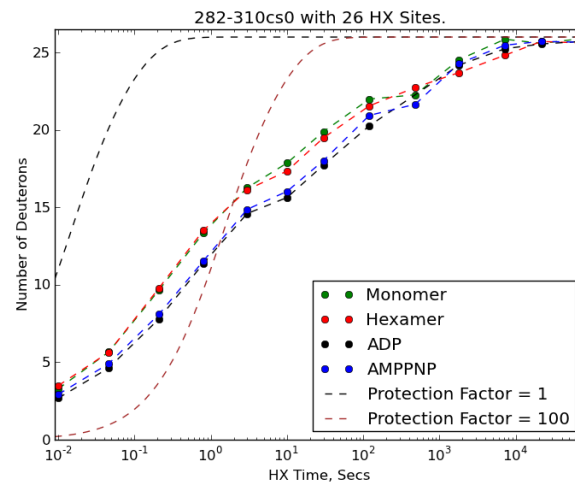
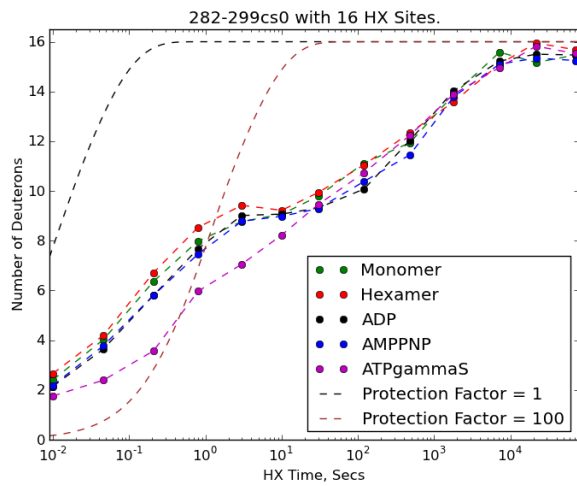


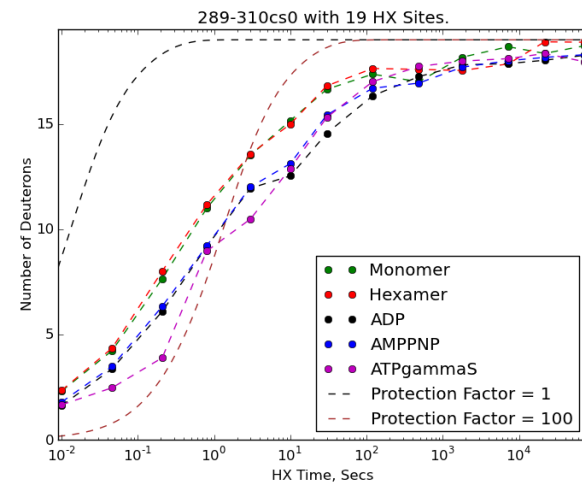
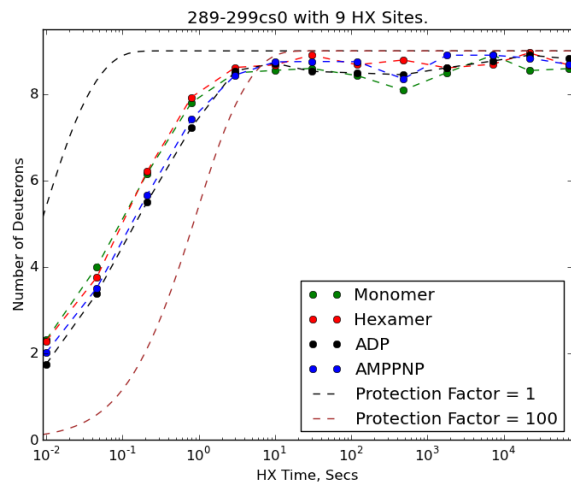
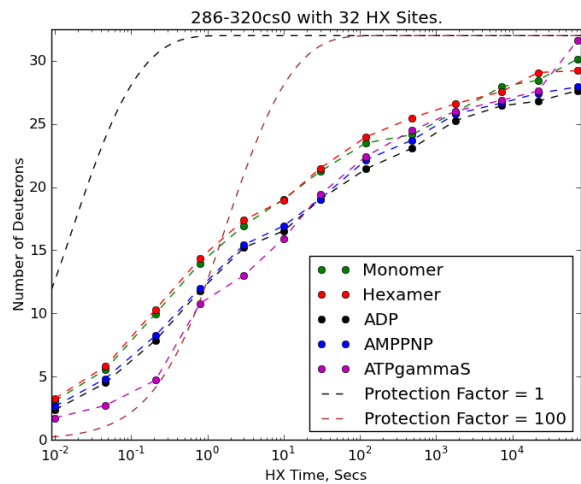
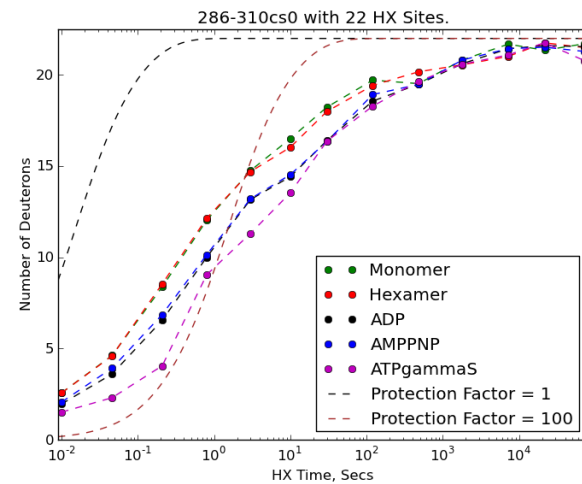
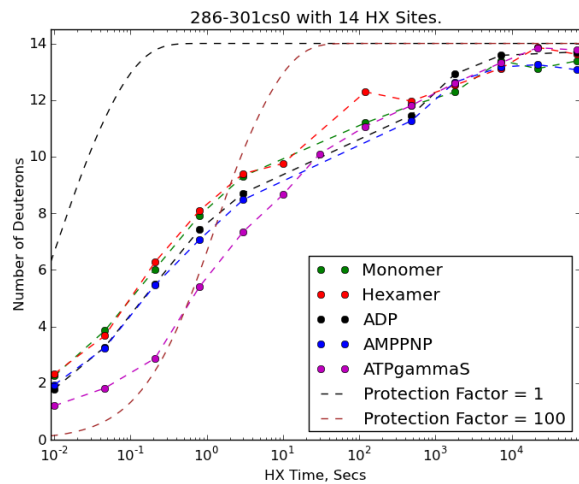
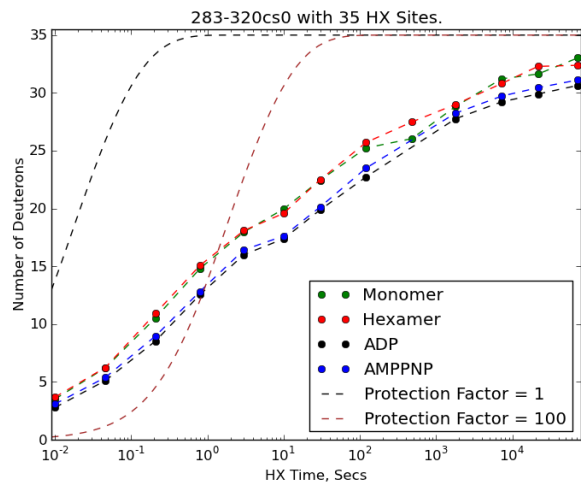




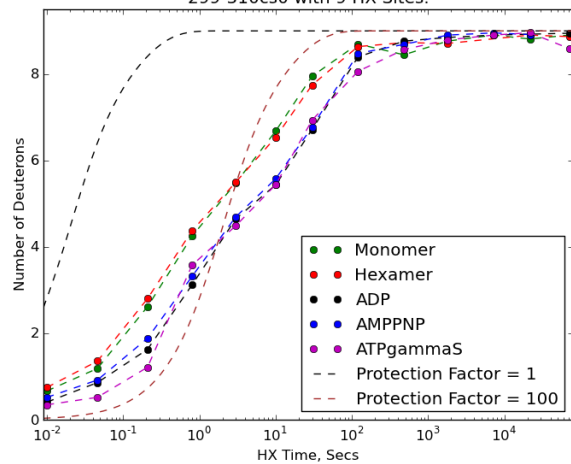




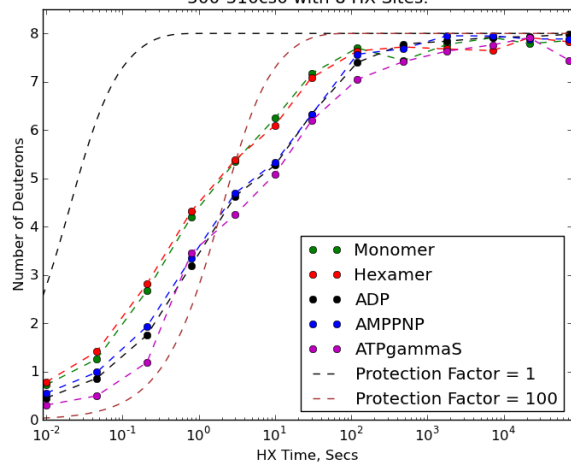




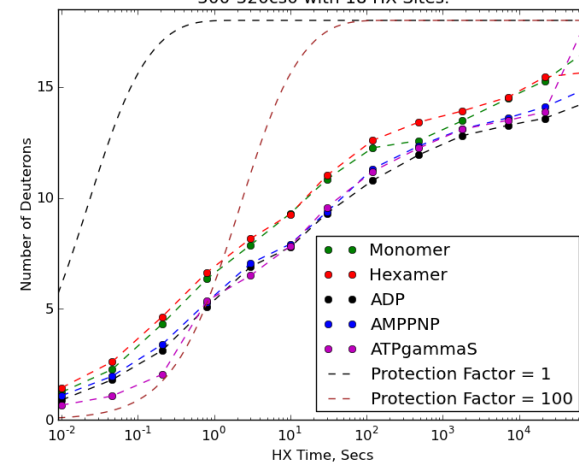
299-310cs0 with 9 HX Sites.



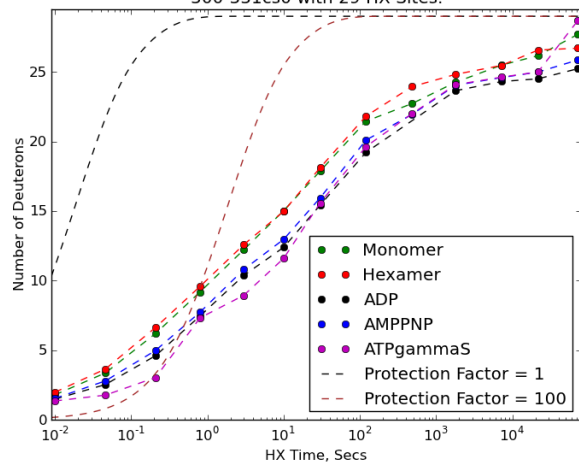
300-310cs0 with 8 HX Sites.



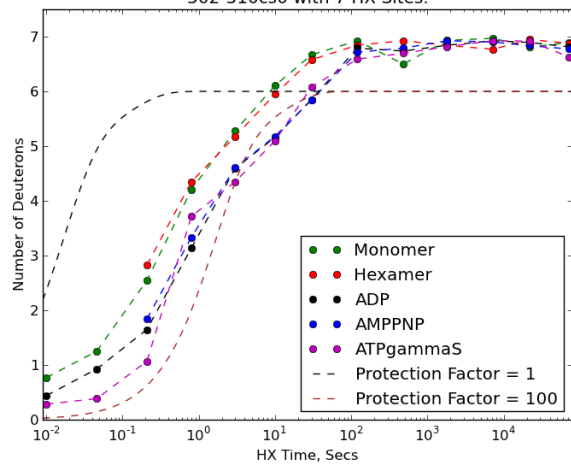
300-320cs0 with 18 HX Sites.



300-331cs0 with 29 HX Sites.



302-310cs0 with 7 HX Sites.



306-320cs0 with 13 HX Sites.

