Comparison of protein H exchange data with local fluctuations observed in perfectly funneled structure based models

Craig PO, Lätzer J, Weinkam P, Hoffman R, Komives E, Wolynes P

Center of Theoretical Biological Physics (CTBP) and Department of Chemistry and Biochemistry, (UCSD), 9500 Gilman Drive -La Jolla, CA 92093 – pocraig@ucsd.edu

ABSTRACT

The principle of minimal frustration assumes that the overall landscape of proteins has evolved to be a rough funnel. When energetic frustration is low enough, topology alone becomes the key factor governing the folding. The structures of transition state ensembles, folding intermediates, and the mechanisms of dimerization, and domain swapping have been well predicted by models where frustration has been removed and topological information about the native state is the sole input. In this work we compare the protein motions simulated using perfectly funneled structure-based models, and amide hydrogen exchange (HDX) measurements under native conditions that report on the thermodynamics and kinetics of protein structures.

1 HDX - local unfolding model

$$[D\text{-}O\text{-}D] + [Protein\text{-}H] \xrightarrow{k_{ex}} [Protein\text{-}D] + [H\text{-}O\text{-}D]$$

$$[Closed\text{-}H] \underset{k_{cl}}{\overset{k_{op}}{\rightleftharpoons}} [Open\text{-}H] \xrightarrow{k_{ch}} [Open\text{-}D]$$

$$Pf = \frac{k_{ch}}{k_{ex}} \qquad k_{ex} = \frac{k_{op} \cdot k_{ch}}{k_{op} + k_{cl} + k_{ch}} \qquad Pf = 1 + \frac{k_{ch} + k_{cl}}{k_{op}}$$

EX1 $k_{cl} < k_{ch}$

$Pf = 1 + \frac{k_{ch}}{k_{op}}$

$k_{cl} >> k_{op}$

$k_{ch} > k_{cl} >> k_{op}$

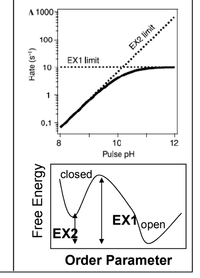
$k_{ex} = k_{op}$

EX2 $k_{cl} >> k_{ch}$

$Pf = 1 + \frac{k_{cl}}{k_{op}}$

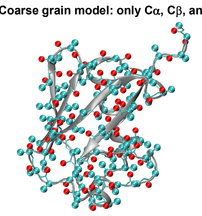
$Pf = \frac{k_{cl}}{k_{op}} = \frac{1}{K_{op}}$

$\Delta G_{op} = kB\ T \ln Pf$



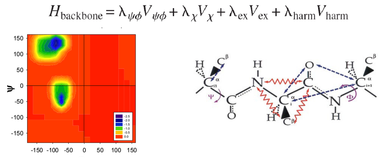
2 Structure based model

Coarse grain model: only Cα, Cβ, and O



$$H = H_{backbone} + H_{na}$$

$$H_{backbone} = \lambda_{\phi\phi} V_{\phi\phi} + \lambda_{\chi} V_{\chi} + \lambda_{ex} V_{ex} + \lambda_{harm} V_{harm}$$



Attractive forces only between residues that make contact in the native structure

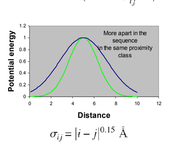
Homogeneous contact potential

Non additivity: forces depend on the structural context. The weight of a pair interaction is modulated by the presence of other residues (p exponent).

$$H_{na} = -\frac{1}{2}\sum_i |E_i|^p$$

$$E_i = \sum_j \epsilon_{ij}(r_{ij}) = -\sum_j \left|\frac{\epsilon}{a}\right|^{1/p} \theta(r_c - r_{ij}^N)\gamma_{ij} \times \exp\left(-\frac{(r_{ij} - r_{ij}^N)^2}{2\sigma_{ij}^2}\right)$$

$$a = \frac{1}{8N}\sum_i \left|\sum_j \gamma_{ij}\theta(r_c - r_{ij}^N)\right|^p$$



$$\sigma_{ij} = |i - j|^{0.15}\ \text{Å}$$

3 Method

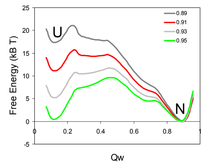
MD simulation - Umbrella sampling Q_w bias

$$Q_w = \frac{2}{(N-1)(N-2)} \sum_{i<j-1} \exp\left(-\frac{(r_{ij} - r_{ij}^N)^2}{2\sigma_{ij}^2}\right)$$

Histogram analysis (WHAM) – Probability of sub states

$$F_{Qwi} = -kB\ T \ln P_{Qwi}$$

Temperature Screening



Sampling was carried out at the Tf (**green curve**), and the results were extrapolated to the experimental temperature (**red curve**) in which the simulated global stability matches the real stability of the protein in the conditions of the HDX experiment

Calculate the probability of the open and closed conformations for each residue from evaluation of local order parameters

Prediction of ln Pf assuming EX2 conditions

$$\ln Pf_{pred} = \ln\left(1 + \frac{k_{cl}}{k_{op}}\right) = \ln\left(1 + \frac{P_c}{P_{cp}}\right)$$

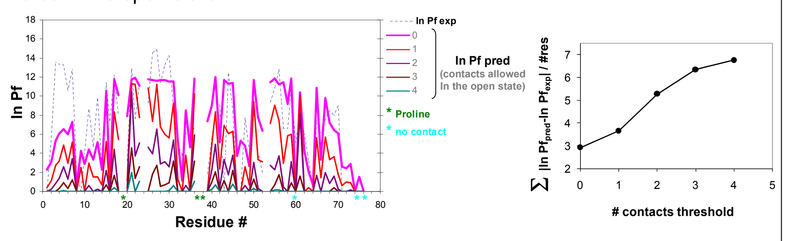
4 Local order parameters used for the structural definition of the open and closed conformations

Accessibility

The accessibility was evaluated by the number of contacts (Qi) of each residue

$$Qi = \sum_j qij \qquad qij = 0.5(1+\tanh(5(r\_cutoff - rij))) \qquad r\_cutoff = 6.5\ \text{Å}$$

Residues with a number of contacts below a certain threshold were considered to be in the open state



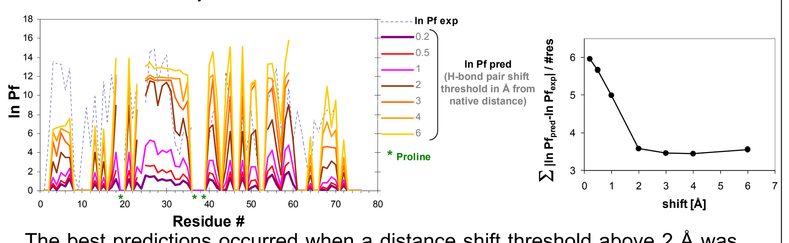
The best predictions occurred when 0 or 1 contact were allowed in the open state. This suggests that an extensive local destabilization is needed for an exchange event to occur.

H-bonding

Residues H-bonded in the native structure were considered to preserve their H-bonding state along the simulations if their pairwise distance was below a certain threshold. In this case they were considered to be in the closed state.

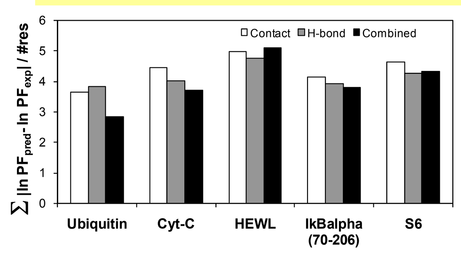
threshold = r_{ijN} + shift

r_{ijN} = native distance between residues i and j



The best predictions occurred when a distance shift threshold above 2 Å was used, suggesting that a significant separation between the H-bond pair has to occur for a successful exchange event

Most predictions improves when the accessibility and H-bonding criteria are combined



Contact criteria
Closed state: #contacts >1

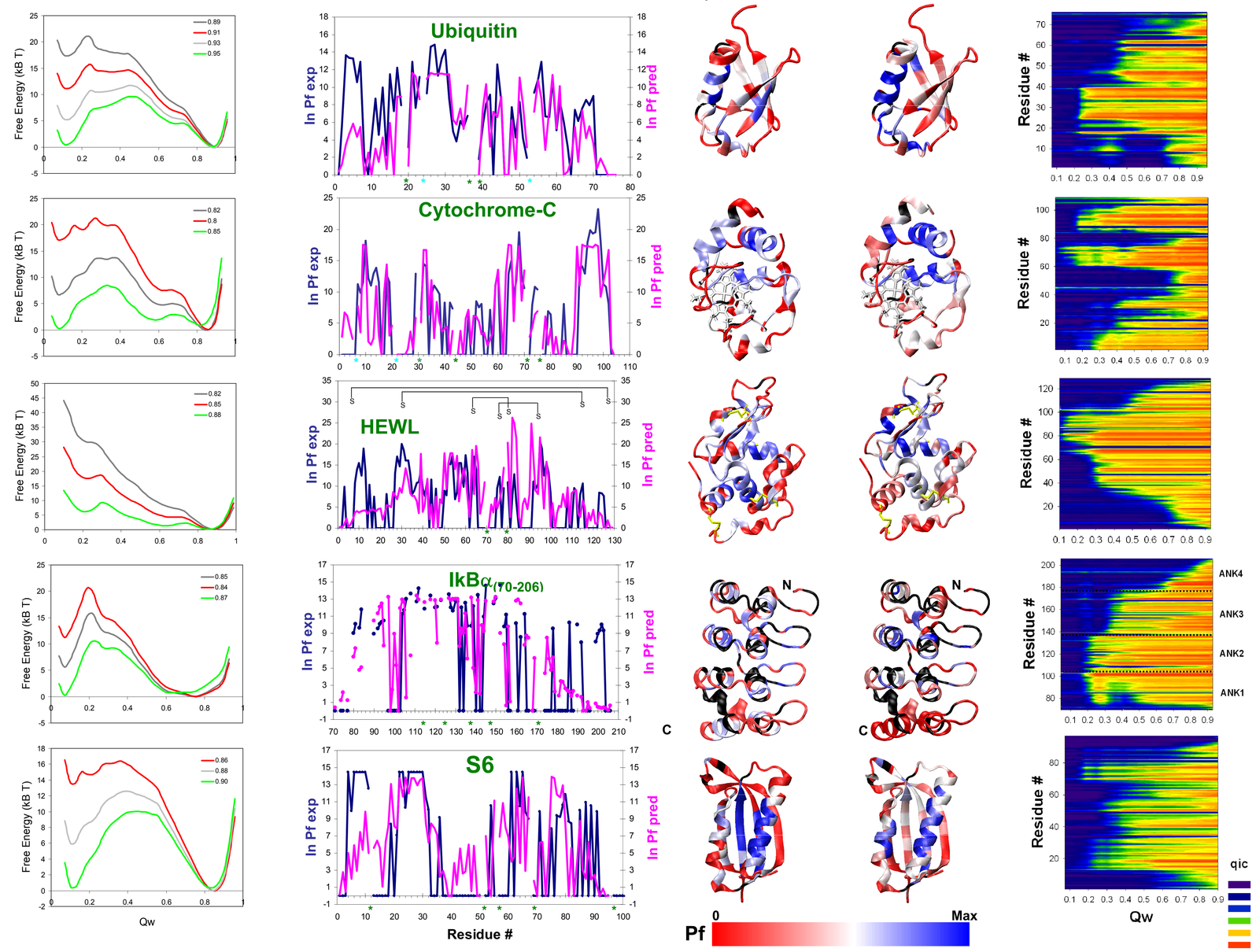
H-bond criteria
Closed state: shift < 2 Å

Combined criteria
Closed state: #contacts >1 or shift < 2 Å

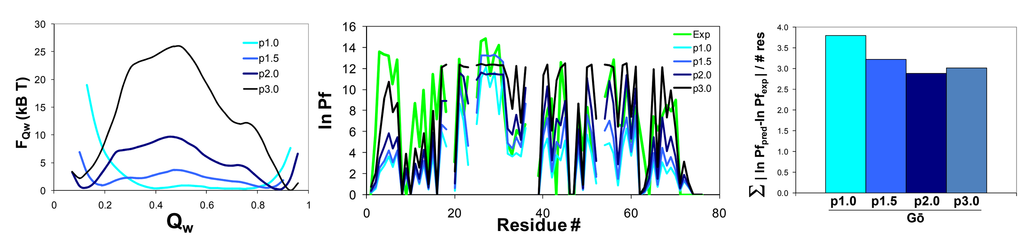
5 Topology is a dominant factor that determines HDX kinetics

Red: Experimental temperature (matches the real global stability under the HDX conditions

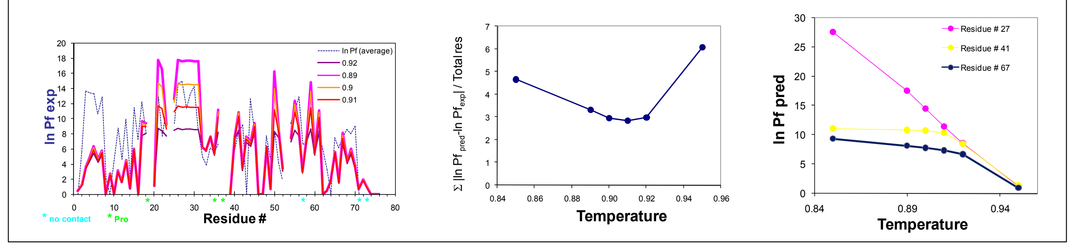
Green: Sampling temperature (Tf)



6 Non additive interactions improves HDX predictions



7 Dependence of HDX predictions with the temperature



CONCLUSIONS

In this work we developed a method for the interpretation of HDX protection factors from perfectly funneled structure based model simulations. The agreement between the predictions and experiments suggests that topology has a dominant role in determining the exchange kinetics. The comparison of different criteria for structurally defining the open and closed states indicates that both accessibility and H-bonding properties should be taken into account in the analysis. The improvement in the predictions produced by the inclusion of non additive forces highlights the relevance of many body interactions in protein dynamics. The method developed will be useful for the interpretation of the exchange process, the parametrization of effective interactions between residues and fine tuning of the energetic terms in molecular dynamic simulations, the analysis of protein motions, and the characterization of the native and partially folded ensembles.