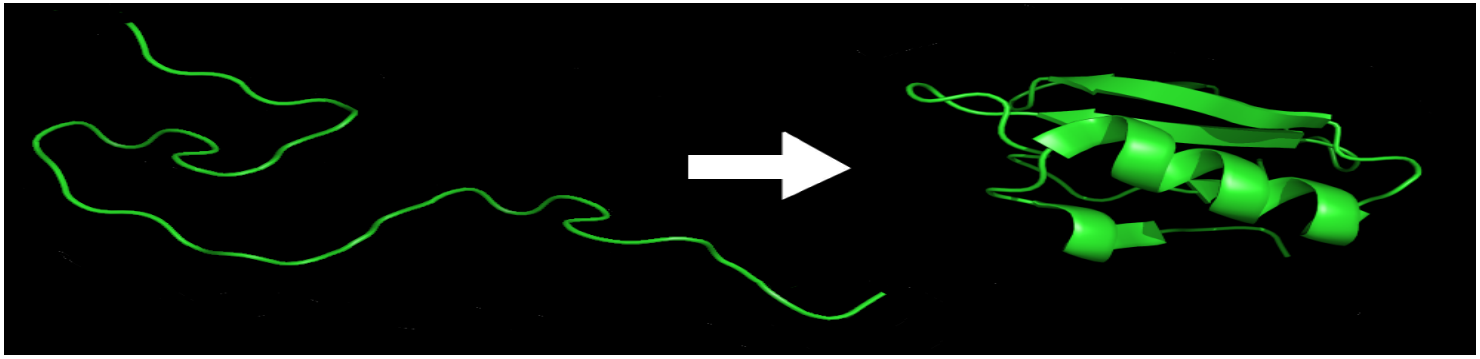
On the Diffusion of Alpha-Helical Proteins in Solvents

Wilson I. Barredo

January 6-9, 2014 JAGNA, BOHOL

Proteins are an essential component to many biological functions and participate in virtually all processes within biological cells.

-> must fold into a functional three-dimensional structure



http://en.wikipedia.org/wiki/Protein_folding

Protein Structures

1. Primary structure- refers to the sequence of amino acids forming the polypeptide chain.
2. Secondary structure- e.g. Alpha helix, beta sheet and turns
3. Tertiary structure- the overall shape of a protein molecule
4. Quaternary structure (Chains) – formed by several proteins.

Predicting the Structure of Protein

E.g.: Folding@Home andRosetta@Home

- To uncover the behavior of a protein during folding.

Two most common methods today: Computational and Theoretical.

Computational-> long computing times is necessary in computer simulation.

-> much longer computing times are needed to reach biologically relevant scales. Expensive

Experimental -> Expensive

Analytical -> White Noise Calculus?? Cheap.

To predict the structure of proteins:

We needed to know their diffusion coefficient and their drift coefficient.

Probability Density for Winding Polypeptides [1]

Fokker-Planck Equation:

$$P(r,\ t+\tau \mid r_0, t) = \left(\frac{1}{2\pi D\tau}\right)^{3/2} \exp\left\{-\frac{1}{2D}\left[\left(\frac{\Delta r}{\tau}\right) - A\right]^2 \tau\right\}.$$

Application of Chapman-Kolmogorov Equation:

$$P(r_1, L \mid r_0, 0) = \int exp\left\{-\frac{3}{2l}\int_0^L \left[\frac{dr}{ds} - \frac{l}{3D} A\right]^2 ds\right\} D[r]$$

Parametrized the path with:

$$r(s) = r_c + \kappa\, \mathbf{B}(s), \quad \Rightarrow \quad dr/ds = \kappa\, \omega(s), \text{ where } \omega(s) = d\mathbf{B}/ds$$

ω(s) is white noise variable

[1] C. C. Bernido, M. V. Carpio-Bernido and M. G. O. Escobido, in Stochastic and Quantum Dynamics of Biomolecular Systems, eds. C. C. Bernido and M. V. Carpio-Bernido (American Inst. of Physics, New York, 2008), pp. 139–148.

Winding Probability for Helical Conformations

$$W(n,L)= R\sqrt{\frac{4\pi}{lL}}\,\frac{exp\left[-\frac{R^2}{lL}\left(2\pi n+\frac{l}{2DR}\int_0^L f(s)ds\right)^2\right]}{\theta_3\left(\frac{1}{4DR}\right)\int_0^L f(s)ds},$$

Besselian drift coefficient of order 2p + 1

$$f(s)=k\,J_{2p+1}(\bar{v}s)$$

$$\int_0^L f(s)ds=(k/v)\left[1-J_0(vL)-2\sum_{m=1}^{p}J_{2m}(vL)\right], \qquad [2]$$

with p≥ 1.

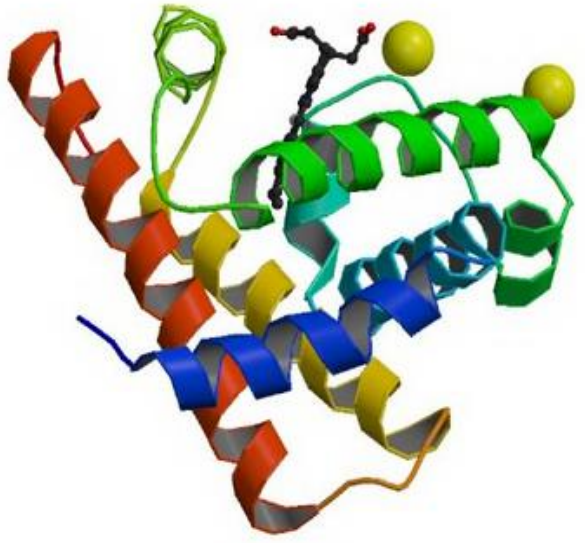
$$W(n,L)\approx R\sqrt{\frac{4\pi}{Ll}}\exp\left\{-\frac{R^2}{Ll}\left[2\pi n+\frac{kl}{2DRv}\times\left(1-J_0(vL)-2\sum_{m=1}^{p}J_{2m}(vL)\right)\right]^2\right\}$$

[2] Aringa, H.P.: Ph.D. Thesis (MSU-Iligan Institute of Technology, 2011).

Application

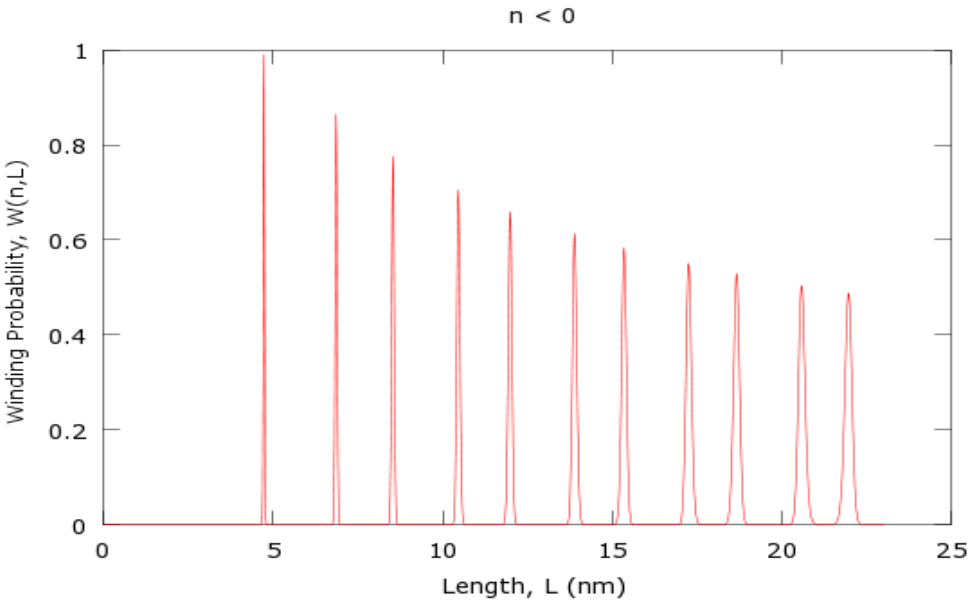
$$W(n,L) \approx R\sqrt{\frac{4\pi}{Ll}}\exp\left\{-\frac{R^2}{Ll}\left[2\pi n + \frac{kl}{2DRv} \times \left(1 - J_0(vL) - 2\sum_{m=1}^{p} J_{2m}(vL)\right)\right]^2\right\}$$

General Properties of Alpha Helical Proteins	
Radius of helix(R)	0.25 nm
Monomer length (*l*)	0.15 nm
# of monomers per helical turn	3.6
Diffusion Coefficient (k/D)	?
Drift Coefficient (v)	?



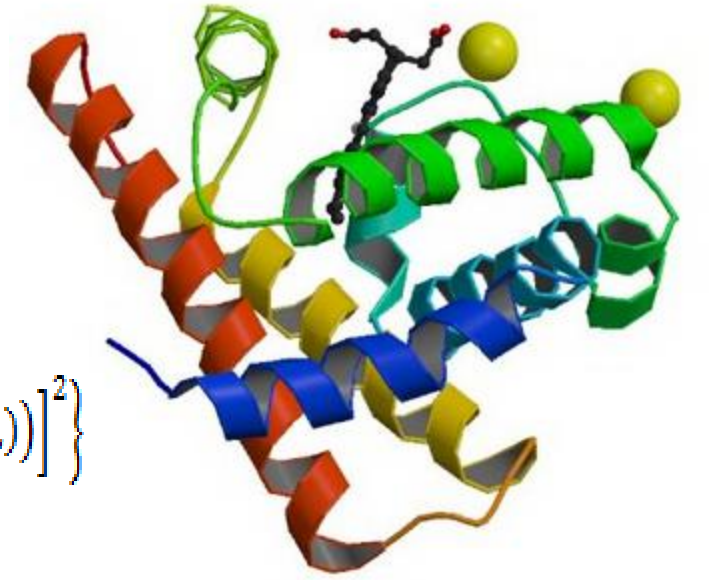
http://www.rcsb.org/pdb/explore/explore.do?structureId=4mbn

Myoglobin (4MBN): Graph of W(-n,L) versus length for v = 1.93/nm and k/D = 1420/nm



Length of 153 residues, alpha helical segments is about 80% of its length or about 123 residues, and has 11 helices.

L = 0.15(153) nm ≈ 23 nm.



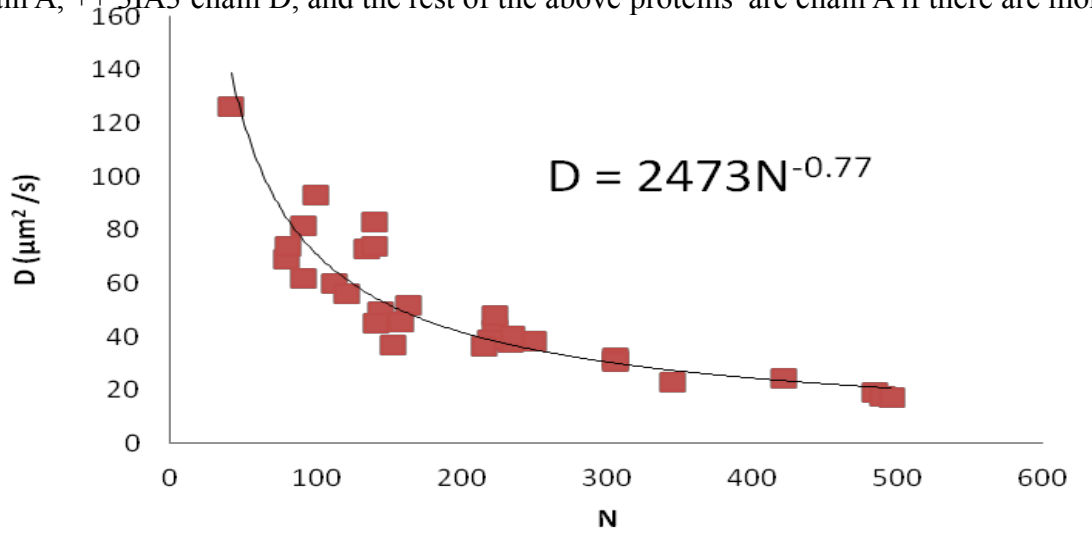
$$W(n,L) \approx R\sqrt{\frac{4\pi}{Ll}}\exp\left\{-\frac{R^2}{Ll}\left[2\pi n + \frac{kl}{2DRv} \times \left(1 - J_0(vL) - 2\sum_{m=1}^{p} J_{2m}(vL)\right)\right]^2\right\}$$

http://www.rcsb.org/pdb/explore/explore.do?structureId=4mbn

Table 1: Properties of alpha helical proteins with the simulated values of k/D and v.

Protein (PBD code)	Length (# of Residues)	% alpha (# of Residues)	# of helices	# of turns	ν (1/nm)	k/D (1/nm)
2K9J	42	57% (24)	1	7	1.30	490
3IA3-A+	91	71% (65)	3	18	1.30	490
2JUW	80	76% (61)	4	17	1.76	640
2I15	135	59% (80)	5	23	1.29	680
2HMZ	113	69% (79)	6	22	1.65	758
3IA3-D++	145	64% (93)	9	26	1.68	800
2MHB	141	73% (104)	9	29	1.79	1120
4MBN	153	80% (123)	11	34	1.93	1420
2O9D	234	71% (167)	12	47	1.36	1290
4E4V	485	63% (310)	31	86	1.48	2690
2YNS	490	64% (314)	33	87	1.53	2759
4BA3	496	64% (319)	34	89	1.55	2839

+ 3IA3 chain A, ++ 3IA3 chain D, and the rest of the above proteins are chain A if there are more than one chain.



Experimental Results

Simon Papadopoulos, Klaus D. Jurgens, and Gerolf Gros. Protein Diffusion in Living Skeletal Muscle Fibers: Dependence on Protein Size, Fiber Type, and Contraction. Biophysical Journal Volume 79 October 2000 2084–2094) and the references within.

TABLE 2 Protein diffusion in muscle and aqueous solution at 22°C

Injected protein	d (nm)	Soleus $D_{cell} \pm$ SE (n)	edl $D_{cell} \pm$ SE (n)	D_{H_2O}	Soleus D_{cell}/D_{H_2O}	edl D_{cell}/D_{H_2O}
Cytochrome *c*	3.1	13.0 ± 0.6 (27)	16.2 ± 0.6* (22)	120	~1/9	~1/7
Myoglobin	3.5	12.5 ± 1.3 (12)	18.7 ± 0.8* (12)	112	~1/9	~1/6
Hemoglobin	5.5	6.3 ± 0.5 (11)	6.2 ± 0.4 (13)	74	~1/12	~1/12
Catalase	10.2	2.6 ± 0.4 (12)	1.9 ± 0.2 (16)	43	~1/17	~1/23
Ferritin	12.2	0.6 ± 0.1 (10)	0.9 ± 0.1 (26)	38	~1/63	~1/42
Earthworm hemoglobin	30	0	0	13	→ 0	→ 0

d, hydrodynamic diameters; D_{cell}, sarcoplasmic D measured in this study. n, number of experiments, where each fiber was used once for microinjection. D is in 10^{-8} cm^2 s^{-1}. D values in dilute aqueous solution (D_{H_2O}) are from Ehrenberg (1957) for cytochrome *c*, Riveros-Moreno and Wittenberg (1972) for myoglobin and hemoglobin, Samejima et al. (1962) for catalase, and Gros et al. (1982) for ferritin and earthworm Hb.

* Significant difference ($p < 0.001$, Student's t-test for unpaired samples) between D in soleus and edl fibers.

Diffusion of Proteins within Cells

Stokes-Einstein Equation $$D = \frac{RT}{6\pi\eta R_h}$$

K. Dill, K, Ghosh and J. Schmit, in *Physical Limits of Cells and Proteomes*, PNAS 44 (108), 17876-17882 (2011).

Based on the analysis of 37,000 proteins:

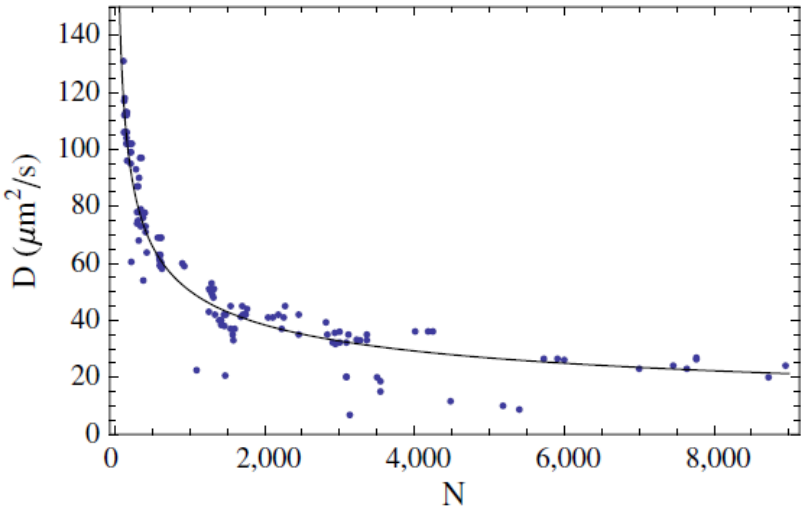
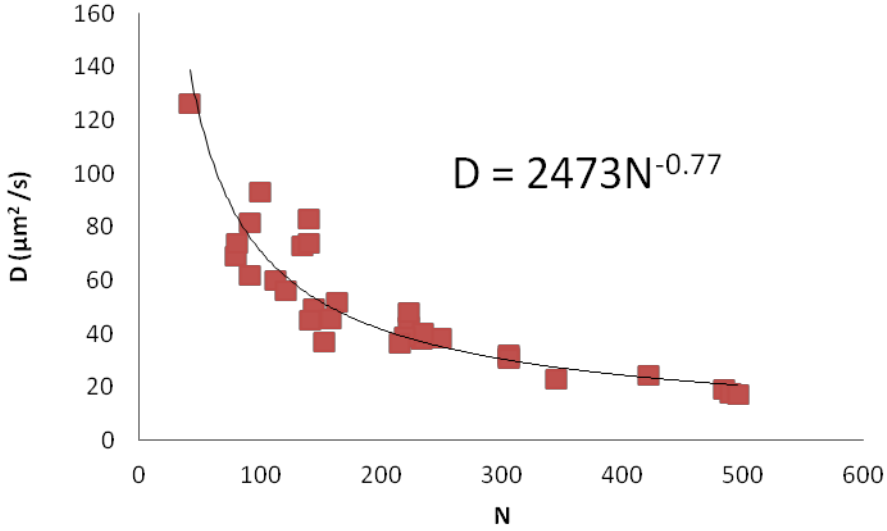


Fig. 5. Dependence of diffusion constant on chain length.



Our Results

Scaling Law:

$$R_h \approx 3.25N^{0.392}$$

$$N = M_w/110 \text{ Da}$$

Predicting k/D using the Empirical Formula

$$k/D \approx A + (\%alpha - \%B)A$$

$$W(n,L) \approx R\sqrt{\frac{4\pi}{Ll}}\exp\left\{-\frac{R^2}{Ll}\left[2\pi n + \frac{kl}{2DRv} \times \left(1 - J_0(vL) - 2\sum_{m=1}^{p} J_{2m}(vL)\right)\right]^2\right\}$$

Table 2: Properties of alpha helical proteins with the predicted values of k/D and v.

Protein (PBD code)	Length (# of Residues)	% alpha (# of Residues)	# of helices	# of turns	ν (1/nm)	k/D (1/nm)
2K9J	42	57% (24)	1	7	1.49	395.67
3IA3-A⁺	91	71% (65)	3	18	1.415	612.92
2JUW	80	76% (61)	4	17	1.83	724.06
2I15	135	59% (80)	5	23	1.29	682.80
2HMZ	113	69% (79)	6	22	1.67	832.47
3IA3-D⁺⁺	145	64% (93)	9	26	1.78	1012.46
2MHB	141	73% (104)	9	29	1.78	1104.50
4MBN	153	80% (123)	11	34	1.91	1346.46
2O9D	234	71% (167)	12	47	1.36	1319.61
4E4V	485	63% (310)	31	86	1.47	2599.31
2YNS	490	64% (314)	33	87	1.55	2772.50
4BA3	496	64% (319)	34	89	1.56	2845.84

Simulated Results vs Predicted

Protein (PBD code)	Length (# of Residues)	% alpha (# of Residues)	# of helices	# of turns	Simulated v (1/nm)	Simulated k/D (1/nm)	Predicted v(1/nm)	Predicted k/D(1/nm)
2K9J	42	57% (24)	1	7	1.30	490	1.49	395.62
3IA3-A^{+}	91	71% (65)	3	18	1.30	490	1.42	612.92
2JUW	80	76% (61)	4	17	1.76	640	1.83	724.06
2I15	135	59% (80)	5	23	1.29	680	1.29	682.80
2HMZ	113	69% (79)	6	22	1.65	758	1.67	832.47
3IA3-D^{++}	145	64% (93)	9	26	1.68	800	1.78	1012.46
2MHB	141	73% (104)	9	29	1.79	1120	1.78	1104.50
4MBN	153	80% (123)	11	34	1.93	1420	1.91	1346.46
2O9D	234	71% (167)	12	47	1.36	1290	1.36	1319.61
4E4V	485	63% (310)	31	86	1.48	2690	1.47	2599.31
2YNS	490	64% (314)	33	87	1.53	2759	1.55	2772.50
4BA3	496	64% (319)	34	89	1.55	2839	1.56	2845.84

Conclusions

1. Our method has successfully modeled the formation of alternating linear-helical segments of a protein by taking into account the interaction of amino acids with its aqueous environment.
2. And our predicted results agreed with the experimental observation that larger protein diffuse slower compared with the smaller one in aqueous solvent.
3. This method can also be used to solve an inverse problem.

Future Plans

1. Investigate different solvent effects.
2. Find the relation or correlation of v or k/D with the amino acid sequence of the folded proteins.
3. etc.

Acknowledgments

Thanks to Dr. Christopher C. Bernido, Dr. Maria Victoria C. Bernido, Dr. Jinky B. Bornales, and Dr. Henry P. Aringa for helpful discussions and guidance in doing this research. I also gratefully acknowledge the financial support from DOST-ASTHRDP.

THANK YOU!