

# Two Methods for Calculating the Size Distribution of Ferritin's Outer Diameter

Xinye Zhao<sup>1</sup>, Junhui Gao<sup>2\*</sup>

<sup>1</sup>Hwa Chong Institution, Singapore, Singapore

<sup>2</sup>American and European International Study Center, Wuxi, China

Email: \*jhgao68@163.com

**How to cite this paper:** Zhao, X.Y. and Gao, J.H. (2018) Two Methods for Calculating the Size Distribution of Ferritin's Outer Diameter. *Computational Molecular Bioscience*, 8, 115-121.

<https://doi.org/10.4236/cmb.2018.83006>

**Received:** June 14, 2018

**Accepted:** July 13, 2018

**Published:** July 16, 2018

Copyright © 2018 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

## Abstract

Ferritin is a kind of iron-storage protein widely found in animals and plants. The dynamic light scattering (Dynamic Light Scattering) method is used in the laboratory to determine the ferritin size. This paper presents two methods for calculating the outer diameter size distribution of ferritin, both of which assume that ferritin is approximately spherical. The ferritin data file was obtained from the PDB website and was calculated using the coordinate data of the amino acids to which the amino acids belong. The first method is based on the calculation of the sphere center; the second method is based on the method of the farthest distance atom pair. The outer diameter size distribution curves obtained by the two methods are basically consistent with the experimental methods. The paper also compares the calculation results and performance of the two methods. Both methods are versatile and can be used to calculate the size distribution of the globular proteins.

## Keywords

Ferritin, Outer Diameter, Density Distribution, Calculation Method

## 1. Background and Introduction

Ferritins have been found to exist almost ubiquitously in biological systems, regulating the storage and release of iron. Molecularly, ferritins are large globular multi-subunit proteins with a central cavity in which a hydrated ferric oxide is mineralised. When produced by multiple living organisms, ferritins vary widely in their primary structures (some share as low as 14% similarity in their amino acid sequences) but share essentially the same quaternary structure. Each hollow globular protein consists of 24 subunits and can store approximately 4500 iron ions. Typically it has internal and external diameters of about 8 and 12 nm, re-

spectively [1].

Clinically, ferritin maintains iron homeostasis and is associated with a wide range of physiologic and pathologic processes. It is predominantly utilized as a serum marker of total body iron stores, which serves a critical role in both diagnosis and management of related diseases such as coronary artery disease, malignancy, and poor outcomes following stem cell transplantation [2].

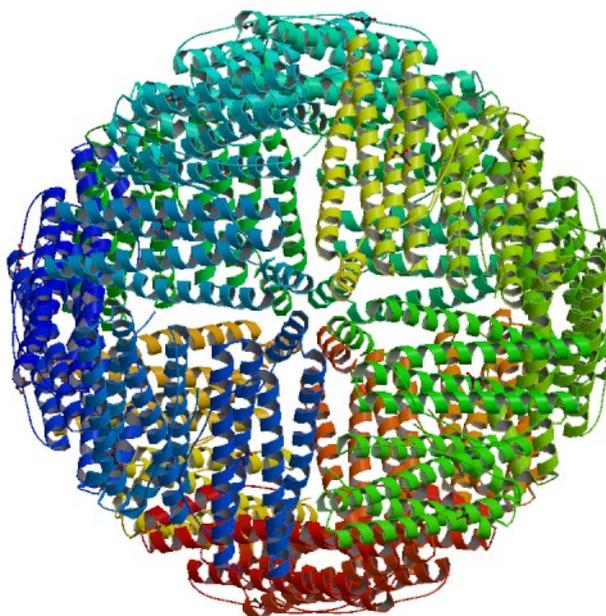
Furthermore, advances in biological nanoparticles have shown promising therapeutic applications. In particular, ferritin can self-organise in the nanometer range while meeting multiple criteria, such as biocompatibility, water solubility and high cellular uptake efficiency with minimal toxicity. They are also highly amenable through genetic and chemical modifications to suit different purposes. This makes ferritin a suitable molecular scaffold for the targeted delivery of the drugs and other molecules by conjugation with specific ligands or for imaging purposes using dyes [3] [4] [5].

The methods introduced in this paper can be applied to related research into ferritin and be further extended to other studies about globular proteins.

## 2. Materials and Methods

### 2.1. Data Sources

Structural information of various types of ferritins can be obtained from Protein Data Bank (PDB) [6]. **Figure 1** shows the biological assembly of ferritin 3kx9 consisting of over 30,000 atoms. It is a mutant of *Archaeoglobus fulgidus* ferritin (Afftn) formed by replacing the amino acid residues Lys-150 and Arg-151 by Alanine. It forms a typical 24-mer structure but with octahedral closed symmetry.



**Figure 1.** Biological assembly of ferritin 3kx9.

## 2.2. The Calculation Method Based on the Center of the Sphere

Outwardly, almost all ferritin is spherical. The calculation method of the center of the ball is simple. First, read the X, Y, Z coordinates of all the atoms in the PDB file which are labeled as ATOM, and then calculate the average value of all X, Y, Z, and the calculated average value is the coordinate value of the center of the ball. Because ferritin is not spherical, we can only calculate the density distribution of its outer diameter. By keeping the atoms larger than the center of the sphere, doubling each of these distances, we get a number of samples representing the size of the outer diameter.

## 2.3. Calculation Method Based on the Furthest Distance Atom Pair

We first describe this method in theory. The atom B, which is the furthest away from the atomic A, is calculated, and then the atom C, which is the furthest away from the atom B, is calculated. This C may be A, or it may not be A. In this way, iterative computation is carried out until no new atom appears.

Now we describe the actual calculation process: traversing all the atoms in the ferritin molecule, put in List 0, calculate the atom in the farthest distance from each atom in List, put in List 1, remove the repeating atoms in List 1, and get a new List 2. In List 2, calculate the atoms that are the furthest away from each atom, put them in List 3, remove the duplicate atoms in List 3, and get the new List 4. The distance between the atomic pairs with the furthest distance in List 4 is taken as a sample of the outer diameter.

## 2.4. The Combination of the Two Methods

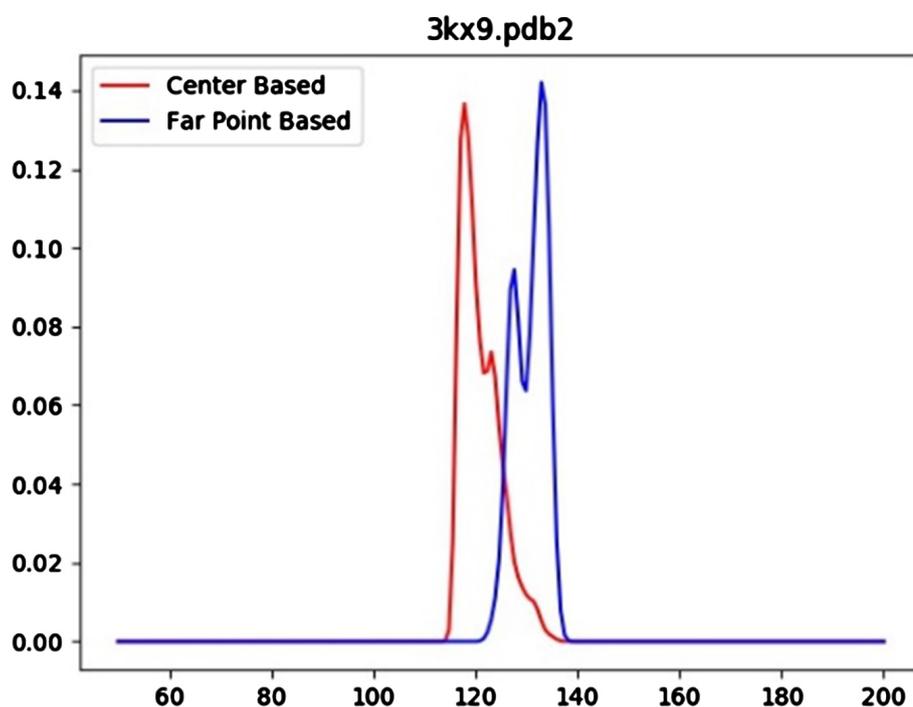
First, calculate the coordinates of the center of the sphere, and keep the atoms that are larger in the center of the sphere. This step is the same as the first calculation method. The next steps are the same as the second calculation method. For each atom, calculate the atom that is the furthest from this atom and calculate the distance between them as one of the approximate diameters. Calculate all such atom pairs.

## 2.5. Calculation of Outer Diameter Distribution

After obtaining a batch of data on the outside diameter, we can program to get the distribution of the outside diameter size. The specific approach is to use Python programming, using `gaussian_kde` and its PDF density distribution function in the `scipy.stats` package, calculate the density distribution of the outer diameter size, and draw the density distribution map with `matplotlib`.

## 3. Results

Taking 3kx9 as an example, the outer diameter density distribution curve calculated by the two methods is shown in **Figure 2**. Among them, the red curve is based on the calculation method of the sphere center, and the blue curve is based on the calculation method of the farthest distance atom pair.



**Figure 2.** Size distribution of ferritin 3kx9.

In the figure above, the abscissa is the size of the outside diameter, the unit is nanometer, and the ordinate is the corresponding density value. The area under the entire curve is 1. In the red curve, the area with the highest density is at 11, 12 and 12, and the blue curve is the area with the highest density at 13, 14 and 14 points.

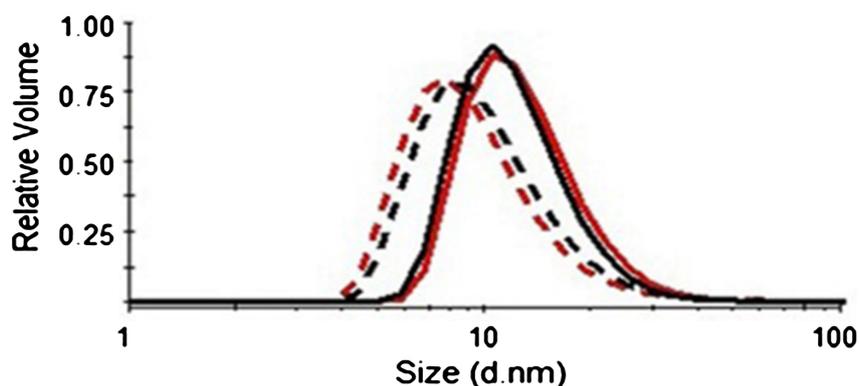
## 4. Discussions

### 4.1. Comparison of Outer Diameter Calculation Methods and Experimental Methods

Dynamic Light Scattering (DLS) is normally used in lab to determine the size distribution profile of ferritin and other nanoparticles in suspension. It has become a common characterization method in nanotechnology due to its accuracy, speediness and reproducibility.

When light hits small particles, the light scatters in all directions (Rayleigh scattering). Due to small molecules in solutions undergoing Brownian motion, the distance between the scatterers in the solution is constantly changing and thus the scattering intensity fluctuates over time. According to Stokes-Einstein equation, faster dynamics due to smaller particles results in more rapid fluctuation of scattering intensity. The size distribution is thus generated by machine through analysis of this correlation (**Figure 3**).

Compared to mathematical method, DLS involves preparation of pure protein solution of suitable concentration, certain instruments, and accurate operation of experiments. On the other hand, calculations are not necessarily performed in lab and the data obtained is very close to laboratory results.



**Figure 3.** Size distribution profile obtained through DLS.

#### 4.2. Comparison of Calculation Results between Outer Diameter Calculation Methods

As described above, when calculating the outside diameter, method 1 reserves atoms that are larger than the center of the sphere. We call it the method based on the center of the sphere. The second method does not need to calculate the center of the sphere. We call it the method based on the remote point. What is the difference between these two methods?

Taking ferritin 5v5k as an example, after the method acquires the center of the sphere, we first calculate the point farthest from the center of the sphere, assuming that the distance is  $Dis$ , and then keep those atoms that are greater than  $0.8 Dis$ ,  $0.85 Dis$ , and  $0.9 Dis$  from the center of the sphere, respectively. Then the density distribution calculation is performed, and the obtained graphs are the red curves in the sub-graphs (a), (b), and (c) of **Figure 4**, respectively. The blue curve is based on the method of the longest distance atom pair.

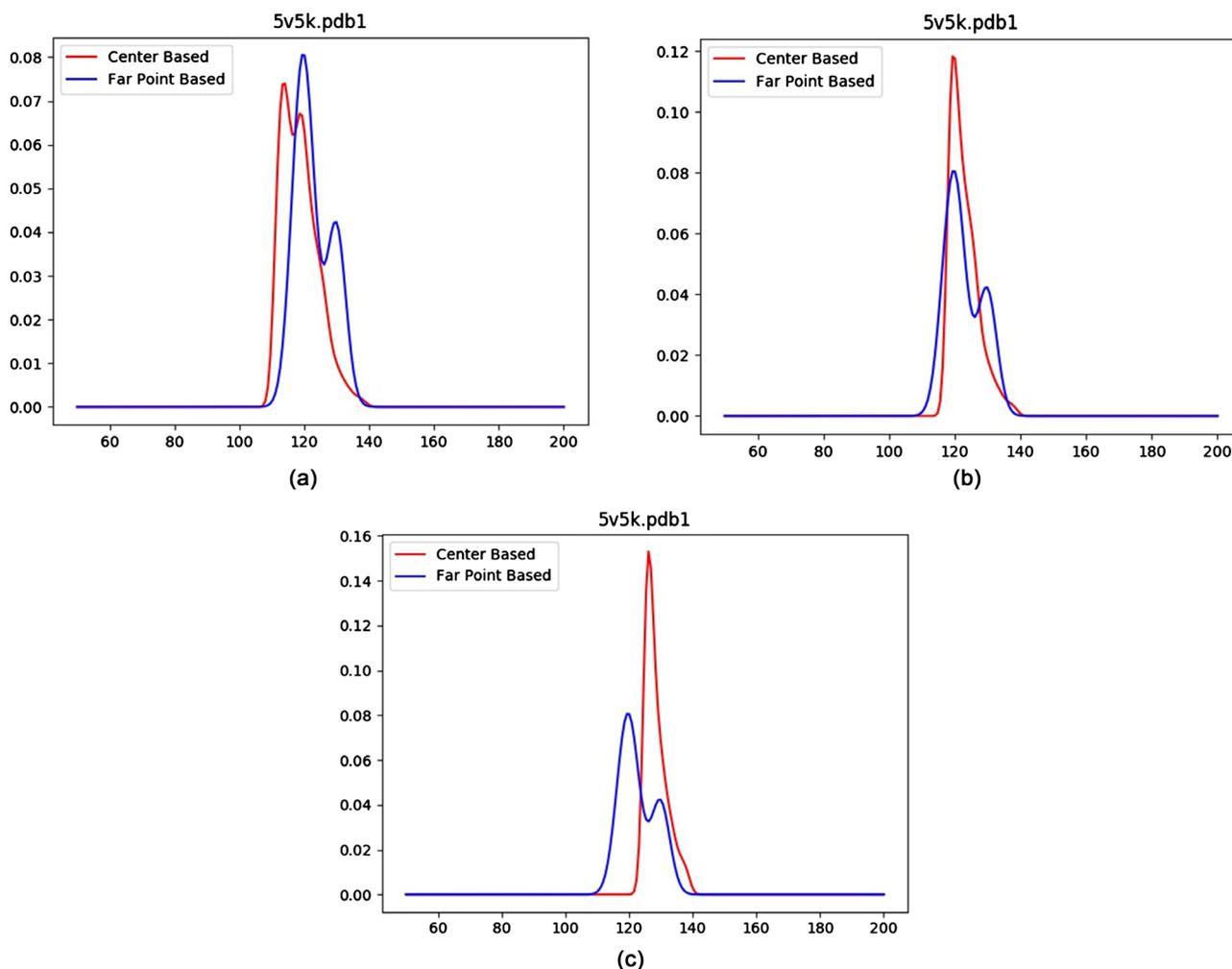
From the figure below, we can see that in **Figure 4(b)**, the results obtained by the two methods are the closest. In **Figure 4(a)**, the outer diameter value based on the center of the sphere is smaller than that based on the far point, and in **Figure 4(c)**, the outer diameter value based on the center of the sphere is greater than that based on the far point.

#### 4.3. Comparison of Calculation Performance between Outer Diameter Calculation Methods

Now let's analyze and compare the performance of the two calculation methods.

The first method is based on the calculation method of the center of the sphere. First, to find the coordinates of the center of the circle, we need to do  $N$  additions and one division; then, we need to calculate the distance between each atom and the center of the sphere, and we need to perform  $N$  calculations; finally, we need to set the critical value (such as  $0.8 Dis$ ). For  $N$  comparisons, the samples needed to calculate the outer diameter distribution were obtained. The computational complexity is proportional to  $N$ .

The second method is based on the calculation of the longest distance atom pair. First, we need to calculate the distance between all atom pairs, and we



**Figure 4.** Comparison between calculation methods (ferritin number: 5v5k). (a) 0.8 of the maximum; (b) 0.85 of the maximum; (c) 0.9 of the maximum.

need to do  $N^2$  operations; then, remove the repeated atoms, and the calculation amount is also proportional to  $N^2$ ; finally, the distance operation between the furthest atoms is also proportional to  $N^2$ . So the computational complexity is proportional to  $N^2$ .

Taking 3kx9 as an example, the calculation method based on the center of the sphere takes about 30 seconds, and the calculation method based on the longest distance atom pair takes about half an hour, and the two methods require a time difference of two orders of magnitude. The corresponding computing environment is CPUi5-7200U, memory 16 G, MS Windows 10.

We believe that the calculation method based on the center of the sphere is simple and rapid, but the threshold value is not certain; the calculation method based on the longest distance atom pair is more accurate, but the calculation is more complicated and the time required for calculation is obviously increased.

## 5. Conclusions

In this paper, we propose several methods for calculating the outer diameter size

distribution of ferritin. Since almost all the outer surfaces of ferritin are similar to the spherical surface, we first proposed a calculation method based on the center of the sphere. However, we also proposed a more elaborate calculation method based on the longest distance atom pair, and a combination of the two methods.

We use 3kx9 as an example to compare the calculated results with experimental data. The density curves of the two are basically consistent. In addition, we use 5v5k as an example to compare the results of the two calculation methods; we also use 3kx9 as an example to compare the performance of the two calculation methods.

These methods are versatile and can be used to calculate the outer diameter size distribution of globular proteins.

## References

- [1] Liu, B.J. and Zhang, W.B. (2009) Structure, Function and Expression Manipulation of the Iron Protein. *Feed Industry*, **4**, 42-47.
- [2] Knovich, M.A., *et al.* (2009) Ferritin for the Clinician. *Blood Reviews*, **23**, 95-104. <https://doi.org/10.1016/j.blre.2008.08.001>
- [3] Lee, L.A. and Wang, Q. (2006) Adaptations of Nanoscale Viruses and Other Protein Cages for Medical Applications. *Nanomedicine*, **2**, 137-149. <https://doi.org/10.1016/j.nano.2006.07.009>
- [4] Kwon, K.C., *et al.* (2014) Proteinticle/Gold Core/Shell Nanoparticles for Targeted Cancer Therapy without Nanotoxicity. *Advanced Materials*, **26**, 6436-6441. <https://doi.org/10.1002/adma.201401499>
- [5] Lee, B.-R., *et al.* (2016) Engineered Human Ferritin Nanoparticles for Direct Delivery of Tumor Antigens to Lymph Node and Cancer Immunotherapy. *Scientific Reports*, **6**, 35182. <https://doi.org/10.1038/srep35182>
- [6] <http://www.rcsb.org>