Fidabio

LINKING STRUCTURE AND FUNCTION

USE PDB DATA AS THE BASELINE FOR YOUR FIDABIO CHARACTERIZATION

VERSION 2.1 Author: Emil G. P. Stender, Senior Application Scientist, FIDA Biosystems Aps.

- Optimal data reliability
- In-solution correlation with X-ray, Cryo-EM structural data and predicted structures (AlphaFold)
- End-to-end quantitative workflow
- Integrated in the Fidabio Assay Design Software

INTRODUCTION

In biophysics, the detection of molecular size changes is a strong, quantitative source of information about the structure of a protein and its complexes. However, correlating a protein's structure to size measures, like Hydrodynamic radius (Rh), requires consideration of structural features of the protein and its assemblies based on either experimental structures or predicted ones by e.g. AlphaFold [1]. To establish a strong and reliable size baseline, we have developed the Fidabio PDB (Protein Data Bank) Correlator. In the examples below we use the Fidabio PDB Correlator to predict the in solution Rh of bovine β-lactoglobulin, human serum albumin, bovine serum albumin and chicken lysozyme (Figure 2). Subsequently, using Fida 1 equipped with a UV fluorescence detector, the Rh of label-free proteins is experimentally determined and compared with the predicted Rh values. The workflow is illustrated in Figure 1. In this way, we can make accurate correlations between structural data and in-solution Rh measurements. Such an approach would enable validation of structures predicted by AlphaFold [1]. The Fidabio PDB Correlator is also applicable when you are working with fluorescently labelled proteins and using a fluorescence detection at colored wavelengths.

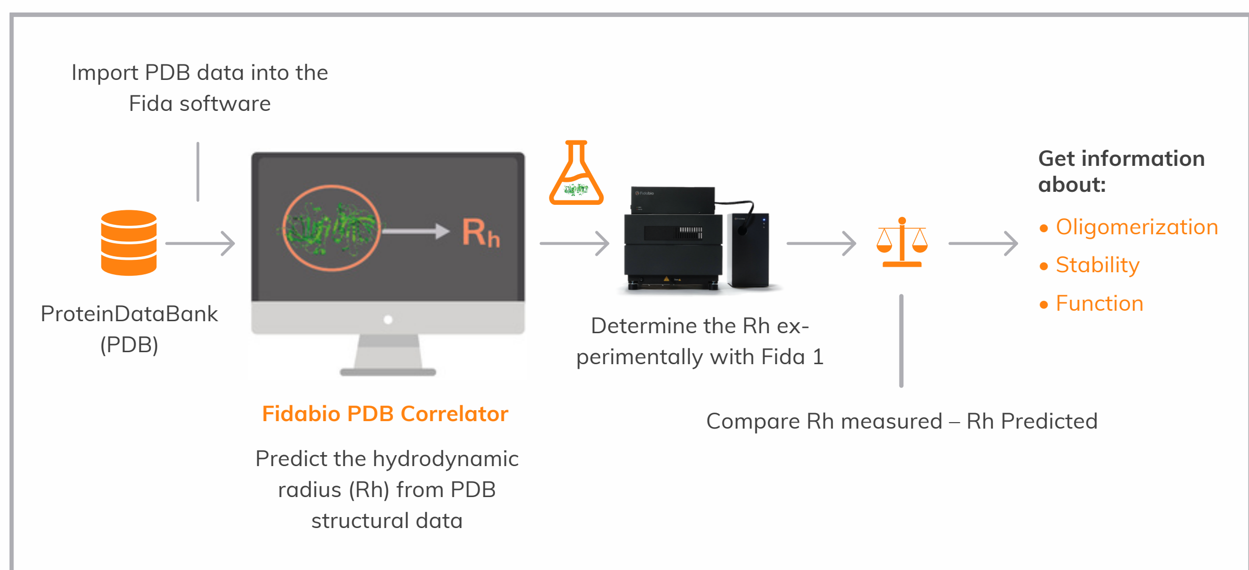

Figure 1. Fidabio PDB Correlator workflow.

FIDA MEASURING PRINCIPLE

Flow Induced Dispersion Analysis is a capillary-based microfluidic method, exploiting that flow rate in the center of the capillary is faster than the one at the edges of the capillary. The resulting radial concentration gradients at the front and the tail of the dispersion results in diffusion of your chosen indicator, which enables a "first principle" biophysical measurement of size. The FIDA measurement of size is broadly applicable for studying biomolecular stability, interactions, etc. To learn more, visit fidabio.com.

MATERIAL & METHODS

Fida 1 instrument with Fidabio 280 nm LED fluorescence detection for sizing experiments (Fida Biosystems ApS). Fidabio standard capillary (i.d.: 75 µm, LT: 100 cm, Leff: 84 cm). Bovine β-lactoglobulin (BLG), Human Serum Albumin (HSA), Bovine Serum Albumin (BSA) and Chicken lysozyme (LYZ) were all dissolved in buffer (1 x PBS). The buffer was used as Analyte. Sample analysis was performed by filling the capillary with the analyte, followed by an injection of preincubated indicator, which was mobilized towards the detector with analyte at 400 mbar. LYZ was measured on a capillary coated with HS reagent (Fida Biosystems ApS). The Fidabio PDB Correlator was used to correlate hydrodynamic radius and molecular weight. Crystal structures were acquired from the Protein Data Bank (https://www.rcsb.org/). Removal of crystallographic water, co-crystalized molecules and additional PDB file cleanup and visualization was done using PyMOL 2.5.0 (Schrödinger LCC). Data analysis was done using the Fida Software. All standard deviation is given as the experimental error of nine replicates.

RESULTS

Label free Rh – Structural correlation using the Fidabio PDB Correlator The Fida 1 instrument configured with a Fidabio UV fluorescence detector provides a label free and absolute measurement of hydrodynamic radius (Rh). It was used to measure the size of Bovine β-lactoglobulin (BLG), Human Serum Albumin (HSA), Bovine Serum Albumin (BSA) and Chicken lysozyme (LYZ). BLG is a homodimer at neutral pH, the rest are well characterized monomeric proteins [2-5]. Using the Fidabio PDB Correlator, the PDB data are loaded Into the Fida Software which subsequently returns the predicted mass and Rh. For large monomeric proteins, the predicted mass is within 2 % of the theoretical value (Table 1 and Appendix). For small proteins, however, it is recommended that PDB files are cleaned up, when they contain structures of co-crystallized molecules. Figure 2 illustrates the process of cleaning up the structure using PyMOL.

Protein	PDBCode	Apparent Rh (nm)	Predicted Rh (nm)	Predicted Mw (kDa)	Theoretical Mw
LYZ	1IEE	1.91 ± 0.01	1,97	14,51	14,31
BLG	1BEB	2,77 ± 0.06	2,85	36,21	36,58
HSA	1AO6	3.81 ± 0.04	3,5	65,76	66,47
BSA	4F5S	3,53 ± 0.04	3,52	66,77	66,43

Table 1. Predicted and measured (Apparent) values of hydrodynamic radius (Rh), predicted and theoretical molecular weight (Mw) and PDB codes. The predicted value is generated by loading the cleaned up PDB file into the Fidabio PDB correlator. The experimental values are measured by Fida 1, and the error is the standard deviation of nine replicates. Also see Appendix.

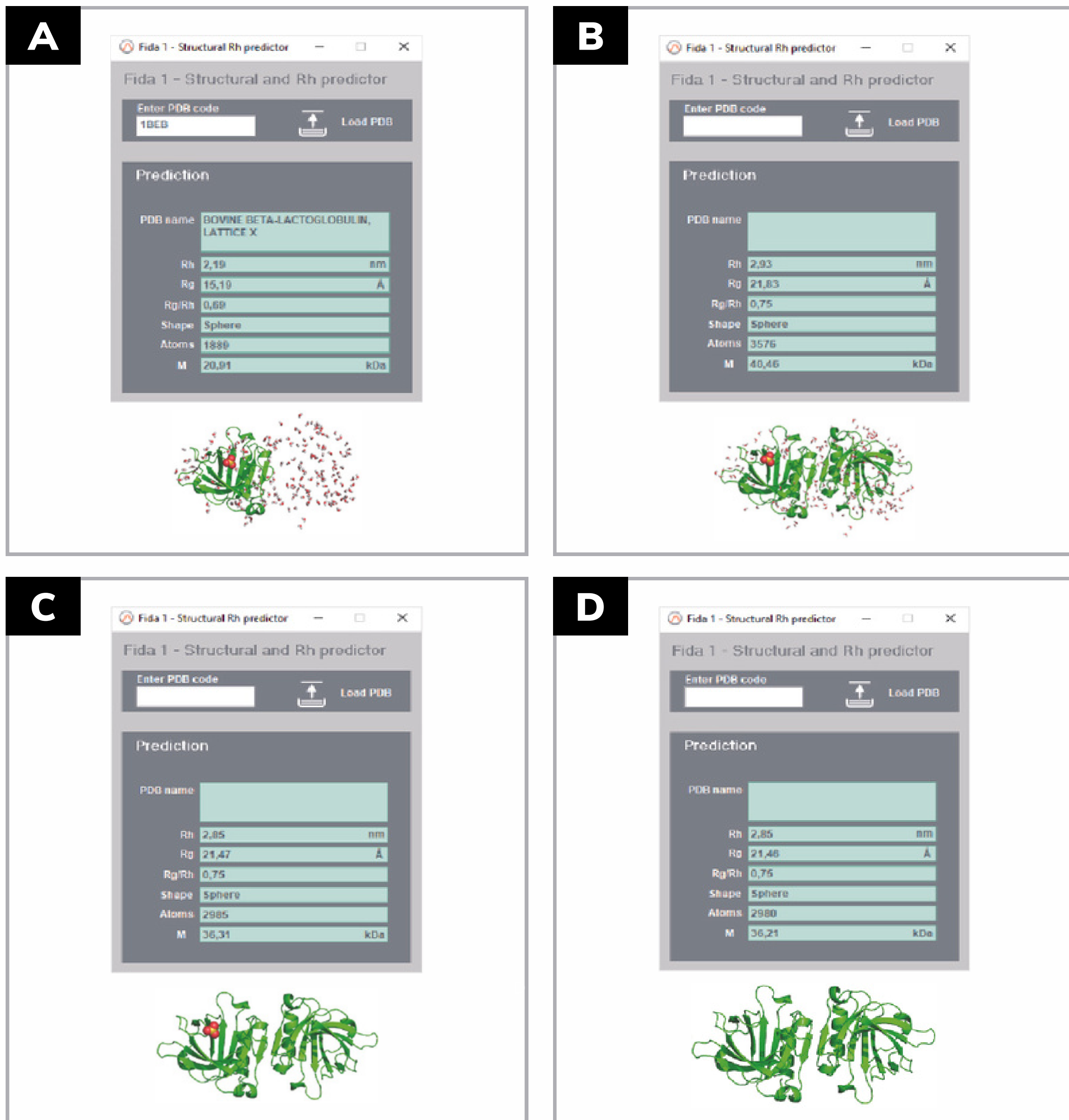

Figure 2. Fidabio PDB Correlator (above), visualized PDB files used for correlation (down).
(A) PDB file for BLG (1BEB) loaded directly from PBD using the Fidabio PDB Correlator
(B) PDB file generated using the PyMOL Molecular Visualization System
(C) PDB file with crystallographic waters removed
(D) Removal of co-crystalized molecules and additional cleanup for correct oligomeric state. In this particular case no Rh difference occurs with the removal of co-crystalized molecules (Rh=2,85 nm in both steps C and D). However, this is an important process for many crystal structures that have rather large co-crystalized molecules which can influence the Rh predictions.

Without clean-up the apparent Rh of LYZ, BSA and HSA is within 10% of the experimental Rh. The correlation can be improved by removing co-crystalized molecules, leading to 0.3% accuracy (theoretical values compared to the experimental data) (Table 1). BLG is predicted to have an Rh of 2.19 nm, but the experimental value is 28% off (Appendix-Table 1)

Pointing to an in-solution structure different from a monomer. Cleaning up the file and using a homo-dimer structure yields an Rh of 2,85 nm, which is within 2,8% of the experimental value of 2.77 ± 0.06 nm. This indicates that BLG is a homodimer, which has also been extensively shown in literature [2].

CONCLUSIONS

The data show that oligomeric states and hydrodynamic radius of a given protein can be identified label free in solution using very small amounts of sample (~40 nL/run at 0.1 g/L). The Fidabio PDB Correlator provides an accurate correlation of protein hydrodynamic radius (<9% error) and can be used to for correlating hydrodynamic radius to structural data or AlphaFold structural predictions.

REFERENCES

1. Jumper, J et al. Highly accurate protein structure prediction with AlphaFold., 2021, Nature.

2. Brownlow S. et al.; Bovine beta-lactoglobulin at 1.8 A resolution--still an enigmatic lipocalin., 1997, Structure; Vol 5, p. 481-495.

3. Sugio S. et al.; Crystal structure of human serum albumin at 2.5 A resolution., 1999; Protein Engineering Vol 12: 439–446.

4. Bujacz, A.; Structures of bovine, equine and leporine serum albumin., 2012; Acta Crystallography D, Vol 68: 1278-1289

5. Vaney, M.C. et al.; High-resolution structure (1.33 A) of a HEW lysozyme tetragonal crystal grown in the APCF apparatus. Data and structural comparison with a crystal grown under microgravity from SpaceHab-01 mission. 1996 Acta Crystallography D Biological Crystallography, Vol 52: 505-517..

APPENDIX

Table A1. Predicted (4th column) and measured values (3rd column) of Rh. -H2O: removal of crystallographic waters. -CC: removal of co-crystalized molecules. AC: additional cleanup, modifying files for correct oligomerization. Orange: >10% difference between predicted and measured values. Green: <10% difference between predicted and measured values (accepted difference).

Protein	PDB	App Rh (nm)	Pred Rh (nm)	Pred Mw (kDa)	Theor Mw	- H_2O	-CC	AC
BLG	1BEB	2,77 ± 0.06	2,19	20,91	18,29*	N	N	N
HSA	1AO6	3.81 ± 0.04	3,5	65,81	66,47	N	N	N
BSA	4F5S	3,53 ± 0.04	3,55	67,74	66,43	N	N	N
LYZ	193L	1.91 ± 0.01	2,05	16,99	14,31	N	N	N
1st Clean-up								
BLG	1BEB	2,77 ± 0.06	2,93	36,31	18,29*	Y	N	N
HSA	1AO6	3.81 ± 0.04	4,54	131,52	66,47	Y	N	N
BSA	4F5S	3,53 ± 0.04	4,74	133,61	66,43	Y	N	N
LYZ	1IEE	1.91 ± 0.01	1,97	14,57	14,31	Y	N	N
2nd Clean-up								
BLG	1BEB	2,77 ± 0.06	2,85	36,31	18,29*	Y	Y	N
HSA	1AO6	3.81 ± 0.04	4,54	131,52	66,47	Y	Y	N
BSA	4F5S	3,53 ± 0.04	4,74	133,46	66,43	Y	Y	N
LYZ	1IEE	1.91 ± 0.01	1,97	14,51	14,31	Y	Y	N
3rd Clean-up								
BLG	1BEB	2,77 ± 0.06	2,85	36,21	36,58*	Y	Y	Y
HSA	1AO6	3.81 ± 0.04	3,5	65,76	66,47	Y	Y	Y
BSA	4F5S	3,53 ± 0.04	3,52	66,77	66,43	Y	Y	Y

*BLG is a homodimer

Fida Biosystems ApS | Generatorvej 6 | 2860 Søborg | Tel.: +45 53 72 78 70 | www.fidabio.com