Biophysical Journal, Volume *120*

Supplemental Information

Conformational Ensembles of Antibodies Determine Their Hydrophobicity

Franz Waibl, Monica L. Fernández-Quintero, Anna S. Kamenik, Johannes Kraml, Florian Hofer, Hubert Kettenberger, Guy Georges, and Klaus R. Liedl

Conformational Ensembles of Antibodies Determine Their Hydrophobicity

Effect of Conformation and Protonation

Franz Waibl[1] (FW), Monica L. Fernández-Quintero[1] (MLF), Anna S. Kamenik[1] (ASK), Johannes Kraml[1] (JK), Florian Hofer[1] (FH), Hubert Kettenberger[2] (HK), Guy Georges[2] (GG), Klaus R. Liedl[1*] (KRL)
1: Institute of General, Inorganic and Theoretical Chemistry, and Center for Molecular Biosciences Innsbruck (CMBI), University of Innsbruck, Innsbruck, Austria
2: Roche Pharma Research and Early Development, Large Molecule Research, Roche Innovation Center Munich, Penzberg, Germany

Supporting Information

1. Combining ΔG(solv) of cluster representatives

Given a set of conformations (in this case, cluster representatives), for each of which the free energy of solvation ΔGsolv has been estimated using GIST, the free energy of solvation of the ensemble is given as

$$\Delta G_{solv} = k_B T \ln \int e^{\frac{\Delta G_{solv}(q)}{k_B T}} p(q) dq$$

Where k_B is Boltzmann's constant, T is the temperature, q denotes the cartesian coordinates of the respective conformation, $p(q)$ is the probability of the conformation (or, in the case of cluster representatives, the population of the respective cluster), and $\Delta G_{solv}(q)$ is the hydration free energy of the conformation.
This equation places a high weight on conformations with a positive free energy of hydration. Conceptually, this stems from the fact that ΔG_{solv} is not purely a property of the simulated (solution) state, but a difference between two different states, the reference state being defined by an isolated solute in gas phase and all the water being transferred to the bulk. In the present publication, we estimate cluster probabilities in the liquid state. Since conformations with a high (not very negative) ΔG_{solv} are most favorable in the gas phase, they are the most likely to participate in the equilibrium between solution and gas states, and therefore need to be given a higher weight. If our cluster probabilities had been determined in the gas state, we would need to use a different equation.

We begin our derivation by writing ΔG_{solv} in terms of partition sums:

$$\frac{\Delta G_{solv}}{k_B T} = -\ln \frac{Z^{sol}}{Z^u Z^v} = \ln \frac{Z^u Z^v}{Z^{sol}} \quad (1)$$

Where Z^{sol}, Z^v, and Z^u represent the partition sums of the solution, the pure solvent, and the solute in gas phase, respectively.

Writing out the terms of the partition sums gives:

$$Z^u = \int e^{-\frac{H^{uu}(q_u)}{k_B T}} dq_u \quad (2)$$

$$Z^v = \int e^{-\frac{H^{vv}(q_v)}{k_B T}} dq_v \quad (3)$$

$$Z^{sol} = \int\int e^{-\frac{H^{uu}(q_u)+H^{vv}(q_v)+H^{uv}(q_u,q_v)}{k_B T}} dq_v dq_u \quad (4)$$

Where q_u denotes the solute degrees of freedom, qv the solvent degrees of freedom, and H^{uu}, H^{vv}, and H^{uv} denote the solute-solute, solvent-solvent, and solute-solvent energy contributions.

The probability of finding a solute conformation $p(q_u)$ in a molecular dynamics simulation is:

$$p(q_u) = \frac{\int e^{-\frac{H^{uu}(q_u)+H^{vv}(q_v)+H^{uv}(q_u,q_v)}{k_B T}} dq_v}{Z^{sol}} \quad (5)$$

Assuming that H^{uu} is independent of the solvent degrees of freedom, we write this as:

$$p(q_u) = \frac{\int e^{-\frac{H^{vv}(q_v)+H^{uv}(q_u,q_v)}{k_B T}} dq_v \, e^{-\frac{H^{uu}(q_u)}{k_B T}}}{Z^{sol}} \quad (6)$$

From each GIST calculation, we obtain an estimate of the hydration free energy of a single solute conformation, i.e., the change in free energy upon placing a solute *in this conformation* from the gas phase into the solution phase. We write this as:

$$e^{-\frac{\Delta G_{solv}(q_u)}{k_B T}} = \frac{\int e^{-\frac{H^{vv}(q_v)+H^{uv}(q_u,q_v)}{k_B T}} dq_v}{Z^v} \quad (7)$$

Here, the partition sum of the solute Z^v is practically introduced by referencing our GIST results to a pure solvent box.

Inserting the righthand side of Eq. 7 into Eq. 6 gives us:

$$p(q_u) = \frac{e^{-\frac{\Delta G_{solv}(q_u)}{k_B T}} e^{-\frac{H^{uu}(q_u)}{k_B T}} Z^v}{Z^{sol}} \quad (8)$$

To arrive at an expression of the total ΔGsolv, we first note that we do not have an expression for the (gas phase) solute ensemble Zu. We therefore seek to express Eq. 2 in terms of other quantities. We do so by separating the term involving Huu in Eq. 8 and inserting it to Eq. 2. We arrive at:

$$Z^u = \frac{\int p(q_u) e^{\frac{\Delta G_{solv}(q_u)}{k_B T}} Z^{sol}}{Z^v} dq_u \quad (9)$$

Note the positive sign before $\Delta G_{solv}(q_u)$ that arises from taking its reciprocal value. Z^u and Z^{sol} are already integrals and therefore independent of q_u and can be separated.

Inserting Eq. 9 in Eq. 1 results in:

$$\frac{\Delta G_{solv}}{k_B T} = \ln \int p(q_u) e^{\frac{\Delta G_{solv}(q_u)}{k_B T}} dq_u \quad (10)$$

2. Used PDB codes

SI Table 1: PDB codes used as starting structures for the molecular dynamics (both GaMD and CpHMD) simulations.

Antibody	PDB
Adalimumab	4NYL
alemtuzumab	1BEY
Anifrolumab	4QXG
Atezolizumab	5X8L
bapineuzumab	4OJF
Basiliximab	1MIM
Belimumab	5Y9J
Bevacizumab	1BJ1
Bevacizumab	6BFT
Bimagrumab	5NGV
Briakinumab	5N2K
Canakinumab	4G6J
Certolizumab	5WUX
Cetuximab	1YY8
Crenezumab	5VZY
Daclizumab	3NFP
Drozitumab	4OD2
Eculizumab	5I5K
Efalizumab	3EOA
epratuzumab	5VL3
gantenerumab	5CSZ
gevokizumab	4G6M
Golimumab	5YOY
Ibalizumab	3O2D
Infliximab	4G3Y
Ipilimumab	5TRU
ixekizumab	6NOV
lebrikizumab	4I77
matuzumab	3C09
motavizumab	3IXT
muromonab	1SY6
natalizumab	4IRZ
necitumumab	6B3S
nivolumab	5WT9
ofatumumab	3GIZ
olokizumab	4CNI
omalizumab	2XA8
onartuzumab	4K3J
panitumumab	5SX5
pembrolizumab	5GGS
pertuzumab	1S78
pinatuzumab	6AND
ponezumab	3U0T
ranibizumab	1CZ8
rituximab	6VJA
sifalimumab	4YPG
tanezumab	4EDW
tralokinumab	5L6Y
trastuzumab	1N8Z
tremelimumab	5GGV
urelumab	6MHR
ustekinumab	3HMW

3. Experimental details

In the present work, we use experimental data from a dataset by Jain et al. (1) Here, we shortly describe the experimental conditions of the assays that we mention in the main text. For more detailed information, the reader is referred to the original works.

Hydrophobic Interaction Chromatography:
The methodology described in reference (2) was used. 5 µg (1 mg/mL) of sample was analyzed using a Sepax Proteomix HIC butyl-NP5 column over 20 min. A linear gradient from mobile phase A (1.8M ammonium sulfate and 0.1M sodium phosphate at pH 6.5) to mobile phase B (0.1 M sodium phosphate at pH 6.5) was employed with a flow rate of 1 mL/min. The UV adsorbance was monitored at 280 nm.

Standup Monolayer Adsorption Chromatography:
The methodology described in reference (3) was used. 2 µg of sample was analyzed using a Zenix SEC-300 column. A 0.15M sodium phosphate buffer at pH 7.0 was used as mobile phase, with a flow rate of 0.35 mL/min.

Cross Interaction Chromatography:
The methodology described in reference (4) was used. 5 µg of sample was analyzed using a 1-mL HiTrap column with ~30mg of human serum polyclonal antibodies coupled to it. A PBS buffer was used as mobile phase, with a flow rate of 0.1 mL/min.

4. Effect of the ROC cutoff and the cutoff for $\Delta G_{unfavorable}$

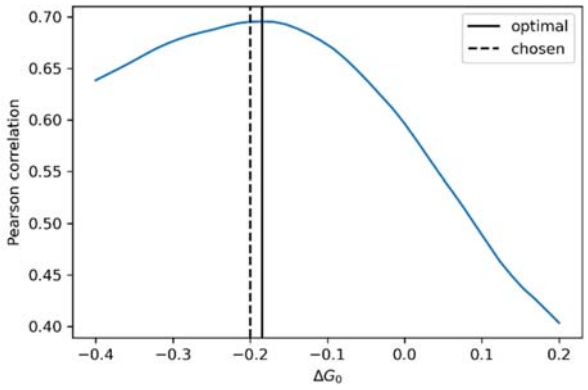

Figure 1: Pearson correlation between experimental HIC retention times of the PDB set and $\Delta G_{unfavorable}$, depending on the x offset in the cutoff function (corresponds to ΔG_0 in Equation 8).The optimal value (the x position with the highest correlation), as well as the rounded value used in this study, are shown as vertical lines. The input data was the ΔG_{solv} values obtained after GaMD sampling of the PDB set. Computing $\Delta G_{unfavorable}$ using the chosen ΔG_0 leads to the data shown in Figure 1E.

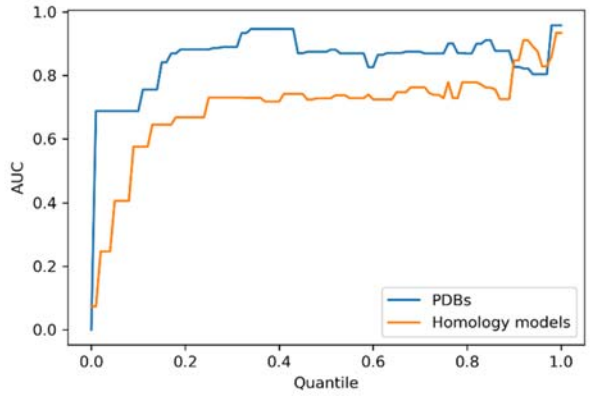

Figure 2: AUC value for the separation between strongly-binding and weakly-binding antibodies in the PDB set (blue) and in the homology models set (orange), depending on the quantile at which the border between strongly-binding and weakly-binding was drawn. The separation was performed based on the $\Delta G_{unfavorable}$ values shown in Figure 1E.

1. Jain, T., T. Sun, S. Durand, A. Hall, N. Houston, J. Nett, B. Sharkey, B. Bobrowicz, I. Caffry, Y. Yu, Y. Cao, H. Lynaugh, M. Brown, H. Baruah, L. Gray, E. Krauland, Y. Xu, M. Vasquez, and K. Wittrup. 2017. Biophysical properties of the clinical-stage antibody landscape. Proceedings of the National Academy of Sciences of the United States of America 114(5):944-949. Article.

2. Estep, P., I. Caffry, Y. Yu, T. W. Sun, Y. Cao, H. Lynaugh, T. Jain, M. Vasquez, P. M. Tessier, and Y. D. Xu. 2015. An alternative assay to hydrophobic interaction chromatography for high-throughput characterization of monoclonal antibodies. Mabs 7(3):553-561.

3. Kohli, N., N. Jain, M. L. Geddie, M. Razlog, L. H. Xu, and A. A. Lugovskoy. 2015. A novel screening method to assess developability of antibody-like molecules. Mabs 7(4):752-758.

4. Jacobs, S. A., S. J. Wu, Y. Q. Feng, D. Bethea, and K. T. O'Neil. 2010. Cross-Interaction Chromatography: A Rapid Method to Identify Highly Soluble Monoclonal Antibody Candidates. Pharm Res-Dordr 27(1):65-71.