

In Silico Process Development via Computational Modeling: Insights into Molecular Biophysics to Advance and Improve Biologics Purification

Francis Insaiddoo¹, John Welsh¹, Karol Lacki², Thomas Linden³, and David Roush¹

¹Merck & Co Inc

²Avitide, Inc

³Takeda Bio Development Center Ltd

May 26, 2020

Abstract

The goal of this research is to leverage computational molecular biophysics to guide process development, reduce experimental burden and focus purification activities on feasible targets. Here, we distill a complex separation problem (e.g. chromatographic retention of monoclonal antibodies) into a tangible model (ligand/protein complex), which is computationally feasible while preserving enough detail (atomistic level for interaction site) to support industrially relevant separation challenges. Computational docking, coupled with molecular dynamics simulation, produces results that are directionally consistent with chromatography for proteins (mAb). This approach is generalizable and can be applied to a range of ligands (AEX, CEX, and Mixed Mode). A detailed model of the chromatography base matrix (agarose) was constructed to obtain a biophysical understanding of potential protein/base matrix interactions. The base matrix was then modified *in silico* with ligands over a range of ligand densities representative of commercial chromatography resins to generate an agarose/ligand complex. A generic approach was developed to model the impact of avidity and ligand density on mAb/ligand interaction. The results revealed that increasing ligand density mask contributions of base matrix binding. Increasing the number of ligands that can interact with mAb results in more favorable free energy of binding or ΔG (more negative) with a limited incremental increase in ΔG by increasing N (number of ligands per agarose cluster) above three. Additionally, for protein/ligand interactions at each binding site, not all ligands contribute equally to the binding affinities or interaction energies and a redistribution of binding interactions/energies occur as N increases. These observations yield insights into the impact of avidity on retention (macroscopic affinity measurement via k'). The generic approach described in this manuscript can be leveraged to inform resin selection and design as well as targeted ligand selection/purification development in a rational manner.

Introduction

Technological advancement in cell culture has led to an increase in therapeutic protein titers over the last decade. This advancement has shifted the attention towards downstream processes as the bottleneck in the manufacture and production of biopharmaceuticals (Guiochon & Beaver; Hanke & Ottens, 2014). For biologics, antibodies are extremely complex and provide an almost unlimited design space to engineer binding with target molecules (antigens) (Maier & Labute). This results in a diverse array of biophysical properties and a challenge for separation. The relative ease with which new antibody candidates are generated and low toxicity of their degradation products make this class of molecules an excellent therapeutic agent (Tiller et al., 2008). Yet, the structural complexity of these molecules and diverse biophysical properties (including conformational flexibility) pose a significant challenge to the selection of candidates and development of appropriate purification modalities. An approach is proposed that can be applied to determine feasibility of separation of specific degraded products (e.g. oxidation, deamidation sites) *via in silico* screening of ligands

to a targeted location. This *in silico* screening approach is analogous to the experimental developability (Lorenz et al., 2014; Yang et al., 2013) approach and hence can be viewed as *in silico* developability.

The process of developing therapeutic antibodies require the generation of many variants. The number of molecules generated make empirical biophysical characterization of each molecule difficult, resource intensive, and time consuming (Jarasch et al., 2015; Shah et al., 2015). In-depth biophysical characterization of antibodies requires a significant amount of protein mass to determine stability profiles, purification, and formulation conditions. In addition, this process can be very time consuming even with the advent of high-throughput screening (HTS) approaches (Petroff et al., 2016). Thus, the ability to use *in silico* computational modeling will not only save time, but also optimize candidate selection for the best therapeutic candidate and accelerate the path from discovery to first-in-human clinical trials. Computational approaches may also be able to assess mechanisms and pathways that are not readily accessible experimentally and hence provide additional insights into molecular development.

Recently published work has demonstrated the utility of computational algorithms to isolate antibody complementary determining region (CDR) loops and associated structural features that infer biological and biophysical properties when bound to receptors and antigens (Morea, Lesk, & Tramontano, 2000). Further, computational modeling algorithms to generate 3D structure of any novel antibody based on currently available protein databank (PDB) structures are becoming increasingly routine (Morea, Leplae, & Tramontano, 1998; Morea, Tramontano, Rustici, Chothia, & Lesk, 1997). *In silico* docking studies have yielded results that are consistent with relative chromatographic retention (k'), and surface properties for a range of biologics (Insaiddo et al., 2015). The next step is to combine these advances in antibody homology modeling and *in silico* computational docking to understand chromatographic separation and select appropriate ligands for bioprocess development for a specific separation.

In this paper, we investigate the biophysical principles governing protein-chromatographic ligand interactions. We show *in silico* docking studies in combination with molecular dynamics simulations can inform bio-process development and biologics purification. Specifically, we connect antibody biophysical properties to ligand selection and impact of ligand density on retention and selectivity.

In silico models of agarose fibers functionalized with chromatographic ligands showed increased binding affinity as a function of increasing number of interacting ligands (N). The observed increase is consistent with experimental correlation between resin ligand density and k' . These results have implications not only for bioprocess development but also for resin design and our understanding of the principal criteria for biologics design for therapeutic efficacy and successful downstream development.

Currently, the selection of lead candidates from a range of highly potent efficacious antibodies relies on limited downstream process development data (Jarasch et al., 2015; Lorenz et al., 2014). Even with high throughput screening, it would still require significant resources to fully screen and develop the operational design space for the process. This paper maps the interaction between chromatographic ligands and monoclonal antibodies at an atomic level using a general, extendable computational approach. This *in silico* approach to process development allows us to discern the impact of subtle differences on candidate selection and process development.

Results and Discussion

Antibody Structure and Surface Properties

Molecular operating environment (MOE) from Chemical Computing Group (Montreal, Canada) (Maier & Labute), was used to generate homology models of three antibodies (referred to here as mAb-1, mAb-2, and mAb-3). The isoelectric point (pI), percent surface area distribution of hydrophobic and electrostatic surface properties (positive and negative patches) were calculated. Patch analysis of these molecules indicate that the total patches of the molecules are relatively constant (Table 1). Yet, the patch distribution for the complementary determining region (CDR) varied amongst the three molecules (Table 1). This aligns with the fact that the constant regions of the molecules are 98% identical. The variability of the target antigen

and recognition site accounts for the diversity in the CDR region, even though it forms a small fraction of the entire molecule (approximately 10%). Thus, any variations between the molecules resulting from the CDR are diminished when averaged over the entire molecule (e.g. pI). This effectively masks the specific contribution of the CDR to chromatographic retention and separation if their contribution is not decoupled from the entire molecule (Figure 1).

Ligand Density and Resin Surface Design

Previous studies demonstrated a strong correlation between chromatographic ligand docking binding affinities (surrogate for ΔG) and chromatographic separation (k') (Insaïdoo et al., 2015). While these results may hold true for a smaller biological molecule, the complexity of possible ligand interaction sites on an antibody makes such an approach challenging (Hanke & Ottens, 2014). To address the complexity in chromatographic separation, a molecular model of agarose ligand complex was formulated to capture a broader range of ligand densities and interaction sites. The ligands were attached to the resin surface with a separation distance of 10-14 Å, ranging from one attached ligand to six ligands per resin area. From our internal calculation, these agarose models span the range of low to medium and high-density resin that is commercially available (Figure 2A).

Protein-Chromatographic Ligand Interaction

For chromatographic separation, ligands are expected to interact with different regions of the molecules based on their surface properties and inherit affinity for a given region. To answer the question of preferred protein-ligand interactions that lead to separation, different docking studies were performed. Multiple single ligand docking assessments were performed to the entire antibody surface, the Fab region, the CDR region, and the Fc region. The contact sites for protein/ligand and associated affinities were ranked based on interaction energy. Figure 1A shows that the region with the highest (more negative ΔG) protein-ligand interaction energies for the top three binding sites differed for the three antibodies. There is a correlation between *in silico* binding affinities (ΔG) and experimentally determined protein retention. Previous studies have demonstrated that a single post translational modifications (PTM) in the CDR region not only affect antibody potency but can have significant effect on chromatographic separation (Kern, Mende, Denefeld, Sackewitz, & Chelius, 2014). Thus, chromatographic ligand binding to CDR region can elucidate potential chromatographic separation (Figure 1B).

Effect of Ligand Density on Chromatographic Retention

Despite the correlation between single chromatographic ligand binding affinities and experimental data, chromatographic separation is much more complex than a single ligand-protein interaction. A single ligand binding utilizes both the head group and ligand backbone that attaches to the resin (Figure 2 B and C). Yet, on a chromatographic scale, ligand density, separation distance from resin to head group, multiple interaction effect (potential avidity and binding optimization), and compensatory protein surface properties could all be critical contributions to protein retention and separation (Hanke & Ottens, 2014). The flexibility of the spacer arm could also contribute to interaction of the ligand and the protein. To understand this mechanism and demonstrate the effect of the aforementioned factors on protein binding, we utilized an agarose surface model which was functionalized with different chromatographic ligands. This mimic a patch of chromatographic resin and how it interacts with proteins. In addition, the ligands were attached to the resin surface with a separation distance of 10-14 Å, ranging from one attached ligand to six ligands per section of resin (supplementary figure 1). These agarose models span the range of low to medium and high-density resins that are commercially available (Figure 2A). The entire agarose-ligand complex could scan the surface of the antibody using the sub-region classification outlined above to identify potential binding sites and cumulative binding affinities – effectively a combination of avidity and binding/interaction energies. A control binding study was done using only the agarose resin to investigate the contribution of the agarose to the calculated protein-ligand binding affinities as agarose has been reported to support protein purification previously in the literature (Crone, 1974).

Figure 3 shows a plot of the change in binding affinity over the un-functionalized agarose. For each chro-

matographic modality, their *in silico* binding affinities (ΔG) to the antibody-ligand complex increased as the number of attached ligands per agarose increased. Further, the different chromatographic ligands had different binding affinities with Capto MMC showing the strongest interactions (Figure 3).

Individual ligand Contribution to Increased Binding with Increasing Ligand Density

We next examined the structural conformation adopted by the agarose-ligand complex to determine the compensatory binding conformations adopted by both the protein and the ligand to achieve the desired interaction. Table 2 indicates that as the number of ligands per agarose increases, the contribution of each individual ligand to the overall binding score decreases. This is because contributions from the backbone of the ligand are sterically hindered from binding to the antibody and increasing the interaction energy. At the highest ligand density (six ligands per agarose), not all the ligands contact the protein equally or at all. In addition, the ligands that made contact differed in their binding score due to differences in the protein interaction site. Thus, once a strong interaction is made by one ligand, the remaining ligands need to interact with proximal sites and hence limit optimal interaction for the remaining ligands. For the low ligand density agarose, the docking can adjust to incorporate the ligand backbone or agarose in the final binding conformation and affinity. Figure 4A, shows the binding site in the context of the overall antibody molecule. As the number of ligands attached to the agarose molecule increases, the contribution of the agarose and ligand backbone to the total binding affinity decreases (Figure 2B). In doing so, the ligand backbone cannot contribute to the total and individual ligand binding affinities. Further, beyond two ligand per agarose, there is a sub-optimal utilization of the attached ligand. This means that in making optimal contacts to the protein one side of the ligand complex is directed away from the protein and cannot make contact, thus excluding one ligand each time.

Ligand Density Docking Score Correlate Directionally to Experimental Retention studies

In a packed bed, the number of possible protein-ligand interaction sites will exceed one. Yet, for each given patch of resin with a specific ligand density, there is an optimal number of possible interactions and associated interaction energy. Data from the computational docking studies suggest that there is an optimal number of possible interactions at each ligand density. To address this assumption, the interaction energy for the agarose ligand complex was compared to experimentally determined retention data. From the retention studies, the capacity factor (k') for mAb-1 was determined for the same ligand (Capto MMC) and at different ligand densities (low, medium, and high). Figure 5 indicates that there is a directional correlation between the experimentally determined values of k' and the binding affinities determined *in silico*. As the ligand density increased the binding affinities increased and the k' also increased. Figure 5A, shows that the calculated increase in *in silico* binding affinities for the agarose-ligand complex are greater than the baseline agarose interaction to the protein and higher than a single ligand docking.

To examine the effect of local perturbations and solvent components on binding affinity, we performed a droplet molecular dynamics simulation of the ligand binding site with a solvation shell ratio of 10 Å. Using droplet molecular dynamics simulation allowed for the exploration of dynamics near and around the docking site. This method was preferred over a solvation box because of the size and complexity of simulating a fully solvated antibody system. Constraints were applied to residues outside the solvation droplet to dampen out dynamics as we move outside the water droplet. Figure 6 shows that during the course of a 30 ns unconstrained atomistic molecular dynamics simulation (Insaiddo et al., 2011), the six-ligand agarose remained bound to the antibody with an average affinity of 18.26 (± 1.05) kcal/mol. This value is very close to the original induced fit binding score for Capto MMC bound to mAb1 (19.08 kcal/mol), indicating that the binding interactions retained consistent.

Conclusion

In conclusion, our data shows that computational modeling can be used to gain meaningful insight into structural properties of an antibody that is critical for process development. This high through-put structural information can be used as an additional parameter to select antibodies that will lead to successful development downstream. While the overall structural features are important, a better mechanistic understanding

of protein-chromatographic resin is gained by coupling *in silico* docking studies to molecular dynamics and experimental data. We demonstrated using this approach that different regions of an antibody contribute at different levels to the overall column retention and selectivity. This is because different high affinity binding sites were observed base on the ligand density of the agarose-ligand complex. Further, we addressed the impact of ligand density on overall protein-resin binding and how the attached ligands to the resin compensate to maximize its interaction with the protein. Higher overall binding affinities are achieved at higher ligand densities by a cumulative effect of lower individual binding affinities (avidity) (Figure 4). This computational approach to evaluating protein ligand interaction has broader implications for biologics development (from lead candidate selection through purification to formulation) and commercial chromatographic resin design and head groups selection and optimization.

Materials and Methods

Antibody Modeling

Homology modeling of antibodies used in this study was generated by using BLAST algorithm (National Center for Biotechnology Information) to compare our sequence against the non-redundant protein data bank (PDB) for Fab and Fc structures that have a high level of sequence similarity with the target molecule. High similarity structures were found for the framework region and CDR loop regions separately. Once pdb files are identified that best represent the molecule of interest, MOE software was used to graft the structures together and mutated to the new targeted sequence. The final structure was energy minimized. The final structure phi and psi angles were plotted to ensure that they fall within what is allowed for proteins (reference Ramachandran plot) (Ramachandran, Ramakrishnan, & Sasisekharan, 1963). The final Fab and Fc structure were superimposed on a full antibody structure, the linker sequence was connected, and disulfide bonds were formed. The overall structure is energy minimized to reduce backbone and side chain strain using steepest decent algorithm.

Docking Studies

Docking and protein structure preparation was performed on Molecular Operating Environment (MOE 2015.10 and 2016.08) from Chemical Computing Group, Canada. Protein-Ligand docking was designed such that the ligands are placed in the site using the Triangle Matcher method, and ranked with the London ΔG (Wright, Anderson, Shadnia, Durst, & Katzenellenbogen, 2013) scoring function. The initial poses and score were then refined for energy minimization in the binding pocket and rescored using GBVI/WSA ΔG (Corbeil, Williams, & Labute, 2012) scoring function under the Amber12:EHT forcefield set. The target protein and ligands were allowed to sample alternative conformations to allow for induced-fit binding of protein-ligand complex. The highest docking score conformation was used in molecular dynamics (MD) simulation and a final binding affinity reported from the ensemble average using GBVI/WSA ΔG calculation. Accessible solvent area (ASA) was calculated from a single molecular structure using a lattice and a 0-potential contour of the van der Waals potential. Using the median of each potential value as the iso-contour cutoff, the surface was then extended to adjacent grid triangles to smoothen the patch transition (Jorgensen, Maxwell, & TiradoRives, 1996; Wildman & Crippen, 1999). Docking scores were determined for antibody molecules 1, 2, and 3 all IgG1 (mAb1 (pI = 7.59), mAb2 (pI = 7.96), and mAb3 (pI = 8.38)) and protonated at pH 5.5.

Retention Mapping Assessment

Retention mapping was conducted with mAb-1 for all resins using 1 mL HiTrap columns with a GE Healthcare ÄKTA avant 25 system. Protein was loaded with a 4 min residence time, and protein concentration was monitored through A280 measurements. The feed material (purified antibody) was used as the absorbance at 100% breakthrough.

Molecular Dynamics Simulation

Molecular dynamics (MD) simulations were performed and analyzed with the Molecular Operating Environment (MOE) (Chemical Computing Group, Canada) and NAMD 2.7 suite

[<http://www.ks.uiuc.edu/Research/namd>] (Phillips et al., 2005). For molecular dynamics simulation with the antibody-ligand complexes, starting coordinates were from homology modeled structures. Hydrogen atoms were added to the starting structures using the LigX protonation tool of the MOE 2016.08 suite. Using the solvation algorithm in MOE, a 10 Å region around the protein-ligand complex was immersed in a TIP3P water droplet such that protein atoms outside the droplet were fixed to dampen long range dynamic effect. Sodium cations were added for electroneutrality. Additional Sodium and Chloride ions were added to a final concentration of 0.15 M. This resulted in systems consisting of 18,780 atoms for the mAb1, 17,788 atoms for mAb2, and 16,469 atoms for mAb3. Dynamics simulations were then performed using NAMD, with parameters from the Amber10:EHT forcefield set. Equilibration consisted of 15,000 steps of conjugate gradient energy minimization, followed by 50 ps of MD with restraints applied to the protein to equilibrate the water. The systems were then warmed to 300 K over a total of 18 ns with a 50 K increment and equilibration at each temperature step. This was followed by unrestrained production run of 30 ns, which is enough for the determination of empirically relevant binding affinities (Salvalaglio, Zamolo, Busini, Moscatelli, & Cavallotti, 2009). The SHAKE algorithm was used, allowing a 2 fs time-step. Long-range electrostatics were treated via the particle mesh Ewald method. Trajectory analysis was carried out with the MOE analysis tool and in-house scripts.

Reference

- Corbeil, C. R., Williams, C. I., & Labute, P. (2012). Variability in docking success rates due to dataset preparation. *J Comput Aided Mol Des*, *26* (6), 775-786. doi:10.1007/s10822-012-9570-1
- Crone, H. D. (1974). Ion-Exclusion Effects on the Chromatography of Acetylcholinesterase and other Proteins on Agarose Columns at Low Ionic Strength. *Journal of Chromatography*, *92* (1974).
- Guiochon, G., & Beaver, L. A. Separation science is the key to successful biopharmaceuticals. *J Chromatogr A*, *1218* (49), 8836-8858. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=21982447
- Hanke, A. T., & Ottens, M. (2014). Purifying biopharmaceuticals: knowledge-based chromatographic process development. *Trends Biotechnol*, *32* (4), 210-220. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=24630477
- Insaidoo, F. K., Borbulevych, O. Y., Hossain, M., Santhanagopalan, S. M., Baxter, T. K., & Baker, B. M. (2011). Loss of T cell antigen recognition arising from changes in peptide and major histocompatibility complex protein flexibility: implications for vaccine design. *J Biol Chem*, *286* (46), 40163-40173. doi:10.1074/jbc.M111.283564
- Insaidoo, F. K., Rauscher, M. A., Smithline, S. J., Kaarsholm, N. C., Feuston, B. P., Ortigosa, A. D., . . . Roush, D. J. (2015). Targeted purification development enabled by computational biophysical modeling. *Biotechnol Prog*, *31* (1), 154-164. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=25482184
- Jarasch, A., Koll, H., Regula, J. T., Bader, M., Papadimitriou, A., & Kettenberger, H. (2015). Developability assessment during the selection of novel therapeutic antibodies. *J Pharm Sci*, *104* (6), 1885-1898. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=25821140
- Jorgensen, W. L., Maxwell, D. S., & TiradoRives, J. (1996). Development and testing of the OPLS all-atom force field on conformational energetics and properties of organic liquids. *Journal of the American Chemical Society*, *118* (45), 11225-11236. doi:10.1021/Ja9621760
- Kern, W., Mende, R., Denefeld, B., Sackewitz, M., & Chelius, D. (2014). Ion-pair reversed-phase

high performance liquid chromatography method for the quantification of isoaspartic acid in a monoclonal antibody. *J Chromatogr B Analyt Technol Biomed Life Sci*, 955-956 , 26-33. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=24631807

Lorenz, T. P., Fiaux, J. P., Heitmann, D. P., Gupta, K. P., Kocher, H. P. P., Knopf, H. P., & Hartmann, S. P. (2014). Developability Assessment of Biologics by Integrated Biologics Profiling. *American Pharmaceutical Review* .

Maier, J. K., & Labute, P. Assessment of fully automated antibody homology modeling protocols in molecular operating environment. *Proteins*, 82 (8), 1599-1610. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=24715627

Morea, V., Leplae, R., & Tramontano, A. (1998). Protein structure prediction and design. *Biotechnol Annu Rev*, 4 , 177-214. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9890141

Morea, V., Lesk, A. M., & Tramontano, A. (2000). Antibody modeling: implications for engineering and design. *Methods*, 20 (3), 267-279. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=10694450

Morea, V., Tramontano, A., Rustici, M., Chothia, C., & Lesk, A. M. (1997). Antibody structure, prediction and redesign. *Biophys Chem*, 68 (1-3), 9-16. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=9468606

Petroff, M. G., Bao, H., Welsh, J. P., van Beuningen-de Vaan, M., Pollard, J. M., Roush, D. J., . . . Linden, T. O. (2016). High throughput chromatography strategies for potential use in the formal process characterization of a monoclonal antibody. *Biotechnol Bioeng*, 113 (6), 1273-1283. doi:10.1002/bit.25901

Phillips, J. C., Braun, R., Wang, W., Gumbart, J., Tajkhorshid, E., Villa, E., . . . Schulten, K. (2005). Scalable molecular dynamics with NAMD. *Journal of Computational Chemistry*, 26 (16), 1781-1802. doi:DOI 10.1002/Jcc.20289

Ramachandran, G. N., Ramakrishnan, C., & Sasisekharan, V. (1963). Stereochemistry of polypeptide chain configurations. *J Mol Biol*, 7 , 95-99. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=13990617

Salvalaglio, M., Zamolo, L., Busini, V., Moscatelli, D., & Cavallotti, C. (2009). Molecular modeling of Protein A affinity chromatography. *Journal of Chromatography A*, 1216 (50), 8678-8686. doi:DOI 10.1016/j.chroma.2009.04.035

Shah, K. A., Clark, J. J., Goods, B. A., Politano, T. J., Mozdierz, N. J., Zimmisky, R. M., . . . Love, K. R. (2015). Automated pipeline for rapid production and screening of HIV-specific monoclonal antibodies using pichia pastoris. *Biotechnol Bioeng*, 112 (12), 2624-2629. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=26032261

Tiller, T., Meffre, E., Yurasov, S., Tsuiji, M., Nussenzweig, M. C., & Wardemann, H. (2008). Efficient generation of monoclonal antibodies from single human B cells by single cell RT-PCR and expression vector cloning. *J Immunol Methods*, 329 (1-2), 112-124. Retrieved from http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=17996249

Wildman, S. A., & Crippen, G. M. (1999). Prediction of physicochemical parameters by atomic contributions. *Journal of Chemical Information and Computer Sciences*, 39 (5), 868-873. doi:10.1021/Ci9903071

Wright, J. S., Anderson, J. M., Shadnia, H., Durst, T., & Katzenellenbogen, J. A. (2013). Experimental versus predicted affinities for ligand binding to estrogen receptor: iterative selection and rescoring of docked poses systematically improves the correlation. *J Comput Aided Mol Des*, 27 (8), 707-721. doi:10.1007/s10822-013-9670-6

Yang, X., Xu, W., Dukleska, S., Benchaar, S., Mengisen, S., Antochshuk, V., . . . Ambrogelly, A. (2013). Developability studies before initiation of process development: improving manufacturability of monoclonal antibodies. *MAbs*, 5 (5), 787-794. doi:10.4161/mabs.25269

FIGURE LEGENDS:

FIGURE 1 Surface representation of antibodies indicating different binding sites for chromatographic ligands. (A) Shows the highest affinity sites for Capto MMC to three different antibodies. The typical interaction energy values for the top three binding sites range from -5 to -7 kcal/mol. The region with the highest interaction energy differs as a function of molecule. (B) The CDR region showed the highest interaction for Capto MMC ligand outside of the top three binding sites for each molecule. The interaction energies ranged from -4 to -5 kcal/mol. In addition, the interaction site, interacting residues, and surface area explored are diverse as should be expected for a hypervariable region.

FIGURE 2 Structure of agarose base matrix and attached ligands.(A) A surface representation of the agarose base matrix and the attached ligands. The ligands are spaced evenly around the resin from a single ligand attachment to six ligands per resin. (B) A cartoon representation of the attachment site of a chromatographic ligand to resin. The head group of the ligand is separated from the base by a linker (backbone). (C) Structure of the ligands used in this study. CaptoTM adhere and CaptoTM MMC are produced by GE Healthcare, USA and CEX (HyperDTM F) is produced by PALL Life Science, USA.

FIGURE 3 Comparison of the interaction energy between different agarose-ligand densities against mAb-1 compared to un-functionalized agarose. The mAb-ligand docking was performed at pH 5.5 and compared between modalities. All agarose-ligand complex had an improved interaction energy over un-functionalized agarose (base matrix). The general trend is that, increasing the number of ligands increase the interaction energy.

FIGURE 4 The interaction site of functionalized agarose (agarose-ligand complex) bound to mAb-1. (A) A close-up view of the interaction site between agarose-ligand complex bound to antibody (mAb-1). There is a range of interaction energies that is attributed to the head group of each attached ligand making interactions with specific sites instead of the entire ligand (head group and backbone) (B) Highest affinity interaction site of agarose-ligand complex in the context of the entire antibody. A solvation droplet was designed to cover the area in blue to simulate any changes in interaction energy as a function of molecular dynamics.

FIGURE 5 Comparison of experimental binding affinities with *in silico* interaction energies for mAb-1. (A) Interaction energy of Free Ligand (MMC) (Capto MMC ligand un-attached to a base matrix), agarose, and agarose with Capto MMC attached (Agarose MMC also called agarose-ligand complex). As the number of Capto MMC ligands attached to the agarose increase, the overall interaction energy of the functionalized agarose increases. (B) Chromatographic retention (k') determined from retention mapping assessment with increasing ligand density shows a similar trend as docking interaction energies.

FIGURE 6 Molecular dynamics simulation of mAb-1 ligand complex compared to docking score. The interaction energies from the molecular simulation was slightly lower than the docking score but remained constant for the entire length of the simulation.

Hosted file

Tables.docx available at <https://authorea.com/users/325488/articles/453482-in-silico-process-development-via-computational-modeling-insights-into-molecular-biophysics-to-advance-and-improve-biologics-purification>

Hosted file

Figures.docx available at <https://authorea.com/users/325488/articles/453482-in-silico-process-development-via-computational-modeling-insights-into-molecular-biophysics-to-advance-and-improve-biologics-purification>