frontiers | Frontiers in Molecular Biosciences

Check for updates

Comparison of hydrophobicity scales for predicting biophysical properties of antibodies

Franz Waibl[1], Monica L. Fernández-Quintero[1], Florian S. Wedl[1], Hubert Kettenberger[2], Guy Georges[2] and Klaus R. Liedl[1]*

[1]Department of General, Inorganic and Theoretical Chemistry, University of Innsbruck, Innsbruck, Austria, [2]Large Molecule Research, Roche Pharma Research and Early Development, Roche Innovation Center Munich, Penzberg, Germany

While antibody-based therapeutics have grown to be one of the major classes of novel medicines, some antibody development candidates face significant challenges regarding expression levels, solubility, as well as stability and aggregation, under physiological and storage conditions. A major determinant of those properties is surface hydrophobicity, which promotes unspecific interactions and has repeatedly proven problematic in the development of novel antibody-based drugs. Multiple computational methods have been devised for in-silico prediction of antibody hydrophobicity, often using hydrophobicity scales to assign values to each amino acid. Those approaches are usually validated by their ability to rank potential therapeutic antibodies in terms of their experimental hydrophobicity. However, there is significant diversity both in the hydrophobicity scales and in the experimental methods, and consequently in the performance of in-silico methods to predict experimental results. In this work, we investigate hydrophobicity of monoclonal antibodies using hydrophobicity scales. We implement several scoring schemes based on the solvent-accessibility and the assigned hydrophobicity values, and compare the different scores and scales based on their ability to predict retention times from hydrophobic interaction chromatography. We provide an overview of the strengths and weaknesses of several commonly employed hydrophobicity scales, thereby improving the understanding of hydrophobicity in antibody development. Furthermore, we test several datasets, both publicly available and proprietary, and find that the diversity of the dataset affects the performance of hydrophobicity scores. We expect that this work will provide valuable guidelines for the optimization of biophysical properties in future drug discovery campaigns.

1 Introduction

In recent years, antibodies and related formats have emerged as a major class of novel therapeutic proteins (Strohl and Strohl, 2012; Kaplon et al., 2020; Kaplon and Reichert, 2021; Mullard, 2021; Kaplon et al., 2022) with more than 120 approved therapeutic antibodies up to now (Raybould et al., 2020). Antibodies are characterized by their unique binding properties and their broad applicability. In particular, therapeutic antibodies have revolutionized the treatment of various diseases, such as cancer (Scott et al., 2012) and autoimmune diseases (Chan and Carter, 2010). Therapeutic antibodies are usually also the fastest answer to new medical challenges and viral threats, which has become apparent in the current SARS-CoV-2 pandemic (Chvatal-Medina et al., 2021).

In the development of therapeutic antibodies, it is important to avoid problems regarding the stability, aggregation, solubility, and immunogenicity. One driving force of those problems is the tendency of hydrophobic, i.e., apolar, regions on the surface to form interactions with each other. This phenomenon is called the hydrophobic effect and is driven by the release and entropy increase of water from the hydrophobic surface into bulk solution (Southall et al., 2002).

A plethora of methods have been used to investigate hydrophobicity of antibodies in-silico. A summary of those methods will be given in the subsection titled "In-silico methods." However, it is unclear which methods and which hydrophobicity scales perform well at predicting a given experimental metric of hydrophobicity. Here, we compare between different methods and different scales to predict retention times from hydrophobic interaction chromatography (HIC). We hope that our findings will be useful to guide future efforts at in-silico optimization of potential biopharmaceuticals.

1.1 Hydrophobicity

Hydrophobicity is one of the most important predictors of developability when designing antibody-based drugs (Lauer et al., 2012; Hebditch et al., 2019). Large scientific efforts have been directed towards reducing hydrophobicity without affecting the binding capability (Jain et al., 2017a; Jain et al., 2017b; Raybould et al., 2019; Jo et al., 2020). It has been shown that hydrophobicity as well as surface charges contribute to self-aggregation in IgG-type antibodies (Esfandiary et al., 2015; Das et al., 2022). Hydrophobicity also leads to faster clearance in antibody-drug conjugates (ADCs) (Lyon et al., 2015). While some antibodies exhibit high hydrophobicity in their folded state, partial unfolding can lead to exposure of additional hydrophobic sidechains and accelerate aggregation (Amin et al., 2014).

Hydrophobicity of monoclonal antibodies (mAbs) is routinely quantified using Hydrophobic Interaction Chromatography (HIC) (Haverick et al., 2014; Wang et al., 2016), while their solubility may be assessed using, e.g., PEG precipitation assays (Gibson et al., 2011; Sormanni et al., 2017). Aggregation of mAbs is commonly assessed using size exclusion chromatography (Brusotti et al., 2018) or dynamic light scattering experiments combined with incubation under stress conditions like elevated temperature (Amin et al., 2014).

Within this work, we test different hydrophobicity scores for their ability to predict HIC retention times. We use both publicly available datasets (Jain et al., 2017b) and Roche-internal data. The novel HIC retention times are shown in the Supplementary Material. For antibodies from public sources, the sequence is shown together with the retention times.

1.2 Hydrophobicity of antibodies

Hydrophobic interactions of biomolecules are typically mediated by one or few hydrophobic surface patches. Experimentally, the interaction of proteins with a column in Hydrophobic Interaction Chromatography (HIC) is different for proteins with homogeneous or inhomogeneous hydrophobicity profiles (Mahn et al., 2009).

Here, we test hydrophobicity scores based on the whole surface as well as scores based only on the hydrophobic surface regions. Physically, the first option implies a process where the whole surface is desolvated and directly contacts the HIC column, while the latter implies that only the hydrophobic regions are desolvated. From literature, it is expected that the second process more closely describes the process of HIC.

Research on statistical mechanics of the hydrophobic effect suggests that hydrophobic patches or moieties need to exceed a certain minimum size to exhibit the full effect on the water properties (Acharya et al., 2010; Huang and Chandler, 2000; Acharya et al., 2010). Furthermore, the behavior of hydrophobic surfaces also depends on the experimental conditions. Additionally, hydrophobic interactions between biomolecules depend on many other effects such as shape complementarity between the interaction partners (Mahn et al., 2005), or entropic penalties due to reduced conformational flexibility in the bound state.

Aggregation propensity of biomolecules is often discussed in terms of the related concept of Aggregation Prone Regions (APRs). An APR is a part of a protein structure or sequence that has a high tendency to form aggregates when exposed to the protein surface (Wang et al., 2009). This concept has been linked to highly hydrophobic surface regions (Chennamsetty et al., 2010), but also to the propensity to unfold or form amyloid β-sheets (Willbold et al., 2021).

The treatment of hydrophobic patches in in-silico methods will be discussed below.

1.3 In-silico methods

A multitude of methods have been devised to predict the hydrophobicity of antibodies in-silico. Generally, they may be divided into structure-based and sequence-based approaches.

Sequence-based methods, such as CamSol (Sormanni et al., 2015) and many others (Conchillo-Sole et al., 2007; Tartaglia and Vendruscolo, 2008; Walsh et al., 2014), offer the highest computational speed, such that large numbers of sequences can be scanned. Furthermore, they may be used even in cases where the three-dimensional structure of a protein is unknown.

On the other hand, methods that incorporate the three-dimensional protein structure may yield more accurate results. Examples include AggScore (Sankar et al., 2018), the Spatial Aggregation Propensity (SAP) method (Chennamsetty et al., 2009), as well as AggreScan3D (Zambrano et al., 2015). When the structure is not known experimentally, these methods can still be applied in combination with in-silico structure prediction. However, we have shown previously (Waibl et al., 2021) that homology models are often insufficiently accurate to describe the surface hydrophobicity, at least when using descriptors based on molecular dynamics simulation of the surrounding water.

Many in-silico methods favor large hydrophobic surfaces by searching for continuous hydrophobic patches (Lijnzaad et al., 1996; Chemical Computing Group ULC, 2020). Alternatively, the hydrophobicity of nearby atoms can be incorporated using so-called hydrophobic potentials (Heiden et al., 1993). In this approach, hydrophobicity is mapped to the protein surface *via* a distance weighting function. Hydrophobicity scores are then computed either by summing the surface values or by searching for patches above a certain cutoff. This approach favors large hydrophobic patches, because the effect of a single hydrophobic atom can be negated by a more hydrophilic surrounding, while the values in a patch of several hydrophobic atoms add up favorably.

Another approach is to assign additional hydrophobicity to each atom or residue based on the hydrophobicity of nearby atoms (Manavalan and Ponnuswamy, 1978; Gromiha et al., 2013). Newer approaches, such as the Spatial Aggregation Propensity (SAP) (Chennamsetty et al., 2009; Lauer et al., 2012) or AggreScan3D (Zambrano et al., 2015) combine this approach with terms for the solvent-accessible surface area (SASA) to avoid contributions from the hydrophobic core of the protein. In the case of SAP, this is done by computing the sum of hydrophobicity values of surface-exposed side-chain atoms within a pre-defined cutoff radius (Chennamsetty et al., 2009; Chennamsetty et al., 2010).

While a lot of research has been conducted to elucidate the hydrophobic effect using all-atom explicit solvent simulations (Acharya et al., 2010; Waibl et al., 2021), those methods have not been widely adopted in biopharmaceutical research due to high computational demand.

1.4 Hydrophobicity scales

Both sequence-based and structure-based methods often treat hydrophobicity as an innate property of the atoms or amino acids that constitute the protein. The individual values are tabulated in hydrophobicity scales and can be based on experimental measurements or more detailed calculations.

A plethora of different hydrophobicity scales has been devised. Lienqueo et al. (2002) investigated the ability of different hydrophobicity scales to predict HIC retention times over a wide range of different proteins. They classify hydrophobicity scales into three groups. Firstly, direct scales are based on the transfer free energy of each amino acid between phases of different polarity, retention time in RP-HPLC, or other properties such as polarity or geometry. Secondly, indirect scales are based on the solvent-accessible surface area (SASA) or other spatial distributions of amino acids in a protein. Lastly, mixed scales are derived from several of those properties or incorporate previous hydrophobicity scales.

A review of experimental hydrophobicity scales has been presented by Biswas et al. (2003). The usage of hydrophobicity scales to predict HIC retention is summarized by Mahn et al. (2009), and applications on interactions between proteins and lipid membranes have been reviewed by MacCallum and Tieleman (2011). Another review focused on secondary structure prediction has been presented by Simm et al. (2016).

While most hydrophobicity scales provide hydrophobicity parameters on a per-residue basis, some scales have also been parameterized per-atom. Notable examples include the Wildman and Crippen (1999) parameters, which are designed to predict octanol-water partitioning coefficients (logP) and molecular refractivities of small molecules, as well as the scale by Eisenberg and Mclachlan (1986), which is fitted to the transfer free energy of amino acids between the outer and inner regions of a protein. While those scales clearly offer a higher structural resolution, they might also introduce additional inaccuracies, since properties of amino acids are not exactly equal to the sum of their atomic contributions.

In this study we aim to compare various hydrophobicity scales and descriptors regarding their ability to predict antibody surface hydrophobicity.

Since there are too many hydrophobicity scales in literature to test them all, we selected a representative set for the purpose of this work. The scales were selected based on three properties: they are widely used (Eisenberg, Kyte-Doolittle, Crippen, Wimley-White), they are specifically aimed at predicting HIC or RP-HPLC (Jain, Meek, Miyazawa), or they have been previously used for hydrophobicity of antibodies (Black-Mould). Additionally, we added the scale by Rose as an example of a scale based purely on the change in SASA on folding. In Table 1, we present the selected scales and classify them in terms of experimental data and their resolution (atomic or residue-based). In the Results

TABLE 1 Comparison of hydrophobicity scales that will be used in the present work.

Scale	Resolution	Principle	Notes	References
Bandyopadhyay-Methler	Residue	Local environment, Rekker coefficients (Rekker and Kort, 1979)		Bandyopadhyay and Mehler (2008)
Black-Mould	Residue	Rekker coefficients (Rekker and Kort, 1979)		Black and Mould (1991)
Eisenberg	Residue	Consensus of 5 previous scales		Eisenberg et al. (1982)
Jain	Residue	HIC retention		Jain et al. (2017a)
Kyte-Doolittle	Residue	Consensus of ΔG (water—vapor) and surface accessibility		Kyte and Doolittle (1982)
Meek	Residue	RP-HPLC retention	At pH 7.4	Meek (1980)
Miyazawa	Residue	Surface accessibility		Miyazawa and Jernigan (1985)
Rose	Residue	Fraction of surface area buried while folding		Eisenberg et al. (1982), Rose et al. (1985)
Wimley-White	Residue	ΔG (water—lipid bilayer)	Interface Scale	Wimley and White (1996)
Crippen	Atomic	LogP of small molecules		Wildman and Crippen (1999)
Eisenberg-dG	Atomic	Surface accessibility		Eisenberg and Mclachlan (1986)

section, we will compare these scales in terms of their ability to predict biophysical properties of monoclonal antibodies.
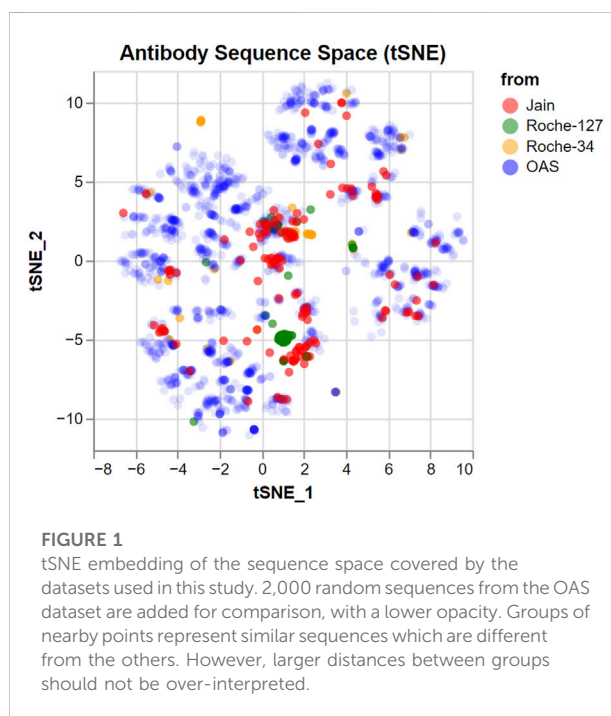
1.5 Aim of this work

In the early stages of development, a fast in-silico method to estimate the hydrophobicity of antibodies is often desired. In this work, we test the performance of several existing methods combined with multiple hydrophobicity scales by predicting relative hydrophobicity of antibodies and comparing to experimental data from HIC. We compare the predictivity of those methods between multiple datasets with different sequence variability to test whether our scores work better for closely related sequences or for a broader selection of antibodies. We find that choosing an appropriate hydrophobicity scale is crucial to obtain good agreement with experiments. Additionally, the choice of the scoring function, the origin of the homology models, as well as conformational sampling, can influence the results. We also demonstrate that better correlation with experimental results can be expected using datasets with high sequence similarity, also in cases where no crystal structures are available.

2 Results

2.1 Comparison between the datasets

In this work, we investigate several sets of antibodies, which are described in detail in the Methods section. In short, the dataset by Jain et al. (2017b) (called the "Jain" dataset below) dataset contains 127 variable domains of antibodies which were approved or undergoing clinical trials at the time of publication. These variable domains were grafted on a

FIGURE 1
tSNE embedding of the sequence space covered by the datasets used in this study. 2,000 random sequences from the OAS dataset are added for comparison, with a lower opacity. Groups of nearby points represent similar sequences which are different from the others. However, larger distances between groups should not be over-interpreted.

constant IgG1 Fc domain for consistency. We further split this into a group where crystal structures are available (Jain-PDB) and one where we rely on homology models generated using DeepAb (Ruffolo et al., 2021) (Jain-models). The Roche-34 dataset contains a diverse set of Roche-internal in addition to publicly available antibodies. It overlaps with the Jain-Models in 3 cases. The Roche-127 contains a group of 127 closely related Roche-internal antibodies, for which no crystal structures are available.

To investigate how much of the antibody sequence space is spanned by our input datasets, we produced a tSNE (van der
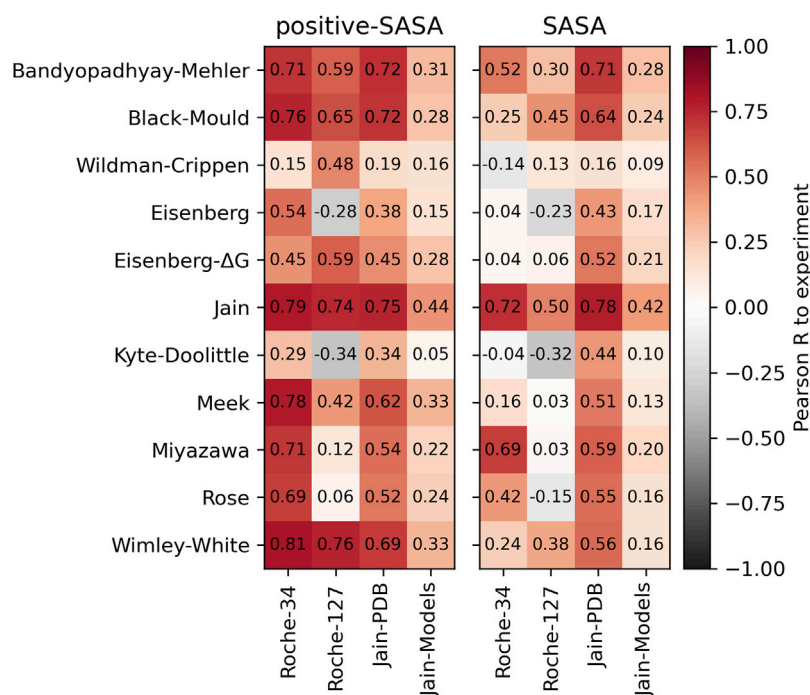
FIGURE 2
Pearson correlation values obtained by applying the positive surface score (left) and the direct surface score (right) on the different datasets and comparing to the experimental HIC data.

Maaten and Hinton, 2008) embedding of all investigated sequences combined with 2,000 randomly chosen sequences from the Observed Antibody Space database (Kovaltsuk et al., 2018; Olsen et al., 2022). The computational details are described in the Methods. The tSNE embedding produces a two-dimensional coordinate representation based on a distance matrix and aims to reproduce primarily the small distances. This means that groups of points in the embedding represent highly similar antibodies, while the distance between such groups is not necessarily representative of their similarity.

The result is a two-dimensional representation of the antibody sequence space, as shown in Figure 1. It can be used to visually distinguish between datasets that span a wide portion of the sequence space and datasets with a higher internal similarity.

We find that the antibodies in both the Jain and the Roche-34 dataset are generally unrelated, although there are some groups of similar antibodies in both cases. In contrast, the Roche-127 dataset has higher similarity, such that most antibodies fall into two groups, with few antibodies outside of those groups.

In Supplementary Table S4, we show the germline annotation [generated using ANARCI (Dunbar and Deane, 2016)] of the antibodies in all public datasets. Furthermore, we show a similarity matrix of all sequences in Supplementary Table S5.

It shows that the Roche-127 set is significantly less diverse than the other datasets.

2.2 Performance of different scales

For each of our datasets, we compute the SASA score as well as the positive-SASA score of all antibodies. The scores are discussed in more detail in the Methods section. In short, the SASA score (defined as $S_{surf} = \sum_i^{atoms} h_i \times A_i$) is the sum of all solvent accessible surface areas (SASAs) A_i of the amino acids in the Fv, multiplied by the respective hydrophobicity values h_i. The positive-SASA score is the same, but only taking the hydrophobic amino acids into account.

In the left panel of Figure 2, we show the Pearson correlation between the positive-SASA score and the respective experimental HIC values. We find good correlations using hydrophobicity scales that are optimized towards HIC or other RP-HPLC data, such as the Jain or Meek scales. Furthermore, we find good correlation values using the Wimley-White scale. The Miyazawa scale, which has often been associated with HIC prediction in literature, performs well on the Roche-127 set but not on the Roche-34 set. The relatively old scales by Eisenberg as well as Kyte and Doolittle do not perform well in our analysis. Furthermore, we find
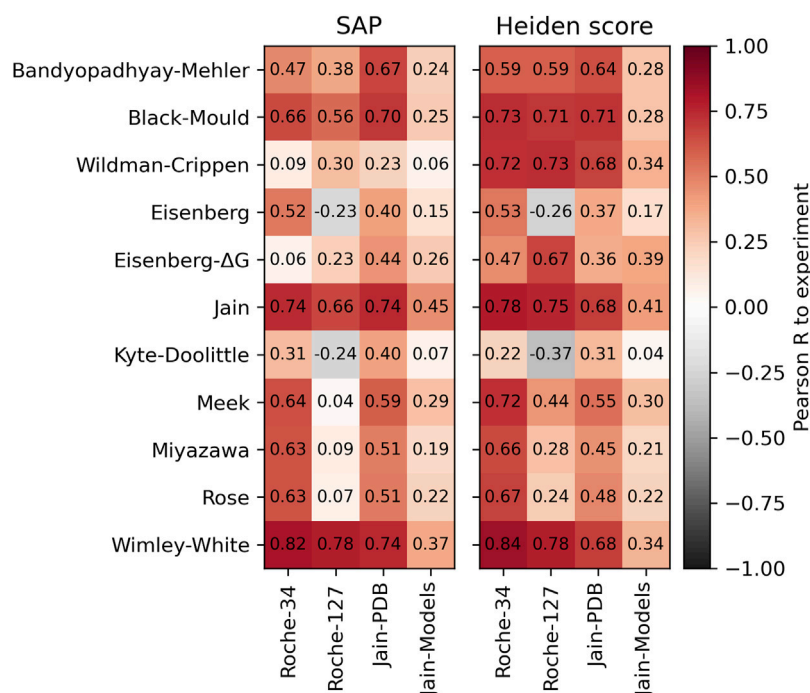
FIGURE 3
Performance of the SAP method (left) and the Heiden method (right) with different hydrophobicity scales. The Pearson correlation was calculated with respect to the HIC retention times of the respective dataset.

significantly worse predictivity of all scales when using homology models in the Jain-Models dataset.

In the right panel, we show the same analysis using the total surface score. While the general trends are the same, the overall correlation is significantly lower. The best-performing scale in this analysis is the Jain scale, which is little surprising, since it was optimized to predict HIC retention.

2.3 Other methods

In addition to the simple surface scores, we created our own implementation of the Spatial Aggregation Propensity (SAP) method (Voynov et al., 2009). This method is originally used in combination with the Black-Mould hydrophobicity scale (Black and Mould, 1991), but we also combine it with several other hydrophobicity scales. In addition, we also test our own implementation of the Heiden method (Heiden et al., 1993). Since this method can reasonably work with atomic hydrophobicity scales, it is expected that the scales by Wildman and Crippen (1999) as well as the atomic hydrophobicity scale of Eisenberg and Mclachlan (1986) perform better. The result is shown in Figure 3. The scores are described in detail in the *Methods* section.

Again, we find good correlations to the experimental HIC values using the Wimley-White scale. Furthermore, both methods work reasonably well with their original scales, which is the Black-Mould scale for SAP and the Wildman and Crippen scale for the Heiden method. We note that the Heiden method works especially well with atomic hydrophobicity scales such as the Crippen or Eisenberg-ΔG scale.

While the highest predictivity is found using the Wimley-White scale, we note that atomic hydrophobicity scales have the further advantage of a higher spatial resolution, permitting more detailed analysis of the surface properties. This will be used for the spatial analysis of cavities below.

2.4 Performance of homology modelling packages

To compare the efficiency of different homology modelling packages, we performed our calculations on the Roche-127 dataset using models created by MoFvAb, MOE, and DeepAb, and compare the results. We find that MoFvAb and DeepAb perform very similarly, while MOE performs slightly worse in combination with most residue-based scales. However, when using the SASA score (the right
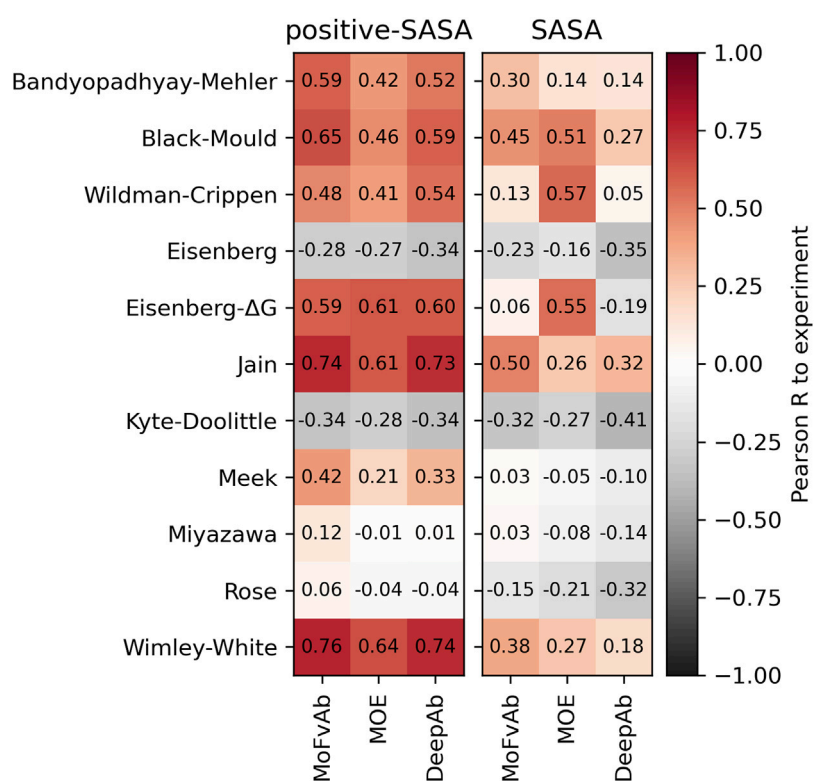
FIGURE 4
Performance of different homology modelling packages on the Roche-127 set in combination with the positive-SASA (left panel) and SASA (right panel) scores as well as different hydrophobicity scales. The Pearson correlation was calculated with respect to the HIC retention times of the Roche-127 dataset.

panel in Figure 4), MOE performs better in combination with the atom-based Crippen and Eisenberg-ΔG scales, as well as the residue-based Black-Mould scale. This might originate from a different side chain packing of MOE compared to the other modelling packages.

2.5 Effect of sampling through molecular dynamics

To test whether sampling of conformational ensembles through short molecular dynamics simulations improves the description of hydrophobicity, we performed 200 ns Gaussian accelerated Molecular Dynamics (GaMD) (Miao et al., 2015) simulations of each antibody in the Roche-127 set, using the MoFvAb models as starting structures. We then computed the positive-SASA score and the Heiden score and compare them to the experimental values. The result is shown in Figure 5.

As in our previous work (Waibl et al., 2021), we find no systematic improvements of the predictivity due to the sampling. The difference between hydrophobicity scales is

clearly bigger than that due to the conformational sampling. We find improvements in the atomic Wildman-Crippen and Eisenberg-ΔG scales, while the effect on most residue-based scales is small. This suggests that the sampling of sidechain conformations might be better than that of global motions within the protein.

2.6 Cavity effects

When visualizing structures of antibodies where the predicted hydrophobicity exceeds the experimental one, we often find that the structures contain cavities or pockets. Those cavities can represent, for example, binding pockets of antibodies that bind small molecules, or they can occur due to unfavorable sidechain packing in the homology modelling process. In Figure 6, we show examples of a chemically meaningful and an erroneous example of cavities, highlighting the respective structures in the correlation plots.

The 1L7T crystal structure (Valjakka et al., 2002) contains the unbound structure of an anti-testosterone Fab

fragment. The binding pocket is clearly visible, albeit smaller than in the corresponding bound structure (PDB code 1VPO). While the 1L7T structure is scored very high by the Heiden method, the low experimental HIC retention time indicates that a portion of the hydrophobic surface (probably the binding pocket) is inaccessible to the HIC column.
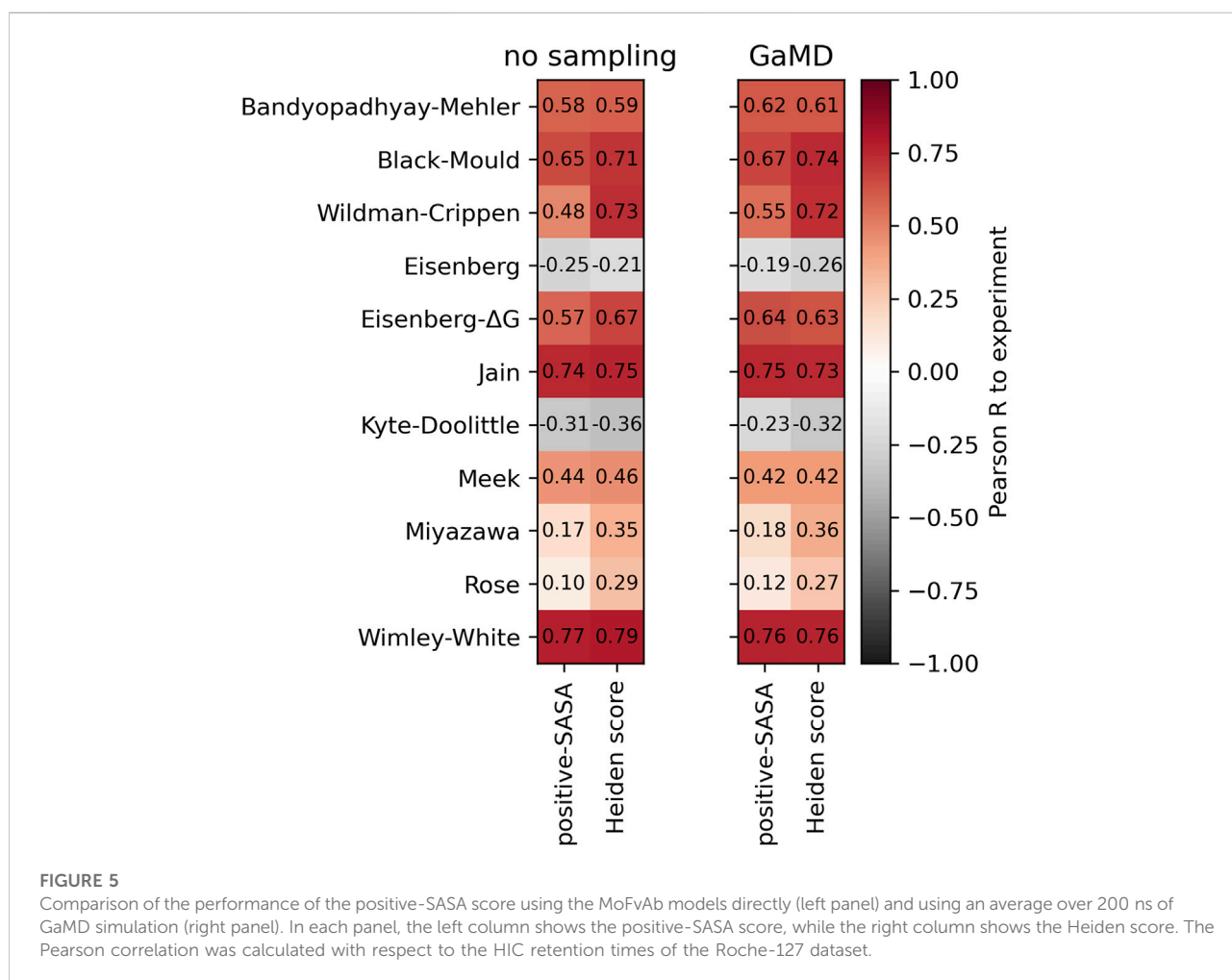
In the MoFvAb model of mAb_3L, the CDR regions are modelled as a very rugged surface, with several small cavities and hydrophobic residues pointing towards the solvent. However, this might be due to non-optimal side chain packing in the homology modelling process. Again, comparison to the experimental values indicates that not all regions of the modelled hydrophobic surface are exposed to the HIC column.

We also visualized several antibodies where the predicted hydrophobicity is significantly lower than the experimental one, to check whether they would show a particularly smooth surface. In contrast to this expectation, we find several instances of broad cavities, which might be able to fit a phenyl or butane sidechain of the HIC stationary phase.

Two examples (mAb_17L and golimumab) are shown in Figure 7.

3 Discussion

In the cases we studied, it is always better to predict HIC retention based on only the positive amino acid contributions. Even the Jain scale performs worse using the total surface score instead of the positive surface score. The only exception is the Jain dataset in combination with the Jain scale, which is not surprising since this scale is optimized to predict HIC retention. However, the Jain scale performs significantly worse on other datasets when using the SASA score (which includes the negative hydrophobicity values), which indicates poor transferability of the hydrophilic values. However, the transferability of hydrophobic values seems to be better, since the positive-SASA score using the Jain scale performs well also on the Roche-34 and Roche-127 datasets. We find a similar trend also for other hydrophobicity scales, which indicates

FIGURE 5
Comparison of the performance of the positive-SASA score using the MoFvAb models directly (left panel) and using an average over 200 ns of GaMD simulation (right panel). In each panel, the left column shows the positive-SASA score, while the right column shows the Heiden score. The Pearson correlation was calculated with respect to the HIC retention times of the Roche-127 dataset.

FIGURE 6
Examples for structures where cavities lead to over-prediction of the surface hydrophobicity. Left column: scatter plots showing the Heiden score vs. the experimental HIC values, for the Roche-34 set (top) and the Roche-127 set (bottom). In each plot, one antibody is marked, and the respective surface is shown at the right side. Green surface corresponds to hydrophobic regions (Heiden score > 0) and blue surface corresponds to hydrophilic regions (Heiden score < 0).

that the poor transferability of hydrophilic values in HIC prediction is a general phenomenon, while the hydrophobic values seem more robust.
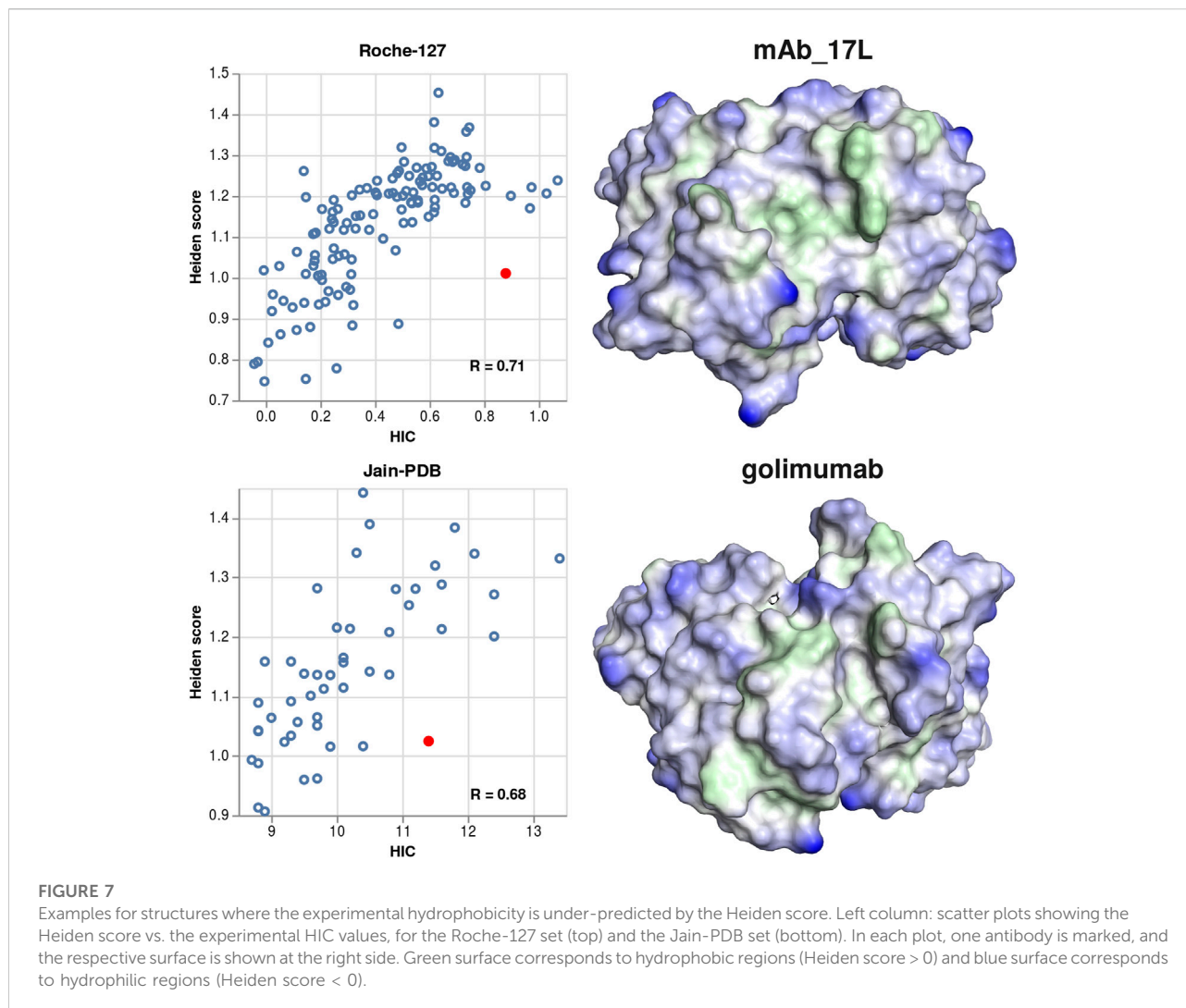
This finding is consistent with the idea that protein-column interactions in HIC are dominated by the most hydrophobic surface regions (Mahn et al., 2005). We have shown previously that the charged amino acids form very strong enthalpic interactions with the surrounding water (Schauperl et al., 2016). However, these amino acids do not have a large impact on HIC retention, since they keep their hydration shell when the protein binds to a HIC column, while only the more hydrophobic amino acids are dehydrated.

We therefore conclude that HIC measurements are controlled essentially by the most hydrophobic surface regions. This is consistent with previous findings by Mahn et al. (2005), who achieved high predictivity using a molecular docking approach to identify the most probable interaction

region on the surface of ribonucleases, and scored them by the hydrophobicity of this region. We expect that this interaction region frequently coincides with a strongly hydrophobic surface region.

When visualizing antibodies where the hydrophobicity is over-predicted compared to experiment, we sometimes find cavities in the surface. These cavities can be, for example, binding pockets of antibodies that bind to small molecules, or they can be due to inaccurate sidechain packing in the homology models. In both cases, they lead to an increased hydrophobic surface. Since this is not reflected by a higher experimental HIC retention time, we assume that those pockets do not interact with the stationary phase.

We also show two examples of antibodies where the hydrophobicity is under-predicted compared to the experiment. In contrast to the over-predicted antibodies, the cavities found in those structures are broader or not very

FIGURE 7
Examples for structures where the experimental hydrophobicity is under-predicted by the Heiden score. Left column: scatter plots showing the Heiden score vs. the experimental HIC values, for the Roche-127 set (top) and the Jain-PDB set (bottom). In each plot, one antibody is marked, and the respective surface is shown at the right side. Green surface corresponds to hydrophobic regions (Heiden score > 0) and blue surface corresponds to hydrophilic regions (Heiden score < 0).

deep. We expect that such broader hydrophobic cavities can form specific interactions with the stationary phase, thus leading to a high HIC retention time although there is only a relatively small hydrophobic surface. This finding is also consistent with the idea that binding to HIC columns is dominated by the most hydrophobic region.

In general, methods based on the solvent-accessible surface area (SASA) perform rather poorly in combination with atomic hydrophobicity scales. This is expected due to the underlying assumption of additivity of the hydrophobic contribution: Correlations between the hydrophobic effect of neighboring atoms are likely stronger than those between neighboring residues. The surface projection by Heiden et al. (1993) provides a smoother way of projecting hydrophobicity scales onto the protein surface, leading to an improvement in combination with atomic scales.

Our results using the Spatial Aggregation Propensity (SAP) algorithm are similar to those using the positive-SASA score.

While we used a cutoff radius R of 5 Å throughout the main text, Supplementary Figure S1 shows the comparison between R of 5 or 10 Å. In almost all cases, the cutoff of 5 Å performs better.

We find that homology models generated with MoFvAb and DeepAb perform very similar in terms of predicting experimental HIC values. On the other hand, models generated by MOE perform slightly worse in combination with most residue-based hydrophobicity scales, while outperforming them when combined with atomic scales. One possible explanation would be that MOE is more accurate at predicting side chain orientation, while large-scale contributions to the structure, such as inter-domain orientation, are predicted better by MoFvAb and DeepAb.

Our findings show that it is most difficult to work with homology models of highly diverse data sets. Several hydrophobicity scales provide good correlations with the Roche-127 dataset, which consists of homology models with

high sequence similarity. We also find good correlations using the Jain-PDB set and the Roche-34 set, which have large sequence diversity but consist mostly of crystal structures. Only the HIC data of the Jain-models set is poorly predicted, likely because there is a high sequence diversity combined with homology models, which are less reliable than crystal structures. This means that there are many differences between structures, while the SASA of each amino acid is not completely reliable. In that case, summation of errors can lead to large errors in the total hydrophobicity score.

We expect the CDR-H3 loop to be the most challenging part of the structure to predict. Despite the enormous recent advances in predicting antibody structures (Ruffolo et al., 2021; Abanades et al., 2022), this loop remains difficult to predict accurately due to its high flexibility and unchallenged diversity in length, sequence, and structure. As the CDR-H3 loop is situated in the center of the antibody binding site, it influences the conformations of the neighboring CDR loops and consequently plays a critical role for predicting and quantifying surface hydrophobicity (Regep et al., 2017; Fernandez-Quintero et al., 2020).

The agreement between hydrophobicity scores and HIC retention times depends strongly on the hydrophobicity scale. The Kyte-Doolittle and Eisenberg scales perform poorly, probably because they were aimed at quantities different from HIC retention times. In contrast, the Jain scale, which is parameterized to predict HIC retention times, performs much better. Furthermore, the Wimley-White scale, which represents the free energy change when transferring pentapeptides between water and a lipid bilayer, also performs very well. A common feature of the Jain and Wimley-White scales is that they assign high hydrophobicity to aromatic residues and especially to tryptophan.

The Meek and Miyazawa scales perform comparably well on datasets which contain crystal structures (Roche-34 and Jain-PDB), but not on datasets that contain only homology models (Roche-127 and Jain-models). This contrasts with the better predictions obtained using the Jain and Wimley-White scales. The reason might be related to differences in the amino acid composition of the datasets, but also to the lower hydrophobicity assigned to tryptophan and tyrosine in the Meek and Miyazawa scales. Especially tyrosine is often found in antibody CDR regions. Since the CDRs also encompass the highest sequence diversity within an antibody, their amino acid composition is a main contributor to their surface properties (Yugandhar and Gromiha, 2014).

Thus, we conclude that several factors must be considered to predict HIC-based hydrophobicity using hydrophobicity scales. The choice of the hydrophobicity scale is crucial. There are several scales which correlate well with experimental HIC data. However, the Kyte-Doolittle and Eisenberg scales produce very poor results, even though they are still widely used. Furthermore, it is important to be aware about the sequence diversity of the dataset in question, as well as the reliability of the available structures. When predicting the hydrophobicity of highly diverse antibodies, very reliable structures—such as crystal structures—are required, while less diverse datasets can be effectively described by homology models.

4 Theory and methods

4.1 Hydrophobicity scales

Scales were normalized by adding a constant such that Gly has a value of zero and scaled such that hydrophobic residues are positive and the variance of the values is 1. The variance was calculated as the average of h^2, where h denotes the individual values relative to Gly. The resulting values are shown in Table 2.

4.2 Datasets

The following datasets were investigated in this work:

4.2.1 Roche-34

This dataset contains 34 antibodies for which HIC measurements of full-length IgGs have been performed. Of this dataset, 14 have crystal structures deposited in the PDB (Berman et al., 2000). Further three antibodies have been previously published and have a name in the standard antibody nomenclature, but do not have crystal structures. The other 17 structures are Roche-internal. The relative HIC retention times of this dataset, as well as identifiers and sequences of the published antibodies, are shown in the Supplementary Table S1.

4.2.2 Roche-127

This dataset contains 127 Roche-internal antibodies for which HIC measurements of full-length IgGs have been performed. Since no crystal structures are available for those antibodies, homology models were created using MoFvAb, as described below. Relative HIC retention times are shown in Supplementary Table S2.

4.2.3 Jain-PDBs

This dataset is a subset of the publicly available dataset by Jain et al. (2017b), containing 49 structures for which crystal structures are available from the PDB. The HIC measurements from the original publication were used as reference data for this dataset. The used PDB codes are shown in Supplementary Table S3.

4.2.4 Jain-Models

This dataset contains 77 antibodies from the dataset by Jain et al. (2017b), for which no crystal structures were found

in the PDB. Homology models were created using DeepAb (Ruffolo et al., 2021).

4.3 tSNE

tSNE (t-distributed stochastic neighbor embedding) (van der Maaten and Hinton, 2008) is a dimensionality reduction technique that aims to preserve the local structure, i.e., the distance information between close-lying data points. Here, we apply it to generate a two-dimensional representation of the antibody sequence space, based on a distance matrix generated using Clustal Omega (Sievers et al., 2011; Sievers and Higgins, 2018; Sievers et al., 2020). We used the tSNE implementation in Scikit-Learn (Pedregosa et al., 2011) with a "perplexity" setting of 500. The value of 500 was chosen because it is significantly bigger than the size of the Roche-127 set but small compared to the total number of datapoints. If there are groups of similar antibodies with a size larger than the perplexity, the relation of this group to other groups is lost almost completely.

While t-SNE excels at reproducing the relation between adjacent datapoints (i.e., similar sequences) in the two-dimensional embedding, the relations between more distant sequences are less reliable. It has been shown that the initialization method is crucial to obtain embeddings that also

reproduce some of the global structure (Kobak and Linderman, 2021). While literature states (Kobak and Linderman, 2021) that initialization using Laplacian Eigenmaps (LE) (Belkin and Niyogi, 2002) is superior to random initialization at preserving global structure, this approach performed very poorly for our dataset, producing consistently lower Pearson correlations between the original distance matrix and the two-dimensional distances. We therefore chose to use random initialization, but repeat the calculation 10 times and use the best embedding as judged by the Kullback-Leibler divergence (Kullback and Leibler, 1951).

4.4 Starting structures

For the Jain-PDB dataset, equilibrated versions of the crystal structures of the Jain dataset were taken from our previous work (Waibl et al., 2021).

All homology models were created starting from the respective VH and VL sequences. The program settings were as follows:

4.4.1 MoFvAb

The settings for MoFvAb were as in the original work by Bujotzek et al. (2015).

TABLE 2 The per-residue hydrophobicity scales that were used in this work.

Residue	BaMe	BlMo	Ei	KyDo	Me	Ro	WiWh	Ja	Mi
ALA	0.75	0.37	0.15	0.76	0.07	0.18	−0.20	0.06	0.40
ARG	−0.02	−1.52	−3.09	−1.41	0.11	−0.71	−0.41	−0.32	−0.15
ASN	−0.16	−0.79	−1.29	−1.06	0.11	−0.80	−0.51	0.13	−0.37
ASP	−0.50	−1.43	−1.42	−1.06	−1.08	−0.89	−1.53	−0.43	−0.43
CYS	2.60	0.55	−0.19	1.00	−0.90	1.69	0.31	0.55	1.65
GLN	−0.11	−0.76	−1.37	−1.06	−0.63	−0.89	−0.71	0.46	−0.29
GLU	−0.54	−1.40	−1.26	−1.06	−2.23	−0.89	−2.51	−0.72	−0.40
GLY	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
HIS	0.57	−1.00	−0.90	−0.96	−0.46	0.53	−0.20	0.03	0.30
ILE	2.19	1.34	0.92	1.68	1.83	1.42	0.35	1.54	2.08
LEU	1.97	1.34	0.60	1.44	1.16	1.16	0.71	1.54	1.91
LYS	−0.90	−0.67	−2.03	−1.20	0.01	−1.78	−0.59	−1.07	−0.74
MET	1.22	0.73	0.16	0.79	0.63	1.16	0.30	0.49	2.14
PHE	1.92	1.52	0.73	1.10	1.74	1.42	1.43	2.48	2.18
PRO	0.72	0.64	−0.37	−0.41	0.80	−0.71	−0.55	0.44	−0.29
SER	0.11	−0.43	−0.68	−0.14	0.16	−0.53	−0.15	0.07	−0.19
THR	0.47	−0.15	−0.55	−0.10	0.36	−0.18	−0.16	0.16	0.00
TRP	1.51	1.16	0.34	−0.17	1.96	1.16	2.33	2.81	1.52
TYR	1.36	1.16	−0.23	−0.31	0.80	0.36	1.19	1.84	0.68
VAL	1.88	1.00	0.61	1.58	0.36	1.25	−0.08	0.97	1.51

Abbreviations: BaMe, Bandyopadhyay-Mehler; BlMo, Black-Mould; Ei, Eisenberg; KyDo, Kyte-Doolittle; Me, Meek; Ro, Rose; WiWh, Wimley-White; Ja, Jain; Mi, Miyazawa.

4.4.2 MOE

Fab Models were created using the antibody modeling protocol in MOE 2020.09. Templates were selected automatically from the built-in database, and the resulting structures were extended to FAb fragments. MOE models were used as exported from MOE. All protonation states were kept the same, histidines with a hydrogen in δ-position were renamed to HID, and histidines with a hydrogen in the ε-position were renamed to HIE. CYS residues in a disulfide bond were renamed to CYX. Hydrogen atoms were removed and re-added in standard positions using the reduce program (Word et al., 1999).

4.4.3 DeepAb

DeepAb (Ruffolo et al., 2021) was downloaded from GitHub at 3 Nov 2021, and used with PyRosetta 4 release 293 (Chaudhury et al., 2010). The pre-trained model was used, and calculations were run without GPU acceleration. All settings were left at their default values. The protonation was kept as in the DeepAb output, histidines and cysteines were renamed in the same way as for the MOE models. We note that the protonation only affects the atomic hydrophobicity scales, since no simulations were performed starting from the DeepAb models.

4.5 Gaussian accelerated molecular dynamics simulations

GaMD simulations were performed starting from the MoFvAb models of the Roche-127 dataset. C_H1/C_L domains were added for the simulations but omitted in all analyses.

The models were protonated using the Protonate3D protocol (Labute, 2009) in MOE (Chemical Computing Group ULC, 2020) to obtain consistent protonation patterns.

Gaussian accelerated Molecular Dynamics (GaMD) (Miao et al., 2015) simulations were performed using the same protocol as described previously (Waibl et al., 2021). The ff14SB force field (Maier et al., 2015) was used in combination with the TIP3P water model (Jorgensen et al., 1983). Energy statistics for GaMD were collected during a number of MD frames equal to 4 times the number of atoms in the system, rounded up to the next picosecond (Case et al., 2019). Then, GaMD was run using a dual boost. SHAKE (Ryckaert et al., 1977) was used on all bonds including hydrogen. The integration timestep was 2 fs, using a Langevin thermostat (Adelman and Doll, 1976) at 300 K with a collision frequency of 2 ps^{-1} and a Monte Carlo barostat (Åqvist et al., 2004) with one volume change attempt per 100 steps.

For post-processing, one frame per nanosecond was used, resulting in 200 representative frames. Energies were collected every picosecond and were used to reweight the probabilities of representative frames using cumulative expansion to the second order (Miao et al., 2014).

4.6 Hydrophobicity scores

All scores were calculated in Python using a series of in-house Python scripts.

Residue-based hydrophobicity scales were assigned based on the residue name. All protonation states of histidine were considered equally except for scales that contain separate values.

The atoms in each of the 20 amino acids were assigned types according to the Crippen scale using the SMILES matching functionality in RDKit (Landrum et al., 2020). The hydrogen atoms bound to aromatic nitrogen in TRP and HIS residues were set to the H3 type ("amine"), rather than H2 ("alcohol"). All other types were used as assigned by RDKit. After the initial assignment, atom types were assigned by a simple lookup table using the residue name and atom name as a key.

The solvent-accessible surface area (SASA) was computed using the Shrake-Rupley algorithm (Shrake and Rupley, 1973) implemented in MDTraj (McGibbon et al., 2015).

The direct surface score S_{surf} was computed by multiplying the SASA of each atom A_i by the hydrophobicity value obtained from a scale, h_i. This score is conceptually consistent with an experiment where all surface-exposed residues are desolvated, thereby contributing to the overall hydrophobicity.

$$S_{surf} = \sum_{i}^{atoms} h_i \times A_i \tag{1}$$

The positive surface score S_{pos} was defined in the same way, except that all negative h_i values were set to zero. This is consistent with an experiment where only the hydrophobic surface regions are desolvated, while the hydrophilic regions remain in contact with water. It is expected (Mahn et al., 2009) that this matches the experimental conditions of HIC more closely.

The Spatial Aggregation Propensity (SAP) (Chennamsetty et al., 2010) was computed as:

$$SAP_i = \sum_{j, r_{ij} < R}^{side\ chain\ atoms} \frac{A_j}{A_j^{residue}} h_j \tag{2}$$

where r_{ij} is the distance between atoms i and j, and R is the cutoff radius, chosen as 5 Å in this study to be consistent with the original work. $A_j^{residue}$ is the average sidechain SASA of a residue capped with N-methyl and acetyl groups in TIP3P water, computed from a 100 ns cMD simulation using the ff14SB Amber force field and TIP3P. The SASA was again computed using MDTraj.

The resulting hydrophobicity score was computed as:

$$S_{SAP} = \sum_{i}^{atoms} \max(SAP_i, 0) \tag{3}$$

The final *SAP* score is computed including atoms that are not solvent-exposed. However, the solvent-exposure is already considered by using A_j in the individual atom scores. This is consistent with the original literature (Chennamsetty et al., 2009; Lauer et al., 2012).

For the score by Heiden et al. (1993), we first computed a solvent-excluded surface (SES), using the post-processing routines of our previous work (Waibl et al., 2022) (available from https://github.com/liedllab/gisttools) to compute the lowest possible distance to a solvent molecule on a 3D grid, and then creating an isosurface at the solvent radius of 1.4 Å using the marching cubes algorithm in scikit-image (van der Walt et al., 2014). For each vertex k of the surface, we compute the molecular lipophilicity potential (MLP) using:

$$MLP_k = \frac{\sum_{j,r_{jk}<R}^{atoms} g(r_{jk})h_j}{\sum_{j,r_{jk}<R}^{atoms} g(r_{jk})}, \quad g(r) = \left[\exp\left(\alpha\left(r - \frac{R}{2}\right)\right) + 1\right]^{-1}$$

(4)

This is essentially a weighted average over nearby atoms, with $g(r)$ as the weighting function. $g(r)$ is a (mirrored) logistic function with height 1, steepness α and midpoint $R/2$, where R is the cutoff radius. Consistent with the original implementation, we choose α as 1.5 Å^{-1} and R as 5 Å. We compute a hydrophobicity score as

$$S_{Heiden} = \sum_{k}^{vertices} \max(MLP_k, 0) \times A_k$$

(5)

where A_k is the SASA of vertex k, which is calculated by splitting the area of each triangle to the three vertices. The sum over all positive vertex scores is used to define the Heiden score. By visualizing the hydrophobic surface regions, we find the same hydrophobic patches as defined by MOE. The patch area, however, is not the same due to slight differences in the definition of the molecular surface and the way the patches are searched.

4.7 Hydrophobic interaction chromatography

Apparent hydrophobicity was determined essentially as described previously (Jarasch et al., 2015), by injecting 20 μg of sample onto a HIC-Ether-5PW (Tosoh) column equilibrated with 25 mM Na-phosphate, 1.5 M ammonium sulfate, pH 7.0. Elution was performed with a linear gradient from 0 to 100% buffer B (25 mM Na-phosphate, pH 7.0) within 60 min. Retention times were compared to protein standards with known hydrophobicity.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

FW conducted research and drafted the manuscript. MF-Q supervised the drafting of the manuscript. FSW contributed in analyzing data. HK and GG supported planning the research and provided experimental data. KL supervised the research.

Conflict of interest

HK and GG are employees of Roche Diagnostics GmbH.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary Material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fmolb.2022.960194/full#supplementary-material

SUPPLEMENTARY TABLE S1
Relative HIC retention times for the antibodies of the Roche-34 dataset. For publicly available antibodies, the sequences and PDB codes are also reported.

SUPPLEMENTARY TABLE S2
Relative HIC retention times for the antibodies of the Roche-127 dataset.

SUPPLEMENTARY TABLE S3
PDB codes that were used for the Jain dataset. Where no PDB code is shown, homology models are shown. HIC data and sequences were obtained from (Jain et al. 2017b).

SUPPLEMENTARY TABLE S4
Germline annotation for all investigated antibodies.

SUPPLEMENTARY TABLE S5
Similarity matrix of all investigated antibodies.

References

Abanades, B., Georges, G., Bujotzek, A., and Deane, C. M. (2022). ABlooper: Fast accurate antibody CDR loop structure prediction with accuracy estimation. *Bioinformatics* 38 (7), 1877–1880. doi:10.1093/bioinformatics/btac016

Acharya, H., Vembanur, S., Jamadagni, S. N., and Garde, S. (2010). Mapping hydrophobicity at the nanoscale: Applications to heterogeneous surfaces and proteins. *Faraday Discuss.* 146, 353–365. doi:10.1039/b927019a

Adelman, S. A., and Doll, J. D. (1976). Generalized Langevin equation Approach for atom-solid-surface scattering - general formulation for classical scattering off harmonic solids. *J. Chem. Phys.* 64 (6), 2375–2388. doi:10.1063/1.432526

Amin, S., Barnett, G. V., Pathak, J. A., Roberts, C. J., and Sarangapani, P. S. (2014). Protein aggregation, particle formation, characterization & rheology. *Curr. Opin. Colloid & Interface Sci.* 19 (5), 438–449. doi:10.1016/j.cocis.2014. 10.002

Åqvist, J., Wennerström, P., Nervall, M., Bjelic, S., and Brandsdal, B. O. (2004). Molecular dynamics simulations of water and biomolecules with a Monte Carlo constant pressure algorithm. *Chem. Phys. Lett.* 384 (4-6), 288–294. doi:10.1016/j. cplett.2003.12.039

Bandyopadhyay, D., and Mehler, E. L. (2008). Quantitative expression of protein heterogeneity: Response of amino acid side chains to their local environment. *Proteins* 72 (2), 646–659. doi:10.1002/prot.21958

Belkin, M., and Niyogi, P. (2002). Laplacian eigenmaps and spectral techniques for embedding and clustering. *Adv. Neural Inf. Process. Syst.* 14 (14), 585–591.

Berman, H., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T., Weissig, H., et al. (2000). The protein data bank. *Nucleic Acids Res.* 28 (1), 235–242. doi:10.1093/nar/28.1.235

Biswas, K. M., DeVido, D. R., and Dorsey, J. G. (2003). Evaluation of methods for measuring amino acid hydrophobicities and interactions. *J. Chromatogr. A* 1000 (1-2), 637–655. doi:10.1016/s0021-9673(03)00182-1

Black, S. D., and Mould, D. R. (1991). Development of hydrophobicity parameters to analyze proteins which bear post- or cotranslational modifications. *Anal. Biochem.* 193 (1), 72–82. doi:10.1016/0003-2697(91)90045-U

Brusotti, G., Calleri, E., Colombo, R., Massolini, G., Rinaldi, F., and Temporini, C. (2018). Advances on size exclusion chromatography and applications on the analysis of protein biopharmaceuticals and protein aggregates: A mini review. *Chromatographia* 81 (1), 3–23. doi:10.1007/s10337-017-3380-5

Bujotzek, A., Fuchs, A., Qu, C. T., Benz, J., Klostermann, S., Antes, I., et al. (2015). MoFvAb: Modeling the Fv region of antibodies. *Mabs* 7 (5), 838–852. doi:10.1080/19420862.2015.1068492

Case, D. A., Ben-Shalom, I. Y., Brozell, S. R., Cerutti, D. S., Cheatham, T. E., III, Cruzeiro, V. W. D., et al. (2019). *Amber 2019.* San Francisco: University of California.

Chan, A. C., and Carter, P. J. (2010). Therapeutic antibodies for autoimmunity and inflammation. *Nat. Rev. Immunol.* 10 (5), 301–316. doi:10.1038/nri2761

Chaudhury, S., Lyskov, S., and Gray, J. J. (2010). PyRosetta: A script-based interface for implementing molecular modeling algorithms using rosetta. *Bioinformatics* 26 (5), 689–691. doi:10.1093/bioinformatics/btq007

Chemical Computing Group ULC (2020). *Molecular operating environment (MOE).* Montreal, QC, Canada, H3A 2R7: Chemical Computing Group ULC, S.S.W., Suite #910.

Chennamsetty, N., Voynov, V., Kayser, V., Helk, B., and Trout, B. L. (2009). Design of therapeutic proteins with enhanced stability. *Proc. Natl. Acad. Sci. U. S. A.* 106 (29), 11937–11942. doi:10.1073/pnas.0904191106

Chennamsetty, N., Voynov, V., Kayser, V., Helk, B., and Trout, B. L. (2010). Prediction of aggregation prone regions of therapeutic proteins. *J. Phys. Chem. B* 114 (19), 6614–6624. doi:10.1021/jp911706q

Chvatal-Medina, M., Mendez-Cortina, Y., Patino, P. J., Velilla, P. A., and Rugeles, M. T. (2021). Antibody responses in COVID-19: A review. *Front. Immunol.* 12, 633184. doi:10.3389/fimmu.2021.633184

Conchillo-Sole, O., de Groot, N., Aviles, F., Vendrell, J., Daura, X., and Ventura, S. (2007). Aggrescan: A server for the prediction and evaluation of "hot spots" of aggregation in polypeptides. *Bmc Bioinforma.* 8, 65. doi:10.1186/1471-2105-8-65

Das, T. K., Chou, D. K., Jiskoot, W., and Arosio, P. (2022). Nucleation in protein aggregation in biotherapeutic development: A look into the heart of the event. *J. Pharm. Sci.* 111 (4), 951–959. doi:10.1016/j.xphs.2022.01.017

Dunbar, J., and Deane, C. M. (2016). Anarci: Antigen receptor numbering and receptor classification. *Bioinformatics* 32 (2), 298–300. doi:10.1093/bioinformatics/btv552

Eisenberg, D., and Mclachlan, A. D. (1986). Solvation energy in protein folding and binding. *Nature* 319 (6050), 199–203. doi:10.1038/319199a0

Eisenberg, D., Weiss, R. M., Terwilliger, T. C., and Wilcox, W. (1982). Hydrophobic moments and protein-structure. *Faraday Symposia Chem. Soc.* 17, 109–120. doi:10.1039/fs9821700109

Esfandiary, R., Parupudi, A., Casas-Finet, J., Gadre, D., and Sathish, H. (2015). Mechanism of reversible self-association of a monoclonal antibody: Role of electrostatic and hydrophobic interactions. *J. Pharm. Sci.* 104 (2), 577–586. doi:10.1002/jps.24237

Fernandez-Quintero, M. L., Heiss, M. C., Pomarici, N. D., Math, B. A., and Liedl, K. R. (2020). Antibody CDR loops as ensembles in solution vs. canonical clusters from X-ray structures. *MAbs* 12 (1), 1744328. doi:10.1080/19420862.2020.1744328

Gibson, T. J., Mccarty, K., Mcfadyen, I. J., Cash, E., Dalmonte, P., Hinds, K. D., et al. (2011). Application of a high-throughput screening procedure with PEG-induced precipitation to compare relative protein solubility during formulation development with IgG1 monoclonal antibodies. *J. Pharm. Sci.* 100 (3), 1009–1021. doi:10.1002/jps.22350

Gromiha, M. M., Pathak, M. C., Saraboji, K., Ortlund, E. A., and Gaucher, E. A. (2013). Hydrophobic environment is a key factor for the stability of thermophilic proteins. *Proteins* 81 (4), 715–721. doi:10.1002/prot.24232

Haverick, M., Mengisen, S., Shameem, M., and Ambrogelly, A. (2014). Separation of mAbs molecular variants by analytical hydrophobic interaction chromatography HPLC Overview and applications. *Mabs* 6 (4), 852–858. doi:10.4161/mabs.28693

Hebditch, M., Roche, A., Curtis, R. A., and Warwicker, J. (2019). Models for antibody behavior in hydrophobic interaction chromatography and in self-association. *J. Pharm. Sci.* 108 (4), 1434–1441. doi:10.1016/j.xphs.2018.11.035

Heiden, W., Moeckel, G., and Brickmann, J. (1993). A new approach to analysis and display of local lipophilicity hydrophilicity mapped on molecular-surfaces. *J. Comput. Aided. Mol. Des.* 7 (5), 503–514. doi:10.1007/BF00124359

Huang, D. M., and Chandler, D. (2000). Temperature and length scale dependence of hydrophobic effects and their possible implications for protein folding. *Proc. Natl. Acad. Sci. U. S. A.* 97 (15), 8324–8327. doi:10.1073/pnas.120176397

Jain, T., Boland, T., Lilov, A., Burnina, I., Brown, M., Xu, Y. D., et al. (2017a). Prediction of delayed retention of antibodies in hydrophobic interaction chromatography from sequence using machine learning. *Bioinformatics* 33 (23), 3758–3766. doi:10.1093/bioinformatics/btx519

Jain, T., Sun, T., Durand, S., Hall, A., Houston, N., Nett, J., et al. (2017b). Biophysical properties of the clinical-stage antibody landscape. *Proc. Natl. Acad. Sci. U. S. A.* 114 (5), 944–949. doi:10.1073/pnas.1616408114

Jarasch, A., Koll, H., Regula, J. T., Bader, M., Papadimitriou, A., and Kettenberger, H. (2015). Developability assessment during the selection of novel therapeutic antibodies. *J. Pharm. Sci.* 104 (6), 1885–1898. doi:10.1002/jps.24430

Jo, S., Xu, A., Curtis, J. E., Somani, S., and MacKerell, A. D. (2020). Computational characterization of antibody-excipient interactions for rational excipient selection using the site identification by ligand competitive saturation-biologics approach. *Mol. Pharm.* 17 (11), 4323–4333. doi:10.1021/acs.molpharmaceut.0c00775

Jorgensen, W. L., Chandrasekhar, J., Madura, J. D., Impey, R. W., and Klein, M. L. (1983). Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* 79 (2), 926–935. doi:10.1063/1.445869

Kaplon, H., Chenoweth, A., Crescioli, S., and Reichert, J. M. (2022). Antibodies to watch in 2022. *Mabs* 14 (1). doi:10.1080/19420862.2021.2014296

Kaplon, H., Muralidharan, M., Schneider, Z., and Reichert, J. M. (2020). Antibodies to watch in 2020. *Mabs* 12 (1). doi:10.1080/19420862.2019.1703531

Kaplon, H., and Reichert, J. M. (2021). Antibodies to watch in 2021. *Mabs* 13 (1). doi:10.1080/19420862.2020.1860476

Kobak, D., and Linderman, G. C. (2021). Initialization is critical for preserving global data structure in both t-SNE and UMAP. *Nat. Biotechnol.* 39 (2), 156–157. doi:10.1038/s41587-020-00809-z

Kovaltsuk, A., Leem, J., Kelm, S., Snowden, J., Deane, C. M., and Krawczyk, K. (2018). Observed antibody space: A resource for data mining next-generation sequencing of antibody repertoires. *J. Immunol.* 201 (8), 2502–2509. doi:10.4049/jimmunol.1800708

Kullback, S., and Leibler, R. A. (1951). On information and sufficiency. *Ann. Math. Stat.* 22 (1), 79–86. doi:10.1214/aoms/1177729694

Kyte, J., and Doolittle, R. F. (1982). A simple method for displaying the hydropathic character of a protein. *J. Mol. Biol.* 157 (1), 105–132. doi:10.1016/0022-2836(82)90515-0

Labute, P. (2009). Protonate3D: Assignment of ionization states and hydrogen coordinates to macromolecular structures. *Proteins* 75 (1), 187–205. doi:10.1002/prot.22234

Landrum, G., Tosco, P., Kelley, B., SrinikerGedeck, P., Schneider, N., et al. (2020). *rdkit/rdkit: 2020_03_1 (Q1 2020) Release (Release_2020_03_1)*. Available at: http://www.rdkit.org.

Lauer, T. M., Agrawal, N. J., Chennamsetty, N., Egodage, K., Helk, B., and Trout, B. L. (2012). Developability index: A rapid *in silico* tool for the screening of antibody aggregation propensity. *J. Pharm. Sci.* 101 (1), 102–115. doi:10.1002/jps.22758

Lienqueo, M. E., Mahn, A., and Asenjo, J. A. (2002). Mathematical correlations for predicting protein retention times in hydrophobic interaction chromatography. *J. Chromatogr. A* 978 (1-2), 71–79. doi:10.1016/S0021-9673(02)01358-4

Lijnzaad, P., Berendsen, H. J. C., and Argos, P. (1996). Hydrophobic patches on the surfaces of protein structures. *Proteins* 25 (3), 389–397. doi:10.1002/(SICI)1097-0134(199607)25:3<389::AID-PROT10>3.0.CO;2-E

Lyon, R. P., Bovee, T. D., Doronina, S. O., Burke, P. J., Hunter, J. H., Neff-LaFord, H. D., et al. (2015). Reducing hydrophobicity of homogeneous antibody-drug conjugates improves pharmacokinetics and therapeutic index. *Nat. Biotechnol.* 33 (7), 733–735. doi:10.1038/nbt.3212

MacCallum, J. L., and Tieleman, D. P. (2011). Hydrophobicity scales: A thermodynamic looking glass into lipid-protein interactions. *Trends biochem. Sci.* 36 (12), 653–662. doi:10.1016/j.tibs.2011.08.003

Mahn, A., Lienqueo, M. E., and Salgado, J. C. (2009). Methods of calculating protein hydrophobicity and their application in developing correlations to predict hydrophobic interaction chromatography retention. *J. Chromatogr. A* 1216 (10), 1838–1844. doi:10.1016/j.chroma.2008.11.089

Mahn, A., Zapata-Torres, G., and Asenjo, J. A. (2005). A theory of protein-resin interaction in hydrophobic interaction chromatography. *J. Chromatogr. A* 1066 (1-2), 81–88. doi:10.1016/j.chroma.2005.01.016

Maier, J. A., Martinez, C., Kasavajhala, K., Wickstrom, L., Hauser, K. E., and Simmerling, C. (2015). ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB. *J. Chem. Theory Comput.* 11 (8), 3696–3713. doi:10.1021/acs.jctc.5b00255

Manavalan, P., and Ponnuswamy, P. K. (1978). Hydrophobic character of amino-acid residues in globular proteins. *Nature* 275 (5681), 673–674. doi:10.1038/275673a0

McGibbon, R. T., Beauchamp, K. A., Harrigan, M. P., Klein, C., Swails, J. M., Hernandez, C. X., et al. (2015). MDTraj: A modern open library for the analysis of molecular dynamics trajectories. *Biophys. J.* 109 (8), 1528–1532. doi:10.1016/j.bpj.2015.08.015

Meek, J. L. (1980). Prediction of peptide retention times in high-pressure liquid-chromatography on the basis of amino-acid-composition. *Proc. Natl. Acad. Sci. U. S. A.* 77 (3), 1632–1636. doi:10.1073/pnas.77.3.1632

Miao, Y. L., Feher, V. A., and McCammon, J. A. (2015). Gaussian accelerated molecular dynamics: Unconstrained enhanced sampling and free energy calculation. *J. Chem. Theory Comput.* 11 (8), 3584–3595. doi:10.1021/acs.jctc.5b00436

Miao, Y. L., Sinko, W., Pierce, L., Bucher, D., Walker, R. C., and McCammon, J. A. (2014). Improved reweighting of accelerated molecular dynamics simulations for free energy calculation. *J. Chem. Theory Comput.* 10 (7), 2677–2689. doi:10.1021/ct500090q

Miyazawa, S., and Jernigan, R. L. (1985). Estimation of effective interresidue contact energies from protein crystal-structures - quasi-chemical approximation. *Macromolecules* 18 (3), 534–552. doi:10.1021/ma00145a039

Mullard, A. (2021). FDA approves 100th monoclonal antibody product. *Nat. Rev. Drug Discov.* 20 (7), 491–495. doi:10.1038/d41573-021-00079-7

Olsen, T. H., Boyles, F., and Deane, C. M. (2022). Observed antibody space: A diverse database of cleaned, annotated, and translated unpaired and paired antibody sequences. *Protein Sci.* 31 (1), 141–146. doi:10.1002/pro.4205

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., et al. (2011). Scikit-learn: Machine learning in Python. *J. Mach. Learn. Res.* 12, 2825–2830.

Sormanni, P., Aprile, F. A., and Vendruscolo, M. (2015). The CamSol method of rational design of protein mutants with enhanced solubility. *J. Mol. Biol.* 427 (2), 478–490. doi:10.1016/j.jmb.2014.09.026

Raybould, M. I. J., Marks, C., Krawczyk, K., Taddese, B., Nowak, J., Lewis, A. P., et al. (2019). Five computational developability guidelines for therapeutic antibody profiling. *Proc. Natl. Acad. Sci. U. S. A.* 116 (10), 4025–4030. doi:10.1073/pnas.1810576116

Raybould, M. I. J., Marks, C., Lewis, A. P., Shi, J. Y., Bujotzek, A., Taddese, B., et al. (2020). Thera-SAbDab: The therapeutic structural antibody database. *Nucleic Acids Res.* 48 (D1), D383–D388. doi:10.1093/nar/gkz827

Regep, C., Georges, G., Shi, J. Y., Popovic, B., and Deane, C. M. (2017). The H3 loop of antibodies shows unique structural characteristics. *Proteins* 85 (7), 1311–1318. doi:10.1002/prot.25291

Rekker, R. F., and Kort, H. M. D. (1979). Hydrophobic fragmental constant - extension to a 1000 data point set. *Eur. J. Med. Chem.* 14 (6), 479–488.

Rose, G. D., Geselowitz, A. R., Lesser, G. J., Lee, R. H., and Zehfus, M. H. (1985). Hydrophobicity of amino-acid residues in globular-proteins. *Science* 229 (4716), 834–838. doi:10.1126/science.4023714

Ruffolo, J. A., Sulam, J., and Gray, J. J. (2021). Antibody structure prediction using interpretable deep learning. *bioRxiv* 2005 (2027), 445982. doi:10.1101/2021.05.27.445982

Ryckaert, J.-P., Ciccotti, G., and Berendsen, H. J. C. (1977). Numerical integration of the cartesian equations of motion of a system with constraints: Molecular dynamics of n-alkanes. *J. Comput. Phys.* 23 (3), 327–341. doi:10.1016/0021-9991(77)90098-5

Sankar, K., Krystek, S., Carl, S., Day, T., and Maier, J. (2018). AggScore: Prediction of aggregation-prone regions in proteins based on the distribution of surface patches. *Proteins* 86 (11), 1147–1156. doi:10.1002/prot.25594

Schauperl, M., Podewitz, M., Waldner, B. J., and Liedl, K. R. (2016). Enthalpic and entropic contributions to hydrophobicity. *J. Chem. Theory Comput.* 12 (9), 4600–4610. doi:10.1021/acs.jctc.6b00422

Scott, A. M., Wolchok, J. D., and Old, L. J. (2012). Antibody therapy of cancer. *Nat. Rev. Cancer* 12 (4), 278–287. doi:10.1038/nrc3236

Shrake, A., and Rupley, J. A. (1973). Environment and exposure to solvent of protein atoms - lysozyme and insulin. *J. Mol. Biol.* 79 (2), 351–371. doi:10.1016/0022-2836(73)90011-9

Sievers, F., Barton, G., and Higgins, D. (2020). "Multiple sequence alignment," in *Bioinformatics*. Editors A. Baxevanis, G. Bader, and D. Wishart. 4 ed.

Sievers, F., and Higgins, D. G. (2018). Clustal Omega for making accurate alignments of many protein sequences. *Protein Sci.* 27 (1), 135–145. doi:10.1002/pro.3290

Sievers, F., Wilm, A., Dineen, D., Gibson, T. J., Karplus, K., Li, W. Z., et al. (2011). Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. *Mol. Syst. Biol.* 7, 539. doi:10.1038/msb.2011.75

Simm, S., Einloft, J., Mirus, O., and Schleiff, E. (2016). 50 years of amino acid hydrophobicity scales: Revisiting the capacity for peptide classification. *Biol. Res.* 49, 31. doi:10.1186/s40659-016-0092-5

Sormanni, P., Amery, L., Ekizoglou, S., Vendruscolo, M., and Popovic, B. (2017). Rapid and accurate *in silico* solubility screening of a monoclonal antibody library. *Sci. Rep.* 7 (1), 8200. doi:10.1038/s41598-017-07800-w

Southall, N. T., Dill, K. A., and Haymet, A. D. J. (2002). A view of the hydrophobic effect. *J. Phys. Chem. B* 106 (3), 521–533. doi:10.1021/jp015514e

Strohl, W. R., and Strohl, L. M. (2012). *Therapeutic antibody engineering: Current and future advances driving the strongest growth area in the pharmaceutical industry*. Oxford: WP/Woodhead Publishing.

Tartaglia, G., and Vendruscolo, M. (2008). The Zyggregator method for predicting protein aggregation propensities. *Chem. Soc. Rev.* 37 (7), 1395–1401. doi:10.1039/b706784b

Valjakka, J., Hemminki, A., Niemi, S., Soderlund, H., Takkinen, K., and Rouvinen, J. (2002). Crystal structure of an *in vitro* affinity- and specificity-matured anti-testosterone Fab in complex with testosterone. Improved affinity results from small structural changes within the variable domains. *J. Biol. Chem.* 277 (46), 44021–44027. doi:10.1074/jbc.M208392200

van der Maaten, L., and Hinton, G. (2008). Visualizing Data using t-SNE. *J. Mach. Learn. Res.* 9, 2579–2605.

van der Walt, S., Schonberger, J. L., Nunez-Iglesias, J., Boulogne, F., Warner, J. D., Yager, N., et al. (2014). scikit-image: image processing in Python. *Peerj* 2, e453. doi:10.7717/peerj.453

Voynov, V., Chennamsetty, N., Kayser, V., Helk, B., and Trout, B. (2009). Predictive tools for stabilization of therapeutic proteins. *Mabs* 1 (6), 580–582. doi:10.4161/mabs.1.6.9773

Waibl, F., Fernandez-Quintero, M. L., Kamenik, A. S., Kraml, J., Hofer, F., Kettenberger, H., et al. (2021). Conformational ensembles of antibodies determine their hydrophobicity. *Biophys. J.* 120 (1), 143–157. doi:10.1016/j.bpj.2020.11.010

Waibl, F., Kraml, J., Fernandez-Quintero, M. L., Loeffler, J. R., and Liedl, K. R. (2022). Explicit solvation thermodynamics in ionic solution: Extending grid inhomogeneous solvation theory to solvation free energy of salt-water mixtures. *J. Comput. Aided. Mol. Des.* 36 (2), 101–116. doi:10.1007/s10822-021-00429-y

Walsh, I., Seno, F., Tosatto, S., and Trovato, A. (2014). Pasta 2.0: An improved server for protein aggregation prediction. *Nucleic Acids Res.* 42 (W1), W301–W307. doi:10.1093/nar/gku399

Wang, G., Hahn, T., and Hubbuch, J. (2016). Water on hydrophobic surfaces: Mechanistic modeling of hydrophobic interaction chromatography. *J. Chromatogr. A* 1465, 71–78. doi:10.1016/j.chroma.2016.07.085

Wang, X. L., Das, T. K., Singh, S. K., and Kumar, S. (2009). Potential aggregation prone regions in biotherapeutics A survey of commercial monoclonal antibodies. *Mabs* 1 (3), 254–267. doi:10.4161/mabs. 1.3.8035

Wildman, S. A., and Crippen, G. M. (1999). Prediction of physicochemical parameters by atomic contributions. *J. Chem. Inf. Comput. Sci.* 39 (5), 868–873. doi:10.1021/ci990307l

Willbold, D., Strodel, B., Schroder, G. F., Hoyer, W., and Heise, H. (2021). Amyloid-type protein aggregation and prion-like properties of amyloids. *Chem. Rev.* 121 (13), 8285–8307. doi:10.1021/acs.chemrev.1c00196

Wimley, W. C., and White, S. H. (1996). Experimentally determined hydrophobicity scale for proteins at membrane interfaces. *Nat. Struct. Biol.* 3 (10), 842–848. doi:10.1038/nsb1096-842

Word, J. M., Lovell, S. C., Richardson, J. S., and Richardson, D. C. (1999). Asparagine and glutamine: Using hydrogen atom contacts in the choice of side-chain amide orientation. *J. Mol. Biol.* 285 (4), 1735–1747. doi:10.1006/jmbi.1998.2401

Yugandhar, K., and Gromiha, M. M. (2014). Protein-protein binding affinity prediction from amino acid sequence. *Bioinformatics* 30 (24), 3583–3589. doi:10.1093/bioinformatics/btu580

Zambrano, R., Jamroz, M., Szczasiuk, A., Pujols, J., Kmiecik, S., and Ventura, S. (2015). AGGRESCAN3D (A3D): Server for prediction of aggregation properties of protein structures. *Nucleic Acids Res.* 43 (W1), W306–W313. doi:10.1093/nar/gkv359