Review

Journal of the Brazilian Chemical Society

https://dx.doi.org/10.21577/0103-5053.20240102

J. Braz. Chem. Soc. **2024**, *35*, 10, e-20240102, 1-21

Intrinsically Disordered Malaria Antigens: An Overview of Structures, Dynamics and Molecular Simulation Opportunities

Pamella C. C. da Silva [a] ***and Leandro Martínez*** [*,a]

[a]*Instituto de Química e Centro para Computação em Engenharia e Ciências, Universidade Estadual de Campinas (UNICAMP), 13083-970 Campinas-SP, Brazil*

The proteome of *Plasmodium falciparum* (*Pf*) is abundant in intrinsically disordered proteins (IDPs). Their important roles in the malaria life cycle and the limitations of experimental and computational methods to study this class of proteins hinder the development of antimalarial drugs. At the same time, the growing interest in IDPs and theoretical tools suggest a path for their classification and functional understanding: searching databases with experimental notes and predictions of protein disorder, developing force fields to describe protein flexibility, and using molecular dynamics enhanced sampling techniques to properly sample the IDP conformational diversity. This review discusses possibilities of exploration of *Pf*'s IDPs and their availability in disordered-protein databases to foster molecular modeling studies. The large percentage of intrinsic disorder present in many antigens and their ability to interact with different targets, make these proteins a major class of interest in the area of drug and vaccine development.

Keywords: *Plasmodium falciparum*, molecular dynamics, enhanced sampling methods, antigen-antibody interactions, intrinsically disordered proteins

1. Introduction

The growing interest in the so-called intrinsically disordered proteins (IDPs) stems from their participation in various diseases, such as cardiovascular and neurodegenerative diseases, and cancer.[1] Abundant in nature, IDPs are also found in large proportions in viruses and parasites, having important roles in adhesion, invasion, and membrane localization in the human body.[2] This is the case of the proteins of the *Plasmodium falciparum* (*Pf*), the main agent causative of malaria worldwide.

By definition, IDPs are structurally heterogeneous and do not have a stable and well-defined secondary and tertiary structural elements under physiological conditions.[1,3,4] Their roles in cellular environments and dysfunction in liquid-liquid phase separation is related to debilitating diseases, as this class of proteins are in many biological processes, including heterochromatin formation, nucleus cytoplasmic transport, formation of membrane-free compartments and biomolecular condensates.[5,6] This class of proteins also have functions that are complementary to the roles performed by ordered proteins, such as entropic chains, molecular recognition, display sites and chaperones, which play a part in the folding mechanism of proteins and ribonucleic acids (RNAs).[7]

Significant proportions of *Plasmodium* spp. proteomes are filled with IDPs and disordered regions (IDRs). Computational studies indicate that *Pf* proteins have about 32.7% of disorder, the highest percentage compared to other *Plasmodium* species. This percentage is, on average, the same proportion found in key proteins in the host life cycle.[2,8]

Pf has a limited amino acid repertoire. Above 90% of all proteins on chromosomes 2 and 3 and half of all proteins of *Pf* are more than 60% composed of low complexity regions. The majority of these regions (ca. 90%) are composed of hydrophilic residues. Of these, 20% consist of iterated short oligonucleotides, also called tandem repeat regions (TRR), with areas of poly-asparagine single repeats. These characteristics are much more common in these organisms than in other eukaryotes.[9] Almost 80% of TRRs correspond to regions of disorder.[10] The abundance of IDRs in the *Pf* proteome is correlated with the enrichment of TRRs, since the expansion of these regions can drive the evolution of IDPs and *vice versa*.[11,12]

Feng *et al.*[8] analyzed the proteomes of 20 eukaryotes, 20 archaea, 22 bacteria, and four *Pf* in four life stages.

*e-mail: lmartine@unicamp.br
Editor handled this article: Carlos Maurício R. de Sant'Anna (Guest)

The apicomplexan phylum is unusual among eukaryotes as their proteins contain longer disordered regions. The *Pf* proteome at the sporozoite stage appears to be distinctive in the amount of long repeating regions compared to the other species and life stages of the malarial cycle. Among apicomplexan parasites, greater amounts of disordered regions are observed in mammalian *Plasmodium*. The intrinsic disorder in this species seems to be related to evolutionary mechanisms of immunological evasion and host interaction,[13] thus increasing the challenge of having these proteins as targets for vaccine development.

Despite the many challenges posed by the malaria lifecycle,[14-16] *Pf*-IDPs/IDRs' potential as targets for developing effective treatments and immune responses should not be overlooked.[17] However, understanding IDPs mechanism of interactions with antibodies is necessary. The promiscuous binding of many IDPs to other physiological macromolecules allows them to participate in interactions investigated as potential targets for pharmacological development, as protein-protein (PPI) and protein-deoxyribonucleic acid (DNA) interactions,[18,19] formation of fuzzy complexes or amyloid fibers.[20]

This review investigates the use of computational modeling, particularly molecular dynamics (MD) simulations, to understand the antigen-antibody interactions of *Pf* intrinsically disordered proteins. We review the last 30 years of protein structures and manually curated information available in databases as potential targets for leading vaccines, to understand the impact of research on these proteins on the understanding of the induction of an effective immune response. Because IDPs exhibit sets of loosely defined conformations, it is limitative for experimental techniques to express, purify, and characterize dynamic conformational ensembles, and to obtain consistent conformations of order/disorder.[21] Therefore, MD simulations and other molecular modeling studies on IDPs are reviewed, providing insights into the prediction of disorder, structure, and behavior of IDPs,[22] to complement and expand the scope of understanding of these proteins relative to experimental studies.

2. Intrinsically Disordered Proteins-Historical Background

Highly flexible proteins have been highlighted in the literature for at least 80 years, but it took many years for a consensus of terms and nomenclature to refer to these proteins to be achieved.[23,24] The lack of standardized definitions by researchers imposed a barrier to describing the structural properties of the proteins, grouping information and understand IDPs as a class of interest.[23]

As far as we could investigate, publications on Google Scholar adopting the "Intrinsically Disordered Proteins" terminology began to permeate the literature around 1990-1992.[25] At the end of the last century, however, it was well established that protein disorder was common and functionally relevant.[26] The "Intrinsic disorder" terminology was then consolidated for the understanding of protein function.[27,28]

A search on Google Scholar platform suggests that the IDP terminology has been increasingly adopted by researchers in the last 30 years. This bibliographic review considered articles published between 1990 and 2022 with the filters: (intrinsically OR natively OR inherently) AND (disordered OR unfolded OR unstructured OR flexible) AND (protein OR proteins), and among these papers, we discriminated those with the words "Molecular Dynamics Simulations" AND "malaria" in their scope. These were the choices of terms to avoid ambiguous results where molecular dynamics is used as a literal term and not to refer to a computational method.

Figure 1 indicates a growing interest in the scientific community in understanding the behavior of these proteins, and the contribution of computational works using MD to this field. Each year, the number of papers using IDP terminology grows. This may be due to factors such as the identification of these proteins as a class, improved analysis techniques and amount of high-resolution data on the structure and dynamics of disordered regions available.

The orange line in Figure 1 corresponding to MD related articles indicates around 15% of IDPs articles that also at least discuss MD simulations. The contribution to this field can be greater considering that not all papers adopt exactly the search term "Molecular Dynamics Simulations" in their scope, and that other computational methods can also be considered, showing the relevance of theoretical works for describing and studying these proteins. On average during

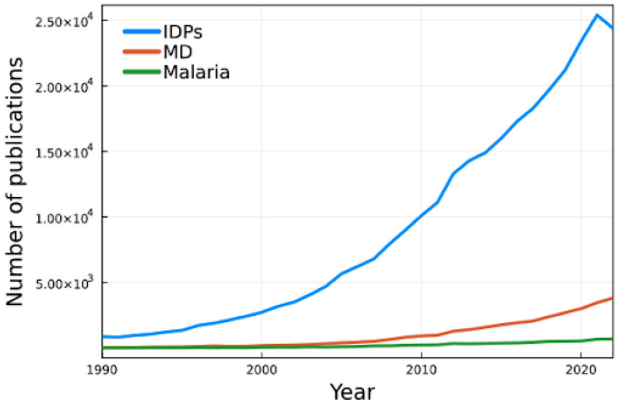

Figure 1. Number of publications from 1990 to 2022 using the "intrinsically disordered proteins" terminology in blue, and related to molecular dynamics, in orange, or malaria, in green.

this 30 years recap, the proportion of articles citing malaria and IDPs was approximately between 1.5 and 2.5% of all IDPs publications. Simulation studies on malaria IDPs, on the other hand, are still relatively poorly represented in this field, with only about 100 articles of the total set satisfying all search criteria in 2022.

3. Databases and Disorder Structure Prediction

The most widespread experimental methods for identifying and characterizing flexible or disordered structures are nuclear magnetic resonance (NMR) spectroscopy, circular dichroism spectroscopy, mass spectrometry, light scattering, fluorescence-based methods like fluorescence resonance energy transfer (FRET), and fluorescence correlation spectroscopy (FCS). Other methods include size exclusion chromatography, UV-Vis spectroscopy, atomic force microscopy (AFM), electron spin resonance (ESR) spectroscopy, fast relaxation imaging (FReI), and analytical ultracentrifugation.[29] In these methods, regions of disorder are observed by the absence of structure, missing signals that impair structure resolution, as in NMR and X-ray, or by the conformational dynamics of the IDPs identified by the methods in light emission, absorption peaks, or even hydrodynamics in chromatography. Also, more advanced bioinformatic tools for the prediction of disorder and function of IDPs are pointed out among the main motives for the growth of interest in the era of unstructured biology.[30] Molecular dynamics simulations can provide complementary information and a comprehensive atomic description of the structure and dynamics of the IDPs. These methods have advanced over the years, and a path can be proposed to guide the research on flexible structures and complement the information inferred or obtained from experimental methods.[31]

As indicated by Piovesan *et al.*,[32] there are two types of biological databases: deposition bases, which are repositories of primary data, such as Protein Data Bank (PDB)[33,34] and UniProt,[35,36] and knowledge bases, which aggregate and visualize data for different kinds of searches and information. Here we highlight the most cited IDP databases (IDPD) that provide specific annotations for IDPs and IDRs (Table 1). A large portion of the disorder information found in primary and frequently used databases has been inferred using experimental methods such as NMR and X-ray crystallography. All IDPD fall under the category of knowledge bases; they provide curated annotations of the most flexible or disordered regions that can be inferred from the absence of electronic density in crystallographic data or from the structural diversity within a set of structures obtained through NMR analysis.

The databases considered in the literature were: DisProt,[37] IDEAL,[38] MobiDB,[39] PED,[40] DIBS,[41] MFIB,[42] DisBind,[43] IDPsBind,[44] FUzDB[45] and D^2P^2.[46] This last one is fundamentally a structure prediction database. Here, we shall categorize the types of information available in the IDPD, such as (C): manually curated experimental annotations from partner databases, derived (D): annotations automatically derived from primary data, e.g., from PDB structures. There is still information obtained by homology (H): annotations propagated by aligning curated regions against ensemble and gene trees; and finally, prediction (P): annotations provided by running software tools from the sequence by evaluating the local amino acid composition. We also considered the type of data stored to understand the focus of the information in the database: sequence, binding regions, and structure. The input, which is a database or method used to cross-feed information, the type of annotation, and the source: all the tools available in the platform, which can be software, a third-party database, or the type of consensus number of entries. This comparison was performed in order to evaluate structures and sequences available for further studies in the context of malaria. IDPs of *Pfs* were searched in the IDPD to determine which ones are the most widely focused by researchers and which ones are less explored.

DisProt, MobiDB, and IDEAL are the three main IDPDs with the greatest amount of available data and scholarly relevance. The Protein Ensemble Database (PED) is concentrated on representing IDPs heterogeneity and dynamics. This is a database focused specifically on conformation ensembles. The conformational sets are produced using computational modeling techniques due to the constraints of NMR, fluorescence resonance energy transfer (FRET), and small-angle X-ray scattering (SAXS) methods. These conformational sets are then manually curated from the literature on disordered structures to ensure their validation.

Database of Disordered Binding Site (DIBS) and MFIB are complementary databases of interaction regions, with DIBS representing the disorder regions that interact with globular partners and MFIB the protein complexes formed by IDPs. These databases consider the PDB structures as evidence documentation.

The only structure of the *Pf* species available in DIBS is Atg8 (PDB:4EOY). It is a disordered motif, from the autophagy family protein, and its interacting peptide with a well-defined tertiary structure; autophagy is also important at the blood cycle since it is required in the transition from sporozoite to erythrocyte and other functions. The protein region involved in the interaction contains a known functional linear disordered motif determined by inference, and the Atg8 ordered domain involved in the interaction

Table 1. Annotations in intrinsically disordered proteins databases

Database	Data stored	Input	Evidence	Annotation	No. of entries	*Pf* entries
DisProt	sequence	UniProt MobiDB	C D	disorder disorder transition	proteins: 2000 residues: 211.300 disorder content: 18.6%	9
IDEAL	sequence	PDB UniProt	C D H	order/disorder	proteins: 995 residues: 33419	–
MobiDB	sequence binding regions structure	UniProtKB DisProt Ideal	C D H P	disorder missing residues mobile residues prediction of disorder	proteins: 189.525.031 residues: 64.278.461.608	5.386
PED	structure	integrative modeling (NMR, SAXS, FRET and MD)	C P	extended denatured coil folded	proteins: 12 residues: 204	–
DIBS	binding regions	DisProt IDEAL PDB UniProt Pfam	C D H	ordered disordered motif	complexes: 1576 sequences: 772	1
MFIB	binding regions	DisProt IDEAL Pubmed Pfam	C D		complexes: 205	–
DisBind	binding regions	PDB	C	disordered	sequences: 226 disordered regions: 428 all binding sites (BS): 4232 BS in disordered regions: 1396	–
IDPsBind	sequence binding regions	PDB	C	disorder	complexes: 9626	4
FuzDB	binding regions	Mobi PDB	C	fuzzy regions	complexes: 404 proteins: 362	–
D^2P^2	sequence	–	P	disorder MoRF regions	sequences 10.429.761	3

Protein Data Bank; NMR: nuclear magnetic resonance; SAXS: small-angle X-ray scattering; FRET: fluorescence resonance energy transfer; MD: molecular dynamics; C: manually curated experimental annotations; D: derived annotations; H: homology annotations; P: prediction annotations.

adopts a stable structure in isolation.[47] The annotations in the database show function, sequence, position of the alpha and beta helices, and related structures in the PDB.

IDPsBind is a database that shows the interaction sites of IDPS with ligands based on PDB structures and using the threshold distance method. There are four annotations of complexes with the *Pf* IDPs, triosephosphate isomerase (disorder content: 11.69%), LytB enzyme (disorder content: 42.8%), erythrocyte membrane protein 1 (*Pf*EMP1) (disorder content: 6.12%) and merozoite surface protein 2 (MSP2) (disorder content: 100%). In this database, sequences, regions of disorder and interaction with the ligand are identified.

IDEAL, PED, MFIB, DisBind and FuzDB have no stored data for *Pf* or even other malaria parasite species. DisProt is a database recognized in the literature for its clarity of information on disordered regions of proteins; it has 9 entries of *Pf* sequences, two of which are from the MSP2 protein. MobiDB indicates 5386 *Pf* annotations in proteome searches, but many of these sequences are not identified or may indicate ambiguity in annotations. When performing a search by organism, MobiDB returns 76 significant structures ranked by the platform's API method.[48] Figure 2 shows their percentage of citations in the literature in Google Scholar, as well as the percentage of disorder indicated in the IDPD. Finding proteins that are frequently mentioned in studies in the literature may be relevant with this type of search. In comparison to the other databases, MobiDB stands out for the amount of complete resources, combining manually curated

reports, experimentally derived information, and structure predictors, including the cutting-edge AlphaFold2 and cross references of the entire UniProtKB and the majority of the other IDPDs.[49]

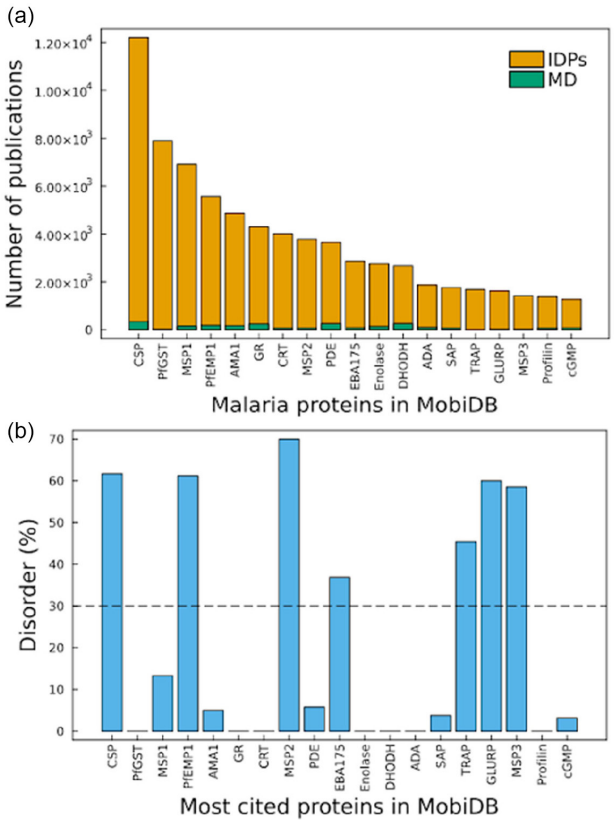

Figure 2. (a) IDPs of *Pf* available in MobiDB that are most often cited in the literature; in yellow, number of papers citing these proteins and in green, the papers that also have MD simulations in their scope for each protein; (b) the percentage of disorder of IDPs of *Pf* available in MobiDB. The dashed black line delimits the annotations with sections that have more than 30% intrinsic disorder.

The literature categorizes proteins with 30% or more disorder as IDPs.[50] We are classifying the *Pf* proteins in this search in publications and databases using this percentage disorder criterion, which indicates that more than 30% of the structural residues exhibit features associated with protein flexibility. Among proteins with annotations in MobiDB, circumsporozoite (CSP), glutathione S-transferase (*Pf*GST), MSP1, *Pf*EMP1, apical membrane antigen 1 (AMA1), glutathione, chloroquine-resistant transporter (CRT), and MSP2 are the most cited in the literature. The CSP, *Pf*EMP1, MSP2, erythrocyte-binding protein 175 (EBA-175), thrombospondin-related adhesive protein (TRAP), the glutamate-rich protein (GLURP), and MSP3 annotations available in MobiDB have a percentage above 30% disordered, and AMA1 and MSP1 have a considerable percentage of IDRs. We can see by the orange region of the bar graph, the quantity still unsatisfactory of the amount of MD studies involving these proteins.

The eight proteins that are listed in DisProt, their disorder content, and the number of cited papers are displayed in Table 2. The majority of them are crucial proteins from the malaria cycle's blood stage; Acp and GcpE are enzymes involved in the biosynthesis of fatty acids and isoprene, respectively.[51,52] *Pf*EMP1 is the most cited protein in articles, while MSP2 has the highest disorder percentage.

DisProt is the major database of manually curated IDPs data from the literature. A DisProt entry corresponds to a protein isoform and unambiguously maps to a UniProt entry. DisProt annotations describe local properties of the protein sequence supported by experimental evidence taken from the literature.[53]

This proliferation of IDPDs may provide more specific information and capture more subtle differences related to the functions of the IDPs. Some have different objectives and focuses that may be related to the interactions they carry out, integrate the maximum number of available tools, as is the case with MobiDB, and even present clarity in the annotation of clutter, as observed in DisProt. This facilitates the observation of consensus in the literature regarding disordered regions, observed and predicted data, molecular function, percentage of disorder, and cross-references, contributing to works that arise from this topic. It is noteworthy that these annotations still remain poorly represented, especially when referring to *Pf*, as manually curated data grows more slowly than experimental data.

Triosephosphate isomerase (*Pf*TIM) is a dimeric glycolytic enzyme of the erythrocytic phase responsible for catalyzing the isomerization of the reversible interconversion of D-glyceraldehyde-3-phosphate into dihydroxyacetone phosphate. *Pf*TIM also plays an important role in gluconeogenesis and fatty acid biosynthesis, being essential for the efficient energy production of virtually all organisms. In *Pf*, this phase of the life cycle is characterized by an increase in glycolytic flux, which is the parasite's main source of metabolic energy.[54]

MSP2 and MSP12 are some of the proteins that cover the merozoite surface (merozoite surface proteins). MSP2 is a GPI-anchored membrane protein that is present mainly in the stages prior to cell invasion. The MSP1 and MSP2 are the most abundant GPI-anchored proteins on the merozoite surface. Most MSPs are rapidly degraded after cellular invasion, but other proteins like the C-terminal portion of MSP1 (namely MSP1-19) and the MSP4 protein persist in the intracellular environment and may have roles in intra-erythrocytic development. MSP2 is supposed to play a role in the process of adhesion to the erythrocyte surface and

Table 2. Protein annotation available in DisProt

Protein	Available in Disprot	Life cycle	Disorder content / %	Papers in Google Scholar	Cross reference
*Pf*TIM	1	en	11.69	102	UniProtKB:Q07412 MobiDB:Q07412 AlphaFold: Q07412
MSP2	2	ab	100	3.780	UniProtKB:P50498 MobiDB:P50498 FuzDB: FC00107 AlphaFold: P50498
			82.58		UniProtKB:P19599 MobiDB:P19599 AlphaFold: P19599
*Pf*EMP1	1	ab	6.12	5.570	UniProtKB:Q25733 MobiDB:Q25733
MSP12	1	ab	13.83	5	UniProtKB:C6KSX0 MobiDB:C6KSX0 AlphaFold: C6KSX0
PTEX150	1	ab	84.29	291	UniProtKB:Q8ILA1 MobiDB:Q8ILA1 AlphaFold: Q8ILA1
AcP	1	en	3.79	1760	UniProtKB:Q8I2X3 MobiDB:Q8I2X3 AlphaFold: Q8I2X3
AMA1	1	ab	12.06	4870	UniProtKB:P50490 MobiDB:P50490 AlphaFold: P50490
GcpE	1	en	40.56	250	UniProtKB:Q8I295 MobiDB:Q8I295 AlphaFold: Q8I295

P*f*TIM: triosephosphate isomerase; MSP2: merozoite surface protein 2; PTEX: *Plasmodium* translocon of exported proteins; AcP: acyl carrier protein; AMA1: apical membrane antigen 1; GcpE: (*E*)-4-hydroxy-3-methylbut-2-enyl diphosphate synthase; ab: asexual blood life cycle; en: enzyme.

have a high propensity for fibril formation, in which the conserved N-terminal domain has a key role.[55]

MSP12 is the archetypal member of the 6-Cys protein family, containing only two s48/45 domains, while other members have as many as 14 such domains. Members of the 6-Cys s48/45 protein family are found on the surface of *Pf* at all stages, *Pf*12 is highly conserved and plays an important functional role in parasite-host cell adhesion or invasion. Studies[56] show that this protein is strongly recognized by immune sera from naturally infected patients.

Plasmodium translocon of exported proteins (PTEX) is essential for the transport of malarial effector proteins across a vacuolar membrane surrounding the parasite to host erythrocytes, but the mechanism of this process remains unknown. PTEX150 is an endogenous, isolated region of PTEX where the EXP2 and PTEX150 structures interdigitate to form a static, funnel-shaped, pseudo-seven-symmetrical, and symmetrical protein conduction channel spanning the vacuolar membrane. These transport proteins are the only known entry point into the host cell for exported proteins and are an attractive target for drugs, as disrupting them blocks delivery of key virulence determinants and induces parasite death.[57]

AMA1 is a surface protein found in more than one stage of the life cycle, acting on sporozoites and merozoites, closely related to parasitic invasion of erythrocytes and hepatocytes, respectively. It is a reorientation protein in the so-called "moving junction", the moment in which the parasite and host membranes are in close contact. Adhesion of the parasite to human cells is followed by a reorientation movement necessary for the apical pole to come into contact with the erythrocyte or hepatocyte surface. During the junction movement, AMA1 acts together with the peptide RON (Rhoptry Neck), an interaction of great importance, since the interruption of the adhesion of AMA1 to the RON prevents the invasion of erythrocytes by the parasite.[58]

Predictors of structural disorder also assist in structural clarification. Frequently, the most flexible regions are subject to modeling studies and MD simulations,[59] that can be combined with low-angle X-ray scattering or NMR data to propose a collection of disordered

conformations.[60] Predictions of IDRs are made from the primary protein sequences, based on the observed enrichment of polar and charged amino acids, and the lower predominance of hydrophobic and aromatic residues. Currently, there are predictors based on neural networks, such as: PONDR (Predictor of Naturally Disordered Regions;[61] VLXT,[62] PONDR VSL2,[63] PONDR VL3,[64] PONDR-FIT,[65] DisEMBL[66] and others based on the physicochemical properties of proteins, which can be highlighted: IUPred,[67] FoldIndex,[68] TopIDP,[69] MobiDB (cited in the database section), PrDOS,[70] MetaDisorder,[71] and GlobPlot.[72] DISOPRED3,[73] part of PSIPRED,[74] was trained on evolutionarily conserved sequence features of missing residue IDRs in X-ray structures and later added two independent predictors of intrinsic disorder, a module that combines intermediate results, and a component that annotates protein-binding IDRs.[73] All predictors, although they still have limitations for generating disordered structural sets, are tested by the Critical Assessment of Techniques for Protein Structure Prediction (CASP) which evaluates the prediction of folded protein structures, making it possible to identify disordered regions.

It is noteworthy that even if these predictors have been trained to identify intrinsic disorder, AlphaFold2 (AF2) was the best ranked so far in CASP14.[75] AF2, as the state-of-the art of protein prediction in the era of machine learning and artificial intelligence in computational chemistry, also brings relevant questions in the field of IDPs. In the context of malaria, a search by taxonomy indicates that the PDB has 1044 resolved *Pf* structures, while AF2 has about 30222 predicted structures of different *Pf* proteins.

The predicted local distance difference test (pLDDT), an AF2 structure scoring measure, classifies a measurement below 50 as very low (orange regions in disorder predictors), between 50 and 70 as low (yellow regions), from 70 to 90 as reliable (light blue), and above 90 as high confidence (dark blue). As discussed by the authors,[76,77] it is possible that regions with very low confidence may indicate that AF2 failed to predict properly or may correspond to regions without a well defined tertiary structure, indicating intrinsic disorder. Many researchers are adopting the second,[78,79] most optimistic, interpretation, due to the disorder percentage in human proteome 37-50% coincides with the fraction of disordered regions in which AF2 classifies as regions of low or very low confidence in prediction. Of course, inferred information or non-resolution of specific regions should not be overestimated as mandatorily a sign of intrinsic disorder. Due to this, for predicted structures, there has been a growing adoption of "resolved/ordered" for regions of high confidence and "unresolved/disordered" for regions of low confidence.[80] pLDDT proved to be a competitive disorder prediction algorithm, having even better results than the predictors conventionally used for IDPs and cited above.[81]

Even more recently, focused on the challenges of predicting regions of intrinsic disorder, the Critical Assessment of Protein Intrinsic Disorder Prediction (CAID) was designed,[82] which is already in its second round.[83] It is a community-based blind study to determine the state of the art in predicting intrinsically disordered regions and the subset of residues involved. Some predictors widely cited in the literature, such as DISOPRED3[73] and IUPred-long,[67] were also in the test. The servers assign scores to each residue to be disordered using DisProt as the main database.

The results also show room for improvement in algorithms focused on IDPs and binding regions. As indicated by the authors, the best methods use deep learning, and there is a need to balance performance with execution time. The methods based on AF2 proved to be good, but they are better at predicting the absence of order than the IDRs indicated in DisProt.

The PDB can assist the recognition of IDRs by the inference of the missing residues from the experimental methods, and more recently based in disordered sequence annotations imported from IUpred2. Also, with AF2 consolidation, more than 1 million computational models are available in PDB allowing the low confidence or disordered regions to be taken into account.

The emergence of increasingly precise methods to study IDPs puts into perspective the need to understand the dynamics of these proteins. Proteins are structures that have complex dynamics and complementary functions between rigid and flexible regions.

4. Epitope Recognition and Antigen-Antibody Interactions

The structural and dynamical factors that play a role in the identification of intrinsically disordered epitopes by their homologous antibodies are still not well understood. The effectiveness of the immune response will depend on how well the available antibody repertoire binds to the accessible epitopes on the surface of the invader, since the molecular recognition process is essential for the neutralization of the pathogen.[84]

In an ideal situation of maximum affinity, the conformation of the epitope in the complex is a perfect structural mold of the cognate antibody binding site (paratope), with a complementary surface charge distribution. So, conformational flexibility is also a determining factor for the affinity and specificity of an

antigen-antibody complex. Disordered regions usually expose consecutive amino acid sequences, which are identified as linear epitopes, unlike folded proteins that expose non-consecutive residues, called discontinuous or conformational epitopes.[85] Disordered epitopes generally have high specificity but low affinity interactions due to the smaller and more specific number of amino acids that can form intermolecular interaction.[86]

When bound to antibodies, linear epitopes can assume conformations that are molded by the interaction sites, and possess a greater density of intermolecular polar contacts. This results from the flexibility of disordered proteins and their consequent ability to better adapt to the shape of their binding partner.[87] Even so, if IDPs features mean an increase in the immune response, it is a subject of discussion.

Several B cell epitopes have been determined to be disordered with the ability to induce personalized immune responses. Furthermore, application of a large dataset of protein antigens revealed that disordered epitopes are authentic targets for recognition by antibodies.[17,88]

The greater exposure and contacts of the linear epitopes chains also relates the solvent-accessible surface area to antigenicity in these cases. The loss of conformational freedom following the interaction has an entropic cost. Due to the desolvation of the proteins and ligands involved, a large change in solvent entropy for IDPs is expected during disorder-to-order transitions, in which a substantial amount of solvent-accessible surface can be buried. Disordered epitopes tend to be recognized by concave parotopes, and the ratio of buried surface area on the antibody to buried surface area on the antigen is lower for this type epitopes, suggesting a more intimate interaction between these antigens and their complementary parotope.[89]

In this case of a mostly hydrophobic buried surface, the change in entropy of the solvent will be very favorable for binding. The resulting increase in solvent entropy will counteract the concomitant decrease in conformational entropy. This phenomenon gives IDP greater freedom to adjust its binding affinities and interact with different targets.[90] Because of this, Dogan *et al.*[90] suggest that specificity and affinity of interaction should be seen on a case-by-case basis for each particular IDP.

In complex formations, IDPs often display a combination of interaction mechanisms. Proteins can follow different binding trajectories that involve multiple steps with energetic barriers. Similarly to what happens in the interaction between proteins and small molecules, the greater the flexibility of the ligand (in this case the epitope), the greater its ability to undergo induced-fit. In IDPs, the separation between the interaction mechanisms is not clear. Conformational selection can be followed by induced fit and mutual adjustments between participating components.[90,91]

The antigenic site of the antibody, the paratope, is formed by the so-called complementarity-determining regions (CDRs). Antigen binding affinity is an important measure of B cell fitness and is largely determined by the structure and composition of the antibody CDRs. Six loops contribute to the binding site, three loops from the heavy chain CDR-H1, CDR-H2 and CDR-H3, and other three from the light chain CDR-L1, CDR-L2 and CDR-L3.[92] CDRs can be identified in fixed regions due to the definition of a canonical structure from the loop length and conserved residues.[93] The CDR-H3 is a loop in which an exception behavior is expected. This loop has a greater variability and plurality of contacts in many observations in the literature. CDR-H3 performs an important mediation in the antigen-antibody interaction, and can adopt a variety of conformations.[94,95]

Antibodies can interact with antigens of different formats.[96] In parallel, Jeliazkov *et al.*,[97] and Lawson *et al.*,[98] showed that the rigidity of antibodies increases with maturation, especially when they are repeatedly exposed to the same antigen. This rigidification evolves providing contacts that can interact quickly with the antigen, forming a three-dimensional parotope. Antibodies can also present protein promiscuity related to flexibility and shallower free energy surfaces, which makes different conformational states accessible on a shorter time scale.[99] Polyspecificity is essentially achieved through minimal structural rearrangement in the paratope complemented with plasticity in the interaction with the antigen.[100]

Pf exhibits enormous genetic variability to escape immune pressure. The 3D structures of the conserved amino acid sequences of the most relevant proteins in host cell invasion can be found in different areas, sometimes 90° to 180° opposite the location of the immunogenic and highly polymorphic regions, making the immune response useless.[101]

Other scaping mechanisms include: different stages of development where the *Pf* changes its morphology and protein functions in the malaria's complex life stages in very short periods of time. Because of this, the occurrence of effective humoral immunity acquired by natural infection is uncertain and still dependent on the molecular redundancy and high polymorphism of these disordered structures.[102]

Protein plasticity and promiscuity, for instance, can facilitate parasite adherence and invasion in the host-parasite relationship. During the *Pf* parasite invasion, molecules can hide or cover functionally relevant structures or domains with some other domains or structures to expose

them only during the last moments of invasion to carry out their fundamental functions.[103]

It is also possible that, as *Pf* IDPs perform functions of invasion of human cells, the challenges in the development of antimalarial treatments are also associated with the characteristics of these proteins. Various studies[89,104] suggest that disordered epitopes can function as a "smoke screen", moving antibodies away from regions that provoke an effective response, guaranteeing the survival of the parasite. The repetitive and disordered residues are antigenic, leading to humoral responses in non-functional regions to combat the disease.

Disordered antigens widely known in the context of malaria, such as MSP2 and CSP, have their continuous epitopes recognized by antibodies.[105,106] The sole approved vaccine against malaria is formulated with RTS-S, a recombinant protein that also has disordered regions of the CSP protein in its constitution.[107,108] Most of the protective antibodies against *Pf*CSP known to date interact with the central Asn-Ala-Asn-Pro (NANP) repeat region. This is also observed in the work of Murugan *et al*.,[105] in which they characterized 200 human monoclonal antibodies induced by immunization with sporozoites, and established that the most potent antibodies bind around a conserved nucleus.[109]

In disordered antigens, conserved epitopes are flanked by disordered polymorphic regions. Transient interactions between these regions of the MSP2 protein and host antibodies may explain the difficulty in establishing broad neutralizing responses to conserved antigens and clarify the possible benefit of these regions for *Pf* survival in the human host. Furthermore, the conformational changes of these epitopes are also related to their antigenicity.[106,110]

The monoclonal antibody 6D8 recognizes a completely conserved nine-residue continuous epitope within the intrinsically disordered malaria antigen, MSP2, but has different affinities for the two allelic forms of this antigen. The work of Krishnarjuna *et al.*[86] also reveals from DM simulations the presence of transient interactions involving polymorphic residues immediately C-terminal to the structurally defined epitope that also involve in multiple transient interactions distributed over a large part of the antibody's accessible surface. Specificity in these cases arises as a consequence of subtle differences in highly dynamic interactions. These findings highlight the importance of determining and choosing *Pf* antigens that can elicit protective immune responses.

Also, changes in surface protein expression at each new stage of *Pf* life cycle require the immune system to produce antibodies that can efficiently be recognized in the different stages of infection prevention, disease prevention, and transmission prevention. Julien and Wardemann[111] studied the different actions of antibodies in each stage. Antibodies that interfere with the development of the asexual erythrocytic stages of *Pf* recognize merozoites or variant surface antigens that are displayed in button-like structures on the surface of infected erythrocytes, such as the blood life cycle proteins, AMA1, MSP1, MSP2, MSP3, EBA-175, GLURP, and Reticulocyte-binding protein Homolog 5 (RH5), available in the IDPDs.[84]

Thus, understanding whether regions of disorder can induce an effective immune response is difficult and prominent for the development of effective vaccines.[89] However, the presence of regions of low complexity and the ability to change conformation should not remove disordered epitopes from the list of candidates for peptide vaccine development. It is possible to say that if a solution is found for the use of IDPs or IDRs in vaccine development, it is possible that the final vaccine efficiency will be improved.

In this aspect, the use of a mixture of ordered and disordered epitopes in a peptide vaccine construct can modulate the recognition of disordered epitopes by antibodies and also overcome overlapping antigenic regions between different strains, supporting the construction of multi-allele vaccines.[10]

As of November 2023, the PDB contains 117 models, of which 12 different *Pf* proteins in complex with antibodies available for affinity studies, they are: Cysteine-Rich Protective Ag (CyRPA), AMA1, CSP, EBA-175, MSP1, MSP2, *Pf*s25, Subtilisin-like Protease 1 (*Pf*SUB1), RH5, *Pf*s48/45, Fab668, *Pf*s230, RIFIN, CRT, EMP1, and *Pf*s230D1M. Table 3 shows experimental PDB structures and epitopes indicated on the domains that are complexed with antibodies. The RAPID server[112] was performed to provide the disorder percentage of the proteins. This server correlates the amino acid sequence with the protein disorder considering the most commonly observed amino acids in regions of protein flexibility. Some peptides, such as the NANP of the CSP and MSP2 proteins, are very short, and not allowing for a proper disorder prediction.

In Table 3 available protein domains have a high percentage of disordered regions (ca. 20 to 30%): AMA1, *Pf*s230, EMP1, *Pf*s230D1M, and The CSP accessible domains are entirely disordered.

In Table 3 as far as has been possible to research, only a few of these PDB IDs have been used in theoretical studies. Some of them are not even focused on malaria as an object of study. In literature, we found that 2J5L, 6B0S, 1OB1, and 5MZA were studied using MD.

The 2J5L structure was used in molecular docking and MD for druggability filtering,[149] and also for verification of the mimicked potential of interactions.[150] This study used protein docking simulations and sequence alignment tools for

Table 3. Structures of the *Pf* antigens in complex with antibodies available in the PDB

Pf antigen	Epitope containing region	Antigen disorder / %	PDBs IDs	Type of antibody	Reference
CyRPA	cysteine-rich protective antigen 6 bladed domain	0.9	5TIH	mouse	113
	cysteine-rich protective antigen 6 bladed domain	10.36	5EZO	mouse	114
	–	1.17	7PHW, 7PI7, 7PI2, 7PHV, 7PI3	mouse	115
AMA1	domain I	5.65	2Q8B, 2Q8A	mouse	116
	domain III	10.0	2J5L	mouse	117
CSP	C-terminal a TSR domain	23.1	6B0S	human	118
	NANP repeat region	–	5BKO	human	–
	–	99.9	6AZM	human	119
	–		6AXK, 6AXL	human	120
	NANP5	99.9	6D11, 6D01	human	121
	–	99.9	6ULE	human	105
	–	50.0	6MB3	human	122
	–	99.9	6UUD	mouse	123
	–	50.0	6MHG	human	122
	NPNA2 peptide, B domain	–	6WFW, 6W00, 6W05, 6WFW, 6WFX, 6WG1, 6WG2	human	124
	NANP4	–	6WFY	human	124
	NANP3	–	6WFZ, 6WG0	human	124
	NANP3	–	6D0X	human	121
	RTS,S/AS01	–	6UC5	human	125
	–	93.33	7RD9, 7LKB, 7LKG, 7RAJ, 7RCS, 7RD3, 7RD4, 7RD9, 7RDA	mouse	126
	–	36.36	7RXL, 7RXI, 7RXJ, 7RXL, 7RXP, 7S0X	human	127
	–	–	6B5L, 6B5M, 6B5N, 6B5O, 6B5P, 6B5R, 6B5S, 6B5T	human	128
	NANP5	99.9	6O23, 6O23, 6O24, 6O25, 6O26, 6O28, 6O29, 6O2A, 6O2B, 6O2C, 6ULE, 6ULF, 6VLN	human	105
	–	–	7K75	human	–
EBA-175	region ii Duffy binding domain	12.13	4QEX, 4K2U	mouse	129
MSP1	MSP1-19	6.06	1OB1	mouse	130
	MSP1 EGF domain 1	6.06	6XQW	human	131
MSP2	C terminal	–	5TBD, 4067	mouse	132
	N terminal	–	4QXT, 4QY8, 4QYO, 4R3S	mouse	106
*Pf*s25	Pvs28 EGF domain	1.64	6B08, 6B0A, 6AZZ, 6B0E, 6B0H, 6B0G	human	133
	Pvs28 EGF domain	2.17	6PHC, 6PHB, 6PHC, 6PHD, 6PHF	human	134
*Pf*SUB1	SUB1 protease Prodomain ProdP9	1.74	4LVN, 4LVO	mouse	135
RH5	Rh5 coiled-coil domain	11.67	5MI0	mouse	136
	3D7	16.17	7PHU	human	115
	Rh5 coiled-coil domain	11.03	4U0Q, 4U0R, 4U1G	human	137
	Rh5 coiled-coil domain	16.17	6RCV, 6RCS, 6RCU, 6RCV	human	138

Table 3. Structures of the *Pf* antigens in complex with antibodies available in the PDB (cont.)

Pf antigen	Epitope containing region	Antigen disorder / %	PDBs IDs	Type of antibody	Reference
*Pf*s48/45	sexual stage antigen s48/45 domain	5.88	6H5N	mouse	139
	6C variant	6.12	7UNB	human	140
	sexual stage antigen s48/45 domain	6.47	6E62, 6E63	mouse	141
	–	3.03	7ZWI	mouse	142
Fab668	junctional peptide		6PBV	human	109
*Pf*s230	–	20.51	7JUM	human	143
RIFIN	V2 domain	17.22	7KHF, 7KFK	human	144
	immunoglobulin domain	16.56	7F9L, 7F9M, 7F9N	human	145
CRT	CRT-like	6.25	6UKJ	human	146
EMP1	Duffy binding domain	21.31	5MZA	human	147
*Pf*s230D1M	–	20.51	6OHG	mouse	148

PDB: Protein Data Bank.

comparisons of protein models structurally recognized by the motifs connections to the receiver. Fernández-Quintero *et al*.,[151] performed the MD of a set of structures, including the PDB ID 6B0S to characterize conformational changes of antibodies and antigen upon binding. In this paper are shown the state probabilities obtained by the Markov-state model based on the docking score. 1OB1 was the initial structure of a MD and docking simulation to understand the interactions of the antigen-antibody complex.[152]

5MZA was used in a rational design of peptide-ligand conjugates focused in malaria.[153] The virtual screening revealed three bioactive natural ligands for *Pf*EMP1 from (NPASS) database, and nine peptide-ligand conjugates were designed with different combinations of peptides and ligands with the suitable non-cleavable triazole linker.

Considering that there are structural models for 117 proteins in the PDB, there is still much to be explored in the field of computational studies to understand the dynamics of malaria proteins. For the most part, the MD simulations are not the main focus of the work in which they are cited. Antigens behavior and interactions with antibodies are crucial issues for the development of vaccines. Several malarial proteins remain potential targets for drug discovery studies. The barriers in analyzing IDPs to be overcome are still issues that need to be addressed.[154]

5. Computational Approaches

Molecular dynamics simulation (MD) methods are important in pharmacological target discovery and vaccine design applications because protein flexibility crucially affects the range of possible target conformational states.[155] In many studies, MD has been used to examine the conformational balance of the H1, H2, and H3 (CDR heavy chains) and the L1, L2, and L3 (CDRs light chains), the conformational freedom of these flexible regions as well as structural variability related to their interactions with antigens.[156,157] MD techniques are used[158] to obtain better conformational sampling particularly of the CDR-H3 and enhanced sampling can aid in obtaining putative structural models.[159,160]

Moreover, molecular and ensemble docking[161] utilize MD to generate conformational sets of structures and calculate complexed pairs configurations in order to anticipate antigen-antibody complexes.[162] Via distinct MD simulations, several antigen and antibody conformations can be produced and subjected to hard docking calculations specially where conformational modifications occur prior to antigen-antibody binding.[163] Also in drug discovery and design of small molecules that target IDPs, conventional MD simulations can be combined with enhanced sampling to generate the conformational ensemble and than carry out conventional virtual screening for highly populated conformations.[155,164]

A molecular dynamics simulation generates a sequence of configurations for an atomic system using an effective interaction potential and, most frequently, the new classical equations of motion. Simulations also require specific initial configurations (momenta and coordinates of the atoms), which can, for proteins, obtained from experimental or computational models.[165] MD simulations are extremely important to sample the configurational space of biomolecular systems and have been extensively used over the past few decades for understanding biomolecular function.[166] Because of the dynamic nature of IDPs, they are best described by an ensemble of structures with multiple thermally accessible conformations. These structures can frequently be simulated with molecular mechanics methods, providing a rich view of the IDP conformational space.[167]

Given the uncommon nature of IDP flexibilities among other proteins, the force field biases or imprecisions can be potentially amplified, as the interaction potentials were mostly parameterized to reproduce the structure and flexibility of folded proteins. Reparameterization strategies including water and solvent models have been adopted to better characterize ensembles without a well-defined tertiary structure. Specific force fields were designed and validated for IDPs in the last decade.[168,169] The choice of solvent also plays a significant role in the dynamics of IDPs, particularly because of the high sensitivity of the conformational equilibrium to protein-water interaction details,[170] associated with the absence of robust hydrophobic centers, and the greater solvent accessible surface.[171]

Widely used force fields as Amber ff99SB,[172] ff03,[173] CHARMM (Chemistry at Harvard Molecular Mechanics-CHARMM22/27,[166] CHARMM36),[174] GROMOS96 (Groningen Molecular Simulation)[175] and OPLS (Optimized Potentials for Liquid Simulations)[176] were developed for structured proteins. Later, implementations for protein flexibility and IDPs were released, such as the CHARMM36m,[177] ff99IDPs[178] and ff14IDPs,[179] which are improvements of CHARMM36, ff99SB and ff14SB, respectively.[180] The main correction in force fields for IDPs includes the backbone dihedral distribution, and the overestimation of the population of secondary structures.

In conventional MDs, a very long time simulation would be required to obtain a proper sampling of IDPs conformational changes.[181] The computational costs for exploring the high conformational diversity of these proteins is very high, and information on the local fluctuations of a single structure is not generally useful. Observable quantities and IDP properties are inherently averages over a highly heterogeneous set of structures, this makes enhanced sampling methods even more necessary than in usual MD simulations studies of folded proteins.[182]

There is indeed some consensus that poor sampling is a greater source of error for MD predictions than the deficiencies of the force fields.[183] At the same time, Hénin *et al.*[184] indicate that the efficiency of enhanced sampling methods depends on the application, the choice of parameters, and the researcher's expertise.

Enhanced sampling methods can be classified by whether or not they adopt collective variables (CV), and by exploring the free energy landscape along the reaction coordinates defined by the CVs. CV-free methods sample biomolecular configurations without a priori knowledge of the path connecting the target states; they are useful for exploring possible structural transition pathways and determining possible unknown intermediate states of systems. In these methods, a random walk in the potential energy space is performed to overcome energy barriers.[185] For example, the methods can exploit thermal fluctuations to increase the frequency of overcoming free energy barriers. Among popular CV-free methods, we can mention in particular Replica Exchange Molecular Dynamics (REMD)[186] and its derivatives Replica Exchange Solute Tempering,[187] Hamiltonian REMD,[188] or methods that modify potential barriers to accelerate the dynamics or accelerate natural system motions, like Accelerated Molecular Dynamics[189] and Self-guided Molecular Dynamics.[190]

In CV-based methods forces are added to the system to constrain or guide the motions in the direction of interest. These techniques can exploit knowledge of the one or both end points of the trajectory states to find minimum free energy paths. Umbrella Sampling,[191] Steered Molecular Dynamics,[192] Adaptive Biasing Force (ABF),[193] Thermodynamic Integration, and Free-Energy Perturbation (FEP),[194] and Metadynamics,[195] and Local-Elevation,[196] are CV-based strategies.

Some strategies compute the forces to be incremented in the system from the sampling of the simulation itself. For instance, ABF consists of a potential arising from the average force that adapts to the resistance that the medium exerts to the evolution of the system along the CVs of choice, similarly in this aspect to Metadynamics and Local-Elevation strategies. Some of these strategies can be applied in non-equilibrium regimes, forcing the system to follow a predetermined CV or an alchemical parameter.[184] These non-equilibrium simulations can be used to obtain qualitative insights into the conformational transformations of interest, to provide initial approximations for equilibrium strategies, or to obtain equilibrium free-energies through the Jarzyinski equality.[197] Biased simulations favor the sampling along predefined coordinates, and from the knowledge of the biasing potentials introduces it is possible to recover the populations of the conformations in the non-perturbed ensemble.[198]

Metadynamics and REMD are used in different flavors according to the system to be analyzed and the necessary application. Furthermore, it is common to use hybrid methods, the combination of methods with enhanced sampling with different principles.[199] It is important to highlight that the choice of methods also depends on the available computational power. In this work we will highlight the exchange of Hamiltonian replicas, which has been one the most successful techniques for the non-biased sampling of conformations of IDPs.

In REMD, changes between parallel simulation replicas at different temperatures are made periodically and accepted or rejected by the Metropolis criterion.[200]

The conformations of the system accessed at different temperatures are exchanged, and the velocities are rescaled to those expected at the new temperature condition.[201] Thus, an advantage of REMD over other generalized ensemble methods is that the exchange probability is quickly determined by the characteristics of the system and its movements as accepted by known statistical weights.[184]

Among the most used REMD variants, we can highlight the exchange of replicas by temperature modification (T-REMD), the Hamiltonian REMD (H-REMD).[202] H-REMD is a more general form of replica sampling involving exchanges between different Hamiltonians; whereas T-REMD accelerates sampling just by varying the temperature parameter. H-REMD has the advantage of improving in dimensions other than temperature with scaling potentials, and the possibility of accelerating subsets of the system without introducing kinetic energy redistribution issues.

US was one of the earliest techniques used to keep track of all atoms' representations while controlling the force field to focus sampling on the relevant areas of the conformational space. US uses a biasing, usually harmonic, potential along a reaction coordinate that restricts the system in an intermediate state of the CV of interest. In this way, a proper sampling of nearby states of the constrained coordinate is obtained. Intermediate steps are covered by a series of windows, in each of which an MD simulation is performed. From the sampled distribution in the system along a reaction coordinate, the free energy changes in each window can be calculated.[203] The windows are then combined by weighted histogram analysis (WHAM)[204] or umbrella integration methods.[184,205,206] The binding free energy can be extracted from the obtained Potential Mean Force (PMF).[207]

Metadynamics and Local Energy Elevation and similar methods that work by introducing energy bumps in the most visited regions of the energy landscape, along a CV of interest. They avoid visited states from being excessively resampled, making it very applicable in systems where ergodicity is impaired by the shape of the energy landscape.[208] Computational resources are directed to search the entire free energy landscape without the need for a very precise description of the potential energy surface, which has advantages that are well applicable to protein folding problems, phase transitions, and conformational changes.[209] An advantage of Local Elevation and Metadynamics over methods such as US is that multiple CVs can be used without the need for a priori knowledge of the final states.[210] Hybrid methods are also frequently applied in these systems, usually using REMD and external biasing potential methods.[184]

Enhanced sampling simulation MD methods are efficiently suitable for studies with *Pf* proteins, but the number of articles that apply them to malaria proteins is still very limited. Despite this, on Google Scholar, we can observe that in the last 30 years there has been a large increase in the number of works using enhanced sampling methods to study IDPs. Figure 3 shows the number of articles that over the years use commonly known enhanced sampling methods in their scope: metadynamics, Umbrella sampling, REMD and IDPs terminology adopted in this review: (intrinsically OR natively OR inherently) AND (disordered OR unfolded OR unstructured OR flexible) AND (protein OR proteins). It is observed that the exchange of replicas is the most used among these, while US and Metadynamics have been experiencing almost equal growth over the years.

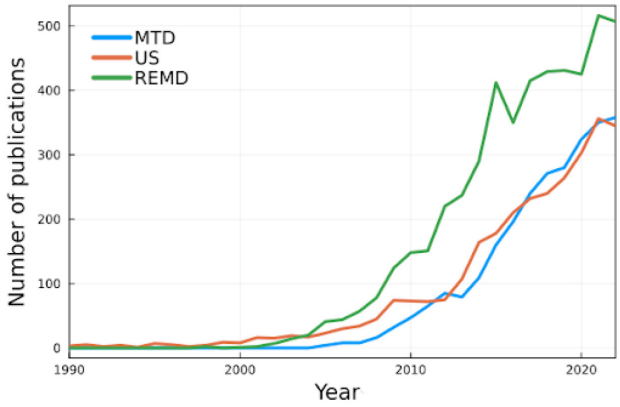

Figure 3. Number of publications from 1990 to 2022 applying enhanced sampling methods to study IDPs: Metadynamics (blue), US (orange) and REMD (green).

Published research that focuses on *Pf* proteins are mostly experimental studies. Among the articles that use enhanced sampling methods, a timid presence of *Pf* IDPs in their scope can be highlighted, appearing as additional information, examples of cross-reaction, or even being used for modeling studies. A bibliographic search of enhanced sampling methods in *Pf* IDPs ends up resulting in false results since the proteins are present in the article keywords but not with the objective of MD application with robust methods and interest in understanding the structural mechanisms of these proteins.

For the most part, MD is only used as an accessory tool within an experimental study, and not used as a robust method for obtaining the microscopic molecular behavior of the proteins.

As the computational study is also an important step in drug discovery, MD with enhanced sampling has much to contribute to the field of IDPs and the development of antimalarial drugs in understanding the molecular

atomic behavior of these highly flexible proteins, their mobility, interaction mechanisms, affinity, flexibility, and intermolecular interactions.

Considering the large amount of IDPs present in the *Pf* in key regions of the life cycle for the development of an effective immune response, the enhanced sampling methods will overcome the diverse energy barriers of the free energy landscape of the IDPs.

The high computational cost of MD is still a limiting factor. A relatively small system with, for example, 50,000 atoms simulated with HREMD with 10 replicas, takes a couple of weeks on a computing node with 128 cores in a modern CPU as of the publication of this article. Still, depending on the available computational cost, even small systems can require a month of calculations to visualize only a few picoseconds of the trajectory.[202,211]

The efficiency of these calculations is based on an interaction of software and hardware, therefore, hybrid general-purpose computing on graphics processing units (GPU) based technologies have been used to improve these methods. Now, GPU-accelerated MD studies are the most efficient way of running enhanced sampling methods. Since 2007, when the Nvidia CUDA technology has been introduced, it is still the best environment for GPU computing and improving the performance of hybrid MD algorithms. The use of GPUs in MD calculations has led to hundredfold reductions in computational costs compared to central processing unit (CPU)-based algorithms, in addition, it allows better adaptation to system sizes using spatial domain decomposition. As a result, MD methods are able to encompass a wider range of biomolecular phenomena and timescales that are useful in drug discovery. In this way, the most commonly used software in MD have already implemented an architecture for calculations using GPUs.[212,213]

The study of IDPs has led to the emergence of new paradigms and, therefore, new challenges in drug discovery. In addition to the points already highlighted, such as the still vast field for MD studies and the improvement of force fields for describing protein flexibility, it is possible to observe the need for curation and deposition of IDPs in IDPDs. This need becomes even greater when referring to *Plasmodium falciparum* IDPs. Important perspectives in the design of vaccines and strategies against malaria can focus on combined studies of MD, antigen-antibody binding, and the development of libraries, especially dedicated to small molecules as targets and *Plasmodium falciparum*. These libraries could offer starting points for developing therapeutic pathways.[214]

Multiconformational selection and affinity, that is, compounds that bind to several groups of conformations with similar affinity, are interesting paths when in the field of dynamic conformational sets. Understanding the IDP-antibody interaction, as discussed previously, is of great importance, also depending on the improvement of methods for simulating complex systems. Antibody serum has already been applied in antimalarial studies and is a promising strategy.[215] Theoretical studies of antibodies against *Plasmodium falciparum* and variations such as nanobodies[216] and intrabodies[217] are also rooms with much space to be explored. Overall, we can state that there is evidence that IDP sets are druggable and recognized by antibodies in forming multiconformational entities, making them a relevant, promising, and current field in the development of antimalarial strategies.

6. Conclusion

The fact that malaria is a neglected tropical disease is reflected in the small number of protein models available and modeling studies performed to date. We observed 1.5 to 2.5% of IDPs articles targeting malaria within its scope. Four databases have information on *Plasmodium falciparum* proteins. DisProt has 9 entries and MobiDB about 5.386 in proteome searches. The structures in the *Pf* available in the PDB in complex with antibodies show that of the 117, only 5 were studied with MD simulations, and in an accessory manner.

In this review, we demonstrate that this is a fertile field for growth in the relatively small number of computational studies needed to understand the atomic behavior of these proteins. Despite the growing interest in the area of IDPs, the protein disorder marks several stages of the malaria life cycle. Its specific relationship with the functions of *Plasmodium falciparum* proteins is not fully elucidated. The study of IDPs is still a challenging and growing field, and with recent advances, it is possible to propose a path to be followed for studies involving these proteins: search in information and structure databases > structure and disorder prediction using the state of the art of available predictors > target analysis that may include: detection of sites, virtual scanning of small molecules, selection of accessions and affinity for conformations, and the identification of CDRs in studies with antibodies > methods of enhanced sampling with the choice of adequate force fields. While manual annotation is not able to compete with the rapid advancement of other methods, it can be understood that there is already material available that is not well suited for understanding the interaction of *Plasmodium falciparum* with antibodies, since the proteins appear timidly in articles and are not the main object of a robust structural description.

MD methods, such as REMD, Metadynamics, and US, and force fields, such as ff99IDPs and ff141IDPs, for IDPs, have already been developed as suitable complements to experimental methods for understanding the dynamics, interactions, and behavior of proteins with their ligands, targets, or different solvents. It is still shown that malaria as an NTD still needs investment in robust computational work, which is a highly important step in the development of drugs and vaccines today. It is also possible to carry out studies that show the behavior of CDRs in interactions with IDPs and the interaction of proteins at different moments of the life cycle with their key peptides. The improvement of force fields and water models, the description of protein function, and the availability of information in databases-all these are open fields for growth and contributions that can help in the advancement of computational chemistry in the study of IDPs.

Acknowledgments

This study was financed in part by the Coordenação de Aperfeiçoamento de Pessoal de Nível Superior-Brasil (CAPES)-Finance Code 001. The authors also acknowledge the financial support of FAPESP (projects 2018/24293-0, 2013/08293-7).

Pamella Cristiny C. da Silva *holds a bachelor's degree in Industrial Chemistry from the Universidade Federal do Pará (UFPA), Brazil. She is a PhD Student in the Institute of Chemistry at the Universidade Estadual de Campinas (UNICAMP). She researches in the field of Molecular Dynamics, enhanced sampling methods and Intrinsically Disordered Proteins (IDPs). In her master's, she worked on the mechanism of antigen-antibody interaction of a Plasmodium falciparum protein, and has a special interest in studies of complex solvation of IDPs in ionic liquids.*

Leandro Martinez *is an associate professor at the Universidade Estadual de Campinas (Unicamp) in Brazil. With a PhD in Chemistry, he specializes in computational biophysics and molecular simulations. His research focuses on the dynamics and interactions of biomolecules in complex solutions, contributing to the understanding of biological systems at the molecular level. Dr. Martinez is the leading author of computational chemistry software packages of high impact, including Packmol and ComplexMixtures.jl. He is also actively involved in teaching and mentoring graduate students.*

References

1. Deiana, A.; Forcelloni, S.; Porrello, A.; Giansanti, A.; *PLoS One* **2019**, *14*, e0217889. [Crossref]
2. Guy, A. J.; Irani, V.; MacRaild, C. A.; Anders, R. F.; Norton, R. S.; Beeson, J. G.; Richards, J. S.; Ramsland, P. A.; *PLoS One* **2015**, *10*, e0141729. [Crossref]
3. Wright, P. E.; Dyson, H. J.; *J. Mol. Biol.* **1999**, *293*, 321. [Crossref]
4. Rangarajan, N.; Kulkarni, P.; Hannenhalli, S.; *PLoS One* **2015**, *10*, e0126729. [Crossref]
5. Uversky, V. N.; Oldfield, C. J.; Midic, U.; Xie, H.; Xue, B.; Vucetic, S.; Iakoucheva, L. M.; Obradovic, Z.; Dunker, A. K.; *BMC Genomics* **2009**, *10*, S7. [Crossref]
6. Borcherds, W.; Bremer, A.; Borgia, M. B.; Mittag, T.; *Curr. Opin. Struct. Biol.* **2021**, *67*, 41. [Crossref]
7. Pajkos, M.; Dosztányi, Z.; *Prog. Mol. Biol. Transl. Sci.* **2021**, *183*, 45. [Crossref]
8. Feng, Z.-P.; Zhang, X.; Han, P.; Arora, N.; Anders, R. F.; Norton, R. S.; *Mol. Biochem. Parasitol.* **2006**, *150*, 256. [Crossref]
9. Brocchieri, L.; *Genome Res.* **2001**, *11*, 195. [Crossref]
10. Davies, H. M.; Nofal, S. D.; McLaughlin, E. J.; Osborne, A. R.; *FEMS Microbiol.* **2017**, *41*, 923. [Crossref]
11. Tompa, P.; Fersht, A.; *Structure and Function of Intrinsically Disordered Proteins*, vol. 1, 1st ed.; CRC Press: New York, US, 2009. [Crossref]
12. Jorda, J.; Xue, B.; Uversky, V. N.; Kajava, A. V.; *FEBS J.* **2010**, *277*, 2673. [Crossref]
13. Pancsa, R.; Tompa, P.; *PLoS One* **2012**, *7*, e34687. [Crossref]
14. Minkah, N. K.; Kappe, S. H. I.; *Nature* **2021**, *595*, 173. [Crossref]
15. Njoroge, M.; Njuguna, N. M.; Mutai, P.; Ongarora, D. S. B.; Smith, P. W.; Chibale, K.; *Chem. Rev.* **2014**, *114*, 11138. [Crossref]
16. Beeson, J. G.; Kurtovic, L.; Dobaño, C.; Opi, D. H.; Chan, J.-A.; Feng, G.; Good, M. F.; Reiling, L.; Boyle, M. J.; *Sci. Transl. Med.* **2019**, *11*. [Crossref]
17. Ameri, M.; Nezafat, N.; Eskandari, S.; *Expert Rev. Vaccines* 2022, *21*, 1. [Crossref]
18. Cumberworth, A.; Lamour, G.; Babu, M. M.; Gsponer, J.; *Biochem. J.* **2013**, *454*, 361. [Crossref]
19. Protter, D. S. W.; Rao, B. S.; Van Treeck, B.; Lin, Y.; Mizoue, L.; Rosen, M. K.; Parker, R.; *Cell Rep.* **2018**, *22*, 1401. [Crossref]
20. Miskei, M.; Horvath, A.; Vendruscolo, M.; Fuxreiter, M.; *J. Mol. Biol.* **2020**, *432*, 2289. [Crossref]
21. Fatafta, H.; Samantray, S.; Sayyed-Ahmad, A.; Coskuner-Weber, O.; Strodel, B.; *Prog. Mol. Biol. Transl. Sci.* **2021**, *183*, 135. [Crossref]

22. He, B.; Wang, K.; Liu, Y.; Xue, B.; Uversky, V. N.; Dunker, A. K.; *Cell Res.* **2009**, *19*, 929. [Crossref]
23. Uversky, V. N.; *Chem. Rev.* **2014**, *114*, 6557. [Crossref]
24. DeForte, S.; Uversky, V. N.; *RSC Adv.* **2016**, *6*, 11513. [Crossref]
25. Uversky, V. N.; Kulkarni, P.; *Biophys. Chem.* **2021**, *279*, 106694. [Crossref]
26. Kragelund, B. B.; Skriver, K.; *Intrinsically Disordered Proteins: Methods and Protocols*, vol. 1, 1st ed.; Humana Press: New Jersey, USA, 2020. [Crossref]
27. Dunker, A. K.; Lawson, J. D.; Brown, C. J.; Williams, R. M.; Romero, P.; Oh, J. S.; Oldfield, C. J.; Campen, A. M.; Ratliff, C. M.; Hipps, K. W.; Ausio, J.; Nissen, M. S.; Reeves, R.; Kang, C.; Kissinger, C. R.; Bailey, R. W.; Griswold, M. D.; Chiu, W.; Garner, E. C.; Obradovic, Z.; *J. Mol. Graph. Model.* **2001**, *19*, 26. [Crossref]
28. Dunker, A. K.; Babu, M. M.; Barbar, E.; Blackledge, M.; Bondos, S. E.; Dosztányi, Z.; Dyson, H. J.; Forman-Kay, J.; Fuxreiter, M.; Gsponer, J.; Han, K.-H.; Jones, D. T.; Longhi, S.; Metallo, S. J.; Nishikawa, K.; Nussinov, R.; Obradovic, Z.; Pappu, R. V.; Rost, B.; Selenko, P.; Subramaniam, V.; Sussman, J. L.; Tompa, P.; Uversky, V. N.; *Intrinsically Disord. Proteins* **2013**, *1*, e24157. [Crossref]
29. Lebowitz, J.; Lewis, M. S.; Schuck, P.; *Protein Sci.* **2002**, *11*, 2067. [Crossref]
30. Tompa, P.; *Curr. Opin. Struct. Biol.* **2011**, *21*, 419. [Crossref]
31. Piovesan, D.; Monzon, A. M.; Tosatto, S. C. E.; *Protein Sci.* **2022**, *31*, e4466. [Crossref]
32. Piovesan, D.; Monzon, A. M.; Quaglia, F.; Tosatto, S. C. E.; *Acta Crystallogr., Sect. D: Struct. Biol.* **2022**, *78*, 144. [Crossref]
33. Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E.; *International Tables for Crystallography* **2006**, 675. [Crossref]
34. RCSB Protein Data Bank (RCSB PDB), https://www.rcsb.org/, accessed in June 2024.
35. The UniProt Consortium.; *Nucleic Acids Res.* **2023**, *51*, D523. [Crossref]
36. UniProt, https://www.uniprot.org/, accessed in June 2024.
37. Hatos, A.; Hajdu-Soltész, B.; Monzon, A. M.; Palopoli, N.; Álvarez, L.; Aykac-Fas, B.; Bassot, C.; Benítez, G. I.; Bevilacqua, M.; Chasapi, A.; Chemes, L.; Davey, N. E.; Davidović, R.; Dunker, A. K.; Elofsson, A.; Gobeill, J.; Foutel, N. S. G.; Sudha, G.; Guharoy, M.; Horvath, T.; Iglesias, V.; Kajava, A. V.; Kovacs, O. P.; Lamb, J.; Lambrughi, M.; Lazar, T.; Leclercq, J. Y.; Leonardi, E.; Macedo-Ribeiro, S.; Macossay-Castillo, M.; Maiani, E.; Manso, J. A.; Marino-Buslje, C.; Martínez-Pérez, E.; Mészáros, B.; Mičetić, I.; Minervini, G.; Murvai, N.; Necci, M.; Ouzounis, C. A.; Pajkos, M.; Paladin, L.; Pancsa, R.; Papaleo, E.; Parisi, G.; Pasche, E.; Barbosa Pereira, P. J.; Promponas, V. J.; Pujols, J.; Quaglia, F.; Ruch, P.; Salvatore, M.; Schad, E.; Szabo, B.; Szaniszló, T.; Tamana, S.; Tantos, A.; Veljkovic, N.; Ventura, S.; Vranken, W.; Dosztányi, Z.; Tompa, P.; Tosatto, S. C. E.; Piovesan, D.; *Nucleic Acids Res.* **2020**, *48*, D269. [Crossref]
38. Fukuchi, S.; Sakamoto, S.; Nobe, Y.; Murakami, S. D.; Amemiya, T.; Hosoda, K.; Koike, R.; Hiroaki, H.; Ota, M.; *Nucleic Acids Res.* **2012**, *40*, D507. [Crossref]
39. Piovesan, D.; Necci, M.; Escobedo, N.; Monzon, A. M.; Hatos, A.; Mičetić, I.; Quaglia, F.; Paladin, L.; Ramasamy, P.; Dosztányi, Z.; Vranken, W. F.; Davey, N. E.; Parisi, G.; Fuxreiter, M.; Tosatto, S. C. E.; *Nucleic Acids Res.* **2021**, *49*, D361. [Crossref]
40. Lazar, T.; Martínez-Pérez, E.; Quaglia, F.; Hatos, A.; Chemes, L. B.; Iserte, J. A.; Méndez, N. A.; Garrone, N. A.; Saldaño, T. E.; Marchetti, J.; Rueda, A. J. V.; Bernadó, P.; Blackledge, M.; Cordeiro, T. N.; Fagerberg, E.; Forman-Kay, J. D.; Fornasari, M. S.; Gibson, T. J.; Gomes, G.-N. W.; Gradinaru, C. C.; Head-Gordon, T.; Jensen, M. R.; Lemke, E. A.; Longhi, S.; Marino-Buslje, C.; Minervini, G.; Mittag, T.; Monzon, A. M.; Pappu, R. V.; Parisi, G.; Ricard-Blum, S.; Ruff, K. M.; Salladini, E.; Skepö, M.; Svergun, D.; Vallet, S. D.; Varadi, M.; Tompa, P.; Tosatto, S. C. E.; Piovesan, D.; *Nucleic Acids Res.* **2021**, *49*, D404. [Crossref]
41. Schad, E.; Fichó, E.; Pancsa, R.; Simon, I.; Dosztányi, Z.; Mészáros, B.; *Bioinformatics* **2018**, *34*, 535. [Crossref]
42. Fichó, E.; Reményi, I.; Simon, I.; Mészáros, B.; *Bioinformatics* **2017**, *33*, 3682. [Crossref]
43. Yu, J.-F.; Dou, X.-H.; Sha, Y.-J.; Wang, C.-L.; Wang, H.-B.; Chen, Y.-T.; Zhang, F.; Zhou, Y.; Wang, J.-H.; *BMC Bioinf.* **2017**, *18*, 206. [Crossref]
44. Sun, C.; Feng, Y.; Fan, G.; *BMC Mol. Cell Biol.* **2022**, *23*, 33. [Crossref]
45. Hatos, A.; Monzon, A. M.; Tosatto, S. C. E.; Piovesan, D.; Fuxreiter, M.; *Nucleic Acids Res.* **2022**, *50*, D509. [Crossref]
46. Oates, M. E.; Romero, P.; Ishida, T.; Ghalwash, M.; Mizianty, M. J.; Xue, B.; Dosztányi, Z.; Uversky, V. N.; Obradovic, Z.; Kurgan, L.; Dunker, A. K.; Gough, J.; *Nucleic Acids Res.* **2012**, *41*, D508. [Crossref]
47. Hain, A. U. P.; Weltzer, R. R.; Hammond, H.; Jayabalasingham, B.; Dinglasan, R. R.; Graham, D. R. M.; Colquhoun, D. R.; Coppens, I.; Bosch, J.; *J. Struct. Biol.* **2012**, *180*, 551. [Crossref]
48. Nightingale, A.; Antunes, R.; Alpi, E.; Bursteinas, B.; Gonzales, L.; Liu, W.; Luo, J.; Qi, G.; Turner, E.; Martin, M.; *Nucleic Acids Res.* **2017**, *45*, W539. [Crossref]
49. Piovesan, D.; Del Conte, A.; Clementel, D.; Monzon, A. M.; Bevilacqua, M.; Aspromonte, M. C.; Iserte, J. A.; Orti, F. E.; Marino-Buslje, C.; Tosatto, S. C. E.; *Nucleic Acids Res.* **2023**, *51*, D438. [Crossref]
50. Colak, R.; Kim, T.; Michaut, M.; Sun, M.; Irimia, M.; Bellay, J.; Myers, C. L.; Blencowe, B. J.; Kim, P. M.; *PLoS Comput. Biol.* **2013**, *9*, e1003030. [Crossref]
51. Adedeji, D.; Hernandez, H.; Wiesner, J.; Köhler, U.; Jomaa, H.; Duin, E. C.; *FEBS Lett.* **2007**, *581*, 279. [Crossref]

52. Waters, N. C.; Kopydlowski, K. M.; Guszczynski, T.; Wei, L.; Sellers, P.; Ferlan, J. T.; Lee, P. J.; Li, Z.; Woodard, C. L.; Shallom, S.; Gardner, M. J.; Prigge, S. T.; *Mol. Biochem. Parasitol.* **2002**, *123*, 85. [Crossref]
53. Quaglia, F.; Hatos, A.; Salladini, E.; Piovesan, D.; Tosatto, S. C. E.; *Curr. Protoc.* **2022**, *2*, e484. [Crossref]
54. Singh, S. K.; Maithal, K.; Balaram, H.; Balaram, P.; *FEBS Lett.* **2001**, *501*, 19. [Crossref]
55. Somé, A. F.; Bazié, T.; Zongo, I.; Yerbanga, R. S.; Nikiéma, F.; Neya, C.; Taho, L. K.; Ouédraogo, J.-B.; *Parasit. Vectors* **2018**, *11*, 323. [Crossref]
56. Tonkin, M. L.; Arredondo, S. A.; Loveless, B. C.; Serpa, J. J.; Makepeace, K. A. T.; Sundar, N.; Petrotchenko, E. V.; Miller, L. H.; Grigg, M. E.; Boulanger, M. J.; *J. Biol. Chem.* **2013**, *288*, 12805. [Crossref]
57. Ho, C.-M.; Beck, J. R.; Lai, M.; Cui, Y.; Goldberg, D. E.; Egea, P. F.; Zhou, Z. H.; *Nature* **2018**, *561*, 70. [Crossref]
58. Lim, S. S.; Yang, W.; Krishnarjuna, B.; Kannan Sivaraman, K.; Chandrashekaran, I. R.; Kass, I.; MacRaild, C. A.; Devine, S. M.; Debono, C. O.; Anders, R. F.; Scanlon, M. J.; Scammells, P. J.; Norton, R. S.; McGowan, S.; *Biochemistry* **2014**, *53*, 7310. [Crossref]
59. Saxena, S.; Durgam, L.; Guruprasad, L.; *J. Biomol. Struct. Dyn.* **2019**, *37*, 1783. [Crossref]
60. Da Vela, S.; Svergun, D. I.; *Curr. Res. Struct. Biol.* **2020**, *2*, 164. [Crossref]
61. Predictor of Natural Disordered Regions, http://www.pondr.com/, accessed in June 2024.
62. Romero, P.; Obradovic, Z.; Li, X.; Garner, E. C.; Brown, C. J.; Dunker, A. K.; *Proteins* **2001**, *42*, 38. [Crossref]
63. Obradovic, Z.; Peng, K.; Vucetic, S.; Radivojac, P.; Dunker, A. K.; *Proteins* **2005**, *61*, 176. [Crossref]
64. Peng, K.; Vucetic, S.; Radivojac, P.; Brown, C. J.; Dunker, A. K.; Obradovic, Z.; *J. Bioinform. Comput. Biol.* **2005**, *3*, 35. [Crossref]
65. Xue, B.; Dunbrack, R. L.; Williams, R. W.; Dunker, A. K.; Uversky, V. N.; *Biochim. Biophys. Acta* **2010**, *1804*, 996. [Crossref]
66. Linding, R.; Jensen, L. J.; Diella, F.; Bork, P.; Gibson, T. J.; Russell, R. B.; *Structure* **2003**, *11*, 1453. [Crossref]
67. Dosztányi, Z.; Csizmok, V.; Tompa, P.; Simon, I.; *Bioinformatics* **2005**, *21*, 3433. [Crossref]
68. Prilusky, J.; Felder, C. E.; Zeev-Ben-Mordehai, T.; Rydberg, E. H.; Man, O.; Beckmann, J. S.; Silman, I.; Sussman, J. L.; *Bioinformatics* **2005**, *21*, 3435. [Crossref]
69. Campen, A.; Williams, R. M.; Brown, C. J.; Meng, J.; Uversky, V. N.; Dunker, A. K.; *Protein Pept. Lett.* **2008**, *15*, 956. [Crossref]
70. Ishida, T.; Kinoshita, K.; *Nucleic Acids Res.* **2007**, *35*, W460. [Crossref]
71. Kozlowski, L. P.; Bujnicki, J. M.; *BMC Bioinf.* **2012**, *13*, 111. [Crossref]
72. Linding, R.; Russell, R. B.; Neduva, V.; Gibson, T. J.; *Nucleic Acids Res.* **2003**, *31*, 3701. [Crossref]
73. Jones, D. T.; Cozzetto, D.; *Bioinformatics* **2015**, *31*, 857. [Crossref]
74. McGuffin, L. J.; Bryson, K.; Jones, D. T.; *Bioinformatics* **2000**, *16*, 404. [Crossref]
75. Tunyasuvunakool, K.; Adler, J.; Wu, Z.; Green, T.; Zielinski, M.; Žídek, A.; Bridgland, A.; Cowie, A.; Meyer, C.; Laydon, A.; Velankar, S.; Kleywegt, G. J.; Bateman, A.; Evans, R.; Pritzel, A.; Figurnov, M.; Ronneberger, O.; Bates, R.; Kohl, S. A. A.; Potapenko, A.; Ballard, A. J.; Romera-Paredes, B.; Nikolov, S.; Jain, R.; Clancy, E.; Reiman, D.; Petersen, S.; Senior, A. W.; Kavukcuoglu, K.; Birney, E.; Kohli, P.; Jumper, J.; Hassabis, D.; *Nature* **2021**, *596*, 590. [Crossref]
76. David, A.; Islam, S.; Tankhilevich, E.; Sternberg, M. J. E.; *J. Mol. Biol.* **2022**, *434*, 167336. [Crossref]
77. Ruff, K. M.; Pappu, R. V.; *J. Mol. Biol.* **2021**, 167208. [Crossref]
78. Anbo, H.; Sakuma, K.; Fukuchi, S.; Ota, M.; *Biology* **2023**, *12*, 182. [Crossref]
79. He, J.; Turzo, S. B. A.; Seffernick, J. T.; Kim, S. S.; Lindert, S.; *J. Phys. Chem. B* **2022**, *126*, 8439. [Crossref]
80. Binder, J. L.; Berendzen, J.; Stevens, A. O.; He, Y.; Wang, J.; Dokholyan, N. V.; Oprea, T. I.; *Curr. Opin. Struct. Biol.* **2022**, *74*, 102372. [Crossref]
81. Wilson, C. J.; Choy, W.-Y.; Karttunen, M.; *Int. J. Mol. Sci.* **2022**, *23*, 4591. [Crossref]
82. Necci, M.; Piovesan, D.; CAID Predictors; DisProt Curators; Tosatto, S. C. E.; *Nat. Methods* **2021**, *18*, 472. [Crossref]
83. Conte, A. D.; Mehdiabadi, M.; Bouhraoua, A.; Miguel Monzon, A.; Tosatto, S. C. E.; Piovesan, D.; *Proteins* **2023**, *91*, 1925. [Crossref]
84. Tan, J.; Piccoli, L.; Lanzavecchia, A.; *Annu. Rev. Immunol.* **2019**, *37*, 225. [Crossref]
85. MacRaild, C. A.; Seow, J.; Das, S. C.; Norton, R. S.; *Pept. Sci.* **2018**, *110*, e24067. [Crossref]
86. Krishnarjuna, B.; Sugiki, T.; Morales, R. A. V.; Seow, J.; Fujiwara, T.; Wilde, K. L.; Norton, R. S.; MacRaild, C. A.; *Commun. Biol.* **2018**, *1*, 58. [Crossref]
87. Mollica, L.; Bessa, L. M.; Hanoulle, X.; Jensen, M. R.; Blackledge, M.; Schneider, R.; *Front. Mol. Biosci.* **2016**, *3*, 52. [Crossref]
88. Olugbile, S.; Kulangara, C.; Bang, G.; Bertholet, S.; Suzarte, E.; Villard, V.; Frank, G.; Audran, R.; Razaname, A.; Nebie, I.; Awobusuyi, O.; Spertini, F.; Kajava, A. V.; Felger, I.; Druilhe, P.; Corradin, G.; *Infect. Immun.* **2009**, *77*, 5701. [Crossref]
89. MacRaild, C. A.; Richards, J. S.; Anders, R. F.; Norton, R. S.; *Structure* **2016**, *24*, 148. [Crossref]
90. Dogan, J.; Gianni, S.; Jemth, P.; *Phys. Chem. Chem. Phys.* **2014**, *16*, 6323. [Crossref]
91. Chakraborty, P.; Di Cera, E.; *Biochemistry* **2017**, *56*, 2853. [Crossref]

92. Sela-Culang, I.; Kunik, V.; Ofran, Y.; *Front. Immunol.* **2013**, *4*, 302. [Crossref]
93. Chothia, C.; Lesk, A. M.; Tramontano, A.; Levitt, M.; Smith-Gill, S. J.; Air, G.; Sheriff, S.; Padlan, E. A.; Davies, D.; Tulip, W. R.; *Nature* **1989**, *342*, 877. [Crossref]
94. Weitzner, B. D.; Dunbrack Jr., R. L.; Gray, J. J.; *Structure* **2015**, *23*, 302. [Crossref]
95. Guloglu, B.; Deane, C. M.; *Front. Immunol.* **2023**, *14*, 1223802. [Crossref]
96. Dunker, A. K.; Oldfield, C. J.; Meng, J.; Romero, P.; Yang, J. Y.; Chen, J. W.; Vacic, V.; Obradovic, Z.; Uversky, V. N.; *BMC Genomics* **2008**, *9*, S1. [Crossref]
97. Jeliazkov, J. R.; Sljoka, A.; Kuroda, D.; Tsuchimura, N.; Katoh, N.; Tsumoto, K.; Gray, J. J.; *Front. Immunol.* **2018**, *9*, 413. [Crossref]
98. Lawson, A. D. G.; MacCoss, M.; Heer, J. P.; *J. Med. Chem.* **2018**, *61*, 4283. [Crossref]
99. Fernández-Quintero, M. L.; Loeffler, J. R.; Kraml, J.; Kahler, U.; Kamenik, A. S.; Liedl, K. R.; *Front. Immunol.* **2018**, *9*, 3065. [Crossref]
100. Kaur, H.; Salunke, D. M.; *IUBMB Life* **2015**, *67*, 498. [Crossref]
101. Patarroyo, M. E.; Aza-Conde, J.; Moreno-Vranich, A.; Pabón, L.; Varela, Y.; Patarroyo, M. A.; *Curr. Issues Mol. Biol.* **2017**, *22*, 65. [Crossref]
102. Fowkes, F. J. I.; Richards, J. S.; Simpson, J. A.; Beeson, J. G.; *PLoS Med.* **2010**, *7*, e1000218. [Crossref]
103. Patarroyo, M. E.; Arevalo-Pinzon, G.; Reyes, C.; Moreno-Vranich, A.; Patarroyo, M. A.; *Curr. Issues Mol. Biol.* **2016**, *18*, 57. [Crossref]
104. Hou, N.; Jiang, N.; Ma, Y.; Zou, Y.; Piao, X.; Liu, S.; Chen, Q.; *Front. Immunol.* **2020**, *11*, 610. [Crossref]
105. Murugan, R.; Scally, S. W.; Costa, G.; Mustafa, G.; Thai, E.; Decker, T.; Bosch, A.; Prieto, K.; Levashina, E. A.; Julien, J.-P.; Wardemann, H.; *Nat. Med.* **2020**, *26*, 1135. [Crossref]
106. Morales, R. A. V.; MacRaild, C. A.; Seow, J.; Krishnarjuna, B.; Drinkwater, N.; Rouet, R.; Anders, R. F.; Christ, D.; McGowan, S.; Norton, R. S.; *Sci. Rep.* **2015**, *5*, 10103. [Crossref]
107. Foquet, L.; Hermsen, C. C.; van Gemert, G.-J.; Van Braeckel, E.; Weening, K. E.; Sauerwein, R.; Meuleman, P.; Leroux-Roels, G.; *J. Clin. Invest.* **2014**, *124*, 140. [Crossref]
108. Flück, C.; Smith, T.; Beck, H.-P.; Irion, A.; Betuela, I.; Alpers, M. P.; Anders, R.; Saul, A.; Genton, B.; Felger, I.; *Infect. Immun.* **2004**, *72*, 6300. [Crossref]
109. Oyen, D.; Torres, J. L.; Aoto, P. C.; Flores-Garcia, Y.; Binter, Š.; Pholcharee, T.; Carroll, S.; Reponen, S.; Wash, R.; Liang, Q.; Lemiale, F.; Locke, E.; Bradley, A.; King, C. R.; Emerling, D.; Kellam, P.; Zavala, F.; Ward, A. B.; Wilson, I. A.; *PLoS Pathog.* **2020**, *16*, e1008373. [Crossref]
110. MacRaild, C. A.; Zachrdla, M.; Andrew, D.; Krishnarjuna, B.; Nováček, J.; Žídek, L.; Sklenář, V.; Richards, J. S.; Beeson, J. G.; Anders, R. F.; Norton, R. S.; *PLoS One* **2015**, *10*, e0119899. [Crossref]
111. Julien, J.-P.; Wardemann, H.; *Nat. Rev. Immunol.* **2019**, *19*, 761. [Crossref]
112. Yan, J.; Mizianty, M. J.; Filipow, P. L.; Uversky, V. N.; Kurgan, L.; *Biochim. Biophys. Acta* **2013**, *1834*, 1671. [Crossref]
113. Chen, L.; Xu, Y.; Wong, W.; Thompson, J. K.; Healer, J.; Goddard-Borger, E. D.; Lawrence, M. C.; Cowman, A. F.; *Elife* **2017**, *6*, e21347. [Crossref]
114. Favuzza, P.; Guffart, E.; Tamborrini, M.; Scherer, B.; Dreyer, A. M.; Rufer, A. C.; Erny, J.; Hoernschemeyer, J.; Thoma, R.; Schmid, G.; Gsell, B.; Lamelas, A.; Benz, J.; Joseph, C.; Matile, H.; Pluschke, G.; Rudolph, M. G.; *Elife* **2017**, *6*, e20383. [Crossref]
115. Ragotte, R. J.; Pulido, D.; Lias, A. M.; Quinkert, D.; Alanine, D. G. W.; Jamwal, A.; Davies, H.; Nacer, A.; Lowe, E. D.; Grime, G. W.; Illingworth, J. J.; Donat, R. F.; Garman, E. F.; Bowyer, P. W.; Higgins, M. K.; Draper, S. J.; *Nat. Commun.* **2022**, *13*, 933. [Crossref]
116. Coley, A. M.; Gupta, A.; Murphy, V. J.; Bai, T.; Kim, H.; Foley, M.; Anders, R. F.; Batchelor, A. H.; *PLoS Pathog.* **2007**, *3*, 1308. [Crossref]
117. Igonet, S.; Vulliez-Le Normand, B.; Faure, G.; Riottot, M.-M.; Kocken, C. H. M.; Thomas, A. W.; Bentley, G. A.; *J. Mol. Biol.* **2007**, *366*, 1523. [Crossref]
118. Scally, S. W.; Murugan, R.; Bosch, A.; Triller, G.; Costa, G.; Mordmüller, B.; Kremsner, P. G.; Sim, B. K. L.; Hoffman, S. L.; Levashina, E. A.; Wardemann, H.; Julien, J.-P.; *J. Exp. Med.* **2018**, *215*, 63. [Crossref]
119. Triller, G.; Scally, S. W.; Costa, G.; Pissarev, M.; Kreschel, C.; Bosch, A.; Marois, E.; Sack, B. K.; Murugan, R.; Salman, A. M.; Janse, C. J.; Khan, S. M.; Kappe, S. H. I.; Adegnika, A. A.; Mordmüller, B.; Levashina, E. A.; Julien, J.-P.; Wardemann, H.; *Immunity* **2017**, *47*, 1197. [Crossref]
120. Oyen, D.; Torres, J. L.; Wille-Reece, U.; Ockenhouse, C. F.; Emerling, D.; Glanville, J.; Volkmuth, W.; Flores-Garcia, Y.; Zavala, F.; Ward, A. B.; King, C. R.; Wilson, I. A.; *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, E10438. [Crossref]
121. Imkeller, K.; Scally, S. W.; Bosch, A.; Martí, G. P.; Costa, G.; Triller, G.; Murugan, R.; Renna, V.; Jumaa, H.; Kremsner, P. G.; Sim, B. K. L.; Hoffman, S. L.; Mordmüller, B.; Levashina, E. A.; Julien, J.-P.; Wardemann, H.; *Science* **2018**, *360*, 1358. [Crossref]
122. Oyen, D.; Torres, J. L.; Cottrell, C. A.; Richter King, C.; Wilson, I. A.; Ward, A. B.; *Sci. Adv.* **2018**, *4*, eaau8529. [Crossref]
123. Thai, E.; Costa, G.; Weyrich, A.; Murugan, R.; Oyen, D.; Flores-Garcia, Y.; Prieto, K.; Bosch, A.; Valleriani, A.; Wu, N. C.; Pholcharee, T.; Scally, S. W.; Wilson, I. A.; Wardemann, H.; Julien, J.-P.; Levashina, E. A.; *J. Exp. Med.* **2020**, *217*, e20200061. [Crossref]
124. Pholcharee, T.; Oyen, D.; Flores-Garcia, Y.; Gonzalez-Paez, G.; Han, Z.; Williams, K. L.; Volkmuth, W.; Emerling, D.; Locke,

E.; Richter King, C.; Zavala, F.; Wilson, I. A.; *Nat. Commun.* **2021**, *12*, 1063. [Crossref]

125. Pholcharee, T.; Oyen, D.; Torres, J. L.; Flores-Garcia, Y.; Martin, G. M.; González-Páez, G. E.; Emerling, D.; Volkmuth, W.; Locke, E.; King, C. R.; Zavala, F.; Ward, A. B.; Wilson, I. A.; *J. Mol. Biol.* **2020**, *432*, 1048. [Crossref]
126. Kratochvil, S.; Shen, C.-H.; Lin, Y.-C.; Xu, K.; Nair, U.; da Silva Pereira, L.; Tripathi, P.; Arnold, J.; Chuang, G.-Y.; Melzi, E.; Schön, A.; Zhang, B.; Dillon, M.; Bonilla, B.; Flynn, B. J.; Kirsch, K. H.; Kisalu, N. K.; Kiyuka, P. K.; Liu, T.; Ou, L.; Pancera, M.; Rawi, R.; Reveiz, M.; Seignon, K.; Wang, L. T.; Waring, M. T.; Warner, J.; Yang, Y.; Francica, J. R.; Idris, A. H.; Seder, R. A.; Kwong, P. D.; Batista, F. D.; *Immunity* **2021**, *54*, 2859. [Crossref]
127. Beutler, N.; Pholcharee, T.; Oyen, D.; Flores-Garcia, Y.; MacGill, R. S.; Garcia, E.; Calla, J.; Parren, M.; Yang, L.; Volkmuth, W.; Locke, E.; Regules, J. A.; Dutta, S.; Emerling, D.; Early, A. M.; Neafsey, D. E.; Winzeler, E. A.; King, C. R.; Zavala, F.; Burton, D. R.; Wilson, I. A.; Rogers, T. F.; *PLoS Pathog.* **2022**, *18*, e1010409. [Crossref]
128. Kisalu, N. K.; Idris, A. H.; Weidle, C.; Flores-Garcia, Y.; Flynn, B. J.; Sack, B. K.; Murphy, S.; Schön, A.; Freire, E.; Francica, J. R.; Miller, A. B.; Gregory, J.; March, S.; Liao, H.-X.; Haynes, B. F.; Wiehe, K.; Trama, A. M.; Saunders, K. O.; Gladden, M. A.; Monroe, A.; Bonsignori, M.; Kanekiyo, M.; Wheatley, A. K.; McDermott, A. B.; Farney, S. K.; Chuang, G.-Y.; Zhang, B.; Kc, N.; Chakravarty, S.; Kwong, P. D.; Sinnis, P.; Bhatia, S. N.; Kappe, S. H. I.; Sim, B. K. L.; Hoffman, S. L.; Zavala, F.; Pancera, M.; Seder, R. A.; *Nat. Med.* **2018**, *24*, 408. [Crossref]
129. Chen, E.; Paing, M. M.; Salinas, N.; Sim, B. K. L.; Tolia, N. H.; *PLoS Pathog.* **2013**, *9*, e1003390. [Crossref]
130. Pizarro, J. C.; Chitarra, V.; Verger, D.; Holm, I.; Pêtres, S.; Dartevelle, S.; Nato, F.; Longacre, S.; Bentley, G. A.; *J. Mol. Biol.* **2003**, *328*, 1091. [Crossref]
131. Thouvenel, C. D.; Fontana, M. F.; Netland, J.; Krishnamurty, A. T.; Takehara, K. K.; Chen, Y.; Singh, S.; Miura, K.; Keitany, G. J.; Lynch, E. M.; Portugal, S.; Miranda, M. C.; King, N. P.; Kollman, J. M.; Crompton, P. D.; Long, C. A.; Pancera, M.; Rawlings, D. J.; Pepper, M.; *J. Exp. Med.* **2021**, *218*, e20200942. [Crossref]
132. Seow, J.; Morales, R. A. V.; MacRaild, C. A.; Krishnarjuna, B.; McGowan, S.; Dingjan, T.; Jaipuria, G.; Rouet, R.; Wilde, K. L.; Atreya, H. S.; Richards, J. S.; Anders, R. F.; Christ, D.; Drinkwater, N.; Norton, R. S.; *J. Mol. Biol.* **2017**, *429*, 836. [Crossref]
133. Scally, S. W.; McLeod, B.; Bosch, A.; Miura, K.; Liang, Q.; Carroll, S.; Reponen, S.; Nguyen, N.; Giladi, E.; Rämisch, S.; Yusibov, V.; Bradley, A.; Lemiale, F.; Schief, W. R.; Emerling, D.; Kellam, P.; King, C. R.; Julien, J.-P.; *Nat. Commun.* **2017**, *8*, 1568. [Crossref]
134. McLeod, B.; Miura, K.; Scally, S. W.; Bosch, A.; Nguyen, N.; Shin, H.; Kim, D.; Volkmuth, W.; Rämisch, S.; Chichester, J. A.; Streatfield, S.; Woods, C.; Schief, W. R.; Emerling, D.; King, C. R.; Julien, J.-P.; *Nat. Commun.* **2019**, *10*, 4328. [Crossref]
135. Withers-Martinez, C.; Strath, M.; Hackett, F.; Haire, L. F.; Howell, S. A.; Walker, P. A.; Christodoulou, E.; Dodson, G. G.; Blackman, M. J.; *Nat. Commun.* **2014**, *5*, 3726. [Crossref]
136. Campeotto, I.; Goldenzweig, A.; Davey, J.; Barfod, L.; Marshall, J. M.; Silk, S. E.; Wright, K. E.; Draper, S. J.; Higgins, M. K.; Fleishman, S. J.; *Proc. Natl. Acad. Sci. U. S. A.* **2017**, *114*, 998. [Crossref]
137. Wright, K. E.; Hjerrild, K. A.; Bartlett, J.; Douglas, A. D.; Jin, J.; Brown, R. E.; Illingworth, J. J.; Ashfield, R.; Clemmensen, S. B.; de Jongh, W. A.; Draper, S. J.; Higgins, M. K.; *Nature* **2014**, *515*, 427. [Crossref]
138. Alanine, D. G. W.; Quinkert, D.; Kumarasingha, R.; Mehmood, S.; Donnellan, F. R.; Minkah, N. K.; Dadonaite, B.; Diouf, A.; Galaway, F.; Silk, S. E.; Jamwal, A.; Marshall, J. M.; Miura, K.; Foquet, L.; Elias, S. C.; Labbé, G. M.; Douglas, A. D.; Jin, J.; Payne, R. O.; Illingworth, J. J.; Pattinson, D. J.; Pulido, D.; Williams, B. G.; de Jongh, W. A.; Wright, G. J.; Kappe, S. H. I.; Robinson, C. V.; Long, C. A.; Crabb, B. S.; Gilson, P. R.; Higgins, M. K.; Draper, S. J.; *Cell* **2019**, *178*, 216. [Crossref]
139. Lennartz, F.; Brod, F.; Dabbs, R.; Miura, K.; Mekhaiel, D.; Marini, A.; Jore, M. M.; Søgaard, M. M.; Jørgensen, T.; de Jongh, W. A.; Sauerwein, R. W.; Long, C. A.; Biswas, S.; Higgins, M. K.; *Nat. Commun.* **2018**, *9*, 3822. [Crossref]
140. McLeod, B.; Mabrouk, M. T.; Miura, K.; Ravichandran, R.; Kephart, S.; Hailemariam, S.; Pham, T. P.; Semesi, A.; Kucharska, I.; Kundu, P.; Huang, W.-C.; Johnson, M.; Blackstone, A.; Pettie, D.; Murphy, M.; Kraft, J. C.; Leaf, E. M.; Jiao, Y.; van de Vegte-Bolmer, M.; van Gemert, G.-J.; Ramjith, J.; King, C. R.; MacGill, R. S.; Wu, Y.; Lee, K. K.; Jore, M. M.; King, N. P.; Lovell, J. F.; Julien, J.-P.; *Immunity* **2022**, *55*, 1680. [Crossref]
141. Kundu, P.; Semesi, A.; Jore, M. M.; Morin, M. J.; Price, V. L.; Liang, A.; Li, J.; Miura, K.; Sauerwein, R. W.; King, C. R.; Julien, J.-P.; *Nat. Commun.* **2018**, *9*, 4458. [Crossref]
142. Ko, K.-T.; Lennartz, F.; Mekhaiel, D.; Guloglu, B.; Marini, A.; Deuker, D. J.; Long, C. A.; Jore, M. M.; Miura, K.; Biswas, S.; Higgins, M. K.; *Nat. Commun.* **2022**, *13*, 5603. [Crossref]
143. Coelho, C. H.; Tang, W. K.; Burkhardt, M.; Galson, J. D.; Muratova, O.; Salinas, N. D.; Alves E Silva, T. L.; Reiter, K.; MacDonald, N. J.; Nguyen, V.; Herrera, R.; Shimp, R.; Narum, D. L.; Byrne-Steele, M.; Pan, W.; Hou, X.; Brown, B.; Eisenhower, M.; Han, J.; Jenkins, B. J.; Doritchamou, J. Y. A.; Smelkinson, M. G.; Vega-Rodríguez, J.; Trück, J.; Taylor, J. J.; Sagara, I.; Healy, S. A.; Renn, J. P.; Tolia, N. H.; Duffy, P. E.; *Nat. Commun.* **2021**, *12*, 1750. [Crossref]
144. Chen, Y.; Xu, K.; Piccoli, L.; Foglierini, M.; Tan, J.; Jin, W.; Gorman, J.; Tsybovsky, Y.; Zhang, B.; Traore, B.; Silacci-Fregni, C.; Daubenberger, C.; Crompton, P. D.; Geiger, R.; Sallusto,

F.; Kwong, P. D.; Lanzavecchia, A.; *Nature* **2021**, *592*, 639. [Crossref]
145. Xie, Y.; Li, X.; Chai, Y.; Song, H.; Qi, J.; Gao, G. F.; *Cell Rep.* **2021**, *36*, 109600. [Crossref]
146. Kim, J.; Tan, Y. Z.; Wicht, K. J.; Erramilli, S. K.; Dhingra, S. K.; Okombo, J.; Vendome, J.; Hagenah, L. M.; Giacometti, S. I.; Warren, A. L.; Nosol, K.; Roepe, P. D.; Potter, C. S.; Carragher, B.; Kossiakoff, A. A.; Quick, M.; Fidock, D. A.; Mancia, F.; *Nature* **2019**, *576*, 315. [Crossref]
147. Lennartz, F.; Adams, Y.; Bengtsson, A.; Olsen, R. W.; Turner, L.; Ndam, N. T.; Ecklu-Mensah, G.; Moussiliou, A.; Ofori, M. F.; Gamain, B.; Lusingu, J. P.; Petersen, J. E. V.; Wang, C. W.; Nunes-Silva, S.; Jespersen, J. S.; Lau, C. K. Y.; Theander, T. G.; Lavstsen, T.; Hviid, L.; Higgins, M. K.; Jensen, A. T. R.; *Cell Host Microbe* **2017**, *21*, 403. [Crossref]
148. Singh, K.; Burkhardt, M.; Nakuchima, S.; Herrera, R.; Muratova, O.; Gittis, A. G.; Kelnhofer, E.; Reiter, K.; Smelkinson, M.; Veltri, D.; Swihart, B. J.; Shimp Jr., R.; Nguyen, V.; Zhang, B.; MacDonald, N. J.; Duffy, P. E.; Garboczi, D. N.; Narum, D. L.; *Commun. Biol.* **2020**, *3*, 395. [Crossref]
149. Nguyen, P. T. V.; Nguyen, G. L. T.; Thi Đinh, O.; Duong, C. Q.; Nguyen, L. H.; Truong, T. N.; *J. Mol. Struct.* **2022**, *1256*, 132520. [Crossref]
150. Beaudoin, C. A.; Jamasb, A. R.; Alsulami, A. F.; Copoiu, L.; van Tonder, A. J.; Hala, S.; Bannerman, B. P.; Thomas, S. E.; Vedithi, S. C.; Torres, P. H. M.; Blundell, T. L.; *Comput. Struct. Biotechnol. J.* **2021**, *19*, 3938. [Crossref]
151. Fernández-Quintero, M. L.; Vangone, A.; Loeffler, J. R.; Seidler, C. A.; Georges, G.; Liedl, K. R.; *Structure* **2022**, *30*, 430. [Crossref]
152. Serrano, M. L.; Gauna, A.; Pérez, H. A.; Squitieri, E.; Medina, J. D.; *J. Comput. Methods Sci. Eng.* **2009**, *9*, 353. [Crossref]
153. Godara, P.; Naik, B.; Meghwal, R.; Ojha, R.; Srivastava, V.; Prajapati, V. K.; Prusty, D.; *Life Sci.* **2022**, *311*, 121121. [Crossref]
154. Burki, T.; *Lancet* **2022**, *400*, 15. [Crossref]
155. Salo-Ahen, O. M. H.; Alanko, I.; Bhadane, R.; Bonvin, A. M. J.; Honorato, R. V.; Hossain, S.; Juffer, A. H.; Kabedev, A.; Lahtela-Kakkonen, M.; Larsen, A. S.; Lescrinier, E.; Marimuthu, P.; Mirza, M. U.; Mustafa, G.; Nunes-Alves, A.; Pantsar, T.; Saadabadi, A.; Singaravelu, K.; Vanmeert, M.; *Processes* **2020**, *9*, 71. [Crossref]
156. Lees, W. D.; Stejskal, L.; Moss, D. S.; Shepherd, A. J.; *Front. Immunol.* **2017**, *8*, 143. [Crossref]
157. Yamashita, T.; *Int. Immunol.* **2018**, *30*, 133. [Crossref]
158. Higashida, R.; Matsunaga, Y.; *Life* **2021**, *11*, 1428. [Crossref]
159. Fernández-Quintero, M. L.; Kraml, J.; Georges, G.; Liedl, K. R.; *mAbs* **2019**, *11*, 1077. [Crossref]
160. Kim, S. T.; Shirai, H.; Nakajima, N.; Higo, J.; Nakamura, H.; *Proteins* **1999**, *37*, 683. [Crossref]
161. Amaro, R. E.; Baudry, J.; Chodera, J.; Demir, Ö.; McCammon, J. A.; Miao, Y.; Smith, J. C.; *Biophys. J.* **2018**, *114*, 2271. [Crossref]
162. Evangelista Falcon, W.; Ellingson, S. R.; Smith, J. C.; Baudry, J.; *J. Phys. Chem. B* **2019**, *123*, 5189. [Crossref]
163. Shirai, H.; Nakajima, N.; Higo, J.; Kidera, A.; Nakamura, H.; *J. Mol. Biol.* **1998**, *278*, 481. [Crossref]
164. Bhattacharya, S.; Lin, X.; *Biomolecules* **2019**, *9*, 146. [Crossref]
165. Karplus, M.; McCammon, J. A.; *Nat. Struct. Biol.* **2002**, *9*, 646. [Crossref]
166. Liwo, A.; *Computational Methods to Study the Structure and Dynamics of Biomolecules and Biomolecular Processes: From Bioinformatics to Molecular Quantum Mechanics,* vol. 8, 2nd ed.; Springer: Berlin, GR, 2018. [Crossref]
167. Showalter, S. A.; *eMagRes.* **2014**, *3*, 181. [Crossref]
168. Shea, J.-E.; Best, R. B.; Mittal, J.; *Curr. Opin. Struct. Biol.* **2021**, *67*, 219. [Crossref]
169. Henriques, J.; Cragnell, C.; Skepö, M.; *J. Chem. Theory Comput.* **2015**, *11*, 3420. [Crossref]
170. Shabane, P. S.; Izadi, S.; Onufriev, A. V.; *J. Chem. Theory Comput.* **2019**, *15*, 2620. [Crossref]
171. Huang, J.; MacKerell Jr., A. D.; *Curr. Opin. Struct. Biol.* **2018**, *48*, 40. [Crossref]
172. Best, R. B.; Hummer, G.; *J. Phys. Chem. B* **2009**, *113*, 9004. [Crossref]
173. Ponder, J. W.; Case, D. A.; *Adv. Protein Chem.* **2003**, *66*, 27. [Crossref]
174. Best, R. B.; Zhu, X.; Shim, J.; Lopes, P. E. M.; Mittal, J.; Feig, M.; MacKerell, A. D.; *J. Chem. Theory Comput.* **2012**, *8*, 3257. [Crossref]
175. Scott, W. R. P.; Hünenberger, P. H.; Tironi, I. G.; Mark, A. E.; Billeter, S. R.; Fennen, J.; Torda, A. E.; Huber, T.; Krüger, P.; van Gunsteren, W. F.; *J. Phys. Chem. A* **1999**, *103*, 3596. [Crossref]
176. Robertson, M. J.; Tirado-Rives, J.; Jorgensen, W. L.; *J. Chem. Theory Comput.* **2015**, *11*, 3499. [Crossref]
177. Huang, J.; Rauscher, S.; Nawrocki, G.; Ran, T.; Feig, M.; de Groot, B. L.; Grubmüller, H.; MacKerell Jr., A. D.; *Nat. Methods* **2017**, *14*, 71. [Crossref]
178. Wang, W.; Ye, W.; Jiang, C.; Luo, R.; Chen, H.-F.; *Chem. Biol. Drug Des.* **2014**, *84*, 253. [Crossref]
179. Song, D.; Wang, W.; Ye, W.; Ji, D.; Luo, R.; Chen, H.-F.; *Chem. Biol. Drug Des.* **2017**, *89*, 5. [Crossref]
180. Maier, J. A.; Martinez, C.; Kasavajhala, K.; Wickstrom, L.; Hauser, K. E.; Simmerling, C.; *J. Chem. Theory Comput.* **2015**, *11*, 3696. [Crossref]
181. Strodel, B.; *J. Mol. Biol.* **2021**, *433*, 167182. [Crossref]
182. Best, R. B.; *Curr. Opin. Struct. Biol.* **2017**, *42*, 147. [Crossref]
183. Shrestha, U. R.; Smith, J. C.; Petridis, L.; *Commun. Biol.* **2021**, *4*, 243. [Crossref]

184. Hénin, J.; Lelièvre, T.; Shirts, M. R.; Valsson, O.; Delemotte, L.; *Living J. Comput. Molecular Sci.* **2022**, *4*, 1583. [Crossref]
185. Miao, Y.; McCammon, J. A.; *Mol. Simul.* **2016**, *42*, 1046. [Crossref]
186. Sugita, Y.; Okamoto, Y.; *Chem. Phys. Lett.* **1999**, *314*, 141. [Crossref]
187. Wang, L.; Friesner, R. A.; Berne, B. J.; *J. Phys. Chem. B* **2011**, *115*, 9431. [Crossref]
188. Jang, S.; Shin, S.; Pak, Y.; *Phys. Rev. Lett.* **2003**, *91*, 058305. [Crossref]
189. Hamelberg, D.; Mongan, J.; McCammon, J. A.; *J. Chem. Phys.* **2004**, *120*, 11919. [Crossref]
190. Wu, X.; Brooks, B. R.; Vanden-Eijnden, E.; *J. Comput. Chem.* **2016**, *37*, 595. [Crossref]
191. Torrie, G. M.; Valleau, J. P.; *J. Comput. Phys.* **1977**, *23*, 187. [Crossref]
192. Izrailev, S.; Stepaniants, S.; Isralewitz, B.; Kosztin, D.; Lu, H.; Molnar, F.; Wriggers, W.; Schulten, K.; *Computational Molecular Dynamics: Challenges, Methods, Ideas,* vol. 1, 1st ed.; Springer: Berlin, GR, 1997. [Crossref]
193. Darve, E.; Rodríguez-Gómez, D.; Pohorille, A.; *J. Chem. Phys.* **2008**, *128*, 144120. [Crossref]
194. Kollman, P.; *Chem. Rev.* **1993**, *93*, 2395. [Crossref]
195. Liao, Q.; *Prog. Mol. Biol. Transl. Sci.* **2020**, *170*, 177. [Crossref]
196. Huber, T.; Torda, A. E.; van Gunsteren, W. F.; *J. Comput. Aided Mol. Des.* **1994**, *8*, 695. [Crossref]
197. Jarzynski, C.; *Phys. Rev. Lett.* **1997**, *78*, 2690. [Crossref]
198. Sinha, S.; Biswas, A.; Mondal, J.; Mandal, K.; *bioRxiv* 2021. [Crossref]
199. Kobayashi, C.; Jung, J.; Matsunaga, Y.; Mori, T.; Ando, T.; Tamura, K.; Kamiya, M.; Sugita, Y.; *J. Comput. Chem.* **2017**, *38*, 2193. [Crossref]
200. Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E.; *J. Chem. Phys.* **1953**, *21*, 1087. [Crossref]
201. Qi, R.; Wei, G.; Ma, B.; Nussinov, R.; *Methods Mol. Biol.* **2018**, *1777*, 101. [Crossref]
202. Bernardi, R. C.; Melo, M. C. R.; Schulten, K.; *Biochim. Biophys. Acta* **2015**, *1850*, 872. [Crossref]
203. Fu, B.; Vendruscolo, M.; *Adv. Exp. Med. Biol.* **2015**, *870*, 35. [Crossref]
204. Kumar, S.; Rosenberg, J. M.; Bouzida, D.; Swendsen, R. H.; Kollman, P. A.; *J. Comput. Chem.* **1992**, *13*, 1011. [Crossref]
205. Harpole, T. J.; Delemotte, L.; *Biochim. Biophys. Acta, Biomembr.* **2018**, *1860*, 909. [Crossref]
206. Kästner, J.; *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2011**, *1*, 932. [Crossref]
207. Mülders, T.; Krüger, P.; Swegat, W.; Schlitter, J.; *J. Chem. Phys.* **1996**, *104*, 4869. [Crossref]
208. Sutto, L.; Marsili, S.; Gervasio, F. L.; *Wiley Interdiscip. Rev. Comput. Mol. Sci.* **2012**, *2*, 771. [Crossref]
209. Wang, W.; *Phys. Chem. Chem. Phys.* **2021**, *23*, 777. [Crossref]
210. Schor, M.; Mey, A. S. J. S.; MacPhee, C. E.; *Biophys. Rev.* **2016**, *8*, 429. [Crossref]
211. Durrant, J. D.; McCammon, J. A.; *BMC Biology* **2011**, *9*, 71. [Crossref]
212. Kondratyuk, N.; Nikolskiy, V.; Pavlov, D.; Stegailov, V.; *Int. J. High Perform. Comput. Appl.* **2021**, *35*, 312. [Crossref]
213. Pandey, M.; Fernandez, M.; Gentile, F.; Isayev, O.; Tropsha, A.; Stern, A. C.; Cherkasov, A.; *Nature Mach. Intell.* **2022**, *4*, 211. [Crossref]
214. *Intrinsically Disordered Proteins: Dynamics, Binding, and Function*, vol 1, 1st ed.; Salvi, N. ed.; Academic Press: Massachusetts, US, 2019.
215. Gaudinski, M. R.; Berkowitz, N. M.; Idris, A. H.; Coates, E. E.; Holman, L. A.; Mendoza, F.; Gordon, I. J.; Plummer, S. H.; Trofymenko, O.; Hu, Z.; Campos Chagas, A.; O'Connell, S.; Basappa, M.; Douek, N.; Narpala, S. R.; Barry, C. R.; Widge, A. T.; Hicks, R.; Awan, S. F.; Wu, R. L.; Hickman, S.; Wycuff, D.; Stein, J. A.; Case, C.; Evans, B. P.; Carlton, K.; Gall, J. G.; Vazquez, S.; Flach, B.; Chen, G. L.; Francica, J. R.; Flynn, B. J.; Kisalu, N. K.; Capparelli, E. V.; McDermott, A.; Mascola, J. R.; Ledgerwood, J. E.; Seder, R. A.; VRC 612 Study Team; *N. Engl. J. Med.* **2021**, *385*, 803. [Crossref]
216. Dietrich, M. H.; Gabriela, M.; Reaksudsan, K.; Dixon, M. W. A.; Chan, L.-J.; Adair, A.; Trickey, S.; O'Neill, M. T.; Tan, L. L.; Lopaticki, S.; Healer, J.; Keremane, S.; Cowman, A. F.; Tham, W.-H.; *Biochem. J.* **2022**, *479*, 2529. [Crossref]
217. Moutel, S.; Nizak, C.; Perez, F.; *Methods Mol. Biol.* **2012**, *907*, 667. [Crossref]

Submitted: January 22, 2024
Published online: June 13, 2024