RESEARCH ARTICLE

Long-time scale simulations of virus-like particles from three human-norovirus strains

Agnieszka G. Lipska[1] | Adam K. Sieradzan[1,2] | Cezary Czaplewski[1,2] | Andrea D. Lipińska[3] | Krzysztof M. Ocetkiewicz[1] | Jerzy Proficz[1] | Paweł Czarnul[4] | Henryk Krawczyk[1,4] | Adam Liwo[1,2]

[1]Centre of Informatics Tri-city Academic Supercomputer and Network (CI TASK), Gdańsk University of Technology, Fahrenheit Union of Universities in Gdańsk, Gdańsk, Poland

[2]Faculty of Chemistry, University of Gdańsk, Fahrenheit Union of Universities in Gdańsk, Gdańsk, Poland

[3]Laboratory of Virus Molecular Biology, Intercollegiate Faculty of Biotechnology, University of Gdańsk and Medical University of Gdańsk, Fahrenheit Union of Universities in Gdańsk, Gdańsk, Poland

[4]Faculty of Electronics, Telecommunications and Informatics, Gdańsk University of Technology, Fahrenheit Union of Universities in Gdańsk, Gdańsk, Poland

Correspondence
Adam Liwo, Faculty of Chemistry, University of Gdańsk, Fahrenheit Union of Universities in Gdańsk, ul. Wita Stwosza 63, 80-308 Gdańsk, Poland.
Email: adam.liwo@ug.edu.pl

Funding information
National Science Centre, Grant/Award Numbers: UMO-2017/27/B/ST4/00926, UMO-2017/26/M/ST4/00044, UMO-2021/40/Q/ST4/00035; Pomorskie Voivodeship Regional Operational Program for 2014-2020, Grant/Award Number: RPPM.01.02.00-22-0001/17; European Funds - Smart Growth, Grant/Award Number: POIR.04.0200-00-D014/20-00

Abstract
The dynamics of the virus like particles (VLPs) corresponding to the GII.4 Houston, GII.2 SMV, and GI.1 Norwalk strains of human noroviruses (HuNoV) that cause gastroenteritis was investigated by means of long-time (about 30 μs in the laboratory timescale) molecular dynamics simulations with the coarse-grained UNRES force field. The main motion of VLP units turned out to be the bending at the junction between the P1 subdomain (that sits in the VLP shell) and the P2 subdomain (that protrudes outside) of the major VP1 protein, this resulting in a correlated wagging motion of the P2 subdomains with respect to the VLP surface. The fluctuations of the P2 subdomain were found to be more pronounced and the P2 domain made a greater angle with the normal to the VLP surface for the GII.2 strain, which could explain the inability of this strain to bind the histo-blood group antigens (HBGAs).

KEYWORDS
coarse-graining, human norovirus, molecular dynamics, UNRES, vaccines, virus-like particles

1 | INTRODUCTION

Virus-like particles (VLPs) are empty virus nanostructures mimicking virions or capsids, which assemble spontaneously from their protein components. They resemble structural, antigenic, and receptor-binding properties of pathogenic viruses,[1] while being non-infectious themselves because of the absence of genomic information. VLPs are often used in experimental studies of virus shape and stability upon mutation.[2–4] Because they are not infectious, they are considered as safe vaccine components.[5,6] Apart from the shape, the dynamics of a virus particle is crucial for its binding to the host cell. Molecular-dynamics simulations are very useful in this regard because they enable us to study the motion of a system at the atomistic (all-atom simulations) or residue (coarse-grained simulations) level[7] to acquire information that is yet unavailable from the state-of-the-art experimental techniques. All-atom simulations provide a more detailed and

accurate picture; however, they are limited to very short time scales. Conversely, coarse-grained simulations, in which several atoms are merged into an extended interaction site, are less accurate but enable us to extend the simulation time scale by about 3 orders of magnitude,[8] which is crucial in investigating biologically-relevant time scales.

Viruses are particularly highly packed systems, due to limitations of their genome size, and therefore, a viral capsid must be dynamic to perform many processes (e.g., assembly, attachment, and entry to the host cell).[9] Hence many molecular dynamics (MD) studies have been carried out on virus capsids and VLPs at the all-atom and coarse-grained levels.[10–13] For example, all-atom simulations of a complete satellite tobacco mosaic virus (9540 amino-acid residues) in explicit water for over 50 ns demonstrated the stability of the entire virion.[14] All-atom simulations of a complete virus capsid of the porcine circovirus type 2 revealed a semipermeable character of the capsid shell for ions and full permeability for water molecules.[15] Atomic force microscopy experiments combined with shape-based coarse-grained MD simulations for the hepatitis B virus (HBV) showed agreement between the simulation and experiment.[16] In that work, the virus-coat protein molecules were coarse-grained by fitting a limited number of sites, each representing 150 atoms, to reproduce their shapes. Later, all-atom simulations of HBV were carried out by Hadden et al.[17] on a microsecond scale, enabling the analysis of capsid flexibility and its implications for biological function. Subsequently, multiscale simulations of this virus, with ultra-coarse graining of the entire dimeric unit to a single particle, and parameterization based on all-atom MD simulations, were carried out by Hagan and coworkers,[18] which enabled these researchers to study virus assembly. A multiscale model of the SARS-CoV-2 coronavirus virion has also been reported.[19] Cieplak and colleagues applied Gō-like models to study mechanostability[20] and assembly[21] of VLPs.

The human noroviruses (HuNoVs) are non-enveloped, positive-sense single-stranded RNA viruses that cause sporadic and epidemic viral gastroenteritis, resulting in ~684 million disease cases and ~212,000 deaths annually.[22] HuNoVs are highly contagious and especially dangerous for young children and for the elderly or immunocompromised patients.[5] For these reasons, the development of norovirus vaccines has been recognized as a top priority by the World Health Organization. There are numerous species of noroviruses categorized in genogroups,[5] of which the genogroups GI, GII, GIV, GVIII, and GIX are responsible for human infections and the GII.4 Houston strain has been circulating as the predominant variant.[23] Their targets on host tissues are usually the histo-blood group antigens (HBGAs). Two proteins compose the mature viral particles of HuNoVs: the major capsid protein VP1 (that encapsidates the genome), and the minor structural protein VP2 (that is suggested to be located internally in the capsid).[24] The shell (and, thereby, the VLP) consists of the VP1 protein exclusively. It contains from 180 to 240 VP1 units and possesses the icosahedral T = 3 symmetry and its diameter is approximately 270–300 Å. The structure of the VLP is illustrated in Figure 1A with the example of the experimental structure of GII.4 Houston strain[3] (180 VP1 units).

The VP1 contains 2 structural domains, S and P, respectively. The S domain is involved in the formation of the icosahedral shell from which the P domains protrude. The two domains are linked by a flexible hinge. Additionally, the N-terminal arm (NTA) ties the chains by a network of interactions. The P domain consists of two subdomains, P1 and P2, respectively.[2] The P2 subdomain is the most exposed and contains the major points of contact with HBGA ligands or with neutralizing antibodies.[25,26] On the other hand, the P1 subdomain binds to the Human Norovirus Receptors (NORO), this process enabling the organism to recognize the pathogen and thus initiate the defense against its attack. The structure of VP1, its division into domains, and the residues of the HBGA and NORO binding sites are illustrated in Figure 1B with the example of the VP1 protein from the GII.4 Houston strain.

A deeper understanding of capsid structure and dynamics may provide a rational basis for developing more effective norovirus vaccines. VLP-based vaccines are currently being developed as the most promising norovirus-vaccine candidates.[27] In addition, a VP1 can also be used as a VLP platform with attached immunogenic fragments of other pathogens or cancer antigens.[28] Computational structural studies can identify the attachment sites such that foreign fragment attachment will not result in the disruption of the VLP structure or changes of its properties.

Even though a number of computational studies have been carried out recently, by means of all-atom molecular dynamics simulations, on the binding of the norovirus units to HBGA[29] and on designing its inhibitors,[30,31] only a single VP1 protein was considered there. To our knowledge, no study has been reported so far on the dynamics of the complete norovirus or its shell. However, virus actions, both regarding binding to the host cell and the ability of the coat to disassemble in order to release the genetic material, as well as the process of maturation, depend on the motion of the entire virus coat, including cooperative motions.[32,33] There is experimental evidence that shape dynamics of the noroviruses strongly affects binding by exposing or hiding the binding sites.[34,35] Consequently, the investigation of cooperative motions even at a less accurate than the all-atom level can lead to valuable insight into the interaction of the virus shell with the immunological machinery of the human body. Therefore, in this work, we investigated the structure and dynamics of the VLPs corresponding to the GII.4 Houston,[3] GII.2 SMV,[4] and the GI.1 Norwalk[2] strains of human noroviruses (HuNoVs). The sequences of these three strains, aligned with the use of the T-Coffee web server,[36] are shown in Figure 2. The shell of each of the three strains consists of 180 VP1 units. The GII.4 and GI.1 strains bind to the HBGAs, while GII.2 SMV does not. These differences are likely to be reflected in the structure and dynamics of the respective VLP particles.

We investigated two variants of each strain, one with a full VP1 sequence, and the other one without the part of the N-terminal segment which is not present in the corresponding experimental structure. In summary, a total of 6 VLP systems have been studied. The lengths of single chains of the full and truncated sequences were 540 and 514 residues for GII.4 (26 N-terminal residues removed), 542 and 517 residues for GII.2 (25 N-terminal residues removed), and
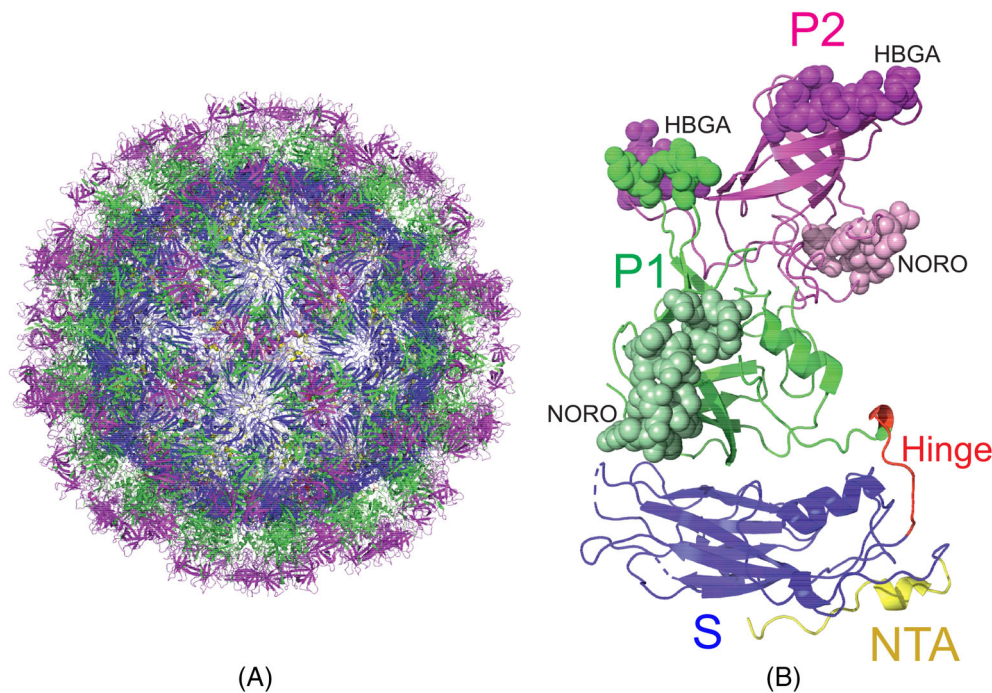
FIGURE 1 (A) The structure of the whole virus-like particle of the GII.4 strain of human norovirus. (B) The structure of the VP1 protein (one subunit) from the GII.4 strain of human norovirus (PDB: 7K6V). The domains and subdomains are indicated with different colors and labeled accordingly. The residues of the NORO and HBGA binding sites are shown as light-colored spheres and the sites are labeled accordingly.

530 and 502 residues for GI.1 (28 N-terminal residues removed), respectively. The aim of using both full VP1 sequences and the sequences truncated at the N-terminus was to determine whether the part of the N-terminal domain of the respective VLP, which is not seen in the experimental structure (and, consequently, is probably unstructured) affects the structural and dynamic properties of the VLPs and, thereby could influence the binding of HuNoVs to the receptors. It should be noted that N-terminal part truncation can be observed in biological systems during recombinant VLP production.[37] Truncation of that part can even change the shape of a VLP.[4] Because the units are tightly packed and the N-terminal segment is located on the inner surface of the VLP, it cannot be determined without carrying out simulations, whether this deletion affects the structure and dynamics of the P1 and P2 subdomains.

Because the systems studied are large, a sufficient time scale can be covered in real time only by means of coarse-grained simulations. Moreover, coarse-grained approaches enable us to run larger-scale simulations of viruses and their components[13] and, save atomic details, give results similar to those of all-atom simulations and to experiment.[10,17,18] In this study, we used the UNited RESidue (UNRES) physics-based model of polypeptide chains developed in our laboratory.[38–42] Owing to high degree of coarse-graining (only 2 interaction sites per amino-acid residue), parallel implementation,[43] and to recent optimization of the parallel code,[44] it enables us to run 1 ns/day simulations for protein systems with size of about 150,000 amino-acid residues.[44] Because of averaging out the degrees of freedom which are not included in the model, this time corresponds to about 1 μs of all-atom-simulation or laboratory time.[8,44] The relationship between the UNRES and the all-atom-simulation or laboratory time was established by comparing the folding times averaged over all folding events found in the MD simulations of a variant of tryptophan cage (PDB: 2JOF) carried out with UNRES[44] with those obtained by all-atom simulations by Lindorff-Larsen et al.[45] and by comparing the folding-rate constants of a series of FBP28 WW domain mutants calculated with UNRES with their experimental counterparts.[46] In all UNRES and all-atom simulations carried out for this purpose the proteins under study folded and unfolded repeatedly.[44–46]

The physics-based character of UNRES makes the simulations reliable and UNRES has already been used in biological applications.[39,42] The latest code optimization that included massive parallelization supported by shared memory multithreaded OpenMP library[44] enabled us to extend the scope of UNRES applications to very large protein systems. Consequently, UNRES appears to be an appropriate coarse-grained model with which to investigate the dynamic structure of HuNoVs to find the differences between different strains and link them with biological activity. Compared to coarse-grained simulations carried out with structure-based models[20,21] or the coarse-grained models, which are parameterized based on the components of a given VLP,[16,18] no bias to the target system, which could affect the results, is involved in UNRES model construction or force-field parameterization. On the other hand, the accuracy of modeling the structural changes and dynamic behavior might be coarser than those of the specifically parameterized models. Therefore, in this study, we refrain from drawing too detailed conclusions and analyze the results at the domain and coarser level.

2 | METHODOLOGY

2.1 | UNRES model of polypeptide chains

All simulations were carried out with the UNRES coarse-grained model of polypeptide chains developed in our laboratory.[38–42]
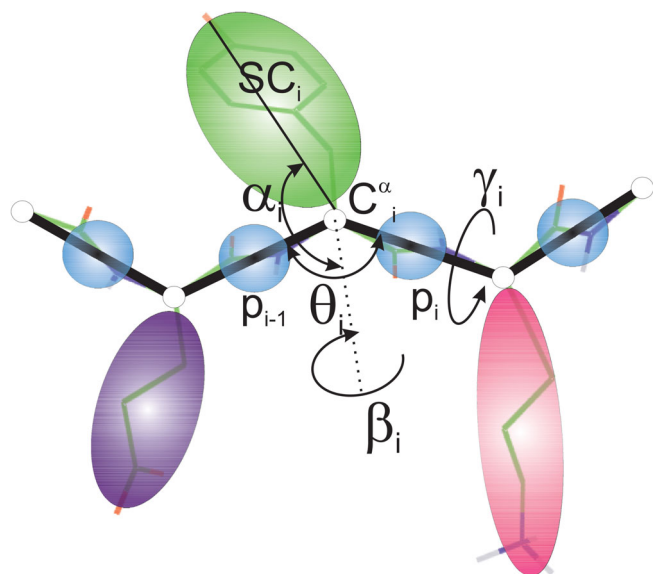
```
7K6V  ----MKMASSDASPSDGSTANLVPEVNNEVMALEPVVGAAIAAPVAGQQNVIDPWIRNNFVQAPGGEFTVSPRNA
6OTF  ----MKMASNDAAPSTDGAAGLVPESNNEVMALEPVAGAALAAPVTGQTNIIDPWIRANFVQAPNGEFTVSPRNA
1IHM  MMMASKDATSSVDGASGAGQLVPEVNASDPLAMDPVAGSSTAVATAGQVNPIDPWIINNFVQAPQGEFTISPNNT

       * *:... : ..    :     .: :*::**.*:: *...:** * *****  ****** ****:**.*:


7K6V  PGEILWSAPLGPDLNPYLSHLARMYNGYAGGFEVQVILAGNAFTAGKIIFAAVPPNFPTEGLSPSQVTMFPHIIV
6OTF  PGEVLLNLELGPELNPYLAHLARMYNGYAGGMEVQVMLAGNAFTAGKLVFAAVPPHFPVENLSPQQITMFPHVII
1IHM  PGDVLFDLSLGPHLNPFLLHLSQMYNGWVGNMRVRIMLAGNAFTAGKIIVSCIPPGFGSHNLTIAQATLFPHVIA

       **::* .   ***.***:* **::****:.*.:.*:::**********:.:.:** *   ..*:  * *:***:*


7K6V  DVRQLEPVLIPLPDVRNNFYHYNQSNDPTIKLIAMLYTPLRANNAGDDVFTVSCRVLTRPSPDFDFIFLVPPTVE
6OTF  DVRTLEPVLLPLPDVRNNFFHYNQKDDPKMRIVAMLYTPLRSNGSGDDVFTVSCRVLTRPSPDFDFTYLVPPTVE
1IHM  DVRTLDPIEVPLEDVRNVLFHNNDRNQQTMRLVCMLYTPLRTGGGTGDSFVVAGRVMTCPSPDFNFLFLVPPTVE

       *** *:*: :** **** ::* *: ::  .::::.*******:...  .* *.*: **:* *****:* :*******


7K6V  SRTKPFTVPILTVEEMTNSRFPIPLEKLFTGPSGAFVVQPQNGRCTTDGVLLGTTQLSPVNICTFRGDVTHIA--
6OTF  SKTKPFTLPILTLGELSNSRFPVSIDQMYTSPNEVISVQCQNGRCTLDGELQGTTQLQVSGICAFKGEVTAHLQD
1IHM  QKTRPFTLPNLPLSSLSNSRAPLPISSMGISPDNVQSVQFQNGRCTLDGRLVGTTPVSLSHVAKIRGTSNG---

       .:*:***:* *.: .::*** *:.:.:   .*. .  ** ****** ** * *** :.     :. ::*  .


7K6V  GTHDYTMNLASQNWNNYDPTEEIPAPLGTPDFVGKIQGVLTQTTRG------DGSTRGHKATVSTGSVHFTPKLG
6OTF  NDHLYNITITNLNGSPFDPSEDIPAPLGVPDFQGRVFGVITQRDKQNAAGQSQPANRGHDAVVPTYTAQYTPKLG
1IHM  ----TVINLTELDGTPFHPFE-GPAPIGFPDLGGCDWHINMTQ---------FGHSSQTQYDVDTTPDTFVPHLG

       :.::. : . :.* * ***:* **: *    :          .    . * *. :.*:**


7K6V  SVQFTTDTNNDLETGQNTKFTPVGVVQDGNSAHQNEPQQWVLPNYSGRTGHNVHLAPAVAPTFPGEQLLFFRSTM
6OTF  QVQIGTWQTDDLKVNQPVKFTPVGLN------DTEHFNQWVVPRYAGALNLNTNLAPSVAPVFPGERLLFFRSYL
1IHM  SIQA------NGIGSGNYVGVLSWISPPSHP-SGSQVDLWKIPNYGSSITEATHLAPSVYPPGFGEVLVFFMSKM

       .:*       :    . .  :    .. : * :*.*..    .:***:* *   ** *:** * :


7K6V  PGCSGYPNMNLDCLLPQEWVLHFYQEAAPAQSDVALLRFVNPDTGRVLFECKLHKSGYVTVAHTG--PHDLVIPP
6OTF  PLKGGYGNPAIDCLLPQEWVQHFYQEAAPSMSEVALVRYINPDTGRALFEAKLHRAGFMTVSSNT--SAPVVVPA
1IHM  PG---PGAYNLPCLLPQEYISHLASEQAPTVGEAALLHYVDPDTGRNLGEFKAYPDGFLTCVPNGASSGPQQLPI

       *       : ******:: *:.* **: .:.**:::::**** * * * :  *::*   .    .    :*


7K6V  NGYFRFDSWVNQFYTLAPMGNGAGRRRAL----
6OTF  NGYFRFDSWVNQFYSLAPMGTGNGRRRIQ----
1IHM  NGVFVFVSWVSRFYQLKPVGTASSARGRLGLRR

       ** * * ***.:** * *:*.. . *
```

FIGURE 2 Comparison of the sequences of the three strains of human norovirus considered in this study. The alignment was obtained with the T-Coffee web server.[36]

In this model, a polypeptide chain is represented by a sequence of α-carbon (C^α) atoms linked by virtual bonds, with peptide groups (p) located halfway between the consecutive C^αs and united side chains attached to the C^αs with the C^α...SC virtual bonds (Figure 3). Only the ps and the SCs are interaction sites, while the C^αs assist in chain-geometry definition. The sites and the corresponding interaction potentials have axial and not spherical symmetry.

FIGURE 3 UNRES model of polypeptide chains. The interaction sites are united peptide groups located between the consecutive α-carbon atoms (light-blue spheres) and united side chains attached to the α-carbon atoms (spheroids with different colors and dimensions). The backbone geometry of the simplified polypeptide chain is defined by the $C^{\alpha}\cdots C^{\alpha}\cdots C^{\alpha}$ virtual-bond angles θ (θ_i has the vertex at C_i^{α}) and the $C^{\alpha}\cdots C^{\alpha}\cdots C^{\alpha}\cdots C^{\alpha}$ virtual-bond-dihedral angles γ (γ_i has the axis passing trough C_i^{α} and C_{i+1}^{α}). The local geometry of the ith side-chain center is defined by the polar angle α_i (the angle between the bisector of the respective angle θ_i and the $C_i^{\alpha}\cdots SC_i$ vector) and the azimuth angle β_i (the angle of counter-clockwise rotation of the $C_i^{\alpha}\cdots SC_i$ vector about the bisector from the $C_{i-1}^{\alpha}\cdots C_i^{\alpha}\cdots C_{i+1}^{\alpha}$ plane, starting from C_{i-1}^{α}). For illustration, the bonds of the all-atom chains, except for those to the hydrogen atoms connected with the carbon atoms, are superposed on the coarse-grained picture. Reproduced with permission from Zaborowski et al., J. Chem. Inf. Model., 55, 2050 (2015). Copyright 2015 American Chemical Society.

The UNRES energy function consists of the site-site (SC-SC, SC-p, and p–p) interaction terms, the local terms corresponding to the energetics of virtual-bonds, backbone virtual-bond angles, backbone virtual-bond-dihedral angles, and side-chain rotamers, and of the multibody or correlation terms that couple the backbone-local and backbone-electrostatic interactions.[40,47] The correlation terms are necessary for the correct modeling of regular secondary structures.[40,47,48] The solvent is implicit in the interaction potentials and the present parameterization corresponds to physiological pH. Details of the UNRES model and force field are available in the references cited.[39–42]

The UNRES effective energy function originates from the potential of mean force of polypeptide chains in water, which is represented[40,47] by a truncated series of Kubo cluster cumulant functions.[49] Consequently, it has the sense of free energy and depends on temperature.[50] The energy terms in the present version of UNRES have been derived based on our recently developed scale-consistent theory of coarse graining,[40] in which the atomistic details are rigorously embedded in the effective interactions potentials. The details of the approach can be found in References [40, 41, 47], and, in Reference [51], which summarizes the scale-consistent methodology and also compares UNRES with the other physics-based coarse-grained models. We used the latest NEWCT-9P variant of the UNRES force field, which had been parameterized by maximum likelihood method[52] with 9 proteins of all secondary-structure types.[41] The calibration was aimed at reproducing the experimental ensembles of the training proteins determined by nuclear magnetic resonance (NMR). The ensembles of three of the training proteins were determined at a range of temperatures encompassing those below the melting temperature, the melting temperature, and those above the melting temperatures. Thus the force field is able to model not only the folded state but also the folding process and major conformational changes.

Owing to non-spherical symmetry of the interaction sites and the respective potentials and the presence of correlation terms in the effective energy function, UNRES is capable of ab initio folding simulations despite high degree of coarse graining, as demonstrated[53–56] in the Community Wide Experiments on the Critical Assessment of Techniques for Protein Structure Prediction (CASP; https://predictioncenter.orghttps://predictioncenter.org). As any heavily coarse-grained model, UNRES yields medium-resolution structures. However, owing to its physics-based nature it results in correct domain-packing topology where knowledge-based methods do not, an example being our prediction of target T1063 in CASP10,[53] for which UNRES yielded coarse resolution of each of the two domains but their correct packing, as opposed to homology-based methods, although the target was highly homologous to proteins with then known structures.

We also successfully used UNRES to model large-scale conformational transitions, for example, that from the closed to the open conformation of bacterial DnaK chaperone.[57] The final open structure obtained in free molecular dynamics simulations had the same topology as the open structure determined later experimentally.[58] Our study of the folding of the FBP28 WW domain and its mutants demonstrated that the calculated folding-rate constants are correlated with the experimental ones, which suggests that UNRES reproduces the experiment not only qualitatively but also quantitatively.[46] It should be noted that, in the examples summarized above, an earlier version of UNRES was used, which produced only coarse β-structure and loop geometry and that the accuracy of UNRES in modeling protein structural details increased considerably, as demonstrated by tests of the latest variant of the force field with benchmark proteins[41] and tests in the CASP13[55] and CASP14[56] experiments. Finally, one of the CASP3 targets (T1099) was the duck hepatitis VLP, of which the models of a minimal non-reducible unit were evaluated. We modeled the whole VLP with UNRES obtaining good results compared to the other groups, which used knowledge-based methods and all-atom models.[56] Thus, UNRES seems to be appropriate for modeling the structure and conformational dynamics of proteins at the residue or a coarser level, which was the purpose of the present study.

To determine the time evolution of the virus-like particles, we used coarse-grained molecular dynamics implementation of UNRES of our previous work.[8,59] These equations contain a non-diagonal albeit constant inertia matrix, which has been reduced to a five-band matrix

in the recent implementation of UNRES.[44] This reduction saves both memory and computing time.

2.2 | Systems studied and simulation procedure

As mentioned in the Introduction, the VLPs of the following three strains of HuNoVs were simulated: (i) the GII.4 Houston strain (PDB: 7K6V, X-ray structure[3]), (ii) the GII.2 SMV strain (PDB: 6OTF,[4] Cryo-EM structure), and (iii) the GI.1 Norwalk strain (PDB: 1IHM, X-ray structure[2]). All strains are highly conserved (Figure 2), the most substantial differences occurring in the P2 subdomain. The segments missing in the experimental structures, especially in their very flexible NTA regions were rebuilt based on the models obtained from AlphaFold.[60] For each strain, two series of simulations were carried out, one with a complete VP1 protein and one with the VP1 protein devoid of the part of the NTA domain not present in the experimental structure (cf. Figure 1B). Thus, a total of 6 simulation series were carried out. Each series consisted of 3 independent canonical UNRES MD runs.

For each system, the simulation procedure consisted of the following three steps: (1) a local energy minimization of the initial structure, (2) a short MD run at $T = 20$ K, and (3) a canonical production run carried out at $T = 300$ K whose duration was 30 ns, which corresponds to 30 µs of laboratory time.[8,44] For the VLPs with complete VP1 structures, which required relaxation of the rebuilt N-terminal segment and its accommodation to the main part of the structure, positional restraints were imposed in step (2) on the C^α atoms and the SC centers, except for those of the rebuilt parts of the flexible NTA domain. The restraints were gradually removed, hence the length of this stage is 80,000 and 20,000 steps for simulations with and without complete VP1 protein, respectively. For all runs the Variable Time Step (VTS) algorithm[59] was used to integrate the equations of motion with the time step of 4.89 fs. The temperature was maintained with the Berendsen thermostat.[61] Structures were saved every 5000 MD steps, while energies and structure-based properties (e.g., the radius of gyration) were saved every 1000 MD steps. Each trajectory of the production part of the simulations took 21–25 wall-clock days with two Intel® Xeon® E5 v3 @ 2,3 GHz, 12-core (Haswell) processors of the Tryton Linux cluster located in the Academic Computer Center in Gdańsk, TASK.

2.3 | Analysis of results

The variations of the radius of gyration (R_{gy}; Equation (1)) and anisotropy (χ; Equation (3)) with simulation time were monitored to determine the change of VLP size and shape during the course of simulations.

$$R_{gy} = \sqrt{\frac{1}{N_{chain} \times n} \sum_{I=1}^{N_{chain}} \sum_{i=1}^{n} \| \mathbf{r}_{il} - \mathbf{r}_C \|^2}, \qquad (1)$$

$$\mathbf{r}_C = \frac{1}{N_{chain} \times n} \sum_{I=1}^{N_{chain}} \sum_{i=1}^{n} \mathbf{r}_{il}, \qquad (2)$$

where N_{chain} is the number of chains of the respective VLP (equal to 180 for all systems studied), n is the number of amino-acid residues per chain, $\mathbf{r}_{il}$ are the Cartesian coordinates of the ith C^α atom of lth chain, and $\mathbf{r}_C$ are the Cartesian coordinates of the center C^α atoms of the VLP.

$$\chi = \frac{\lambda_{max} - \lambda_{min}}{\lambda_{max} + \lambda_{min}} \qquad (3)$$

where λ_{max} and λ_{min} are the maximum and minimum eigenvalue of the distribution tensor ($\mathbf{M}$) of the centers of the chains about the VLP center (Equation (4)).

$$\mathbf{M} = \frac{q}{n \times N} \sum_{I=1}^{N_{chain}} \sum_{J=1}^{N_{chain}} (\mathbf{r}_{C,I} - \mathbf{r}_C)(\mathbf{r}_{C,J} - \mathbf{r}_C), \qquad (4)$$

$$\mathbf{r}_{C,I} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{r}_{il}. \qquad (5)$$

The residue-wise fluctuations of the VLP monomers were calculated as the root mean square fluctuations (RMSF) averaged over monomers and over the second half of snapshots, as given by Equation (6).

$$RMSF_i = \sqrt{\frac{2}{M \times N} \sum_{m=M/2}^{M} \sum_{I=1}^{N_{chain}} \left\| \mathbf{r}_{C_{il}^\alpha}^{(m)} - \left\langle \mathbf{r}_{C_{il}^\alpha} \right\rangle \right\|^2}, \qquad (6)$$

$$\left\langle \mathbf{r}_{C_{il}^\alpha} \right\rangle = \frac{2}{M} \sum_{m=M/2}^{M} \sum_{I=1}^{N_{chain}} \mathbf{r}_{C_i^\alpha}^{(m)}, \qquad (7)$$

where $RMSF_i$ is the RMSF of the ith residue, M is the total number of snapshots and $\left\langle \mathbf{r}_{C_{il}^\alpha}^{(m)} \right\rangle$ is the vector of the average Cartesian coordinates of the ith C^α atom of chain l. All chains are superposed before computing the $RMSFs$.

To determine the dominant motions of the monomers, we did the principal component analysis (PCA) of the variance–covariance matrix of C^α coordinates (Equation (8)), by using ProDy (http://prody.csb.pitt.edu).[62] The PCA was carried out for monomers and all monomers from all trajectories were considered (180 copies per snapshot).

$$cov_{ij} = \frac{2}{M \times N} \sum_{m=M/2}^{M} \sum_{I=1}^{N_{chain}} \left(\mathbf{r}_{C_{il}^\alpha}^{(m)} - \left\langle \mathbf{r}_{C_{il}^\alpha} \right\rangle \right) \circ \left(\mathbf{r}_{C_{jl}^\alpha}^{(m)} - \left\langle \mathbf{r}_{C_{jl}^\alpha} \right\rangle \right) \, i = 1, 2, \ldots n, j$$
$$= 1, 2, \ldots n. \qquad (8)$$

To study the internal motion of the monomers and the motion of the P2 subdomain (that protrudes from the VLP surface), we analyzed the angle Θ_D between the centers of the S, P1, and P2 domains and the angle Θ_N between the center of the VLP, the center of P1, and the center of P2; the latter angle is close to the angle between the normal to the VLP surface and P2 axis. Both angles are illustrated in Figure 4 and defined by Equations (9) and (10), respectively.
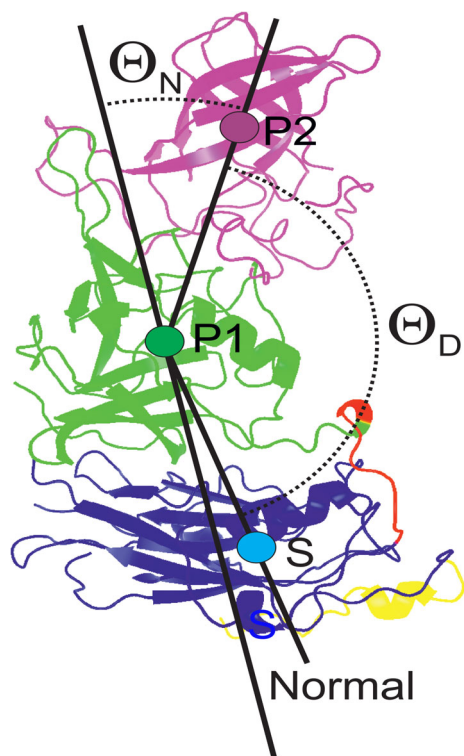
FIGURE 4 Definition of the angle Θ_N between the line linking VLP center and the center of the P1 subdomain (effectively the normal to the VLP surface) and that linking the centers of the P1 and P2 subdomains and the angle Θ_D between the centers of the S, P1, and P2 domains. The centers of the domains/subdomains are shown as spheres.

$$\cos\Theta_D = \frac{(\mathbf{C}_S - \mathbf{C}_{P1}) \circ (\mathbf{C}_{P2} - \mathbf{C}_{P1})}{\| \mathbf{C}_S - \mathbf{C}_{P1} \| \| \mathbf{C}_{P2} - \mathbf{C}_{P1} \|}, \tag{9}$$

$$\cos\Theta_N = \frac{(\mathbf{C}_{P1} - \mathbf{C}) \circ (\mathbf{C}_{P2} - \mathbf{C}_{P1})}{\| \mathbf{C}_{P1} - \mathbf{C} \| \| \mathbf{C}_{P2} - \mathbf{C}_{P1} \|}, \tag{10}$$

$$\mathbf{C}_X = \frac{1}{n_X} \sum_{i \in \{X\}} \mathbf{r}_{C_i^\alpha}, \tag{11}$$

where C_X, $X = S$, P1 or P2, denotes the Cartesian coordinates of the center of domain X, n_X and $\{X\}$ are the number of residues in that domain and the set of indices of the residues that belong to it, respectively, $\mathbf{C}$ defined by Equation (2) denotes the coordinates of the center of the entire VLP and the chain. The chain and snapshot indices were omitted.

The correlation coefficients between the Θ_N angles of different monomers, which were used to determine the concerted motions of those, are defined by Equation (12).

$$\rho(\cos\Theta_I, \cos\Theta_J) = \frac{\langle(\cos\Theta_{N,I} - \langle\cos\Theta_{N,I}\rangle)(\cos\Theta_{N,J} - \langle\cos\Theta_{N,J}\rangle)\rangle}{\sqrt{\langle\cos\Theta_{N,I} - \langle\cos\Theta_{N,I}\rangle\rangle^2 \langle\cos\Theta_{N,J} - \langle\cos\Theta_{N,J}\rangle\rangle^2}}, \tag{12}$$

where averaging is carried out over the snapshots. We used all snapshots except for the 200 initial ones in the analysis of Θ_D and Θ_N.
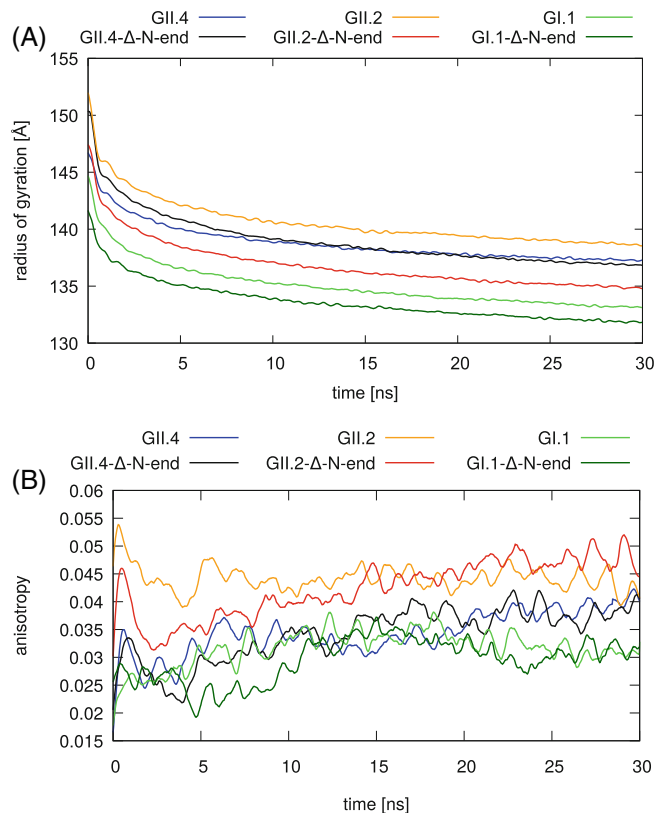

FIGURE 5 Time dependence of the radius of gyration (A) and anisotropy (B) of the GII.4 strain with (blue) and without the N-terminal segment (black), the GII.2 strain with (orange) and without the N-terminal segment (red), the GI.1 strain with (green) and without the N-terminal segment (dark green), averaged over 3 MD trajectories for each system.

To determine the variation of the exposure of the HBGA- and the NORO-antigen binding sites with simulation progress, the solvent accessible surface area (SASA) of each residue for each snapshot was computed by using the algorithm developed by Richmond,[63] which was taken from the `surfatom` subroutine of the TINKER package[64] and adapted to UNRES MD program in our earlier work.[8] The side-chain-center coordinates were used. Because the glycine residue in the UNRES model has no side chain, the respective contribution to SASA was set at 0. The SASA values were averaged over all 180 chains.

3 | RESULTS AND DISCUSSION

3.1 | Time evolution of the VLPs

In all simulations, we observed some shrinking of the VLP, the radius of gyration decreasing by about 13 Å (about 10% of the initial radius). Plots of the radius of gyration vs. time for all six series of runs, each plot averaged over all 3 trajectories of a run, are shown in Figure 5A. However, as can be seen from Figure 5A, shrinking is only moderate, which means that the structure of the particle is not destroyed. This

FIGURE 6 Illustration of the first principal component of the motion of the VP1 of the GII.4 strain with (A) and without the N-terminal segment (B). The directions and relative magnitudes of the motion of the C^α atoms are shown with arrows. All snapshots, all trajectories of a given series and all VP1 units were taken to calculate the variance–covariance matrices.

(A)

(B)
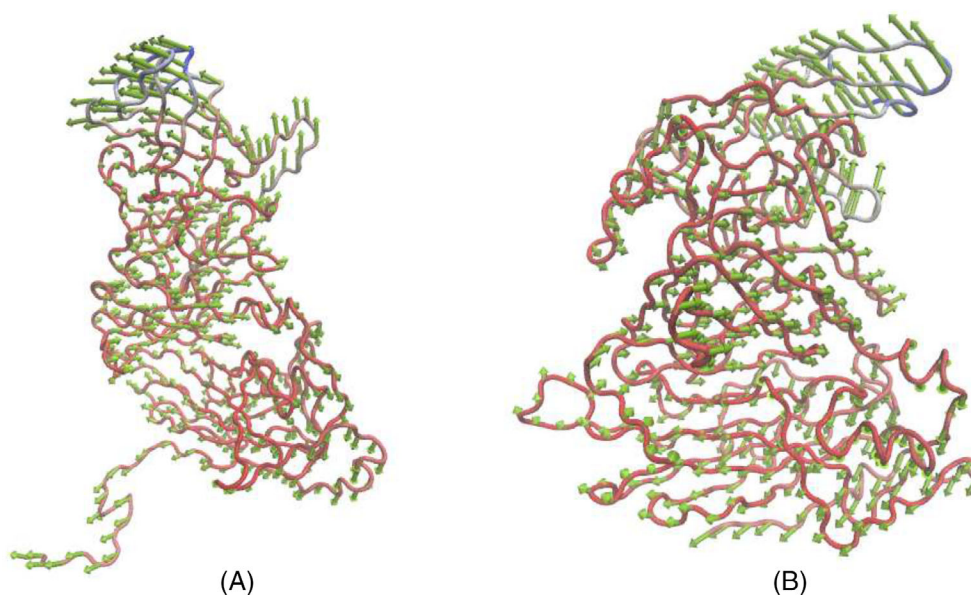
conclusion is supported by the change of anisotropy of the particle. The respective plots, for all simulation series, each averaged over all 3 trajectories of a series, are shown in Figure 5B. As can be seen from the plots, the anisotropy remains small throughout the simulations. It can also be noted that there are no remarkable differences in the long-time behavior of radius of gyration and anisotropy between the VLPs composed of full VP1 proteins and those devoid of the N-terminal section except that equilibrium seems to be reached faster when the N-terminal section is not present. Interestingly, the anisotropy is small but not zero, as also observed in the all-atom simulations of the HBV VLP.[17]

A video of the first trajectory of the VLP of the full-sequence GII.4 variant is included in the Supplementary Material.

3.2 | Internal motion of VP1 chains

The mobility of the P2 domain of VP1 certainly is an important factor determining its interactions and, ultimately, binding to HBGA or NORO. Therefore, we analyzed the motion of individual chains of VP1, as described in section "Analysis of results". PCA analysis showed that the dominant motion is the bending of the top part of the protein, which results in wagging motions of the P2 domain with respect to VLP surface, as illustrated in Figure 6A, B with the example of the GII.4 Houston strain with the full sequence and with the sequence devoid of NTA, respectively. Similar wagging motion was observed in the all-atom simulations of HBV VLP.[17] The data from the first trajectory of each of the two variants of the GII.4 VLP were used in the PCA analysis shown in Figure 6. The N-terminal part also moves significantly (Figure 6A); however, it constitutes the inner part of the VLP. The motions of the VP1s from the remaining 2 trajectories of the GII.4 VLP with the full sequence are shown in Figure S1A and S1B, while those for the sequence without the N-terminal segment are shown in Figures S1C and S1D of the Supplementary Material, respectively. The first principal components of the motions of the

monomers of the GII.2 and GI.1 VLPs are shown in Figures S2A-S2F and S3A-S3F of the Supplementary Material, respectively. As can be seen from Figures 6 and S1–S3, P2 domain bending with respect to the rest of the protein is dominant and is very similar for all the three strains analyzed regardless of whether the NTA is or is not present.

While there are no remarkable differences between strains in the form of the dominant motion of VP1, there are in its extent. This is visualized in Figure 7A, B for the VLPs composed of the respective VP1s with and without the N-terminal segment, respectively, which shows superposed plots of the RMSF of all six simulation series gathered from the second half of all trajectories. It can be seen that the P2 domain of the GII.2 SMV strain moves to the greatest extent. On the other hand, the presence or absence of NTA does not seem to have a significant effect on the extent of motion. As stated in the Introduction, the GII.2 strain does not bind to the HBGAs.

To analyze the dominant "bending" domain motion found by PCA in a more tangible and quantitative manner, we computed the distributions of the cosine of the Θ_D angle between the centers of the S, P1, and P2 domains (see Equation (9) and Figure 4 for definition). The plots of the initial distributions and those collected from all snapshots and all trajectories of the respective systems are shown in Figure 8A–8C, while the distributions collected from all snapshots for the three strains with and without N-end, respectively, are shown in Figure 8D, E. It can be seen from panels A–C of the Figure that the distributions become more diffuse and slightly shift towards more positive value (i.e., the Θ_D angle becomes less extended). From Figure 8D, E it can be seen that there are no remarkable differences between the distributions corresponding to different strains in the simulations carried out with full VP1 sequences. On the other hand, the distribution for the GII.2 strain without the N-terminal section is shifted to the right compared to those for the GII.4 and GI.1 strain and the distribution for the GII.4 strain is shifted to the left compared to that for the GI.1 strain. This observation suggests that the absence of the N-terminal section differentiates the strains regarding domain motion.

3.3 | Motion of the P2 subdomain with respect to VLP surface

Because the P2 subdomain contains most of the HBGA binding sites (Figure 1B), its location with respect to the surface is important for binding. Therefore, we calculated the angle Θ_N between the vector from VLP center and that from the center of the P1 (which is approximately the normal to the VLP surface) to that of the P2 subdomain (Equation (10) and Figure 4). The probability distribution functions of $\cos\Theta_N$ for the initial structure and averaged over all but the 200 first snapshots of all trajectories for each series of runs are shown in Figure 9A–C, each panel corresponding to a given strain with and without the N-terminal section, respectively, while the distributions for all strains with and without the N-terminal section are compared in panels D and E of the Figure. It can be seen from panels A–C of Figure 9 that the distributions became broader and shifted to the left from the initial ones, which means that the P2 domain drifts away from the normal to the VLP surface. As in the case of $\cos\Theta_D$ distributions, the differences of the distributions for the strains stripped of the N-terminal section are more remarkable than those for the full-sequence VLPs (Figure 9D, E). Remarkably, in both cases the distribution for the GII.2 strain is most shifted to the left, thus bringing the P2 domain closer to VLP surface and probably making it less accessible for HBGA binding, which conforms with the lack of HBGA binding potency of the GII.2 strain.

The angle Θ_N of the VP1 units changes during the course of the simulations. Sample plots of the time dependence of Θ_N for the GII.4 Houston strain are shown in Figure S4 of the Supplementary Material. To determine if and how the changes of Θ_N of the units are correlated, we computed the correlation coefficients of the time dependence of $\cos\Theta_N$ of all the VP1 units (Equation (12)). This correlation occurs for all trajectories; a sample heat map is shown in Figure 10A. Unit indices were sorted to group the units mostly correlated with each other. This was, in turn, accomplished by selecting, for a given simulation, the unit for which the sum of correlation coefficients with the other units was the largest and sorting the indices of the units in the descending order of the correlation coefficients with the selected unit. As can be seen from Figure 10A, two clear groups are observed with intra-groups correlation but inter-group anticorrelation. Figure 10B, in which the VLP coarse-grained representation with only unit centers is shown, the centers being colored according to the correlation coefficients with the selected unit (that has the largest sum of correlation coefficients) shows clear clusters of the correlated units. One group (colored red, saturation dependent on the correlation coefficient) are those units correlated with the selected unit, the other group (colored blue, saturation dependent on anticorrelation) contains the units anticorrelated with the units of the first group but correlated among themselves. Clusters of correlated units can be identified, some of which are linear and resemble allosteric pathways that rely the wagging motion of the P2 domain with respect to the normal to the VLP surface. A minor group (light-blue, light-red or white color) contains non-correlated units. This result suggests that the motion of the VP1 units with
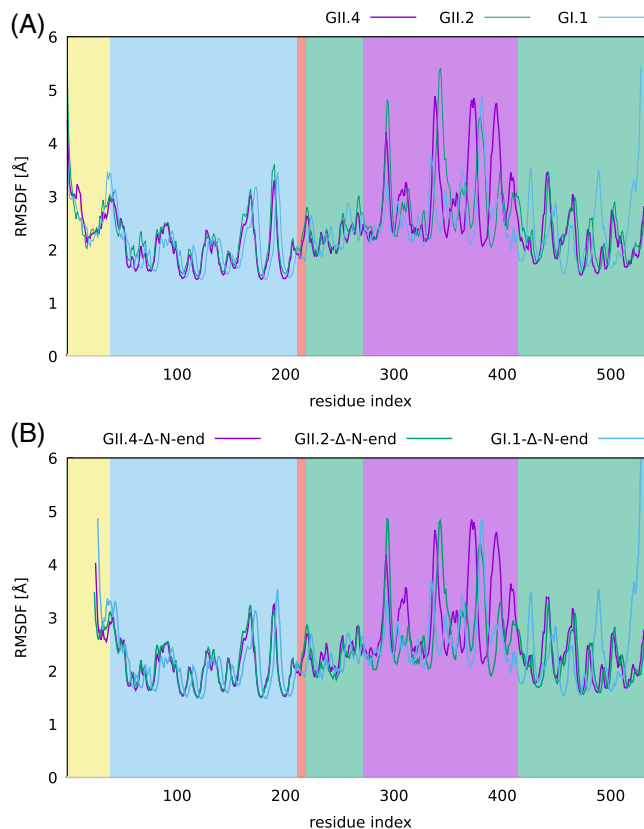

FIGURE 7 Root-mean-square fluctuations (RMSF) of the C^α atoms of the consecutive residues of the VP1 proteins of the VLP of the GII.4, GII.2, and GI.1 strains averaged over all monomers, all snapshots and all trajectories of a given system, corresponding to simulations with (A) and without (B) the N-terminal segment.

respect to the VLP surface is highly concerted. On the other hand, this seems to be a general feature of VLP dynamics, which is similar to all strains, regardless of whether the N-terminal section is or is not present. As mentioned in the Introduction, shape dynamics affects exposing or hiding the binding sites[34,35] and can, therefore, affect the binding of the VP1 protein to the receptors. All correlation plots are shown in Figures S5A–S5F (the GII.4 strain), S6A–S7F (the GII.2 strain), and S7A–S7F (the GI.1 strain) of the Supplementary Material, respectively.

The correlated motions can also be seen when inspecting the video of the first trajectory of the full-sequence GII.4 VLP, which is included in the Supplementary Material.

3.4 | Solvent accessible surface area analysis

To determine how the areas accessible to the respective receptors of the HBGA- and NORO-antigen-binding sites of the three norovirus strains studied change during the course of the simulations, if there are differences between the strains, and if the presence of N-terminal segment changes the area, we computed the SASA over all simulations for a given kind of VLP, as described in section "Analysis of
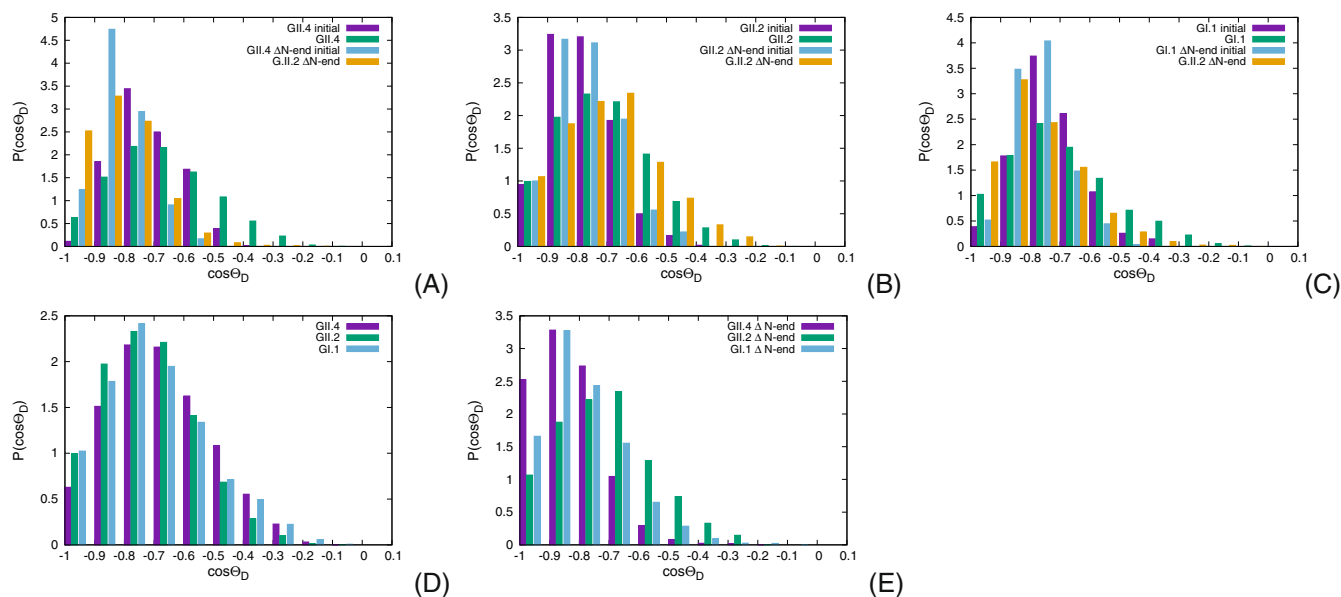
FIGURE 8 Histograms of the probability density distribution of $\cos\Theta_D$ (the cosine of the angle between the centers of the S, P1, and P2 domains; Figure 4). The initial distributions and those drawn from all but the 200 initial snapshots are shown in panels A–C, while the distributions for the GII.4, GII.2 and GI.1 strains with and without the N-terminal section are shown in panels D and E, respectively. The graph colors correspond to different distributions, labels shown in the plot area.
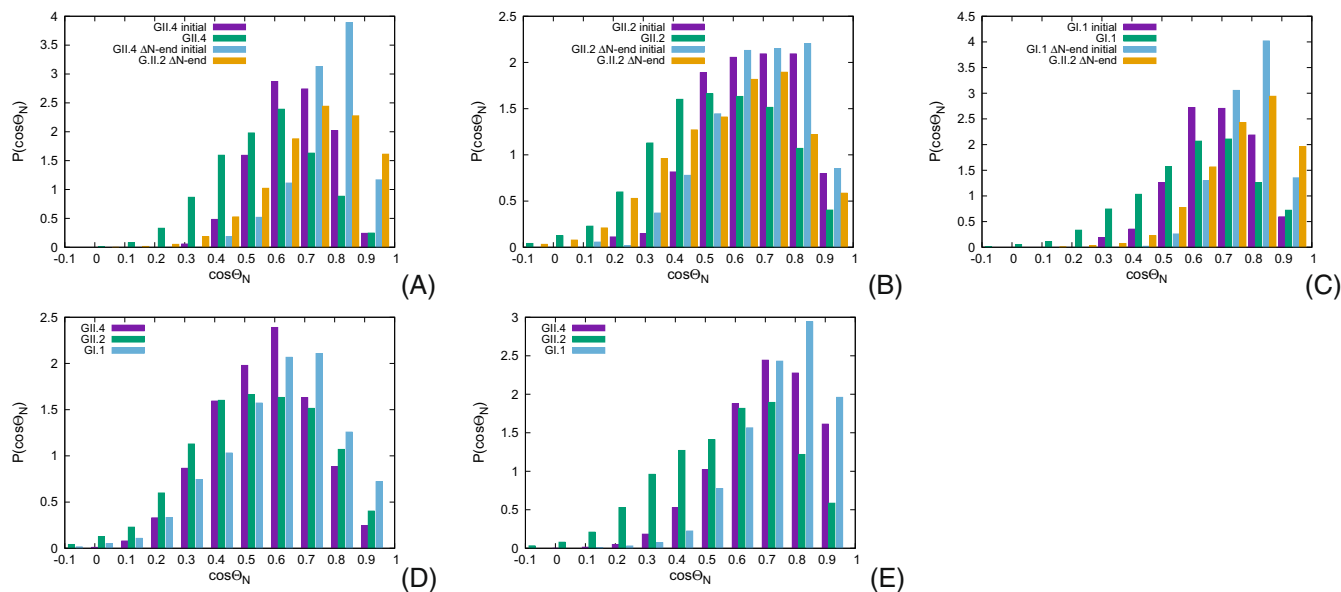

FIGURE 9 Histograms of the probability density distribution of $\cos\Theta_N$ (the cosine of the angle between the line linking the VLP center with the center of the P1 domain and that linking the P1 and P2 domains; Figure 4). The initial distributions and those drawn from all but the 200 initial snapshots are shown in panels A–C, while the distributions for the GII.4, GII.2 and GI.1 strains with and without the N-terminal section are shown in panels D and E, respectively. The graph colors correspond to different distributions, labels shown in the plot area.

results". The SASA plots are shown in Figure 11A, B, for systems with and without the N-terminal segments, respectively. As can be seen from the Figure, for all systems the SASA shrinks initially to stabilize after about 20 ns, in agreement with the variation of the radius of gyration (Figure 5A).

It can be seen from Figure 11A that the GII.2 strain that does not bind to HBGA has the largest area which could be available for the HBGA receptor (842.8 Å²). Conversely, the GII.4 and the GI.1 strains that bind HBGA have smaller SASA of that site (697.205 Å² and 620.575 Å², respectively). The differences in the affinity for HBGA
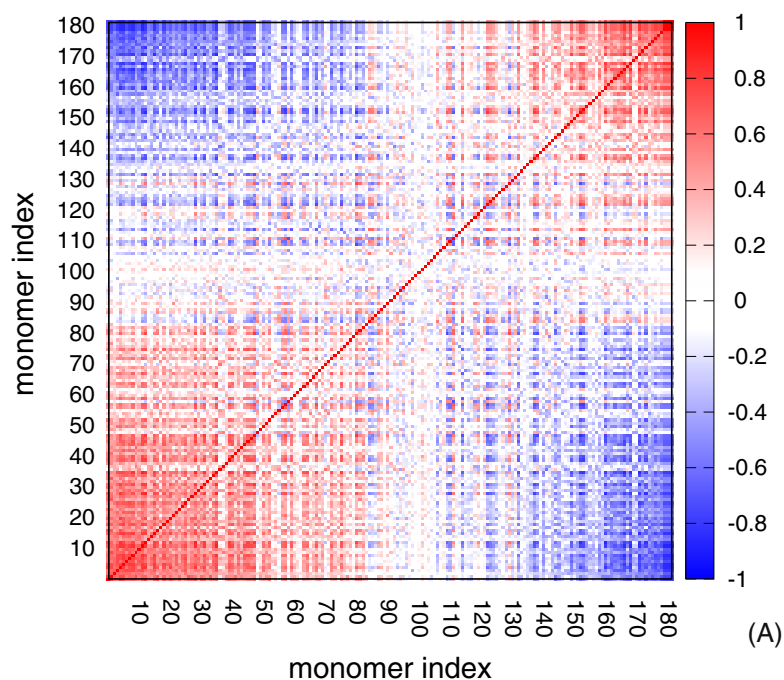
(A)

FIGURE 10 Heat map of the correlation between the time series of $\cos\Theta_N$ of different units averaged over all but 200 first snapshots of one of the trajectories of the GII.2 strain (A) and the same heat map superposed on the VLP represented by the centers of the VP1 units in two projections (B). Two major groups of units, each having correlated $\cos\theta_N$, while the inter-group $\cos\theta_N$ time series are anticorrelated can be seen in panel A. The putative allosteric paths corresponding to group 1 and group 2 running along the equator of the respective projection can be seen in the left and right sections of panel B, respectively.

(B)
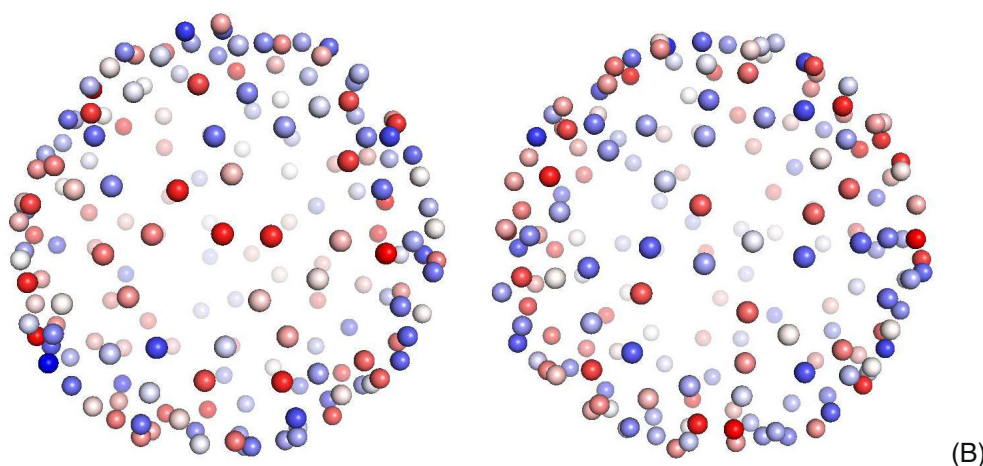
could therefore be related to mobility and steric properties rather than to SASA. In fact, the P2 domain of the GII.2 strain makes a greater angle with the normal than those of the two other strains (Figure 9), which can result in unfavorable binding geometry. Moreover, the fluctuations of the P2 domain are more pronounced for this strain (cf. Figure 7), which can additionally make the binding more difficult.

The NORO-antigen-binding site, on the other hand, has the greatest SASA for the GII.4 strain (188.5 Å^2), compared to the GII.2 (138.2 Å^2) and GI.1 (126.8 Å^2) strains. For the GII.4 strain, SASA also is remarkably larger for the VLPs without the N-terminal segments compared to that with the full sequence (Figure 11B). The SASA of the NORO-antigen-binding sites of the GII.2 and GI.1 strains is smaller for the truncated compared to the full sequences (Figure 10B). These results suggest that the presence of the N-terminal part could influence the dynamic structures of the strains of HuNoV VLPs and, consequently, influence their interactions with the receptor.

4 | CONCLUSIONS

The results of long-time simulations of the virus-like particles corresponding to GII.4, GII.2, and GI.1 strains of the human norovirus studied in this work, with and without the N-terminal segments, suggest that molecular dynamics with UNRES can be a useful tool to study the behavior of such systems in relation to their interactions with the immune system. In all simulations the symmetry of the VLP was largely preserved, the anisotropy being very small, and only a moderate shrinking of the particle was observed (Figure 5). The laboratory time scale of the simulations was about 30 μs, which enabled us to draw meaningful conclusions.

The simulations enabled us to determine the character of the local and global motion of the P2 domain that protrudes from the surface. The local motions are the wagging motions of P2 with respect to the P1 domain and, consequently wagging motions with respect to
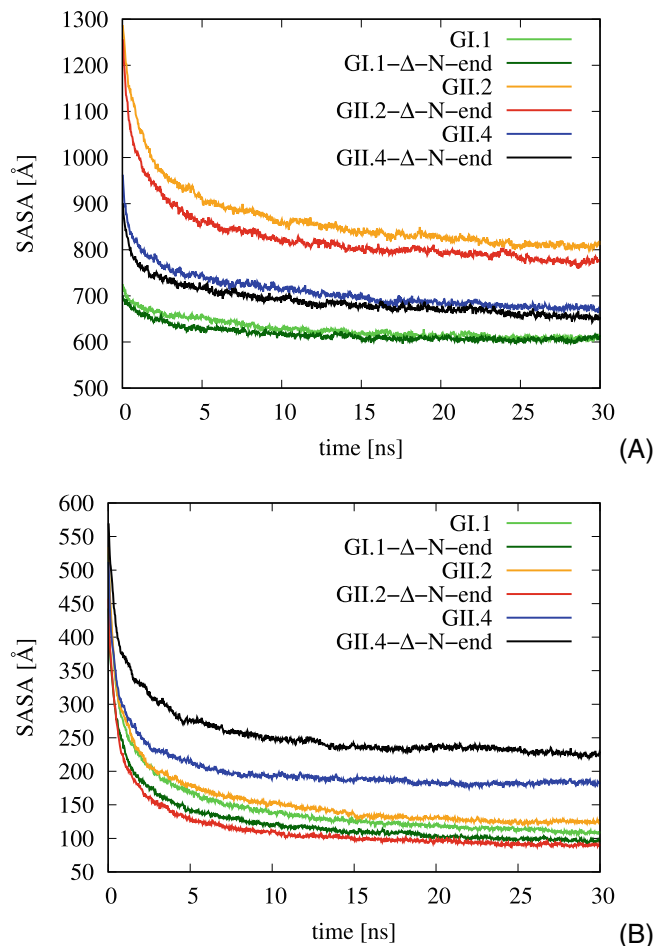
FIGURE 11 Time dependence of the total solvent accessible surface areas (SASAs) of residues comprising the HBAGA (A) and NORO-antigen (B) binding sites of the GII.4 strain with (blue) and without (black) N-terminal segment, the GII.2 strain with (orange) and without (red) the N-terminal segment, and the GI.1 strain with (green) and without (dark green) N-terminal segment. Each curve is an average over the 3 trajectories for the respective system.

the VLP surface (Figure 6 and S1 – S3 of the Supplementary Material), making them more or less accessible to the receptors. The fluctuations of the P2 domains are most pronounced for the VLP corresponding to the GII.2 strain (Figure 7) that, in contrast to the GII.4 and GI.2 strains, does not bind to the histo-blood group antigen (HBGA). Moreover, the P2 domain makes the largest angle Θ_N (Figure 4) with the normal to the surface for this strain (Figure 9 E), this suggesting that it is least accessible to HBGA. The absence of the N-terminal domain, even though it is located at the inner surface of the VLP makes the deviation of the P2 from the normal to the VLP surface even more pronounced and also differentiates the orientation of the P2 domain with respect to the VLP surface for GII.2 and GI.2.

An interesting feature of the motion of the P2 domain with respect to the surface is their concerted character. Clear two groups can be distinguished, inside which the motion is correlated but anticorrelated between the groups. This suggests the emergence of

allosteric pathways along which the changes are relayed (Figure 10 and S5–S7 of the Supplementary Material).

The results of our study suggest that, for each of the three strains studied, removing the putatively unstructured part of the N-terminal segment not present in the experimental structure influences the structure and dynamics of the VLPs. The distributions of the Θ_N angles of all three strains become different after removing this segment (Figure 9 E). Moreover, the surface of the NORO-antigen-binding site of the GII.4 strain increases significantly when the sequence of the virus-shell protein is truncated (Figure 11B). Consequently, the N-terminal segments could influence the receptor-binding properties of HuNoVs.

To our knowledge, this work is the first simulation study of the dynamics of the complete VLPs of the human-norovirus strains. Although the achieved time scale is significant, simulation of events such as virus-particle assembly or its binding to receptors requires still longer time scales. To meet these demands, we are now working on porting UNRES to the Graphical Processor Units.

DATA AVAILABILITY STATEMENT

The structures corresponding to the snapshots from MD trajectories and the numerical values of the quantities displayed in Figures 5–11 and Figures S1–S7 of the Supplementary Material are available from the authors upon request.

ORCID

Agnieszka G. Lipska https://orcid.org/0000-0002-8390-8473
Adam K. Sieradzan https://orcid.org/0000-0002-2426-3644
Cezary Czaplewski https://orcid.org/0000-0002-0294-3403
Andrea D. Lipińska https://orcid.org/0000-0002-8115-9499
Krzysztof M. Ocetkiewicz https://orcid.org/0000-0002-4472-6116
Jerzy Proficz https://orcid.org/0000-0003-2975-9339
Paweł Czarnul https://orcid.org/0000-0002-4918-9196
Henryk Krawczyk https://orcid.org/0000-0003-0436-6264
Adam Liwo https://orcid.org/0000-0001-6942-2226

REFERENCES

[1] E. A. Almand, M. D. Moore, L.-A. Jaykus, *Front. Microbiol.* **2017**, *8*, 2549.

[2] B. V. Prasad, M. E. Hardy, T. Dokland, J. Bella, M. G. Rossmann, M. K. Estes, *Science* **1999**, *286*, 287.

[3] L. Hu, S. E. Crawford, R. Czako, N. W. Cortes-Penfield, D. F. Smith, J. Le Pendu, M. K. Estes, B. V. Prasad, *Nature* **2012**, *485*, 256.

[4] J. Jung, T. Grant, D. R. Thomas, C. W. Diehnelt, N. Grigorieff, L. Joshua-Tor, *Proc. Natl. Acad. Sci. U. S. A.* **2019**, *116*, 12828.

[5] E. Robilotti, S. Deresinski, B. A. Pinsky, *Clin. Microbiol. Rev.* **2015**, *28*, 134.

[6] S. Esposito, N. Principi, *Front. Immunol.* **2020**, *11*, 1383.

[7] S. Kmiecik, D. Gront, M. Kolinski, L. Wieteska, A. E. Dawid, A. Kolinski, *Chem. Rev.* **2016**, *116*, 7898.

[8] M. Khalili, A. Liwo, A. Jagielska, H. A. Scheraga, *J. Phys. Chem. B* **2005**, *109*, 13798.

[9] A. K. Jana, E. R. May, *Curr. Opin. Virol.* **2020**, *45*, 8.

[10] P. L. Freddolino, A. Y. Shih, A. Arkhipov, Y. Ying, Z. Zheng, K., in *Coarse-Graining of Condensed Phase and Biomolecular Systems* (Ed: G. Voth), CRC Press, Boca Raton, **2008**, p. 299.

[11] J. R. Perilla, B. C. Goh, C. K. Cassidy, B. Liu, R. C. Bernardi, T. Rudack, H. Yu, Z. Wu, K. Schulten, *Curr. Opin. Struct. Biol.* **2015**, *31*, 64.

[12] E. Tarasova, D. Nerukh, *J. Phys. Chem. Lett.* **2018**, *9*, 5805.

[13] P. E. Jones, C. Pérez-Segura, A. J. Bryer, J. R. Perilla, J. A. Hadden-Perilla, *Curr. Opin. Virol.* **2021**, *50*, 128.

[14] P. L. Freddolino, A. S. Arkhipov, S. B. Larson, A. McPherson, K. Schulten, *Structure* **2006**, *14*, 437.

[15] E. Tarasova, I. Korotkin, V. Farafonov, S. Karabasov, D. Nerukh, *J. Mol. Liq.* **2017**, *245*, 109.

[16] A. Arkhipov, W. H. Roos, G. J. Wuite, K. Schulten, *Biophys. J.* **2009**, *97*, 2061.

[17] J. A. Hadden, J. R. Perilla, C. J. Schlicksup, B. Venkatakrishnan, A. Zlotnik, K. Schulten, *eLife* **2018**, *7*, e32478.

[18] F. Mohajerani, B. Tyukodi, C. J. Schlicksup, J. A. Hadden-Perilla, A. Zlotnik, M. F. Hagan, *ACS Nano* **2022**, *27*, 13845.

[19] A. Yu, A. J. Pak, P. He, V. Monje-Galvan, L. Casalino, Z. Gaieb, A. C. Dommer, R. E. Amaro, G. A. Voth, *Biophys. J.* **2021**, *120*, 1097.

[20] M. Cieplak, M. D. Robbins, *PLoS One* **2013**, *8*, e63640.

[21] K. Wolek, M. Cieplak, *J. Phys. Condens. Matter* **2017**, *29*, 474003.

[22] C. P. Mattison, C. V. Cardemil, A. J. Hall, *Expert Rev. Vaccines* **2018**, *17*, 773.

[23] P. Chhabra, M. de Graaf, G. I. Parra, M. C.-W. Chan, K. Green, V. Martella, Q. Wang, P. A. White, K. Katayama, H. Vennema, M. P. G. Koopmans, J. Vinjé, *J. Gen. Virol.* **2019**, *100*, 1393.

[24] S. Vongpunsawad, B. Venkataram Prasad, M. K. Estes, *J. Virol.* **2013**, *87*, 4818.

[25] L. C. Lindesmith, E. F. Donaldson, A. D. LoBue, J. L. Cannon, D.-P. Zheng, J. Vinje, R. S. Baric, *PLoS Med.* **2008**, *5*, e31.

[26] L. C. Lindesmith, E. F. Donaldson, R. S. Baric, *J. Virol.* **2011**, *85*, 231.

[27] L. Hu, W. Salmen, R. Chen, Y. Zhou, F. Neill, J. E. Crowe, R. L. Atmar, M. K. Estes, B. Prasad, *Nat. Commun.* **2022**, *13*, 1.

[28] M. Panasiuk, K. Zimmer, A. Czarnota, M. Narajczyk, G. Peszyńska-Sularz, M. Chraniuk, L. Hovhannisyan, S. Żołędowska, D. Nidzworski, A. J. Żaczek, B. Gromadzka, *J. Nanobiotechnol.* **2022**, *20*, 1.

[29] M. Yildiz, A. Kocak, *J. Comput. Biol.* **2019**, *26*, 962.

[30] O. Ebenezer, N. Damoyi, M. Shapi, *Front. Chem.* **2021**, *9*, 753427.

[31] E. Behmard, A. Ghasemian, E. Barzegari, A. Farjadfar, A. Kouhpayeh, S. Najafipour, *J. Mol. Graph. Model.* **2023**, *118*, 108345.

[32] T. Domitrovic, N. Movahed, B. Brian, T. Matsui, Q. Wang, P. Doerschuk, J. E. Johnson, *J. Mol. Biol.* **2013**, *425*, 1488.

[33] M. G. Matheu, *Arch. Biochem. Biophys.* **2013**, *531*, 65.

[34] G. S. Hansman, D. W. Taylor, J. S. McLellan, T. J. Smith, I. Georgiev, J. R. H. Tame, S.-Y. Park, M. Yamazaki, F. Gondaira, M. Miki, K. Katayama, K. Murata, P. D. Kwong, *J. Virol.* **2012**, *86*, 3635.

[35] H. Q. Smith, T. J. Smith, *Viruses* **2019**, *11*, 235.

[36] P. Di Tommaso, S. Moretti, I. Xenarios, M. Orobitg, A. Montanyola, J.-M. Chang, J.-F. Taly, C. Notredame, *Nucleic Acids Res.* **2011**, *39*, W13.

[37] A. Bertolotti-Ciarlet, A. A. White, R. Chen, B. V. Prasad, M. K. Estes, *J. Virol.* **2002**, *76*, 4044.

[38] A. Liwo, S. Oldziej, M. R. Pincus, R. J. Wawak, S. Rackovsky, H. A. Scheraga, *J. Comput. Chem.* **1997**, *18*, 849.

[39] A. Liwo, M. Baranowski, C. Czaplewski, E. Gołaś, Y. He, D. Jagieła, P. Krupa, M. Maciejczyk, M. Makowski, M. A. Mozolewska, A. Niadzvedtski, S. Ołdziej, H. A. Scheraga, A. K. Sieradzan, R. Ślusarz, T. Wirecki, Y. Yin, B. Zaborowski, *J. Mol. Model.* **2014**, *20*, 2306.

[40] A. K. Sieradzan, M. Makowski, A. Augustynowicz, A. Liwo, *J. Chem. Phys.* **2017**, *146*, 124106.

[41] A. Liwo, A. K. Sieradzan, A. G. Lipska, C. Czaplewski, I. Joung, W. Żmudzińska, A. Halabis, S. Ołdziej, *J. Chem. Phys.* **2019**, *150*, 155104.

[42] A. K. Sieradzan, C. Czaplewski, P. Krupa, M. A. Mozolewska, A. S. Karczyńska, A. G. Lipska, E. A. Lubecka, E. Gołaś, T. Wirecki, M. Makowski, S. Ołdziej, A. Liwo, *Modeling the Structure, Dynamics, and Transformations of Proteins with the UNRES Force Field*, Springer US, New York, NY **2022**, p. 399 ISBN 978-1-0716-1716-8.

[43] A. Liwo, S. Ołdziej, C. Czaplewski, D. S. Kleinerman, P. Blood, H. A. Scheraga, *J. Chem. Theor. Comput.* **2010**, *6*, 583.

[44] A. K. Sieradzan, J. Sans-Duñò, E. A. Lubecka, C. Czaplewski, A. G. Lipska, H. Leszczynski, K. M. Ocetkiewicz, J. Proficz, P. Czarnul, H. Krawczyk, A. Liwo, *J. Comput. Chem.* **2023**, *44*, 602.

[45] K. Lindorff-Larsen, S. Piana, R. O. Dror, D. E. Shaw, *Science* **2011**, *334*, 517.

[46] R. Zhou, G. G. Maisuradze, D. Suñol, T. Todorovski, M. J. Macias, Y. Xiao, H. A. Scheraga, C. Czaplewski, A. Liwo, *Proc. Natl. Acad. Sci. U. S. A.* **2014**, *111*, 18243.

[47] A. Liwo, C. Czaplewski, J. Pillardy, H. A. Scheraga, *J. Chem. Phys.* **2001**, *115*, 2323.

[48] A. Kolinski, J. Skolnick, *J. Chem. Phys.* **1992**, *97*, 9412.

[49] R. Kubo, *J. Phys. Soc. Jpn.* **1962**, *17*, 1100.

[50] A. Liwo, M. Khalili, C. Czaplewski, S. Kalinowski, S. Ołdziej, K. Wachucik, H. A. Scheraga, *J. Phys. Chem. B* **2007**, *111*, 260.

[51] A. Liwo, C. Czaplewski, A. K. Sieradzan, E. A. Lubecka, A. G. Lipska, Ł. Golon, A. Karczyńka, P. Krupa, M. A. Mozolewska, M. Makowski, R. Ganzynkowicz, A. Giełdoń, M. Maciejczyk, in *Computational Approaches for Understanding Dynamical Systems: Protein Folding and Assembly*, Vol. 170 (Eds: B. Strodel, B. Barz), Academic Press, London **2020**, p. 73.

[52] B. Zaborowski, D. Jagieła, C. Czaplewski, A. Hałabis, A. Lewandowska, W. Żmudzińska, S. Ołdziej, A. Karczyńska, C. Omieczynski, T. Wirecki, A. Liwo, *J. Chem. Inf. Model.* **2015**, *55*, 2050.

[53] Y. He, M. A. Mozolewska, P. Krupa, A. K. Sieradzan, T. K. Wirecki, A. Liwo, K. Kachlishvili, S. Rackovsky, D. Jagieła, R. Ślusarz, C. R. Czaplewski, S. Ołdziej, H. A. Scheraga, *Proc. Natl. Acad. Sci. U. S. A.* **2013**, *110*, 14936.

[54] P. Krupa, M. A. Mozolewska, M. Wiśniewska, Y. Yanping, Y. He, A. K. Sieradzan, R. Ganzynkowicz, A. G. Lipska, A. Karczyńska, M. Ślusarz, R. Ślusarz, A. Giełdoń, C. Czaplewski, D. Jagieła, B. Zaborowski, H. A. Scheraga, A. Liwo, *Bioinformatics* **2016**, *32*, 3270.

[55] E. A. Lubecka, A. S. Karczyńska, A. G. Lipska, A. K. Sieradzan, K. Zięba, C. Sikorska, U. Uciechowska, S. A. Samsonov, P. Krupa, M. A. Mozolewska, Ł. Golon, A. Giełdoń, C. Czaplewski, R. Ślusarz, M. Ślusarz, S. N. Crivelli, A. Liwo, *J. Mol. Graph. Model.* **2019**, *92*, 154.

[56] A. Antoniak, I. Biskupek, K. K. Bojarski, C. Czaplewski, A. Giełdoń, M. Kogut, M. M. Kogut, P. Krupa, A. G. Lipska, A. Liwo, E. A. Lubecka, M. Marcisz, M. Maszota-Zieleniak, S. A. Samsonov, A. K. Sieradzan, M. J. Ślusarz, R. Ślusarz, P. A. Wesołowski, K. Zięba, *J. Mol. Graph. Model.* **2021**, *108*, 108008.

[57] E. I. Golas, G. G. Maisuradze, P. Senet, S. Ołdziej, C. Czaplewski, H. A. Scheraga, A. Liwo, *J. Chem. Theor. Comput.* **2012**, *8*, 1334.

[58] R. Kityk, J. Kopp, I. Sinning, M. P. Mayer, *Mol. Cell* **2012**, *48*, 863.

[59] M. Khalili, A. Liwo, F. Rakowski, P. Grochowski, H. A. Scheraga, *J. Phys. Chem. B* **2005**, *109*, 13785.

[60] J. Jumper, R. Evans, A. Pritzel, T. Green, M. Figurnov, O. Ronneberger, K. Tunyasuvunakool, R. Bates, A. Žídek, A. Potapenko, et al., *Nature* **2021**, *596*, 583.

[61] H. J. C. Berendsen, J. P. M. Postma, W. F. van Gunsteren, A. DiNola, J. R. Haak, *J. Chem. Phys.* **1984**, *81*, 3684.

[62] S. Zhang, J. M. Krieger, Y. Zhang, C. Kaya, B. Kaynak, K. Mikulska-Ruminska, P. Doruker, H. Li, I. Bahar, *Bioinformatics* **2021**, *37*, 3657.

[63] T. J. Richmond, *J. Mol. Biol.* **1984**, *178*, 63.

[64] O. Adjoua, L. Lagardére, L.-H. Jolly, A. Durocher, T. Very, I. Dupays, Z. Wang, T. J. Inizan, F. Célerse, P. Ren, J. W. Ponder, J.-P. Piquemal, *J. Chem. Theory Comput.* **2021**, *17*, 2034.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.