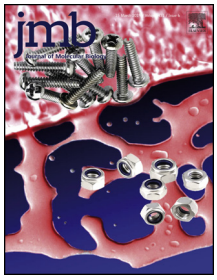

Solution Conformation of Bovine Leukemia Virus Gag Suggests an Elongated Structure

Dominic F. Qualley[1], Sarah E. Cooper[1], James L. Ross[1], Erik D. Olson[2], William A. Cantara[2] and Karin Musier-Forsyth[2]

1 - Department of Chemistry and Biochemistry, *and Center for One Health Studies, Berry College, Mt. Berry, GA 30149, USA*
2 - Department of Chemistry and Biochemistry, *Center for RNA Biology, and Center for Retrovirus Research, Ohio State University, Columbus, OH 43210, USA*

Correspondence to Dominic F. Qualley: *Department of Chemistry and Biochemistry, Berry College, 2277 Martha Berry Hwy NW, Mt. Berry, GA 30149, USA. dqualley@berry.edu*
https://doi.org/10.1016/j.jmb.2019.01.036
Edited by Eric O. Freed

Abstract

Bovine leukemia virus (BLV) is a deltaretrovirus that infects domestic cattle. The structural protein Gag, found in all retroviruses, is a polyprotein comprising three major functional domains: matrix (MA), capsid (CA), and nucleocapsid (NC). Previous studies have shown that both mature BLV MA and NC are able to bind to nucleic acids; however, the viral assembly process and packaging of viral genomic RNA requires full-length Gag to produce infectious particles. Compared to lentiviruses, little is known about the structure of the Gag polyprotein of deltaretroviruses. In this work, structural models of full-length BLV Gag and Gag lacking the MA domain were generated based on previous structural data of individual domains, homology modeling, and flexible fitting to SAXS data using molecular dynamics. The models were used in molecular dynamic simulations to determine the relative mobility of the protein backbone. Functional annealing assays revealed the role of MA in the nucleic acid chaperone activity of BLV Gag. Our results show that full-length BLV Gag has an elongated rod-shaped structure that is relatively rigid, with the exception of the linker between the MA and CA domains. Deletion of the MA domain maintains the elongated structure but alters the rate of BLV Gag-facilitated annealing of two complementary nucleic acids. These data are consistent with a role for the MA domain of retroviral Gag proteins in modulating nucleic acid binding and chaperone activity.

Importance

BLV is a retrovirus that is found worldwide in domestic cattle. Since BLV infection has serious implications for agriculture, and given its similarities to human retroviruses such as HTLV-1, the development of an effective treatment would have numerous benefits. The Gag polyprotein exists in all retroviruses and is a key player in viral assembly. However, the full-length structure of Gag from any virus has yet to be elucidated at high resolution. This study provides structural data for BLV Gag and could be a starting point for modeling Gag–small molecule interactions with the ultimate goal of developing of a new class of pharmaceuticals.

Introduction

Retroviridae is a family of viruses that replicate their single-stranded RNA genome via a double-stranded DNA intermediate [1]. Retroviruses can be divided into different genera, including lentiviruses, such as human immunodeficiency virus type 1 (HIV-1); alpharetroviruses, such as Rous sarcoma virus (RSV); betaretroviruses, including Mason–Pfizer monkey virus; gammaretroviruses, such as Moloney murine leukemia virus (MLV); and deltaretroviruses, such as human T-cell leukemia virus type 1 (HTLV-1). During the retroviral life cycle, newly produced virions undergo a process called maturation, wherein the Gag structural polyprotein is specifically cleaved by a virally encoded protease. Maturation yields, at minimum, three mature proteins: matrix (MA), which remains attached to the inner leaflet of the viral membrane; capsid (CA), which

forms the viral capsid core; and nucleocapsid (NC), a small, generally basic protein that binds to the genomic RNA (gRNA) and facilitates gRNA dimerization, packaging, and reverse transcription [2].

Prior to maturation and viral budding, the full-length Gag protein plays important roles in viral assembly and genome packaging, topics that have been the focus of several review papers [3–11]. A common step in the formation of immature retroviral particles is the oligomerization of Gag molecules to form a spherical lattice at the site of assembly. Retroviruses of different genera use different pathways to accomplish this. HIV-1 Gag has been shown to dimerize in the cytoplasm of infected cells and form higher-order oligomers upon membrane binding, while HTLV-1 Gag is monomeric in the cytoplasm and is capable of membrane targeting at lower concentrations [12–14]. HIV-1 Gag may also form oligomers in the cytoplasm in a concentration-dependent manner; the NC region in particular is important to this process [15]. It has also been proposed that HIV-1 Gag undergoes a conformational switch upon contact with gRNA, as well as when the Gag–RNA complex reaches the plasma membrane [16,17]. By contrast, betaretroviruses assemble in the cytoplasm and then traffic to the plasma membrane prior to budding from the cell [18].

There are currently no existing atomic-resolution structures of full-length retroviral Gag proteins; the intrinsic interdomain flexibility of these proteins makes crystallization challenging, and the size of most Gag proteins is too large for structure determination by conventional NMR spectroscopy. Atomic-resolution structures exist for a number of individual Gag domains from different retroviruses. Using these structures, models of HIV-1 Gag were constructed and tested against data from small-angle neutron scattering (SANS) experiments [19]. The models that best fit the SANS data provide strong evidence that HIV-1 Gag is quite flexible and can adopt a bent conformation in solution where the N- and C-termini are in close proximity. Indeed, MA and NC domains can simultaneously bind to RNA or to model membranes in a bent conformation [16]. Solution NMR experiments of HIV-1 Gag and Gag-derived fragments indicate that the linker regions between domains are intrinsically disordered and dynamic [20,21]. Both ensemble and single-molecule studies have demonstrated that HIV-1 Gag undergoes a conformational change to an extended rod-like structure in the presence of both nucleic acids and model membranes or inositol hexakisphosphate (IP6), a soluble analogue of phosphatidylinositol-(4,5)-bisphosphate (PIP_2) that has been shown to mimic interaction of Gag with the plasma membrane [16,22,23]. Interestingly, IP6 has also been shown to bind to the CA domain of HIV-1 Gag, facilitating the assembly of a hexameric lattice, which is the basis for formation of the mature viral capsid [24,25]. After Gag processing by the virally encoded protease, the NC protein remains bound to the RNA. During reverse transcription, HIV-1 NC facilitates strand transfer reactions via its nucleic acid chaperone activity [2]. The chaperone activity of NC is also evident in the context of the Gag polyprotein, assisting in both RNA dimerization and tRNA annealing to the primer binding site [22,26–29]. The presence of the MA domain lowers the efficiency of HIV-1 Gag's chaperone activity, consistent with observations that MA-RNA binding serves to inhibit non-specific lipid association [22,30–33]. This effect was not observed for RSV Gag, likely reflecting a difference in the nucleic acid binding affinities of HIV-1 and RSV MA [34]. Furthermore, the overexpression of polyphosphoinositide 5-phosphatase IV was shown to have little effect on HTLV-1 assembly, suggesting an alternative pathway for deltaretroviruses [35]. An alternative pathway could involve interaction with intracellular membranes, or preference for a different phospholipid as observed for equine infectious anemia virus (a lentivirus) [36]. One proposed mechanism for HIV-1 assembly involves an intermediate where both NC and MA domains of Gag are simultaneously bound to RNA (gRNA and cellular tRNA, respectively) [11,37]. At the plasma membrane, IP-containing lipids are able to displace RNA from the MA domain, while NC remains bound to gRNA.

A molecular envelope of purified MLV Gag derived from small-angle X-ray scattering (SAXS) data suggests a much more rigid structure than that of HIV-1 Gag [16,19,38,39]. The extended structure is likely the result of proline-rich sequences in its MA and p12 domains and a highly charged “electric wire” region in the CA C-terminal domain [38]. In contrast, experiments with purified RSV Gag show that it is flexible and can exist as an extended form in solution but adopts a more compact shape when bound to membrane [40]. The difference in solution conformation and flexibility between retroviral Gag proteins is in accord with the idea that different retroviral genera follow different assembly pathways, a concept that has been discussed in detail [5,7].

Bovine leukemia virus (BLV) has been shown to have a negative impact on milk productivity of infected cattle [41,42]. More recently, proviral DNA from BLV has been discovered in human breast tissue; a subsequent study found a correlation between the presence of BLV DNA and human breast cancer, although the significance of these findings is under debate [43–50]. BLV has also been suggested as an animal model for HTLV-1 due to their close phylogenetic relationship (both are deltaretroviruses) [51–54]. However, retroviral Gag proteins may functionally differ between viruses in the same genus. For example, HTLV-1 and BLV NC proteins display very different levels of nucleic acid chaperone activity, likely a reflection of BLV NC's closer structural similarity to that of HIV-1 NC [55,56]. On the other hand, the MA domains of Gag from BLV and HTLV-2 were both shown to play a critical role in the packaging of gRNA [57–59]. In HTLV-1, Gag–Gag interactions are critical

for viral assembly and budding; interestingly, the key residues for these interactions reside in the N-terminal domain of CA [60,61]. This suggests similarity to HIV-1 in that residues of CA also play a role in assembly but, unlike HIV-1, occur in the absence of membrane phosphoinositides [24,25,35].

Given the recent findings that possibly link BLV and breast cancer, and the potential use of BLV as an animal model for HTLV-1, learning more about the replication process of BLV is becoming increasingly important. Furthermore, as drug-resistant HIV-1 isolates continue to emerge, there remains a need for development of new inhibitors [62]. Most classes of anti-retroviral drugs focus on the inhibition of virally encoded enzymes such as reverse transcriptase, although viral resistance to these compounds is becoming more common [63]. Small-molecule inhibitors that can effectively disrupt HIV-1 viral particle assembly by inhibition of the interaction of the MA domain with phosphatidylinositol-(4,5)-bisphosphate ($PI(4,5)P_2$) have been reported [64,65]. The discovery of new protein-binding inhibitors often begins with *in silico* screening, followed by biophysical characterization. However, to perform the initial screening, it is essential for a three-dimensional atomic structure to be available. Recently, a HIV-1 Gag model was constructed using homology modeling and Förster resonance energy transfer (FRET) data; the model was then used in molecular dynamic (MD) simulations to gain insights into viral maturation and drug resistance [66]. For deltaretroviruses such as BLV and HTLV-1, there have been no published data regarding the detailed structure of the full-length Gag protein.

In this report, we construct the first molecular model of full-length BLV Gag using currently available structures of the BLV MA and CA domains [67,68], a homology model of BLV NC, and flexible linkers consisting of the authentic amino acid sequences that connect the domains (Fig. 1 and Table 1). The model was compared to a molecular envelope derived from

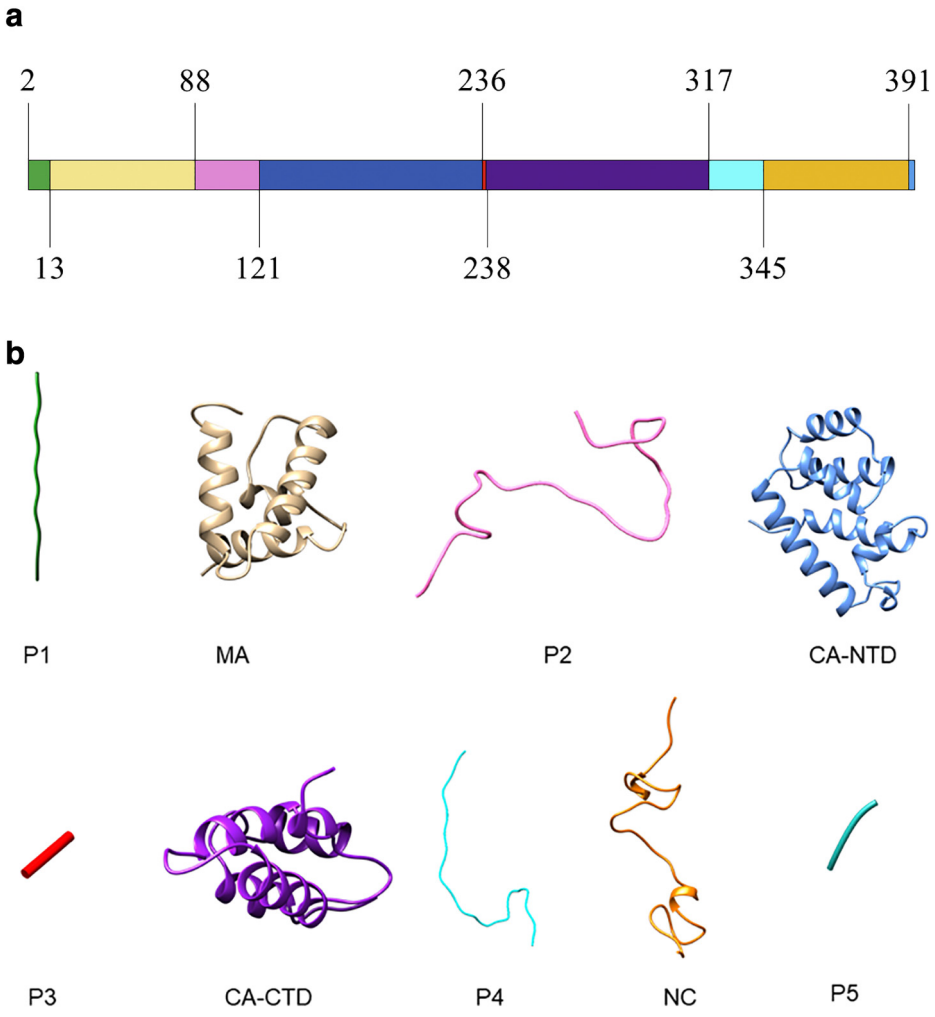

Fig. 1. Domain structure of BLV Gag (a) and segments used to construct the full-length Gag model (b). The origin of each segment is listed in Table 1.

Table 1. Source and composition of the segments used to build the atomic model of BLV Gag

Segment	Residues	Source
P1	2–12	Built with Avogadro
MA	13–87	NMR solution structure of BLV MA [62]
P2	88–120	Built with Avogadro, refined with MODELER
CA_{NTD}	121–235	PDB ID 4PH3 [63]
P3	236–237	Built with Avogadro
CA_{CTD}	238–316	PDB ID 4PH1 [63]
P4	317–344	Built with Avogadro, refined with MODELER
NC	345–390	Built with SWISS-MODEL homology modeling from RSV NC, template PDB ID 2IHX [69]
P5	391–393	Built with Avogadro

SAXS experiments, and MD was used to perform flexible fitting of the model to the envelope. The results suggest that BLV Gag adopts an extended structure in solution and is more similar to MLV Gag and RSV Gag than HIV-1 Gag.

Results

SAXS analysis

SAXS experiments were performed using full-length BLV Gag, as well as BLV CA-NC, to investigate their solution conformation and provide a basis for comparison to other retroviral Gag proteins [19,38,40]. As a solution-based technique, SAXS is amenable to study of a wide range of protein sizes without the requirement of crystallization. The scattering data can be used to produce low-resolution envelopes that approximate the size and shape of a macromolecule. The low squared momentum transfer data (low s^2 region) of the Guinier plots for both constructs were linear, indicating that the samples were not aggregated or experiencing interparticle interactions or repulsions (Fig. S1). To examine the overall globularity of the proteins, Kratky plots were generated from the SAXS data (Fig. 2). A well-folded, globular protein typically yields a bell-shaped peak at lower *q* values, which converges to 0 on the y axis at higher *q* values [71]. Molecules that are largely unfolded yield a Kratky plot that plateaus and does not decrease or continues to increase at higher *q* values. Finally, molecules that are partially unfolded or extended result in Kratky plots with a broad peak that decreases but does not fully converge to zero on the y-axis at higher *q* values. The latter is the case for both BLV Gag and CA-NC, suggesting these multi-domain proteins are extended in solution. These results are also consistent with analytical size-exclusion chromatography (SEC) of BLV Gag which showed anomalous early elution from the column, consistent with a higher molecular mass than expected [72]. However, it should be noted that the concentration of the BLV Gag in the SEC experiments (~0.25 mg/mL) was considerably lower than the concentration used for SAXS.

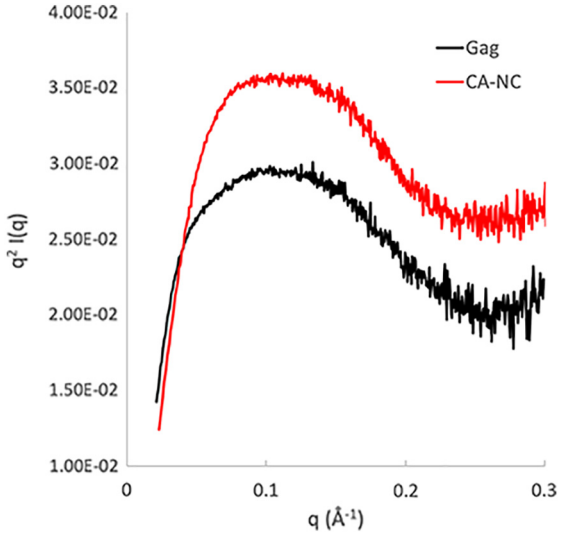

Fig. 2. Kratky plots for BLV Gag and CA-NC. Plots were generated from scattering curves using PRIMUS [70].

The pairwise distance distribution function, *P*(*r*), calculated by Fourier transform of the scattering data, is a histogram that represents the distribution of distances between electrons in a molecule. Examination of the *P*(*r*) curve also provides information on the overall shape of a protein in solution. Globular proteins produce a bell-shaped curve with a single peak that tails off quickly at high *r* values. Elongated proteins are often characterized by asymmetric peaks and longer tail regions at high *r* values. The *P*(*r*) functions calculated for BLV Gag and CA-NC are consistent with elongated structures, in agreement with the Kratky plot (Fig. 3). The resulting radius of gyration (R_g) and

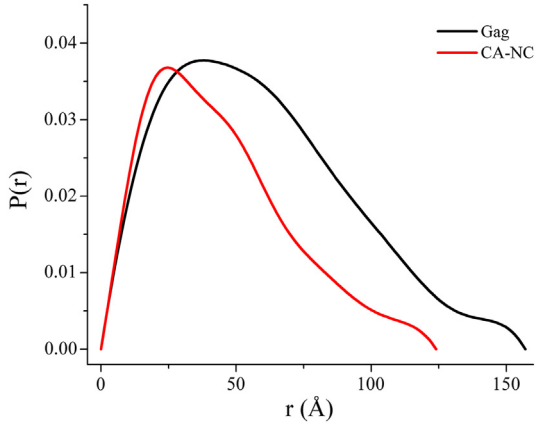

Fig. 3. *P*(*r*) distribution of BLV Gag and CA-NC. The *P*(*r*) data were generated using GNOM analysis software [73].

Table 2. Dimensional parameters for BLV Gag and CA-NC

	SAXS		Calculated	
	R_g (Å)	D_{max} (Å)	R_g (Å)	D_{max} (Å)
Gag	47.6	157	43.2	158
CA-NC	35.4	124	31.3	124

SAXS parameters were calculated using the $P(r)$ distribution from experimental data as described in the Materials and Methods section. Calculated values were derived from the final fitted structures of Gag and CA-NC using the program CRYSOL.

largest intramolecular distance (D_{max}), calculated from the $P(r)$, are shown in Table 2. These parameters were also calculated with the CRYSOL software [74] using the final BLV Gag and CA-NC structures as input and are in good agreement with the experimentally determined values (Table 2). Using the $P(r)$ as input, bead models of BLV Gag and CA-NC were calculated as the average of 23 structural models (Figs. S2–S3) generated by DAMMIF. As expected, the average bead model of full-length BLV Gag depicts an elongated structure with a globular-shaped head region, likely corresponding to the MA domain (Fig. 4a). Although the superimposition of CA-NC onto Gag performed using UCSF Chimera does not result in a perfect overlap, they are similar enough to identify clearly the N- and C-termini of Gag (Fig. 4b). The superposition is consistent with the assignment of MA to the globular domain (Fig. 4, top).

Molecular modeling and MD

We next generated an atomic model of BLV Gag from available structures of individual MA and CA domains, homology modeling of BLV NC, and construction of linker segments as described in the Materials and Methods. SEC–multi-angle light scattering results indicated that although there were small amounts of larger species present, the majority of BLV Gag exists as a monomer under the conditions used for SAXS (Fig. S4). Thus, a monomeric form was used for model construction and MD simulations. The model was refined by fitting the coordinates to the SAXS-derived envelope using the MD Flexible Fitting (MDFF) function of the Nanoscale MD (NAMD) computational software. Surface models of Gag and CA-NC docked into the SAXS envelopes are shown in Fig. 5, with each domain color coded on the surface model. A depiction of the docked Gag model (represented as a surface map generated from the atomic coordinates using Chimera) within its SAXS-generated envelope shows a

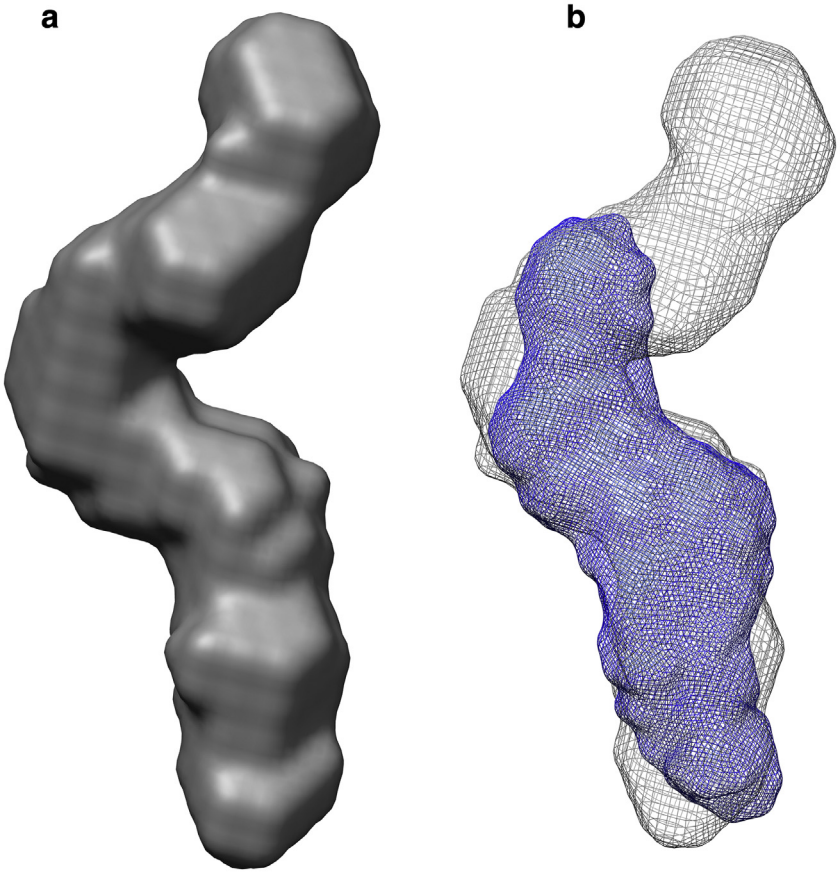

Fig. 4. SAXS-derived molecular envelopes for BLV Gag (a) and CA-NC (b). In panel b, the CA-NC envelope is colored blue and fit to the Gag envelope using the volume fit function of UCSF Chimera [75].

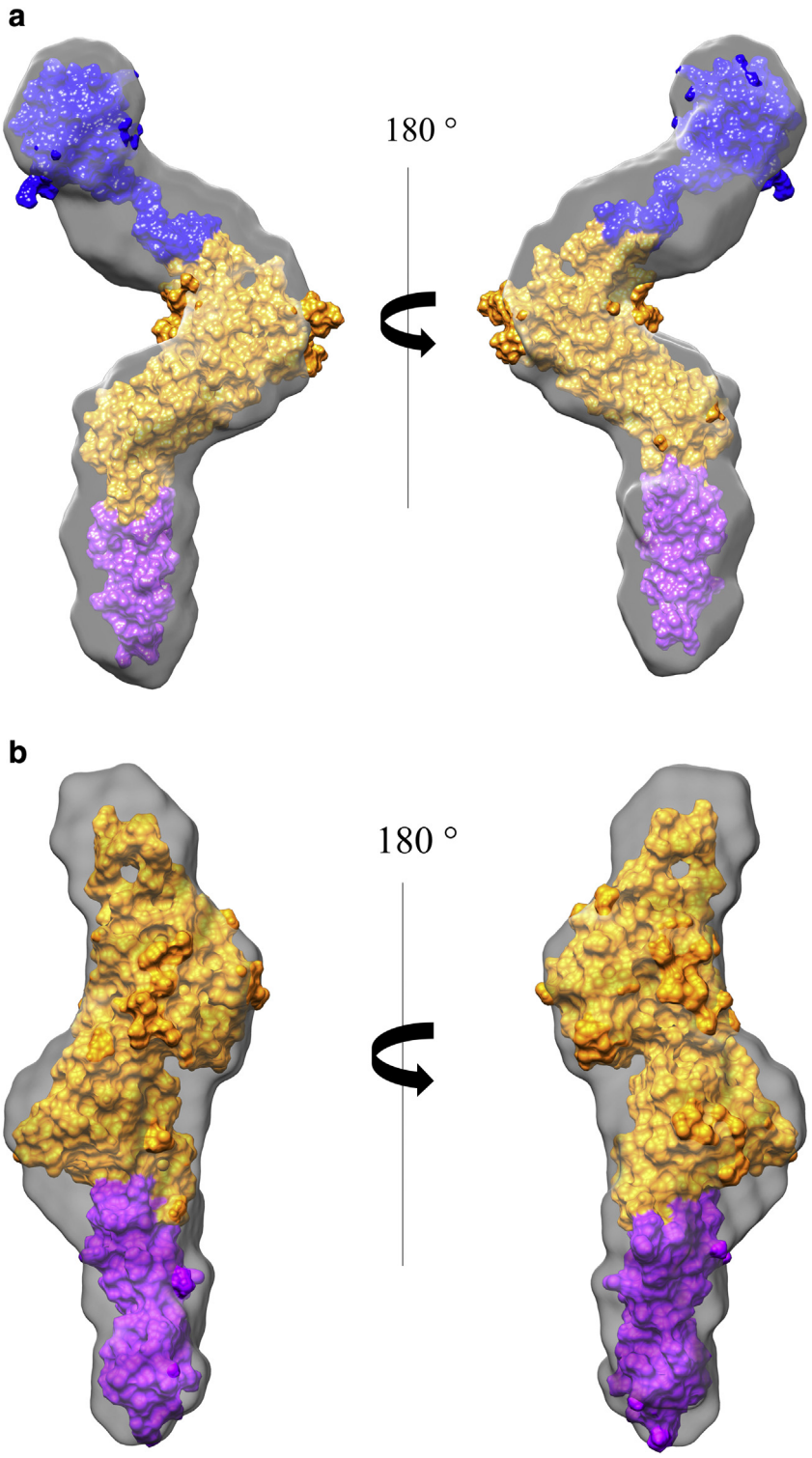

Fig. 5. BLV Gag (a) and CA-NC (b) structures after undergoing flexible fitting to molecular envelopes calculated from SAXS data. Envelopes are transparent gray, and the models are shown as surface representations. The Gag structure is color coded by domain: MA (blue), CA (orange), and NC (purple).

good fit to the experimental data, with a correlation of 0.9029 (Fig. 6a). There are several regions where the model extends beyond the limits of the envelope, but these discrepancies are relatively minor. Likewise, the CA-NC model was docked into its envelope and also is a good fit with a correlation of 0.9331 (Fig. 6b).

The BLV Gag structural model was used in MD simulations to predict the flexibility of the linkers between each domain. The output trajectory was analyzed using the software package TimeScapes. The Agility program uses Gaussian weighting in a sliding window to calculate RMS fluctuations as a

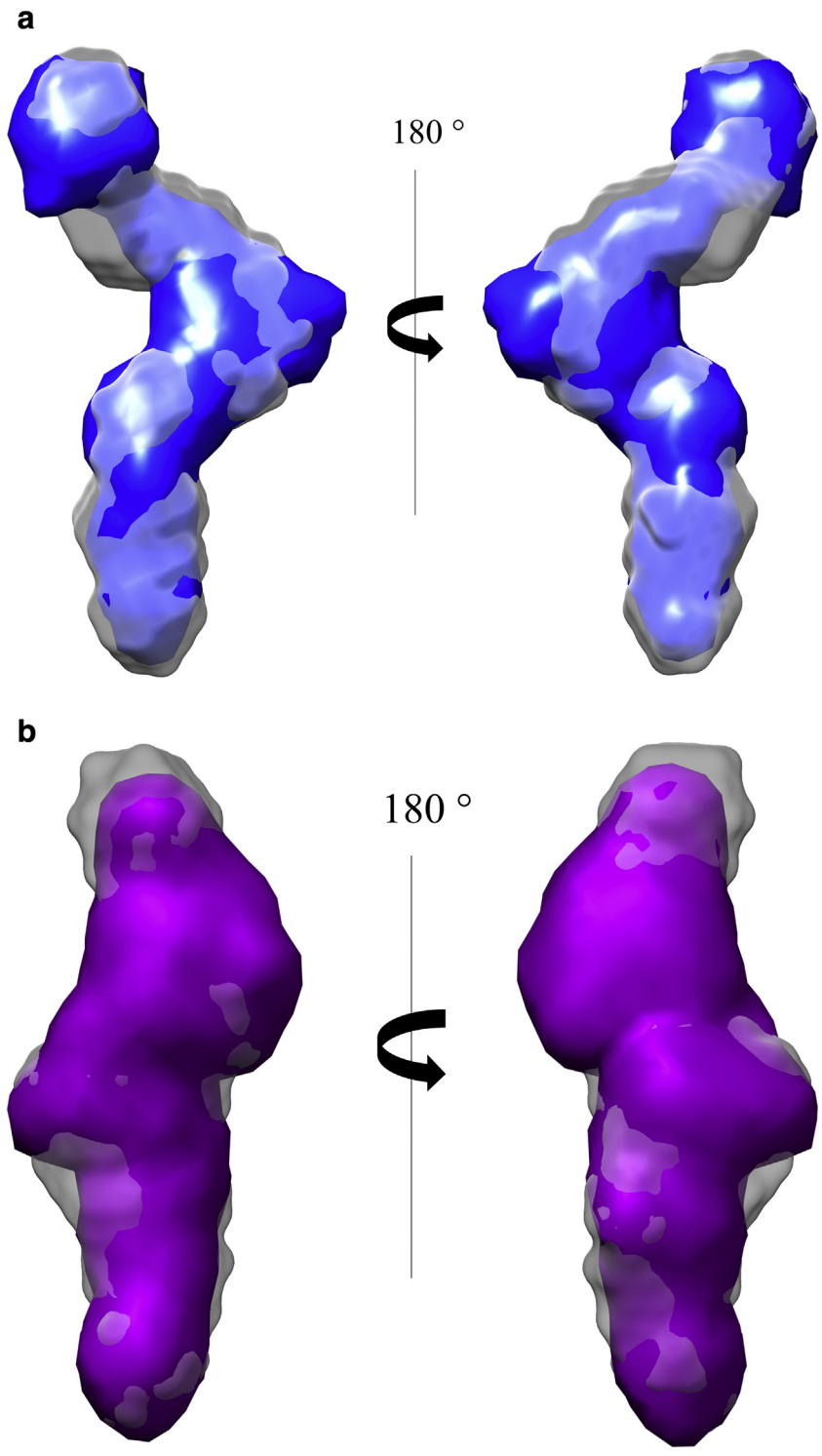

Fig. 6. BLV Gag (a) and CA-NC (b) represented as density maps docked to their respective SAXS envelopes. The density maps (blue for Gag and purple for CA-NC) were generated using the Fit in Map function of Chimera assuming a resolution of 15 Å, which is the resolution of a SAXS envelope generated in SITUS.

global activity function. The output from the Agility computation was used in the Turning program, where the coarse-grained global dynamics were compared to fast, local hinge motions in the dihedral bonds of individual residues. The result is shown as a heat map that depicts flexible areas in red and rigid areas in blue (Fig. 7). As expected, the globular domains are fairly rigid, and the highest degree of flexibility is observed in the unstructured linker between the MA and CA domains.

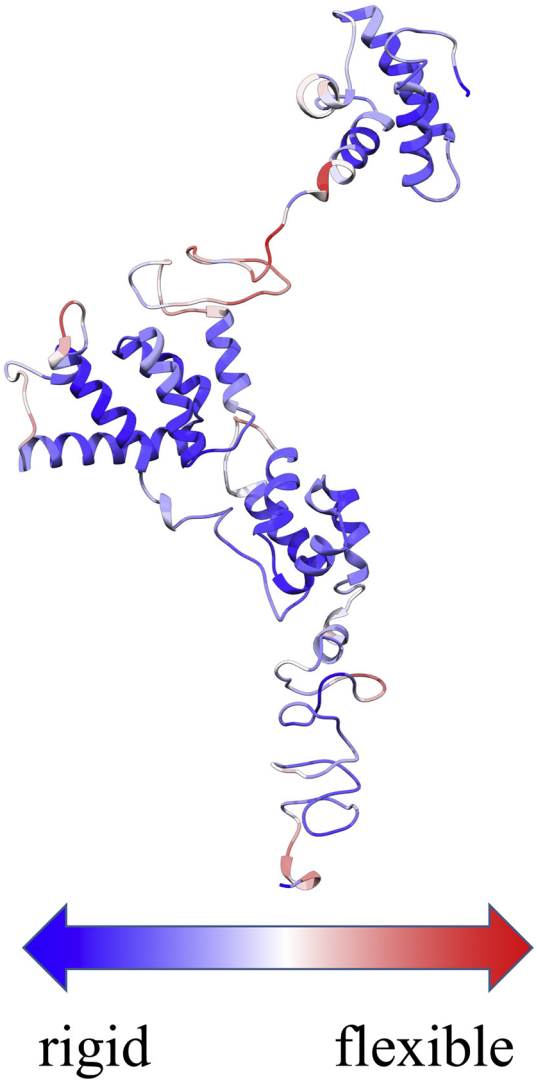

Fig. 7. Flexibility of BLV Gag based on 50 ns of MD simulation. Trajectories were evaluated using the Turning function of the TimeScapes software package [76,77]. More flexible regions are red, and less flexible regions are blue. Flexibility is expressed using RMS fluctuations correlated with slow, global motion as described in the cited literature.

FRET-based annealing assays

Annealing assays using the previously described mini-TAR system with BLV NC were used to measure the chaperone activity of BLV Gag and CA-NC [55]. Interestingly, the annealing kinetics were markedly different (Fig. 8). Similar to BLV NC, the CA-NC annealing data fit to a double-exponential function, whereas full-length BLV Gag-facilitated annealing was much slower and was best described by a single-exponential function (Table S1). In addition, the annealing rate for BLV Gag was lower than both the slow and fast components of CA-NC-facilitated annealing, while Gag facilitated annealing to a greater extent than CA-NC (as evidenced by the greater change in intensity).

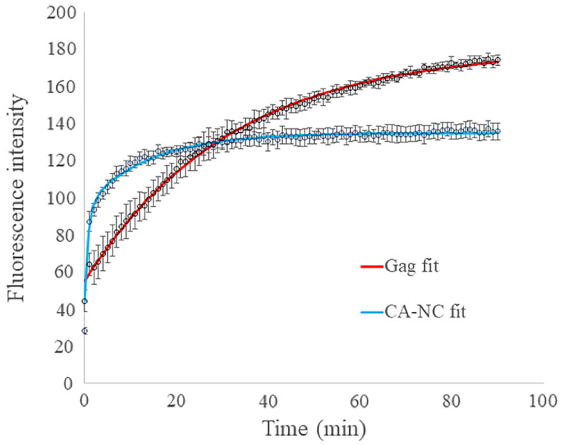

Fig. 8. Time-based annealing of mini-TAR RNA and mini-TAR DNA. Gag data were fit to a single-exponential function, and CA-NC data were fit to a double-exponential function. Error bars represent the standard error of the mean of at least three independent experiments.

Discussion

Using a combination of SAXS and computational methods, we have developed an atomic model of full-length BLV Gag. Analysis of the SAXS data suggests a molecule that is extended in solution. Flexible fitting of an independently generated molecular model to the SAXS envelope resulted in a strong correlation between the theoretical structures and the experimental data. The correlations suggest a better fit for CA-NC than for full-length Gag; this could be explained by the flexibility of the MA–CA linker in Gag, which is one part of the model that does not occupy the amount of space expected compared to the SAXS-generated envelope (Fig. 6a). This aspect is discussed further below. In addition, the SAXS-derived parameters R_g and D_{max} were estimated from the Gag and CA-NC models using the program CRYSOL [74] and were similar to the values derived from the experimental scattering curves (Table 2). CRYSOL is a program that calculates theoretical SAXS curves from all-atom models; thus, the agreement between experimental and theoretical values helps lend validity to the final Gag atomic structure. The SAXS-derived bead model for CA-NC also suggests an extended structure that fits well into the bead model for Gag and allows unambiguous orientation of the domains. MD simulations to determine the regions of relative mobility show that except for the highly flexible linker region between MA and CA, Gag maintains a rigid structure. Interestingly, inspection of the individual bead models (Fig. S2) suggests greater variability in the position of the MA domain relative to the rest of the protein, reinforcing the MD

results. Finally, functional assays using a well-characterized annealing system show a difference in annealing kinetics between full-length BLV Gag and CA-NC, indicating that the presence of the MA domain modulates the nucleic acid interactions and chaperone function of BLV Gag, as previously observed for HIV-1 Gag [22].

We next compared the structure and function of BLV Gag to Gag proteins from other retroviruses. Modeling of HIV-1 Gag from SANS experiments resulted in structures where the relative positions of CA and NC remain relatively constant between models, while the position of the MA domain was variable [19]. Any differences in the region between CA and NC could be attributed to the presence of p2, a short peptide that links these two domains in HIV-1 Gag but is absent in BLV Gag. In addition, several of the best-fitting HIV-1 Gag models show that the MA domain is attached to CA by a long, flexible tether, rather than by a more structured linker, a result that is in agreement with solution NMR experiments [20,21]. On the one hand, BLV Gag is similar to HIV-1 Gag in that the most mobile region is the linker between MA and CA. On the other hand, unlike HIV-1 Gag, which is predicted to exist in a compact conformation in solution, BLV Gag is primarily elongated. In the case of MLV Gag, SAXS experiments have shown that the structure is rigid and rod-shaped [38], while RSV Gag is flexible and extended. Thus, since our data support a relatively rigid and extended conformation for BLV Gag, it appears to be more similar to MLV and HTLV-1 Gag than HIV-1 or RSV Gag.

There are a number of structural elements in MLV Gag that likely have an impact on the global structure, including the presence of p12, the high proline content, and the existence of a stretch of charged residues at the C-terminal end of the CA domain. HIV-1 Gag lacks all of these features; BLV Gag, however, has a slightly higher proline content than MLV Gag (Table 3). Whereas the Gag proteins from BLV and MLV have similar proline content in their MA and CA domains, the NC region of BLV Gag has a significantly higher percentage of prolines (26%) compared to the NC domain of MLV Gag (8%). High proline content is also observed in the C-terminal portion of Gag from other deltaretroviruses (HTLV-1 and -2), as well as in the C-terminal p6 region of HIV-1 Gag (16%). However, HTLV-1 and -2 also contain a large number of acidic residues in the C-terminal region of NC, as does HIV-1 p6. Accordingly, for HTLV-1 and -2 NC, and HIV-1 NCp15 (consisting of NC, p2, and p6), nucleic acid chaperone activity as measured by gel-shift annealing assays was significantly lower than that of HIV-1 NC [56,58,78]. Thus, the large number of proline residues in the NC region of BLV Gag likely contributes to the overall inflexibility of that domain but has little effect, if any, upon its interactions with nucleic acids.

Table 3. Percent composition of proline residues for full-length Gag and Gag domains from BLV, HIV-1, MLV, HTLV-1, HTLV-2, and RSV

	Gag	MA	CA	NC
BLV	13.7	17.6	7.9	26.1
HIV-1	6.6	2.3	7.8	1.8
MLV	11.7	17.6	6.5	8.3
HTLV-1	14.9	19.2	8.9	23.5
HTLV-2	14.5	20.6	8.4	20.5
RSV	8.8	6.5	9.6	9.0

The results of our modeling and functional experiments, taken together with previous data, indicate that the MA and NC domains of BLV Gag are both capable of binding to nucleic acids and that both domains play a role in RNA packaging [55,57,79]. RNA–DNA annealing activity is kinetically slower but goes to a greater level of completion in the case of full-length Gag compared to a Gag that lacks the MA domain. This result was also shown for HIV-1 Gag [22]. Moreover, both HIV-1 and HTLV-2 MA possess nucleic acid chaperone activity as individual domains, with HTLV-2 MA displaying more robust annealing activity [58]; thus, it is reasonable to expect the same for BLV MA, although this has not yet been tested. Given the flexibility in the region linking MA and CA in BLV Gag, it is also likely that BLV Gag is capable of binding to RNA with both domains simultaneously, as previously proposed for HIV-1 Gag [69,80]. Addition of inositol phosphates such as IP6 to HIV-1 Gag has been shown to stimulate annealing activity, presumably by binding to the MA domain and disengaging it from the RNA [22]. Regulation of the chaperone function of Gag in this manner may serve to regulate events in the viral lifecycle including the timing of tRNA primer annealing and reverse transcription. Thus, MA–nucleic acid binding in the context of Gag reduces chaperone function and may explain why the BLV truncation mutant displays faster annealing kinetics than full-length Gag. This effect is not universally observed for all retroviruses; in RSV, full-length Gag is just as efficient at promoting annealing as NC alone [34].

Previous results from analytical SEC have suggested that BLV Gag exists as a monomer at lower concentrations (5 μM), and that it assumes an elongated shape because it elutes slightly earlier than expected based on calibration standards [72]. Thus, the model for BLV Gag was built assuming a monomeric form. It is possible that higher-order complexes may occur within the context of an infected cell, as the crystal structure of mature BLV CA has been shown to be hexameric [68]. However, the R_g and D_{max} calculated from the SAXS data are consistent with a single, monomeric Gag molecule rather than a dimer or larger complex.

In summary, we have developed a semi-empirical, experimentally validated molecular model of full-length BLV Gag. This model will be helpful in future studies aimed at computationally modeling retroviral assembly and budding, and for future studies focusing on the identification of small molecule inhibitors of these processes.

Materials and Methods

Protein expression and purification

Full-length BLV Gag was expressed in Rosetta-2 (DE3) pLysS *Escherichia coli* (EMD Millipore, Billerica, MA) and purified as described previously [72]. In brief, cells were lysed by sonication followed by precipitation of nucleic acids and negatively charged proteins by addition of polyethyleneimine. Cellular debris and precipitate were removed by centrifugation. Saturated ammonium sulfate solution was added (10–20 mL of supernatant) and the resulting precipitate was collected by centrifugation. The precipitate was resuspended in column buffer [50 mM Hepes (pH 7.4), 300 mM NaCl, 5 mM 2-mercaptoethanol, 10% glycerol, 5 mM imidazole] and purified with a Ni^{2+}-agarose column. The concentration of Gag was measured as described [72] followed by the addition of 2.2 M equivalents of $Zn(OAc)_2$. The purified Gag was diluted 1:1 with storage buffer [50 mM Hepes (pH 7.4), 300 mM NaCl, 80% glycerol, 5 mM 2-mercaptoethanol] and stored at −20 °C.

DNA encoding BLV CA-NC (residues 109 to 393 of BLV Gag) was amplified using PCR with the Gag expression vector as a template. The primers incorporated XhoI and BamHI restriction sites at the 5′ and 3′ ends, respectively. After restriction endonuclease digestion and ligation into a pET15a expression vector (EMD Millipore, Billerica, MA) using standard protocols, the clone was confirmed by DNA sequencing (Eurofins Genomics, Louisville, KY). Expression and purification of BLV CA-NC was accomplished using the same procedure reported for BLV Gag [72], except that the ammonium sulfate precipitation step was omitted. The concentration of CA-NC was measured using UV absorbance at 280 nm and an extinction coefficient of 51,920 M^{-1} cm^{-1}.

Prior to SAXS analysis, both proteins were dialyzed into SAXS buffer [50 mM Hepes (pH 7.0), 500 mM NaCl, 5% glycerol, 5 mM 2-mercaptoethanol] using Slide-A-Lyzer Mini dialysis devices (Thermo-Fisher, Waltham, MA). Both Gag and CA-NC were filtered through a 100-kDa MWCO centrifugation device to remove any large complexes or aggregates. Three concentrations of each protein (2.5, 5, and 10 mg/mL) were prepared and stored on ice until analysis.

SAXS analysis

SAXS data were collected on SIBYLS beamline 12.3.1 at the Advanced Light Source at the Lawrence Berkeley National Laboratory. For each sample, a set of 33 frames collected at 0.3-s intervals was recorded, and exposures that showed evidence of radiation damage were not used in further data processing. After buffer subtraction, the data were imported into the PRIMUS program of the ATSAS suite [70]. Scattering profiles free of aggregation and intermolecular repulsion were averaged, and averaged profiles from the 5 and 10 mg/mL samples (for both Gag and CA-NC) were merged. The pair-distribution function $P(r)$ and maximum particle size (D_{max}) were calculated using the GNOM program [73]. The radius of gyration (R_g) was determined from the $P(r)$ distribution. The output from $P(r)$ was used to generate 23 molecular envelopes using the program DAMMIF [81], which were averaged with the program DAMAVER [82].

Gag model construction

The solution NMR structure for the BLV MA domain [67] and crystal structures for the CA domain [68] were used as starting points for the construction of a full-length BLV Gag model. Since the molecular structure of BLV NC has not yet been solved, a homology model was created with the SWISS-MODEL server using the solution structure of RSV NC [Protein Data Bank (PDB) ID: 2IHX] as a template [83,84]. Missing segments were constructed using the program Avogadro and the longer segments P2 and P4 were modeled as random loops using the MODELER plugin of UCSF Chimera [85–87]. Pictures and details regarding the sequences used for each segment of the model are given in Fig. 1 and Table 1, respectively. Each segment was connected sequentially in UCSF Chimera by forming a peptide bond between the N-terminus of one segment and the C-terminus of the preceding segment.

Flexible fitting of Gag model to SAXS-derived envelope

A low-resolution density map of BLV Gag was generated with the PDB2VOL function of the Situs package [88] using the DAMAVER output referenced above. The Gag molecular model underwent rigid-body docking to the density map with the COLORES function of Situs [89], generating several fits. The best fit, determined by the unnormalized correlation coefficients of the output files, was used as a starting structure for flexible fitting.

The MDFF package of the NAMD software was used to perform a flexible fitting of the rigid-docked model to the density map generated by Situs [90,91]. The program Visual MD was used to create protein structure files and PDB structure files containing MDFF potentials, corresponding to the docked Gag model and the Gag density map [92]. Protein structure files are similar to PDB files but contain additional information about the molecular system, including bond connectivity and dihedral angles. The parts of BLV Gag with known structures (MA, CA_{NTD}, and CA_{CTD}) were maintained as rigid domains; the zinc-binding residues and zinc atoms of the NC domain were also restrained to prevent distortion of the zinc finger structures.

Simulations were performed in the presence of explicit solvent. The resulting structure was used for the MD simulations described below. In addition, the structure was used to simulate SAXS curves using the CRYSOL program [74].

MD simulations

MD simulations were performed using the AMBER force field *ff14SB* within the NAMD environment [93,94]. The SAXS-fitted Gag model described above was solvated in a TIP3P water box and neutralized by the addition of Cl^-, followed by a period of energy minimization (50 ps). The system was heated from 10 to 300 K over 50 ps, and was held at 300 K over the length of the simulation (50 ns). Unlike the flexible fitting procedure, the entire protein was allowed to freely move. Trajectories were processed in Visual MD and analyzed using the Agility and Turning functions of the TimeScapes Analytics package, version 1.4.1 [76,77].

Annealing assays

The FRET-based annealing assays were performed essentially as previously described [55]. Mini-TAR DNA was fluorescently labeled with AlexaFluor488 (donor dye) attached by a six-carbon linker to the 5′ end and 4-[4-(dimethylamino)phenylazo]benzoic acid (DABCYL, acceptor) on the 3′ end. The doubly labeled DNA was purchased HPLC-purified from Eurofins USA (Louisville KY). Mini-TAR RNA was from Integrated DNA Technologies (Coralville, IA) and was desalted after synthesis. DNA (0.5 μM) and RNA (13.6 μM) were separately folded in 0.02 M Hepes (pH 7.5) and 0.1 M NaCl by heating at 80 °C for 2 min, 60 °C for 2 min, followed by the addition of $MgCl_2$ to a final concentration of 10 mM. The folding mixtures were then incubated on ice for at least 30 min before use. The annealing reaction mixture containing 20 mM Hepes (pH 7.5), 10 mM NaCl, 10 nM folded DNA, and 0.42 μM BLV Gag or CA-NC protein was allowed to incubate in the fluorimeter at 30 °C for 15 min. After the zero-minute fluorescence intensity was recorded, 0.3 μM folded mini-TAR RNA was added to the annealing mixture to start the reaction. Donor intensity readings were recorded every minute using an excitation wavelength of 495 nm and an emission wavelength of 515 nm. The data were globally fit to a single- or double-exponential function using DynaFit [95].

CRediT authorship contribution statement

Dominic F. Qualley: Conceptualization, Data curation, Formal analysis, Funding acquisition, Writing - original draft, Writing - review & editing. **Sarah E. Cooper:** Data curation, Formal analysis, Writing - original draft, Writing - review & editing. **James L. Ross:** Data curation, Formal analysis, Writing - review & editing, . **Erik D. Olson:** Data curation, Formal analysis, Writing - review & editing. **William A. Cantara:** Data curation, Formal analysis, Writing - review & editing. **Karin Musier-Forsyth:** Funding acquisition, Writing - original draft, Writing - review & editing.

Acknowledgments

We thank Dr. Ioulia Rouzina for helpful discussions and comments on the manuscript. This work was supported by a Faculty Development Grant (FD16-17-19) from Berry College to D.F.Q. and Richards Scholarships to S.E.C. and J.L.R. SAXS data were collected at the Advanced Light Source, a national user facility operated by Lawrence Berkeley National Laboratory on behalf of the Department of Energy, Office of Basic Energy Sciences, through the Integrated Diffraction Analysis Technologies (IDAT) program, supported by DOE Office of Biological and Environmental Research. Additional support comes from the National Institutes of Health project MINOS (R01 GM105404) and R01 GM065056 (to K. M.-F.).

Author Contributions: D.F.Q. designed the study. D.F.Q., S.E.C., J.L.R., E.D.O., and W.A.C. performed experiments and analyzed data. D.F.Q., S.E.C., E.D.O., and K.M.-F. participated in interpretation of the results. D.F.Q., S.E.C., and K.M.-F. wrote the manuscript. All authors contributed to editing of the manuscript.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.jmb.2019.01.036.

Received 10 December 2018;
Received in revised form 28 January 2019;
Accepted 29 January 2019
Available online 4 February 2019

Keywords:
retrovirus;
Gag;
small-angle X-ray scattering;
viral assembly;
molecular dynamics;
matrix

Present address: J.L. Ross, Graduate Division of Cancer Biology, Emory University, Atlanta, GA 30322, USA.

Abbreviations used:
HIV-1, human immunodeficiency virus type 1; RSV, Rous sarcoma virus; MLV, Moloney murine leukemia virus; HTLV-1, human T-cell leukemia virus type 1; MA, matrix; CA, capsid; NC, nucleocapsid; gRNA, genomic RNA; SANS, small-angle neutron scattering; SAXS, small-angle X-ray scattering; BLV, bovine leukemia virus; FRET, Förster resonance energy transfer; MD, molecular dynamics; SEC, size-exclusion chromatography; MDFF, MD Flexible Fitting; NAMD, Nanoscale MD; PDB, Protein Data Bank.

References

[1] J.M. Coffin, S.H. Hughes, H.E. Varmus, Retroviruses, Cold Spring Harbor, NY, 1997.
[2] J.G. Levin, M. Mitra, A. Mascarenhas, K. Musier-Forsyth, Role of HIV-1 nucleocapsid protein in HIV-1 reverse transcription, RNA Biol. 7 (2010) 754–774.
[3] R.A. Dick, V.M. Vogt, Membrane interaction of retroviral Gag proteins, Front. Microbiol. 5 (2014) 187.
[4] K.H. Fogarty, W. Zhang, I.F. Grigsby, J.L. Johnson, Y. Chen, J.D. Mueller, et al., New insights into HTLV-1 particle structure, assembly, and Gag–Gag interactions in living cells, Viruses 3 (2011) 770–793.
[5] E. Hamard-Peron, D. Muriaux, Retroviral matrix and lipids, the intimate interaction, Retrovirology 8 (2011) 15.
[6] J.R. Lingappa, J.C. Reed, M. Tanaka, K. Chutiraka, B.A. Robinson, How HIV-1 Gag assembles in cells: putting together pieces of the puzzle, Virus Res. 193 (2014) 89–107.
[7] J.O. Maldonado, J.L. Martin, J.D. Mueller, W. Zhang, L.M. Mansky, New insights into retroviral Gag–Gag and Gag–membrane interactions, Front. Microbiol. 5 (2014) 302.
[8] W.I. Sundquist, H.G. Krausslich, HIV-1 assembly, budding, and maturation, Cold Spring Harb. Perspect. Med. 2 (2012) a006924.
[9] V. Chukkapalli, A. Ono, Molecular determinants that regulate plasma membrane association of HIV-1 Gag, J. Mol. Biol. 410 (2011) 512–524.
[10] A. Alfadhli, E. Barklis, The roles of lipids and nucleic acids in HIV-1 assembly, Front. Microbiol. 5 (2014) 253.
[11] E.D. Olson, K. Musier-Forsyth, Retroviral Gag protein–RNA interactions: implications for specific genomic RNA packaging and virion assembly, Semin. Cell Dev. Biol. 86 (2018) 129–139.
[12] K.H. Fogarty, S. Berk, I.F. Grigsby, Y. Chen, L.M. Mansky, J. D. Mueller, Interrelationship between cytoplasmic retroviral Gag concentration and Gag–membrane association, J. Mol. Biol. 426 (2014) 1611–1624.
[13] S.B. Kutluay, P.D. Bieniasz, Analysis of the initiating events in HIV-1 particle assembly and genome packaging, PLoS Pathog. 6 (2010), e1001200.
[14] K.H. Fogarty, Y. Chen, I.F. Grigsby, P.J. Macdonald, E.M. Smith, J.L. Johnson, et al., Characterization of cytoplasmic Gag–Gag interactions by dual-color z-scan fluorescence fluctuation spectroscopy, Biophys. J. 100 (2011) 1587–1595.
[15] J. Hendrix, V. Baumgartel, W. Schrimpf, S. Ivanchenko, M.A. Digman, E. Gratton, et al., Live-cell observation of cytosolic HIV-1 assembly onset reveals RNA-interacting Gag oligomers, J. Cell Biol. 210 (2015) 629–646.
[16] S.A. Datta, F. Heinrich, S. Raghunandan, S. Krueger, J.E. Curtis, A. Rein, et al., HIV-1 Gag extension: conformational changes require simultaneous interaction with membrane and nucleic acid, J. Mol. Biol. 406 (2011) 205–214.
[17] J.S. Saad, J. Miller, J. Tai, A. Kim, R.H. Ghanam, M.F. Summers, Structural basis for targeting HIV-1 Gag proteins to the plasma membrane for virus assembly, Proc. Natl. Acad. Sci. U. S. A. 103 (2006) 11364–11369.
[18] J.N. Sfakianos, R.A. LaCasse, E. Hunter, The M-PMV cytoplasmic targeting-retention signal directs nascent Gag polypeptides to a pericentriolar region of the cell, Traffic 4 (2003) 660–670.
[19] S.A. Datta, J.E. Curtis, W. Ratcliff, P.K. Clark, R.M. Crist, J. Lebowitz, et al., Conformation of the HIV-1 Gag protein in solution, J. Mol. Biol. 365 (2007) 812–824.
[20] L. Deshmukh, R. Ghirlando, G.M. Clore, Investigation of the structure and dynamics of the capsid-spacer peptide 1-nucleocapsid fragment of the HIV-1 Gag polyprotein by solution NMR spectroscopy, Angew. Chem. Int. Ed. Engl. 53 (2014) 1025–1028.
[21] L. Deshmukh, R. Ghirlando, G.M. Clore, Conformation and dynamics of the Gag polyprotein of the human immunodeficiency virus 1 studied by NMR spectroscopy, Proc. Natl. Acad. Sci. U. S. A. 112 (2015) 3374–3379.
[22] C.P. Jones, S.A. Datta, A. Rein, I. Rouzina, K. Musier-Forsyth, Matrix domain modulates HIV-1 Gag's nucleic acid chaperone activity via inositol phosphate binding, J. Virol. 85 (2011) 1594–1603.
[23] J.B. Munro, A. Nath, M. Farber, S.A. Datta, A. Rein, E. Rhoades, et al., A conformational transition observed in single HIV-1 Gag molecules during in vitro assembly of virus-like particles, J. Virol. 88 (2014) 3577–3585.
[24] R.A. Dick, K.K. Zadrozny, C. Xu, F.K.M. Schur, T.D. Lyddon, C.L. Ricana, et al., Inositol phosphates are assembly co-factors for HIV-1, Nature 560 (2018) 509–512.
[25] R.A. Dick, D.L. Mallery, V.M. Vogt, L.C. James, IP6 regulation of HIV capsid assembly, stability, and uncoating, Viruses 10 (2018).
[26] Y. Huang, J. Wang, A. Shalom, Z. Li, A. Khorchid, M.A. Wainberg, et al., Primer tRNA3Lys on the viral genome exists in unextended and two-base extended forms within mature human immunodeficiency virus type 1, J. Virol. 71 (1997) 726–728.
[27] Y.X. Feng, S. Campbell, D. Harvin, B. Ehresmann, C. Ehresmann, A. Rein, The human immunodeficiency virus type 1 Gag polyprotein has nucleic acid chaperone activity: possible role in dimerization of genomic RNA and placement of tRNA on the primer binding site, J. Virol. 73 (1999) 4251–4256.
[28] F. Guo, J. Saadatmand, M. Niu, L. Kleiman, Roles of Gag and NCp7 in facilitating tRNA(Lys)(3) annealing to viral RNA in human immunodeficiency virus type 1, J. Virol. 83 (2009) 8099–8107.
[29] A. Roldan, O.U. Warren, R.S. Russell, C. Liang, M.A. Wainberg, A HIV-1 minimal Gag protein is superior to nucleocapsid at in vitro annealing and exhibits multimerization-induced inhibition of reverse transcription, J. Biol. Chem. 280 (2005) 17488–17496.
[30] V. Chukkapalli, S.J. Oh, A. Ono, Opposing mechanisms involving RNA and lipids regulate HIV-1 Gag membrane binding through the highly basic region of the matrix domain, Proc. Natl. Acad. Sci. U. S. A. 107 (2010) 1600–1605.
[31] V. Chukkapalli, J. Inlora, G.C. Todd, A. Ono, Evidence in support of RNA-mediated inhibition of phosphatidylserine-dependent HIV-1 Gag membrane binding in cells, J. Virol. 87 (2013) 7155–7159.
[32] A. Alfadhli, A. Still, E. Barklis, Analysis of human immunodeficiency virus type 1 matrix binding to membranes and nucleic acids, J. Virol. 83 (2009) 12196–12203.

[33] J. Inlora, D.R. Collins, M.E. Trubin, J.Y. Chung, A. Ono, Membrane binding and subcellular localization of retroviral Gag proteins are differentially regulated by MA interactions with phosphatidylinositol-(4,5)-bisphosphate and RNA, MBio 5 (2014), e02202.
[34] T.D. Rye-McCurdy, S. Nadaraia-Hoke, N. Gudleski-O'Regan, J.M. Flanagan, L.J. Parent, K. Musier-Forsyth, Mechanistic differences between nucleic acid chaperone activities of the Gag proteins of Rous sarcoma virus and human immunodeficiency virus type 1 are attributed to the MA domain, J. Virol. 88 (2014) 7852–7861.
[35] J. Inlora, V. Chukkapalli, D. Derse, A. Ono, Gag localization and virus-like particle release mediated by the matrix domain of human T-lymphotropic virus type 1 Gag are less dependent on phosphatidylinositol-(4,5)-bisphosphate than those mediated by the matrix domain of HIV-1 Gag, J. Virol. 85 (2011) 3802–3810.
[36] F. Fernandes, K. Chen, L.S. Ehrlich, J. Jin, M.H. Chen, G.N. Medina, et al., Phosphoinositides direct equine infectious anemia virus Gag trafficking and release, Traffic 12 (2011) 438–451.
[37] P. Bieniasz, A. Telesnitsky, Multiple, switchable protein:RNA interactions regulate human immunodeficiency virus type 1 assembly, Annu. Rev. Virol. 5 (1) (2018) 165–183.
[38] S.A. Datta, X. Zuo, P.K. Clark, S.J. Campbell, Y.X. Wang, A. Rein, Solution properties of murine leukemia virus Gag protein: differences from HIV-1 Gag, J. Virol. 85 (2011) 12733–12741.
[39] S.A. Datta, Z. Zhao, P.K. Clark, S. Tarasov, J.N. Alexandratos, S.J. Campbell, et al., Interactions between HIV-1 Gag molecules in solution: an inositol phosphate-mediated switch, J. Mol. Biol. 365 (2007) 799–811.
[40] R.A. Dick, S.A. Datta, H. Nanda, X. Fang, Y. Wen, M. Barros, et al., Hydrodynamic and membrane binding properties of purified Rous sarcoma virus Gag protein, J. Virol. 89 (2015) 10371–10382.
[41] National Animal Health Monitoring System, Bovine Leukosis Virus (BLV) on U.S. Dairy Operations, 2007, USDA-APHIS-VS-CEAH, Fort Collins, CO, 2008.
[42] S.L. Ott, R. Johnson, S.J. Wells, Association between bovine-leukosis virus seroprevalence and herd-level productivity on US dairy farms, Prev. Vet. Med. 61 (2003) 249–262.
[43] G.C. Buehring, H.M. Shen, H.M. Jensen, K.Y. Choi, D.J. Sun, G. Nuovo, Bovine leukemia virus DNA in human breast tissue, Emerg. Infect. Dis. 20 (2014) 772–782.
[44] G.C. Buehring, H.M. Shen, H.M. Jensen, D.L. Jin, M. Hudes, G. Block, Exposure to bovine leukemia virus is associated with breast cancer: a case–control study, PLoS One 10 (2015), e0134304.
[45] G.C. Buehring, H. Shen, D.A. Schwartz, J.S. Lawson, Bovine leukemia virus linked to breast cancer in Australian women and identified before breast cancer development, PLoS One 12 (2017), e0179367.
[46] K.A. Baltzell, H.M. Shen, S. Krishnamurty, J.D. Sison, G.J. Nuovo, G.C. Buehring, Bovine leukemia virus linked to breast cancer but not coinfection with human papillomavirus: case–control study of women in Texas, Cancer 124 (7) (2017) 1342–1349.
[47] J.S. Lawson, W.K. Glenn, Multiple oncogenic viruses are present in human breast tissues before development of virus associated breast cancer, Infect. Agent Cancer 12 (2017) 55.
[48] J.S. Lawson, B. Salmons, W.K. Glenn, Oncogenic viruses and breast cancer: mouse mammary tumor virus (MMTV), bovine leukemia virus (BLV), human papilloma virus (HPV), and Epstein–Barr virus (EBV), Front. Oncol. 8 (2018) 1.
[49] N.N. Olaya-Galán, A.P. Corredor-Figueroa, T.C. Guzmán-Garzón, K.S. Ríos-Hernandez, S.P. Salas-Cárdenas, M.A. Patarroyo, et al., Bovine leukaemia virus DNA in fresh milk and raw beef for human consumption, Epidemiol. Infect. 145 (2017) 3125–3130.
[50] N.A. Gillet, L. Willems, Whole genome sequencing of 51 breast cancers reveals that tumors are devoid of bovine leukemia virus DNA, Retrovirology 13 (2016) 75.
[51] Y. Aida, H. Murakami, M. Takahashi, S.N. Takeshima, Mechanisms of pathogenesis induced by bovine leukemia virus as a model for human T-cell leukemia virus, Front. Microbiol. 4 (2013) 328.
[52] N. Gillet, A. Florins, M. Boxus, C. Burteau, A. Nigro, F. Vandermeers, et al., Mechanisms of leukemogenesis induced by bovine leukemia virus: prospects for novel anti-retroviral therapies in human, Retrovirology 4 (2007) 18.
[53] G. Gutiérrez, S.M. Rodríguez, A. de Brogniez, N. Gillet, R. Golime, A. Burny, et al., Vaccination against delta-retroviruses: the bovine leukemia virus paradigm, Viruses 6 (2014) 2416–2427.
[54] L. Willems, A. Burny, D. Collete, O. Dangoisse, F. Dequiedt, J.S. Gatot, et al., Genetic determinants of bovine leukemia virus pathogenesis, AIDS Res. Hum. Retroviruses 16 (2000) 1787–1795.
[55] D.F. Qualley, V.L. Sokolove, J.L. Ross, Bovine leukemia virus nucleocapsid protein is an efficient nucleic acid chaperone, Biochem. Biophys. Res. Commun. 458 (2015) 687–692.
[56] D.F. Qualley, K.M. Stewart-Maynard, F. Wang, M. Mitra, R.J. Gorelick, I. Rouzina, et al., C-terminal domain modulates the nucleic acid chaperone activity of human T-cell leukemia virus type 1 nucleocapsid protein via an electrostatic mechanism, J. Biol. Chem. 285 (2010) 295–307.
[57] H. Wang, K.M. Norris, L.M. Mansky, Involvement of the matrix and nucleocapsid domains of the bovine leukemia virus Gag polyprotein precursor in viral RNA packaging, J. Virol. 77 (2003) 9431–9438.
[58] M. Sun, I.F. Grigsby, R.J. Gorelick, L.M. Mansky, K. Musier-Forsyth, Retrovirus-specific differences in matrix and nucleocapsid protein–nucleic acid interactions: implications for genomic RNA packaging, J. Virol. 88 (2014) 1271–1280.
[59] I. Katoh, H. Kyushiki, Y. Sakamoto, Y. Ikawa, Y. Yoshinaka, Bovine leukemia virus matrix-associated protein MA(p15): further processing and formation of a specific complex with the dimer of the 5′-terminal genomic RNA fragment, J. Virol. 65 (1991) 6845–6855.
[60] J.L. Martin, L. Mendonca, R. Marusinec, J. Zuczek, I. Angert, R.J. Blower, et al., Critical role of the HTLV-1 capsid N-terminal domain for Gag–Gag interactions and virus particle assembly, J. Virol. 91 (14) (2018) pii: e00298-17.
[61] J.L. Martin, L.M. Mendonca, I. Angert, J.D. Mueller, W. Zhang, L.M. Mansky, Disparate contributions of human retrovirus capsid subdomains to Gag–Gag oligomerization, virus morphology, and particle biogenesis, J. Virol. 91 (2017).
[62] D.S. Clutter, M.R. Jordan, S. Bertagnolio, R.W. Shafer, HIV-1 drug resistance and resistance testing, Infect. Genet. Evol. 46 (2016) 292–307.
[63] X. Zhang, Anti-retroviral drugs: current state and development in the next decade, Acta Pharm. Sin. B 8 (2018) 131–136.
[64] I. Zentner, L.J. Sierra, A.K. Fraser, L. Maciunas, M.K. Mankowski, A. Vinnik, et al., Identification of a small-molecule inhibitor of HIV-1 assembly that targets the phosphatidylinositol (4,5)-bisphosphate binding site of the HIV-1 matrix protein, ChemMedChem 8 (2013) 426–432.

[65] I. Zentner, L.J. Sierra, L. Maciunas, A. Vinnik, P. Fedichev, M. K. Mankowski, et al., Discovery of a small-molecule antiviral targeting the HIV-1 matrix protein, Bioorg. Med. Chem. Lett. 23 (2013) 1132–1135.
[66] C.T. Su, C.K. Kwoh, C.S. Verma, S.K. Gan, Modeling the full length HIV-1 Gag polyprotein reveals the role of its p6 subunit in viral maturation and the effect of non-cleavage site mutations in protease drug resistance, J. Biomol. Struct. Dyn. (2017) 1–12.
[67] S. Matthews, M. Mikhailov, A. Burny, P. Roy, The solution structure of the bovine leukaemia virus matrix protein and similarity with lentiviral matrix proteins, EMBO J. 15 (1996) 3267–3274.
[68] G. Obal, F. Trajtenberg, F. Carrión, L. Tomé, N. Larrieux, X. Zhang, et al., Conformational plasticity of a native retroviral capsid revealed by x-ray crystallography, Science 349 (2015) 95–98.
[69] A. Rein, S.A. Datta, C.P. Jones, K. Musier-Forsyth, Diverse interactions of retroviral Gag proteins with RNAs, Trends Biochem. Sci. 36 (2011) 373–380.
[70] M.V. Petoukhov, D. Franke, A.V. Shkumatov, G. Tria, A.G. Kikhney, M. Gajda, et al., New developments in the ATSAS program package for small-angle scattering data analysis, J. Appl. Crystallogr. 45 (2012) 342–350.
[71] C.D. Putnam, M. Hammel, G.L. Hura, J.A. Tainer, X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution, Q. Rev. Biophys. 40 (2007) 191–285.
[72] D.F. Qualley, B.L. Boleratz, Expression, purification, and characterization of full-length bovine leukemia virus Gag protein from bacterial culture, Protein Expr. Purif. 93 (2014) 32–37.
[73] D. Svergun, Determination of the regularization parameter in indirect-transform methods using perceptual criteria, J. Appl. Crystallogr. 25 (1992) 495–503.
[74] D. Svergun, C. Barberato, M. Malfois, A. Volkov, P.I. Konarev, A. Petoukhov, et al., CRYSOL—a program to evaluate X-ray solution scattering of biological macromolecules from atomic coordinates, J. Appl. Crystallogr. 28 (1995) 768–773.
[75] C.C. Huang, T.D. Goddard, E.F. Pettersen, G.S. Couch, E.C. Meng, T. Ferrin, UCSF chimera, Abstr. Pap. Am. Chem. Soc. 238 (2009).
[76] J.A. Kovacs, W. Wriggers, Spatial heat maps from fast information matching of fast and slow degrees of freedom: application to molecular dynamics simulations, J. Phys. Chem. B 120 (2016) 8473–8484.
[77] W. Wriggers, K.A. Stafford, Y. Shan, S. Piana, P. Maragakis, K. Lindorff-Larsen, et al., Automated event detection and activity monitoring in long molecular dynamics simulations, J. Chem. Theory Comput. 5 (2009) 2595–2605.
[78] W. Wang, N. Naiyer, M. Mitra, J. Li, M.C. Williams, I. Rouzina, et al., Distinct nucleic acid interaction properties of HIV-1 nucleocapsid protein precursor NCp15 explain reduced viral infectivity, Nucleic Acids Res. 42 (2014) 7145–7159.
[79] D.F. Qualley, C.M. Lackey, J.P. Paterson, Inositol phosphates compete with nucleic acids for binding to bovine leukemia virus matrix protein: implications for deltaretroviral assembly, Proteins 81 (2013) 1377–1385.
[80] J.A. Webb, C.P. Jones, L.J. Parent, I. Rouzina, K. Musier-Forsyth, Distinct binding interactions of HIV-1 Gag to Psi and non-Psi RNAs: implications for viral genomic RNA packaging, RNA 19 (2013) 1078–1088.
[81] D. Franke, D.I. Svergun, DAMMIF, a program for rapid *ab-initio* shape determination in small-angle scattering, J. Appl. Crystallogr. 42 (2009) 342–346.
[82] V.V. Volkov, D.I. Svergun, Uniqueness of ab initio shape determination in small-angle scattering, J. Appl. Crystallogr. 36 (2003) 860–864.
[83] M. Biasini, S. Bienert, A. Waterhouse, K. Arnold, G. Studer, T. Schmidt, et al., SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information, Nucleic Acids Res. 42 (2014) W252-W8.
[84] J. Zhou, R.L. Bean, V.M. Vogt, M. Summers, Solution structure of the Rous sarcoma virus nucleocapsid protein: mu Psi RNA packaging signal complex, J. Mol. Biol. 365 (2007) 453–467.
[85] M.D. Hanwell, D.E. Curtis, D.C. Lonie, T. Vandermeersch, E. Zurek, G.R. Hutchison, Avogadro: an advanced semantic chemical editor, visualization, and analysis platform, J. Cheminform. 4 (2012) 17.
[86] A. Sali, T.L. Blundell, Comparative protein modelling by satisfaction of spatial restraints, J. Mol. Biol. 234 (1993) 779–815.
[87] Z. Yang, K. Lasker, D. Schneidman-Duhovny, B. Webb, C.C. Huang, E.F. Pettersen, et al., UCSF Chimera, MODELLER, and IMP: an integrated modeling system, J. Struct. Biol. 179 (2012) 269–278.
[88] W. Wriggers, P. Chacón, Using Situs for the registration of protein structures with low-resolution bead models from X-ray solution scattering, J. Appl. Crystallogr. 34 (2001) 773–776.
[89] P. Chacón, W. Wriggers, Multi-resolution contour-based fitting of macromolecular structures, J. Mol. Biol. 317 (2002) 375–384.
[90] J.C. Phillips, R. Braun, W. Wang, J. Gumbart, E. Tajkhorshid, E. Villa, et al., Scalable molecular dynamics with NAMD, J. Comput. Chem. 26 (2005) 1781–1802.
[91] L.G. Trabuco, E. Villa, K. Mitra, J. Frank, K. Schulten, Flexible fitting of atomic structures into electron microscopy maps using molecular dynamics, Structure 16 (2008) 673–683.
[92] W. Humphrey, A. Dalke, K. Schulten, VMD: visual molecular dynamics, J. Mol. Graph. 14 (1996) 33–38 (27-8).
[93] D. Pearlman, D. Case, J. Caldwell, W. Ross, T.I. Cheatham, S. DeBolt, et al., AMBER, a package of computer programs for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to simulate the structural and energetic properties of molecules, Comput. Phys. Commun. 91 (1995) 1–41.
[94] J.A. Maier, C. Martinez, K. Kasavajhala, L. Wickstrom, K.E. Hauser, C. Simmerling, ff14SB: improving the accuracy of protein side chain and backbone parameters from ff99SB, J. Chem. Theory Comput. 11 (2015) 3696–3713.
[95] P. Kuzmic, Program DYNAFIT for the analysis of enzyme kinetic data: application to HIV proteinase, Anal. Biochem. 237 (1996) 260–273.