Virology 537 (2019) 53–64

Contents lists available at ScienceDirect

Virology

journal homepage: www.elsevier.com/locate/virology

Mesoscale model of the assembly and cross-linking of HPV virus-like particles

Oleksandr Zavalov[a], Roberto Irizarry[a,*], Matthew Flamm[a], Ryan Marek[b]

[a] US Data Science and Applied Mathematics, GTO, IT, Merck & Co Inc., West Point, PA, 19486, USA
[b] MMD, Merck & Co Inc., West Point, PA, 19486, USA

ARTICLE INFO

Keywords:
Human papillomavirus (HPV)
Virus-like particle (VLP)
VLP reassembly and stabilization
Kinetic Monte Carlo model
Intercapsomeric disulfide bonds
Cross-linking
HPV maturation

ABSTRACT

We present a novel kinetic Monte Carlo model to simulate the real process time-scale of the assembly of Human Papillomavirus (HPV) virus-like particles (VLPs) incorporating the formation of intercapsomeric disulfide bonds. The objective was to develop insights into the underlying mechanisms of HPV VLP assembly and cross-linking during *in vitro* production of the HPV vaccine. The model integrates actual experimental data and detailed information of VLP geometrical structure in microscopic mechanistic steps. The principal novelty of this model is in the concurrent simulation of VLP assembly and cross-linking including a variable for spatial angular arrangement of capsomeres during their assembly that affects the overall rates of VLP assembly and cross-linking. The cross-linking modeled by using the mechanistic probability rules between involved cysteine residues. The model was utilized to better understand the actual process data and check on the hypothesis related to factors affecting the rates of HPV growth and maturation.

1. Introduction

The human papillomavirus (HPV) is a nanosized non-enveloped DNA virus (Garcea and DiMaio, 2007; Buck et al., 2013) with diameter ~50–60 nm (Mach et al., 2006; Zhao et al., 2012). HPV infections, depending on the Papillomaviridae family, may cause genital warts and cervical cancer (Dürst et al., 1983; Bosch and Munoz, 2002; Walboomers et al., 1999; McBride and Münger, 2018; Fields et al., 2007; Burd, 2003). Moreover, HPV might lead to other types of cancer in some group of immunosuppressed individuals (Pastrana et al., 2018; Siegel et al., 2012) and some human papillomavirus-associated head and neck cancers (Gillison, 2004). There are several commercially available HPV vaccines to protect against HPV related maladies (Zhao et al., 2012; Stanley et al., 2006; Schiller and Muller, 2015; Wang and Roden, 2013a). It is well established that the L1 major capsid protein of HPVs can form virus-like particles (VLP) *in vitro*, which are not infectious nanoparticles because they lack the viral DNA inside (Stanley et al., 2006). These VLPs serve as immunogens stimulating both humoral and cellular arms of the human immune system (Mach et al., 2006; Kirnbauer et al., 1992; Noad and Roy, 2003).

An advanced understanding of the mechanisms of VLP assembly process is extremely important for stable and robust *in vitro* vaccine production. The HPV VLP molecular structure has been determined by using X-ray scattering and Cryo-electron microscopy (Baker et al., 1991; Wang and Roden, 2013b). The viral geometry has been studied theoretically in many works (Salunke et al., 1986; Casini et al., 2004; Carter and Saunders, 2007; Modis et al., 2002; Lidmar et al., 2003; Siber, 2006) including a principle of Caspar-Klug (Caspar and Klug, 1962), Lorman and Rochal's approach (Lorman and Rochal, 2008) based on the Landau theory of crystallization, the viral tiling theory (Twarock, 2004, 2006; Angelescu and Linse, 2007; Twarock and Hendrix, 2006) etc. To fulfill the geometric requirements of a T = 7 icosahedron (Zhao et al., 2012; Kirnbauer et al., 1992) with 72 subunits, the capsomere must be capable of different yet stable quasi-equivalent positions relative to neighboring capsomeres. In the ideal structure, 12 capsomeres are in pentavalent positions with five adjacent capsomere neighbors, and 60 capsomeres are in hexavalent positions with six adjacent neighbors (Mach et al., 2006; Kirnbauer et al., 1992; Wang and Roden, 2013b; Salunke et al., 1986; Casini et al., 2004; Carter and Saunders, 2007).

Many theoretical models have been proposed in the literature to simulate and explain the VLP assembly (Angelescu and Linse, 2007; Dykeman et al., 2014; Elrad and Hagan, 2008; Hemberg et al., 2006; Nguyen et al., 2009; Prasad and Schmid, 2012; Whitesides and Boncheva, 2002; Zlotnick and Stray, 2003; Zlotnick and Mukhopadhyay, 2011; Hagan, 2013, 2014; Verdier et al., 2016; Zlotnick, 2005; Moisant et al., 2010; Mateu, 2013; Schwartz et al., 1998; Zandi et al., 2006; Vega et al., 2016; Arkhipov et al., 2006). These models can be roughly categorized into three types according to the

* Corresponding author.
E-mail address: roberto.irizarry2@merck.com (R. Irizarry).

https://doi.org/10.1016/j.virol.2019.08.018
Received 15 May 2019; Received in revised form 16 August 2019; Accepted 19 August 2019
Available online 20 August 2019
0042-6822/

modeling scale used in them (Reddy and Sansom, 2016). The first category encompasses macroscopic modeling approaches (Hagan, 2013), where the assembly is usually modeled as a nucleation step followed by sequential addition of capsomeres to the growing VLP until all capsomeres have added (Dykeman et al., 2014; Zlotnick and Stray, 2003; Zlotnick and Mukhopadhyay, 2011; Moisant et al., 2010; Hagan, 2014; Keef et al., 2006). These models can simulate the real process scales (hours), but they ignore many microscopic details that may affect the kinetics of VLP formation (Wilber et al., 2009). The second category includes very detailed molecular dynamics (MD) simulations (Perilla et al., 2016; Perilla and Schulten, 2017; Hadden et al., 2018). This technique involves molecular-scale details but simulates relatively short time intervals and a limited number of VLPs, usually only one. Notice that all-atom MD simulations have not been applied to study VLP assembly process yet, only to simulations of complete capsids. Another approach here is to use coarse-grained models of capsid assembly (Nguyen et al., 2009; Arkhipov et al., 2006; Rapaport, 2004, 2018; Boyd et al., 2015; May et al., 2012). They allow modeling of microsecond timescales and are mostly used to study the structural dynamics of viral capsids (Arkhipov et al., 2006).

The third category comprises kinetic Monte Carlo (KMC) type models for the assembly and nucleation of nanoparticles (Arkhipov et al., 2006; Gorshkov et al., 2009, 2011; Hagan and Chandler, 2006; Hagan and Elrad, 2010). Such models assume that the particles have a rigid structure. In this work, the size of the scaffold is confined to an icosahedral lattice and limited to 72 capsomeres. Therefore, VLPs of larger size cannot be observed in this study. The process dynamics is described by a set of potential events and their corresponding probabilities (Hemberg et al., 2006). Such models lump many microscopic details in the event-list. Nevertheless, they offer a good compromise between capturing enough microscopic details of structural dynamics while simulating large time scales needed to simulate an *in vitro* process. The model presented in this study belongs to this category.

In addition to the assembly under certain conditions, the HPV capsid undergoes significant changes characterized by morphological transformations and called maturation (Buck et al., 2005; Cardone et al., 2014; Stray et al., 2004). One of the important characteristics of HPV VLP maturation is the formation of disulfide bond network (cross-linking) across its supramolecular structure. For HPV type 16, the majority of cross-linking occurs between Cys 428 and Cys 175, while additional bonding can also occur between Cys 428 and Cys186 (Duda, 1998; Hare and Chan., 1968; Ishizu et al., 2001; Mateu, 2013; Li et al., 1998; Sapp et al., 1998; Twarock, 2006; Wolf et al., 2010; Zhang et al., 2013). A part of the L1 protein has a structure of a flexible hinge (amino acid bridges or “invading arms”). These invading arms propagate within a gap between capsomeres and bring the Cys 428 (located on a-helix h4) into close proximity with Cys 175 of another L1 (Conway et al., 2011; Ishii et al., 2003; Mukherjee et al., 2008; Modis et al. 2002; Sapp and Bienkowska-Haba, 2009). This leads to the formation of disulfide bonds between surrounding L1 proteins.

Under controlled conditions, purified recombinant L1 proteins can be chemically disassembled and reassembled by varying the solvent environment. In this study, by the VLP assembly we understand the actual reassembly process from previously disassembled L1 proteins. Conditions above physiological ionic strength favor VLP reassembly while disulfide maturation favors an oxidizing environment. In the laboratory, step-wise solvent conditions were developed to maximize the separation of the reassembly and disulfide maturation, allowing the former to occur first, prior to solvent exchange into an oxidizing environment. Samples were collected and assayed to monitor both reassembly and disulfide maturation throughout. Reassembly was monitored using size-exclusion chromatography to resolve reassembly intermediates from free L1 to fully formed VLPs. Disulfide maturation was monitored by size-exclusion chromatography after subjecting samples to denaturing and non-reduced conditions. In this way, one can monitor the loss of L1 monomer and the formation of L1 dimers and trimers (Zhao et al., 2012).

To complement and expand upon the laboratory studies, we introduce the developed microscale kinetic Monte Carlo model to enable to concurrent simulation of the particle reassembly and cross-linking maturation process in detail. The design and structure of the model affords one the ability to scan a wider kinetic landscape of these processes, also test and check on hypotheses related to reassembly and particle maturation. The model simulates the process relevant time scale with the experimentally reasonable set of initial parameters.

In Computational model section, we describe the range of initial parameters and modeling details. Also, we provide some additional discussion on the assumptions that were applied to the model. We focus on some key parameters that are responsible for the kinetic properties of the system. Result and discussion section presents the essential results of our modeling and a general discussion on the comparison of the model with the available experimental results.

We would like to reiterate that this is the first study in which the relationship between the kinetic characteristics of the capsid reassembly and the development of its subsequent maturation properties is clearly determined.

2. Computational model

In this work, we present the kinetic Monte Carlo model that describes the coupled dynamics of capsomere reassembly and subsequent cross-linking between two cysteine residues (Cys 175 and Cys 428) on neighboring capsomeres of an HPV VLP. Different from previous KMC models this model incorporates enough details of the microstructure to capture the formation of cross-linked L1–dimers, and L1–trimers, which are relevant measured attributes in the study of VLP reassembly systems. Another novel aspect of the model is in the possibility to explore the direct relationships between VLP growth factors and the subsequent L1 cross-linking evolution. For this purpose, we hypothesized that a relative spatial position of the involved cysteines determines the cross-linking kinetics.

The KMC model is driven by a set of discrete events in the homogeneous system that is considered by Markov type models and can be effectively modeled using the standard Gillespie stochastic simulation algorithm (Gillespie, 1976, 1977; Serebrinsky, 2011). In our work, we implemented the algorithm described by Gibson & Bruck (Gibson and Bruck, 2000). The main inputs to such models at every time step are: 1) a list of events and corresponding event rates and 2) a method to update the system after each event occurs, see Appendix A for details.

Although the model details will be given in the following subsections, the main steps are summarized here:

1. The simulations start with the box of free capsomeres suspended at a concentration specified by the initial conditions.
2. Two free capsomeres can collide and form a dimer with a certain probability (*nucleation step*). Once a dimer is formed a VLP scaffold is assigned to it. This scaffold is a virtual VLP that determines the possible sites where another free capsomeres can be attached.
3. Free capsomeres can attach to VLP scaffolds at allowed sites (neighbors of occupied sites). This step is called *a growth step* and continues until the scaffold is completely filled with 72 capsomeres (intact VLP).
4. Furthermore, the capsomere docking at the scaffold during nucleation and growth was introduced in the model with the capsomere spatial rotation around its center of mass. This rotation and relative position mean to capture energetically favorable and unfavorable capsomeres orientations for attachment and cross-linking.
5. Capsomeres that are neighbors within a VLP can be cross-linked with a certain probability that is a function of the distance between cysteines and the relative capsomere orientation.

Given that the probabilities of the nucleation, growth and cross-

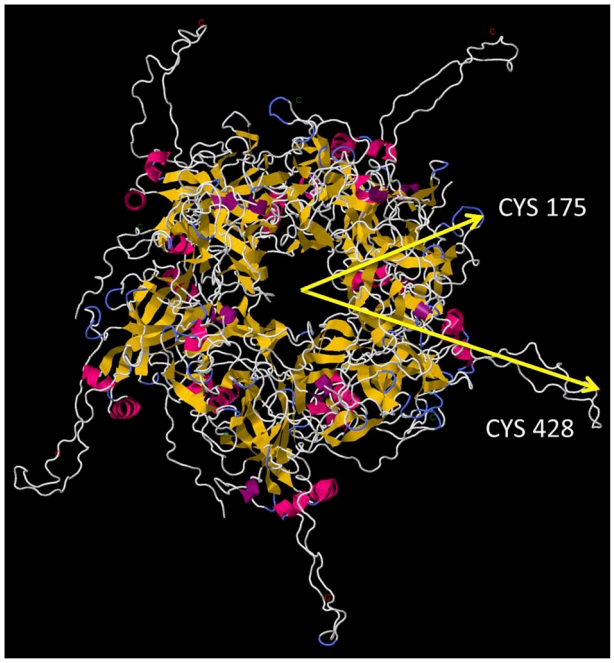

Fig. 1. The illustration of a single HPV capsomere (top view). The image of protein structure of PDB ID: 3J6R (Cardone, G.; Moyer, A. L.; at al. Maturation of the human papillomavirus 16 capsid. MBio, 2014, 5, e01104–01114.) was created with the Jmol, version 14, an open-source interactive viewer for three-dimensional chemical structures. Yellow vectors connect the center of mass of the capsomere with the Cys 175 and Cys 428. These two residues belong to the same L1 protein. Cys 175 is positioned closer to the center in the part that is less movable. Cys 428 is located on a flexible part of L1 called an "invading arm". Therefore, the position of Cys 428 relative to the center of mass of the capsomere varies. That allows it to reach the neighboring L1 proteins of adjacent capsomeres.

linking events depend on the scaffold structure (that approximates a real VLP), the VLP geometry is defined first, and then the probabilities of events are specified for this idealized configuration. With this approach, the modeling objectives are to include more microscopic details than others KMC models, especially around the effect on VLP growth rate and cross-linking. The approximated geometry allows for efficient simulation of a large number of subunits over long process time scales, while incorporating the important microscopic details.

2.1. Geometrical approximation of the forming VLP

The HPV VLP geometrical scaffold used in the model approximates the real physical geometry of a VLP, which has been determined through the HPV electron cryomicroscopy. The geometrical structure is rather unusual. The idealized geometry approximates the structure PDB ID: 3J6R with the resolution 9.1 Å (Cardone et al., 2014). Fig. 1 represents only single pentameric capsomere. The center of mass of each of the 72 capsomeres relative to the center of mass of the VLP was derived from PDB ID: 3J6R and included in a form of Cartesian coordinates in Appendix B.

Since the center of mass of the VLP is located at the origin, these coordinates can also be considered as vectors between the two centers of mass: VLP and capsomere. Each vector will be referred to as the capsomere center of mass vector, $\boldsymbol{R}_{caps}$.

The scaffold geometry also includes the positions of two cysteine residues, Cys 175 and Cys 428, identified from the PDB ID: 3J6R crystal structure. For each of these cysteine residues, vectors were drawn from the center of the VLP to the cysteine location, denoted here as cysteine vectors, and the magnitude of these vectors were measured as, $\boldsymbol{R}_{175}$ or $\boldsymbol{R}_{428}$, respectively. The angles between the cysteine vectors and the corresponding capsomere center of mass vector, $\boldsymbol{R}_{caps}$, were also defined, ϕ_{175} or ϕ_{428}. Therefore, for each L1, each cysteine vector was projected onto a normal plane to the corresponding vector, $\boldsymbol{R}_{caps}$. The angle ϕ_{caps}, between the projected vectors was determined and averaged across all 360 L1 proteins. In this way, five averaged numbers were determined for the location of the two cysteines of interest on an idealized L1 protein. This idealized structure was also equally spaced five times around the capsomere center of mass vector for each capsomere on the VLP (Fig. 2).

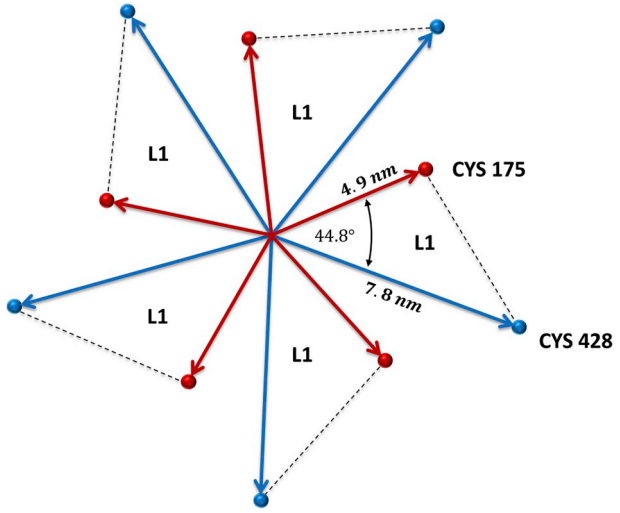

Fig. 2. Schematic illustration of a rigid capsomere structure that has been used in the model. The same geometrical structure was introduced for every capsomere in the system. The average distances from the center of mass of capsomere to Cys 175 and Cys 428 were 4.9 nm and 7.8 nm respectively. These distances have been derived from the HPV crystal structure, PDB ID: 3J6R (Cardone et al., 2014). The average angle between from the center of mass vectors to Cys 175 and Cys 428 of every L1 (also obtained from the structure PDB ID: 3J6R), is 44. 8°.

The average distances from the center of mass of the capsomere to Cys 175 and Cys 428 were derived as 4.9 nm and 7.8 nm respectively (Fig. 2). The average angle between vectors to Cys 175 and Cys 428 of every L1 is 44. 8°. In reality, the capsomere is not a rigid structure. In particular, Cys 428 is located on a flexible invading arm (Buck et al., 2005; Cardone et al., 2014).

2.2. Cross-linking rates probability (effect of capsomere orientation)

With the model's rigid placement of the capsomere centers of mass, we allow the capsomeres to have any arbitrary rotation around the capsomere center of mass vector, $\boldsymbol{R}_{caps}$. The cysteine positions on a given capsomere, while rigidly locked with respect to each other, are rotated around this vector at the certain angle, θ, see Fig. 3. The rotation represents a possible capsomere arrangement in the VLP relative to its neighbors and allows the model to access the effect of different conformational and interaction states on cross-linking.

We hypothesize that the overall probability of cross-linking is related to the distance between cysteine residues of two L1 proteins on opposing capsomeres on the VLP surface.

The distance function, $\sim e^{-\beta r}$, has been used to introduce the dependence on distance. Where $r = |\boldsymbol{R}_{428} - \boldsymbol{R}_{175}|$ is a distance, in *nm*, between a pair of Cys 428 and Cys 175; $\beta \sim 1.4$ is a cysteine interaction factor. The choice of β is intended to describe the flexibility of the invading arm and possible conformational states within this rigid averaged capsomere formulation. The overall rate of cross-linking between any pair of Cys 428 and Cys 175 is given by:

$$rate_{XLink} = k_{XLink} \cdot Exp(-\beta \cdot r) \quad (1)$$

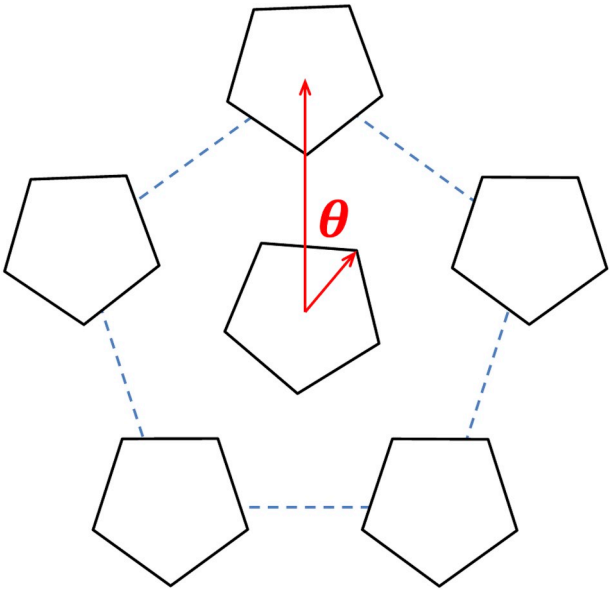

Fig. 3. The scheme represents a possible spatial arrangement of a pentavalent capsomere on the VLP spherical surface relative to its neighbors. The capsomeres are shown as pentagons, since every capsomere has a pentagonal structure comprised of five L1 protein subunits. Capsomere azimuthal angle is denoted by θ. This angle is an important parameter in the model representing a rotational capsomere arrangement.

where k_{XLink} is the cross-linking rate parameter that controls the overall cross-linking rate.

The vast majority of Cys 428 – Cys 175 pairs on the VLP surface are too far apart to cross-link. A screening radius of $D_{screen} \sim 5\,nm$ was employed and any pair exceeding this distance were removed from the list of possible events. This eliminates cross-linking within the same L1, and between different L1s of the same capsomere.

Schematic illustration in Fig. 4 demonstrates a few possible examples of configurations of cross-linked L1–dimers (panels A, B, C) and one L1–trimer (panel D). These examples along with the more complicated ones (L1–multimers) are realized and were observed in this study.

2.3. Effective two-state capsomere orientation in further model simplification

There are potentially infinite spatial combinations of relative angles between two neighboring capsomeres that will lead to infinite combinations of relative angles between cysteines. To further simplify the model, an exploration of the cross-linking rates with respect to relative orientation was performed.

For this procedure, all 72 capsomeres, one by one, were taken with the fixed rotational angle, θ. At the same time, the rotational angles of neighboring capsomeres were randomly selected for each run. This process was repeated 10000 times for rotational angles within the interval: $0 \leq \theta < 360^{\circ}$, with the step 1°, for each capsomere.

The overall probability of cross-linking for every single capsomere has been calculated using Eq. (1) and taking into account all possible interactions between neighboring L1 proteins (Fig. 5, panel A).

This simulation revealed a periodical behavior of the probability function, p, in respect to the rotational angle, θ. There are five periodical angular intervals of $p(\theta)$ with a higher probability of cross-linking, "*H–state*," and five intervals of low cross-linking probability, "*L–state*".

In addition to the case, where a capsomere has all neighbors, other situations with fewer neighboring capsomeres have been also considered (Fig. 5, panel B). In fact, the cross-linking probability demonstrates the same periodical shape with fewer neighbors and even in a presence of only one neighbor. The amplitude of this function, however, strongly depends on the number of neighbors. This periodicity confirms our idea on the existence of more and less favorable rotational capsomere positions.

We hypothesize that the capsomere attachment rate should be a similar periodical function of the rotational angle, $p(\theta)$. Therefore, the model was further simplified by approximating $p(\theta)$ for each capsomere as a step function with two only states: "*H*" and "*L*", as it will be discussed in the next subsections.

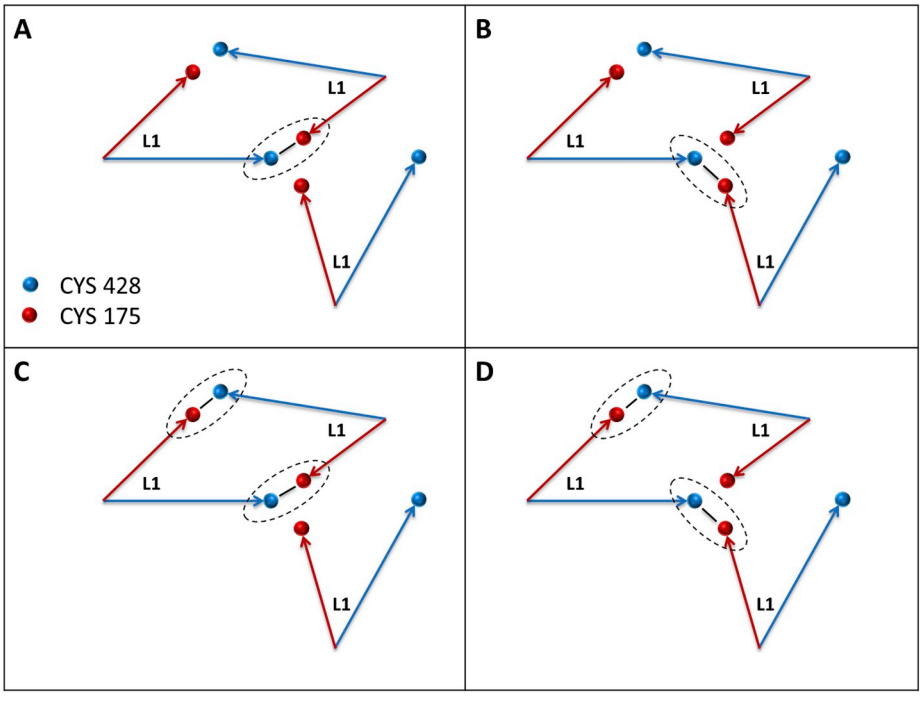

Fig. 4. The schematic illustration of the cross-linking mechanism introduced in the model. It shows three L1 proteins that belong to three different neighboring capsomeres. Cys 428 located on the "*invading arm*" of one capsomere (blue dot) reaches the cysteine Cys 175 of neighboring the L1 (red dot). The panels demonstrate possible opportunities to form the cross-linking (surrounded by black dashed line). There are four panels that represent different types of cross-linked configurations: L1–dimer (**A**, **B**, **C**) and L1–trimer (**D**). Vectors are used to show the distance from the capsomere center of mass vector, $\boldsymbol{R}_{caps}$, to the cysteine residues. Blue vectors are pointed to Cys 428, red vectors – Cys 175.

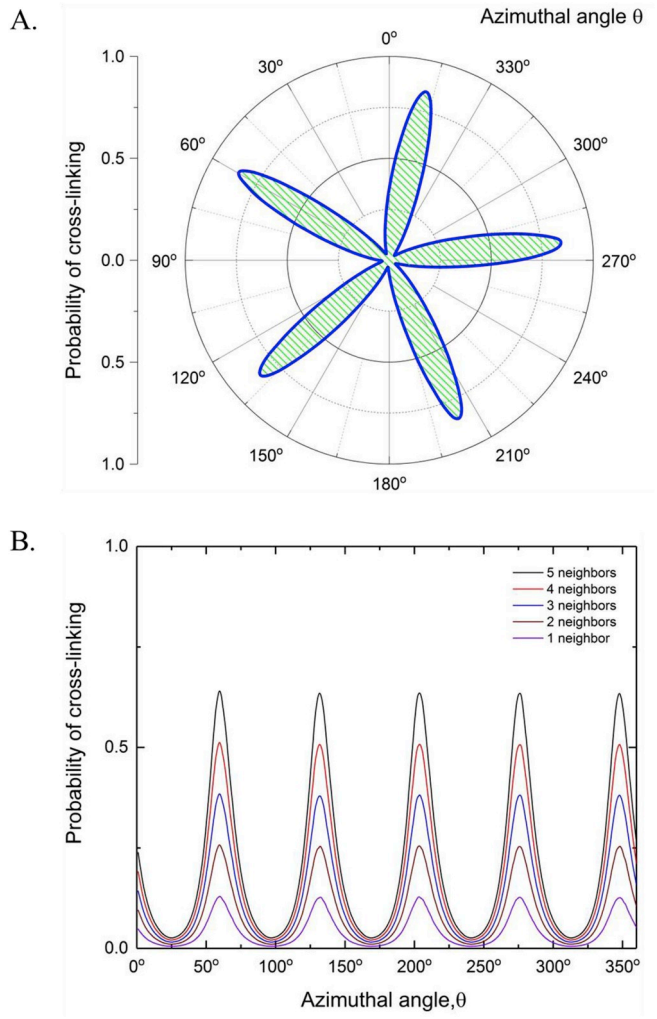

Fig. 5. Panel (A) demonstrates the overall probability of a capsomere, p, as a function of azimuthal angle, θ (see Fig. 3), to be cross-linked calculated as an average value over 10000 simulations while the neighboring capsomeres where randomly rotated. The areas of higher cross-linking probability are highlighted with the green pattern. Panel (B) represents the same probability, $p(\theta)$, but in a case, where a capsomere has a different number of neighbors, from 1 to 5.

2.4. Capsomere dimerization

A capsomere dimer is an initial nucleus or an initial basis of new VLP scaffold. In order to form a dimer, two free capsomeres should be involved in the collision. The capsomere collision frequency estimated with the collision theory is much higher than the observed rate of capsomere dimerization. Thus, the probability that two capsomeres will actually stick together creating a capsomere dimer is low. Actually, two capsomeres must overcome an energy barrier within a certain time interval to be able to aggregate and reach the final bound state. This interval is much shorter than the time of capsomere–capsomere interaction (Vega et al., 2016). Therefore, the attachment of disordered capsomeres into an ordered capsid structure reduces their translational and rotational entropy.

Once the initial nucleus of the VLP scaffold is formed, it grows by subsequent adding other free capsomeres suspended in the volume thus forming higher-order intermediates. At this step, a free capsomere is placed on a VLP scaffold surface filings up a vacant site. This process continues until the VLP is fully assembled.

We consider the following dimerization reaction:

$$Capsomere + Capsomere \xrightarrow{k_{dim}} Capsomere\ Dimer \quad (2)$$

where k_{dim} is the capsomere nucleation (dimerization) rate-constant. This constant regulates the forming of dimers and uptake of free capsomeres as initial "building blocks".

The variation of physical and chemical initial characteristics (e.g. temperature, pH etc.) may significantly change the speed of dimerization. Given all these uncertainties, the value of k_{dim} is determined by fitting of experimental data. The rate of capsomere–capsomere dimerization plays a crucial role in VLP growth dynamics, varying the system conditions from a very low assembled scenario to an undesirable case of "kinetic trapping" (see the Results Section) (Hagan et al., 2011).

In order to estimate the rate of the dimerization event, the following expression has been used:

$$rate_{dim} = k_{dim}[N_{caps}(t)]^2 \quad (3)$$

Here, $N_{caps}(t)$ is the number of free capsomeres available in the system at the time t. It is squared due to the fact that two capsomeres are involved in the dimerization process. Two capsomeres creating a dimer are considered to be initially in the "*H–state*".

2.5. VLP growth

The VLP growth process is modeled as a sequence of free capsomere attachment events, where a single free capsomere occupies a vacant site on the VLP scaffold:

$$Capsomere + Capsomere\ \mathbf{N}mer \xrightarrow{k_{att}} Capsomere\ (\mathbf{N}+1)mer \quad (4)$$

where k_{att} is the VLP growth rate-constant (VLP-growth rate). Similar to the nucleation step, the parameter k_{att} is determined from experimental data.

The vacant site is only available for attaching if it is not occupied and has at least one neighboring capsomere. Once the number of attached capsomere reaches 72, the growth of this particular HPV VLP is completed, and intact VLP is formed.

In our theoretical framework, we assume that the binding forces strongly depend on the particular relative angle between protein subunits. Therefore, the angular arrangement of neighboring capsomeres has a significant effect on the attachment probability.

The attachment rate is given by the following equation:

$$rate_{att} = k_{att} \cdot N_{caps}(t) \cdot [v_n(t) \cdot f(N_H)] \ldots n = 1, \ldots, 6 \quad (5)$$

where v_n is the number of vacant sites with n neighbors ($n \leq 6$), and $f(N_H)$– the function of capsomere–capsomere angular interaction considered in a form:

$$f(N_H) = e^{-\alpha(6-N_H)} \quad (6)$$

where $\alpha = 0.7$ an interaction factor (defined empirically), and N_H– the number of neighboring capsomeres in "H" state.

During attachment, every capsomere on the VLP scaffold is assigned with the rotational angle, θ. This angle is chosen using the following principle:

First, the state of the capsomere is chosen: "H" or "L". The probability of the capsomere to be in an "H" state depends on the number of neighboring capsomeres:

$$P_{H-L} = \gamma \frac{N_H}{N_{Caps} + 1} \quad (7)$$

where $\gamma = 1$ is a coupling parameter that controls the ratio of "*H-L*" capsomeres, N_{caps} is the number of neighboring capsomeres. The parameter γ is important in the VLP growth dynamics since it controls the interplay between the reassembly and cross-linking by determination what fraction of capsomeres in the VLP scaffold will be in "H" or "L" states. If $\gamma = 0$, all attached capsomeres are in "L" state, while intermediate values of γ modulate the coupling between reassembly and cross-linking.

Secondly, the rotational angle θ of capsomere is randomly

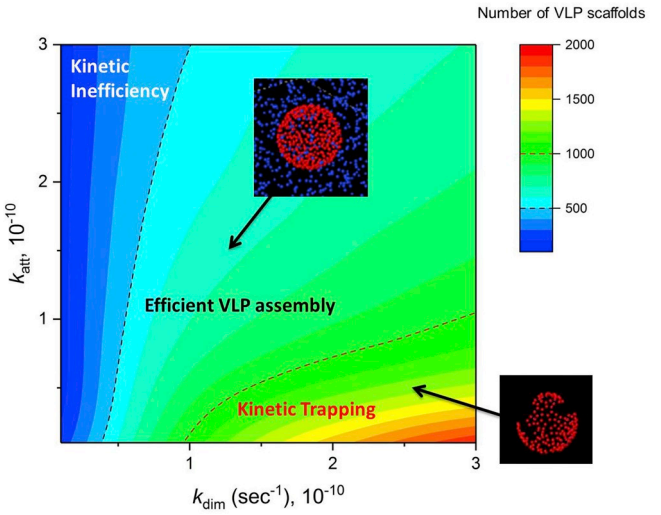

Fig. 6. Color-coded contour plot demonstrates the population of VLP scaffolds (complete and incomplete VLPs) at the end of process time, $t = 50$ hrs. For the selected range of the kinetic rate-constants k_{dim} and k_{att}. The red dashed line marks the "kinetic trapping" area of the excessive VLP development where many of initiated VLP scaffolds are incomplete (see the right insert). The black dashed line shows the region of "kinetic inefficiency" where the number of complete VLPs is less than it should be with the given capsomere concentration. In between there is a regime of "efficient VLP assembly" with mostly complete VLPs (the insert in the center).

calculated within defined "*H*" or "*L*" intervals using the previously described step-function, $p(\theta)$.

The HPV VLP contains 360 of Cys 175 and 360 of Cys 428. The coordinates of these residues are based on the rotational angles, θ, of each capsomere. Once attached every L1 protein of the capsomere is considered to be in the non-cross-linked state and its residues are involved in the "cross-linking" dynamics.

3. Results and discussion

The efficient control of viral capsid reassembly process is important for any virus-like particle system. Moreover, the detailed kinetic analysis of experimental data helps to understand the parameter constraints on self-assembly into closed supramolecular structures (Michaels et al., 2017). To identify the appropriate kinetic parameter space for subsequent simulations, the model has been run varying k_{dim} and k_{att} in the interval: $0.1 - 3.0 \times 10^{-10}$. The total number of VLP scaffolds (intact and incomplete VLPs) at the process time, $t = 50$ Hrs., was stored as a matrix 300×300 and represented in the form of a color-coded contour plot, see Fig. 6. There are three major parameter regions: first – fast dimerization region or "*kinetic trapping*," (Hagan et al., 2011; Michaels et al., 2017) where the number of initiated VLP scaffolds is more than the number of required free capsomeres to complete them; second – fast VLP growth region or "*kinetic inefficiency*" results in slow dimerization and low final number of intact VLPs; third – the efficient assembly region, where the balance is kept between elongation and nucleation processes. Therefore, the efficient VLP reassembly requires a balance between elongation and nucleation processes.

To study the relationship between VLP reassembly and cross-linking, we first established a standard reference model condition to which all other model simulations are compared (Fig. 7). The model was fit to real experimental data to define the overall kinetic parameter optima across each of the VLP growth and cross-linking steps. To avoid confoundedness of fit between the kinetic parameters of reassembly with those of cross-linking, laboratory data were generated under solvent conditions which favor the separation of the reassembly and cross-linking. These kinetic parameters served to establish a benchmark

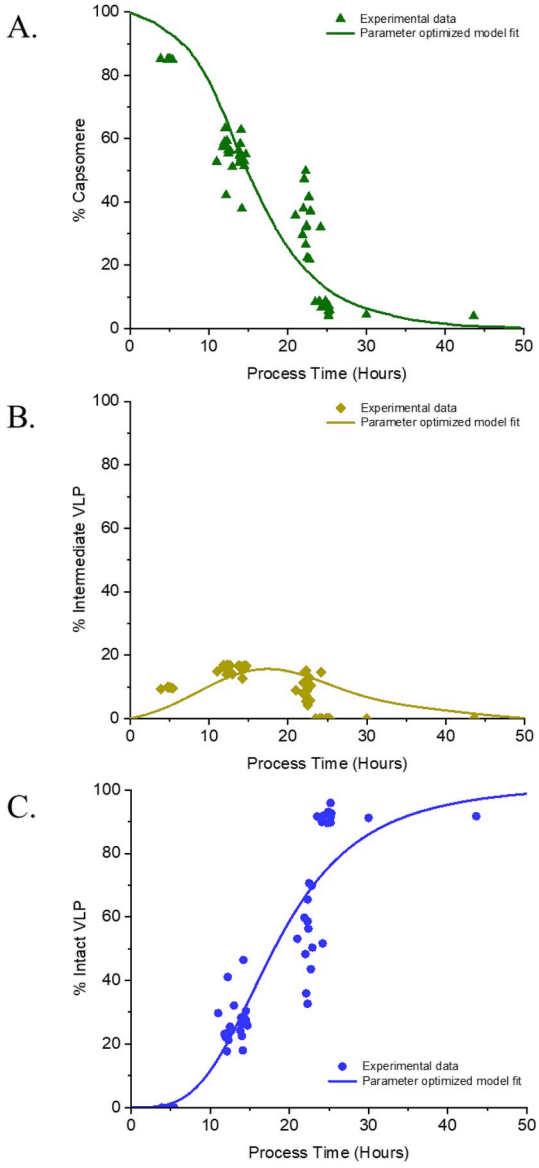

Fig. 7. Fitted reassembly profiles vs. process time: A. Capsomeres, B. Intermediate VLPs, C. Intact VLPs, where lines correspond to modeling, points represent experimental result. All three cases calculated for the same set of fitted kinetic parameters, see Table 1.

reference and relative starting point for subsequent simulations where selected kinetic parameters were varied. The benchmark process timescale continues from 0 up to 50 h and includes two major phases: the VLP reassembly process (0–25 h) and the cross-linking process (25–50 h). The fitted kinetic parameter values used to fit the experimental benchmark case at the first phase are: $k_{dim} = 0.94 \times 10^{-10}$ sec^{-1}, $k_{att} = 0.55 \times 10^{-10}$, $k_x = 0.35$, see also Table 1. At the second, crosslinking phase, the parameter k_x, was multiplied by the factor of 1.92.

The reassembly process in the benchmark case is characterized by the complete VLPs and significantly low concentration of free capsomeres at the end. Starting at 100% in the beginning, the concentration of free capsomeres abates over time due to participating in VLP building (Fig. 7, Panel A). In consequence, the concentration of intermediate VLPs grew first, while many capsomere dimers nucleated, reaching the maximum at the middle of the reassembly time (Fig. 7, Panel B). Then the concentration of intermediates drops down since many of them will be completed. By the end of the reassembly phase,

Table 1
Kinetic parameter values used in the simulations to produce each figure. Here $k_{dim} = 0.94 \times 10^{-10}$ sec^{-1}, $k_{att} = 0.55 \times 10^{-10}$, $k_x = 0.35$ are the kinetic parameter values used to fit the experimental benchmark case.

	Dimerization rates	VLP growth rates	Cross-linking rates	
			$t = 0 - 25$ **hrs**.	$t = 25 - 50$ **hrs**.
Fig. 7	k_{dim}	k_{att}	k_x	$1.92 \times k_x$
Fig. 8	k_{dim}	k_{att}	k_x	$1.92 \times k_x$
Fig. 9 (red curves)	$2 \times k_{dim}$	k_{att}	k_x	$1.92 \times k_x$
Fig. 10 (red curves)	k_{dim}	$2 \times k_{att}$	k_x	$1.92 \times k_x$
Fig. 11 (red curve)	$3 \times k_{dim}$	$0.5 \times k_{att}$	k_x	$1.92 \times k_x$
Fig. 11 (green curve)	$0.5 \times k_{dim}$	$3 \times k_{att}$	k_x	$1.92 \times k_x$
Fig. 12	k_{dim}	k_{att}	k_x	$1.92 \times k_x$

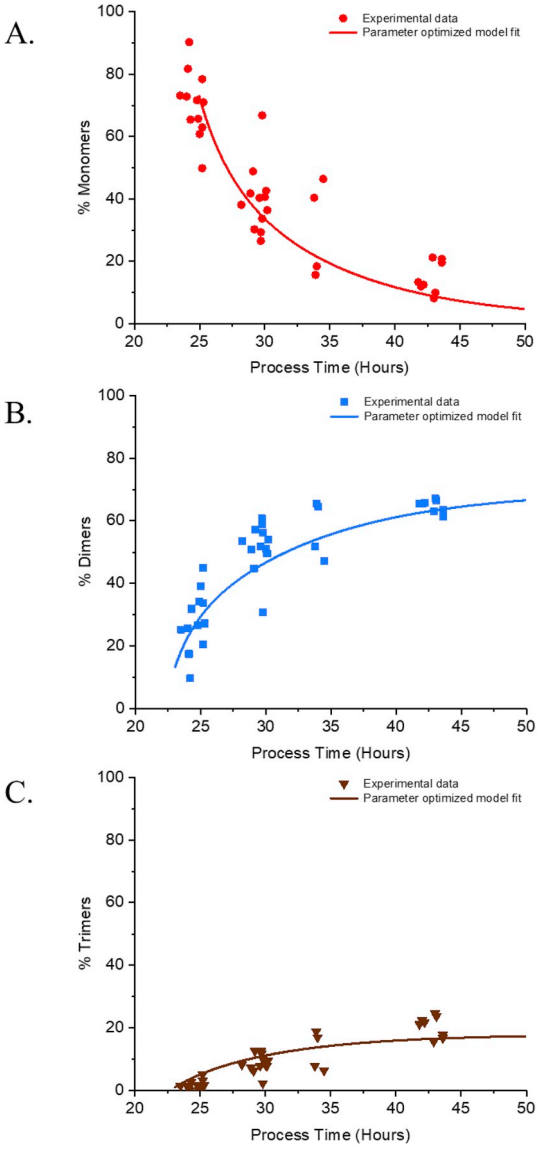

Fig. 8. Fitted cross-linking profiles vs. process time: A. Monomers, B. Dimers, C. Trimers, where lines correspond to modeling, points represent experimental result. All three cases calculated for the same set of fitted kinetic parameters (Table 1).

the concentration of intact VLPs reaches almost 100% (Fig. 7, Panel C). This is a clear indicator that the process has almost finished at 25 h, so it turns into phase two – the cross-linking development.

Separated from the reassembly as described above, the cross-linking process starts with a high number of non-cross-linked L1 monomers. The fitted cross-linking profiles are represented in Fig. 8 (lines – model prediction, points – experimental results), where the following kinetic values of k_{dim}, k_{att}, and $k_x \times 1.92$ were used (see Table 1). The L1 monomers incorporated into VLPs are actively interacting with their neighboring L1s to form disulfide bonds. Consequently, the concentration of non-cross-linked monomers decreases by the end of the process, see Fig. 8, Panel A. The number of L1 dimers in the experiment reaches 60–70% (Fig. 8, Panel B). This number stops growing due to geometrical constraints in the VLP that prevents cross-linking at long distances. In addition to dimers, more complicated structures can be formed such as L1 trimers, see Fig. 4, Panel D. Although, the concentration of trimers is three times lower than dimers, about 20% (Fig. 8, Panel C).

This model explains potential underlying mechanism on the experimental observation that changes in reassembly kinetics only impacts subsequent changes in cross-linking rates, even when the cross-linking process conditions are the same. We performed a test to answer two following important questions: how a change in VLP nucleation only and in VLP growth only would affect the apparent cross-linking rates (despite no change in the actual cross-linking kinetic parameters).

In the first case, see Fig. 9, we present two profiles: reassembly (Panel A) and cross-linking (Panel B). For this simulation, the VLP nucleation rate, k_{dim}, was multiplied by the factor of 2, while the other rates remain the same, see Table 1. The reassembly profile shifts faster, up to 6 h, while the cross-linking profile has an only 2-h shift. The blue curves highlight the fitted case along with experimental data (blue dots), while the shifted case is presented only by the modeling results (red curves). The slope of the curves in both simulations has been changed as well: the increased dimerization caused faster VLP growth and faster cross-linking. In another simulation, Fig. 10, we introduce a case of growth rate shift, where k_{att} was multiplied by the factor of 2, but other rates kept constant (Table 1). Both, reassembly (Panel A) and cross-linking (Panel B) profiles were obtained. Here, the reassembly shift happens faster ~3.3 h over ~1.4 h in cross-linking. However, those shifts are about 2 times less compared to dimerization shift case in Fig. 9.

Summarizing these results, notice that there is a strong correlation between the different condition in reassembly kinetics and the resulting disulfide bonds formation. To identify the limits of this correlation at the extremes of the unbalanced kinetic parameter space, two different tests have been performed, Fig. 11, with the kinetic parameter values represented in Table 1. In the first test, corresponding to “kinetic trapping”, the nucleation rate was much higher than in the fitted case: $3.0 \times k_{dim}$, but the growth rate remained low: $0.5 \times k_{att}$ (see Fig. 11, red curves). In our second test, parameters that belong to the low efficiency region (see Fig. 11, green curves) have been used, $0.5 \times k_{dim}$, and $3 \times k_{att}$. The major cross-linking profiles were introduced in Fig. 11: monomers (Panel A), dimers (Panel B), and trimers (Panel C). One can notice that the cross-linking rate is higher in a “kinetic trapping” (red curves, Fig. 11) case than in fitted (blue curves, Fig. 11). However, we know that the concentration of intact VLPs is lower because many

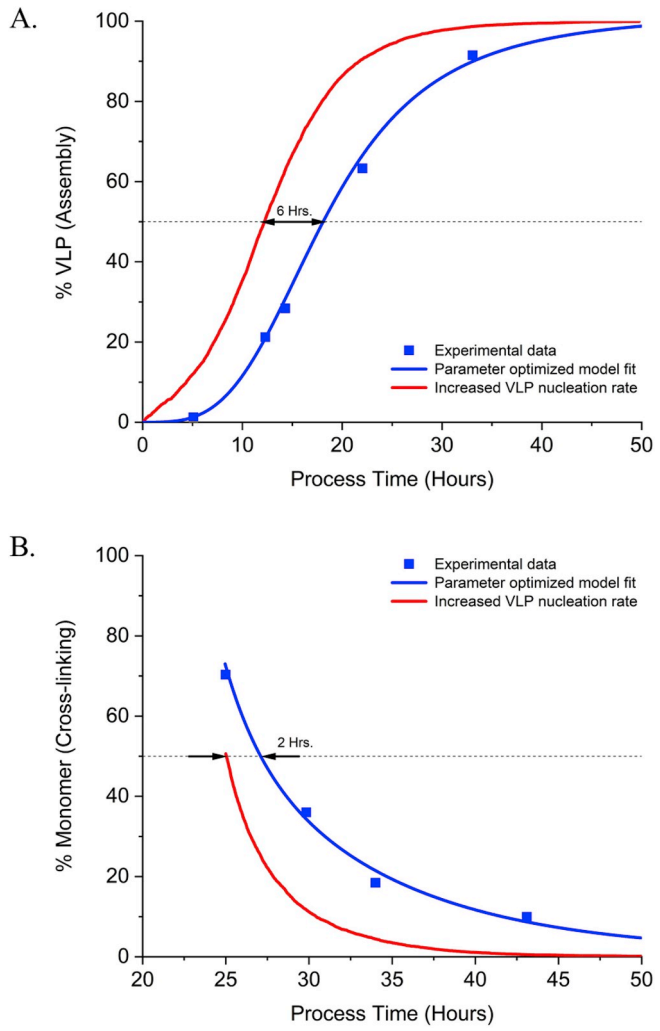

Fig. 9. Impact of increased VLP nucleation rate on reassembly and cross-linking. Reassembly (A. Intact VLPs) and cross-linking (B. Monomers) profiles represent two cases: one is a data fit (blue color lines – modeling and points – experiment), and the second (red lines) corresponds to dimerization shift, where the dimerization rate was significantly increased by the factor of 2, while the growth rate remains the same.

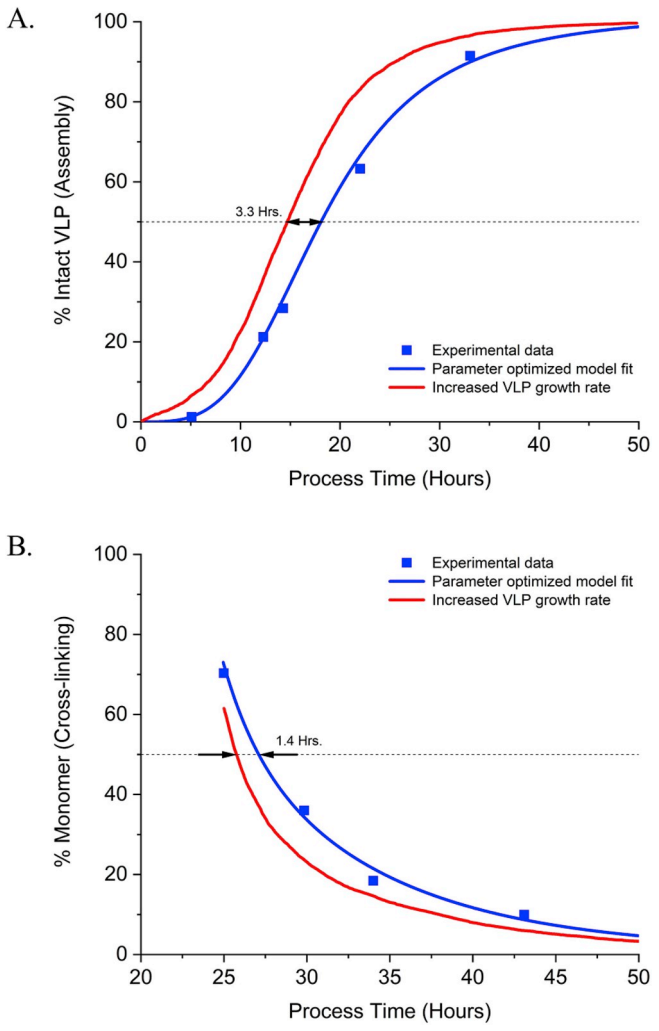

Fig. 10. Impact of increased VLP growth rate on reassembly and cross-linking. Reassembly (A. Intact VLPs) and cross-linking (B. Monomers) profiles for two cases: data fit (blue color lines – modeling and points – experiment), and a case of growth rate shift (red lines), where the growth rate was increased by the factor of 2, while the dimerization rate remains the same.

intermediates are not completed. In contrast to this case, the low-efficiency case demonstrates not only a low concentration of assembled VLPs but also a low level of cross-linking (green curves, Fig. 11). Therefore, this parameter space is characterized by inefficiency in both VLP growth and cross-linking.

The behavior of trimers in Panel C of Fig. 11 is opposite to that in Panel A and B. This result is not immediately intuitive. In case of high dimerization rate, this happens due to VLP geometrical imperfection including many defects that prevents forming complicated structures like trimers, and other multimers, resulting in lower trimer concentration. However, in high growth regime, all intermediate VLPs are fully assembled by the time of cross-linking phase which allows forming a higher number of trimers.

Another important hypothesis was that the overall rates of VLP reassembly and cross-linking are highly depending on an angular spatial arrangement of capsomeres during attachment. To illustrate that we performed two extreme regime tests (Fig. 12). In the first regime, we forced all capsomeres to be only in "*H–state,*" (green curves) during the simulation. In the second – only in "*L–state*" (red curves). Fig. 12 illustrates these results along with the fitted case (blue curves). All three case were performed with the same set of initial parameters: k_{dim}, k_{att}, k_x, see also Table 1. The only changes were related to the capsomere state, which was achieved by varying the value of the parameter γ in equation (7).

The significant displacement of the curves in both extreme cases, as well as their slopes, indicate a high dependence of the reassembly and cross-linking rates on the geometric factors. In our case, it is related to an azimuthal capsomere angle, θ. The displacement obtained in this test practically gives a result which is even higher than in the case of dimerization shift. This suggests that factors which could potentially impact capsomere-capsomere interactions (ionic strength, pH, temperature, etc.), would in turn impact reassembly and cross-linking profiles. Therefore, the proposed theoretical model explains a relationship between the reassembly kinetics and cross-linking. The resulting disulfide bonds formation highly depends on the efficient HPV VLP reassembly process.

4. Conclusions

The developed kinetic Monte Carlo model can serve as a complementary tool to empirical laboratory experimentation to further optimize well established reassembly systems and cross-examine results against expectations as a function of process time. The model includes two major components: VLP reassembly dynamics and the formation of intercapsomeric disulfide bonds (L1 protein cross-linking mechanism),

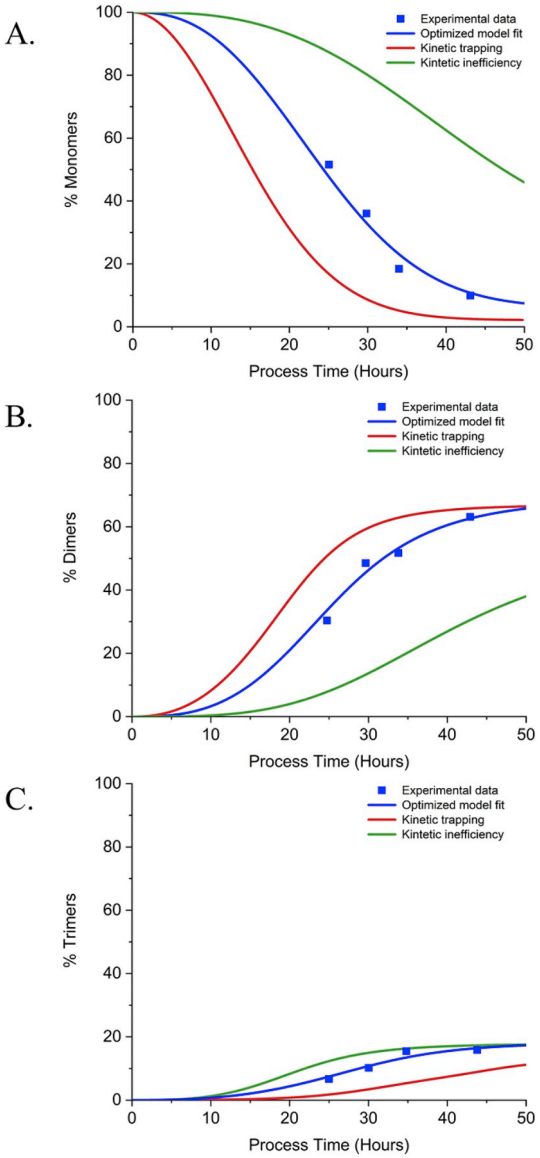

Fig. 11. Demonstration of how inefficient VLP reassembly leads to inefficient cross-linking represented by three profiles: A. Monomers; B. Dimers; C. Trimers. Three cases are shown: fitted (blue curve and points related to experimental data), high VLP nucleation rate (red) corresponds, and high VLP growth rate (green), see Table 1 for the kinetic parameters.

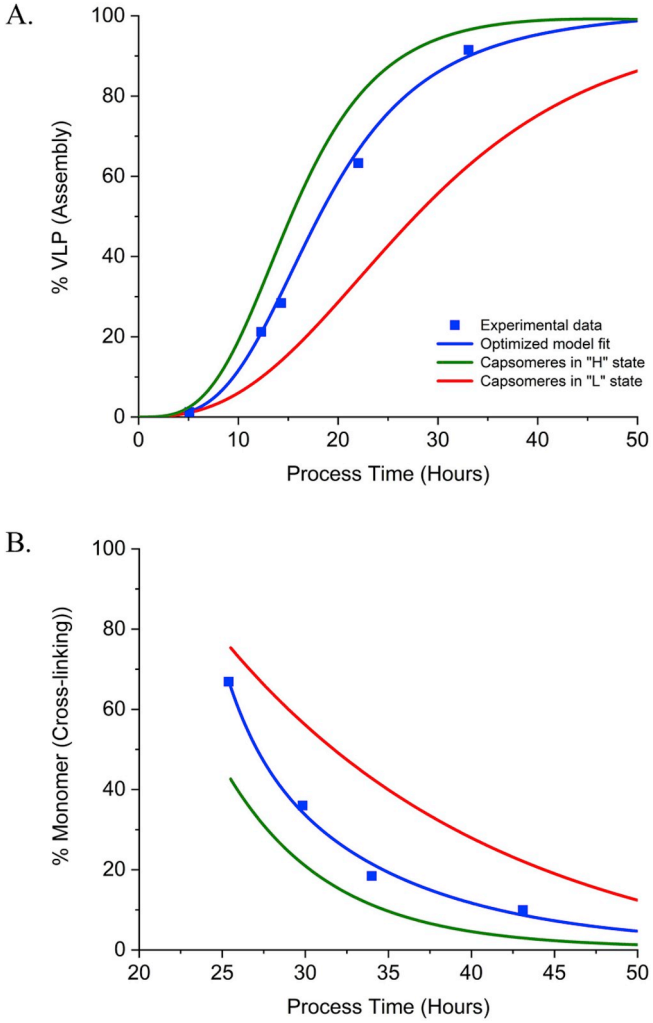

Fig. 12. The illustration of the effect of capsomere geometrical spatial arrangement in the VLP, represented by two profiles: the reassembly (A. Intact VLPs) and cross-linking (B. Monomers) profiles for three different cases: fitted (blue color lines – modeling and points – experiment), and two extreme cases when all capsomeres in the system are in "*H–state*" (green lines) and "*L–state*" (red lines).

incorporating the real VLP geometry and the actual experimental data. We established these phenomena as an ordered series of predictable and interrelated events. We also confirmed our hypothesis that an angular arrangement of capsomeres during reassembly highly affects the overall rates of VLP growth and cross-linking.

In summary, there are many factors (chemical, thermal, mechanical etc.) affect the capsid initiation (capsomere dimerization), growth, and maturation (cross-linking). It is practically impossible to test and include all of them within one study. However, we address these issues to the future research.

Acknowledgments

We thank Dave Wohlpart, Dicky Abraham, Pete DePhillips, and Jon Bendas for enlightening discussions.

Appendix A

The Simulation Procedure

The simulation starts with the initialization of the number of free capsomeres, N_{caps}, in the system. The initial concentration of capsomeres (L1 pentamers) was equal to 0.4 mg/mL, similar to the work (Mach et al., 2006).

Let $a_j(t)$ be a rate at which the event with number j occurs in the system. Given the system state at a time, t, we denote the probability $P_j = \tau \cdot a_j(t)$ that reaction j occurs within the time interval $(t, t + \tau)$. Then we generate the overall sum of the reaction rates, $a_0(t) = \sum a_j(t)$. In this method following two steps are executed in sequence after a final time is reached. Given a current state, at the next step, one has to find the time τ:

The time τ at which the next reaction takes place can be found from the following exponential probability distribution by drawing the first random number, x_1, on the interval [0, 1]:

$$x_1 = (\sum a_j(t))\cdot e^{-\sum a_j(t)\cdot\tau},\ \cdots(\tau \geq 0,\quad 1 \leq j \leq N) \tag{A1}$$

which gives the next time interval in a form

$$\tau = -\frac{logx_1}{\sum a_j(t)} = -\frac{logx_1}{a_0(t)} \tag{A2}$$

Note that the stochastic simulation algorithm (Gillespie, 1976, 1977) uses the rigorous Monte-Carlo approach to generate random pairs (τ, j). Therefore, we choose what reaction takes place by drawing another random number, x_2, from a uniform distribution between 0 and 1. This number defines the next reaction, j, from the list of all possible reaction.

The reaction j is selected with probability:

$$\frac{a_{j-1}}{a_0} < P_j \leq \frac{a_j}{a_0} \tag{A3}$$

The occurrence of the chosen reaction/event changes either the number of free capsomere population, N_{caps}; the number of VLP scaffolds and vacant sites available for attachments; or the number of attached capsomeres accessible for cross-linking. Therefore, we update the capsomere count based on the reaction that occurred, as well as the number of possible reactions, N, and other reaction rates. In case of attachment event, we update the VLP scaffold matrix, which contains the information about all capsomeres in this particular VLP along with their assigned azimuthal angles, θ. Once attached, the L1 proteins of the capsomere are available for the cross-linking with their neighbors. Its Cys 175 and Cys 428 are being included in "cross-linking" dynamics since that moment, increasing value of N. The time is now updated, $t = t+ \tau$, where τ is defined by Eq. (A2). These steps are being repeated until the time, t, reaches the designed process time.

Appendix B

Capsomere number	x (Å)	y (Å)	z (Å)
1	49.8	39.5	251.1
2	109.0	137.7	190.4
3	17.5	215.9	142.1
4	−98.2	166.0	172.9
5	−78.2	57.0	240.3
6	−49.8	−39.5	251.1
7	78.2	−57.0	240.3
8	172.9	−98.2	166.0
9	142.1	17.5	215.9
10	190.4	109.0	137.7
11	215.9	142.1	17.5
12	137.7	190.4	109.0
13	−57.0	240.3	78.2
14	−166.0	172.9	98.2
15	−240.3	78.2	57.0
16	−172.9	98.2	166.0
17	−142.1	−17.5	215.9
18	−109.0	−137.7	190.4
19	−190.4	−109.0	137.7
20	−215.9	−142.1	17.5
21	−137.7	−190.4	109.0
22	−17.5	−215.9	142.1
23	98.2	−166.2	172.9
24	57.0	−240.3	78.2
25	166.0	−172.9	98.2
26	240.3	−78.2	57.0
27	251.1	49.8	39.5
28	39.5	251.1	49.8
29	−39.5	251.1	−49.8
30	−251.1	−49.8	39.5
31	−39.5	−251.1	49.8
32	39.5	−251.1	−49.8
33	17.5	−215.9	−142.1
34	−57.0	−240.3	−78.2
35	−166.0	−172.9	−98.2
36	−172.9	−98.2	−166.0
37	−98.2	−166.0	−172.9
38	−240.3	−78.2	−57.0
39	−215.9	142.1	−17.5
40	−251.1	49.8	−39.5
41	57.0	240.3	−78.2
42	−17.5	215.9	−142.1
43	−109.0	137.7	−190.4
44	−190.4	109.0	−137.7
45	−137.7	190.4	−109.0
46	78.2	57.0	−240.3
47	142.1	−17.5	−215.9
48	49.8	−39.5	−251.1
49	−78.2	−57.0	−240.3

50	−142.1	17.5	−215.9
51	−49.8	39.5	−251.1
52	137.7	−190.4	−109.0
53	109.0	−137.7	−190.4
54	190.4	−109.0	−137.7
55	251.1	−49.8	−39.5
56	215.9	−142.1	−17.5
57	172.9	98.2	−166.0
58	98.2	166.0	−172.9
59	166.0	172.9	−98.2
60	240.3	78.2	−57.0
61	0.0	137.6	222.6
62	0.0	−137.6	222.6
63	222.6	0.0	137.6
64	137.6	222.6	0.0
65	−137.6	222.6	0.0
66	−222.6	0.0	137.6
67	−137.6	−222.6	0.0
68	137.6	−222.6	0.0
69	0.0	−137.6	−222.6
70	−222.6	0.0	−137.6
71	0.0	137.6	−222.6
72	222.6	0.0	−137.6

References

Angelescu, D.G., Linse, P., 2007. Modelling of icosahedral viruses. Curr. Opin. Colloid Interface Sci. 13, 389–394.

Arkhipov, A., Freddolino, P.L., Schulten, K., 2006. Stability and dynamics of virus capsids described by coarse-grained modeling. Structure 14 (12), 1767–1777.

Baker, T.S., Newcomb, W.W., Olson, N.H., Cowsert, L.M., Olson, C., Brown, J.C., 1991. Structures of bovine and human papillomaviruses. Analysis by cryoelectron microscopy and three-dimensional image reconstruction. Biophys. J. 60 (6), 1445–1456.

Bosch, F.X., Munoz, N., 2002. The viral etiology of cervical cancer. Virus Res. 89, 183–190.

Boyd, K.J., Bansal, P., Feng, J., May, E.R., 2015. Stability of Norwalk virus capsid protein interfaces evaluated by in silico nanoindentation. Front. Bioeng. Biotechnol. 3, 103. https://doi.org/10.3389/fbioe.2015.00103.

Buck, C.B., Thompson, C.D., Pang, Y.Y., Lowy, D.R., Schiller, J.T., 2005. Maturation of papillomavirus capsids. J. Virol. 79, 2839–2846.

Buck, C.B., Day, P.M., Trus, B.L., 2013. The papillomavirus major capsid protein L1. Virology 445, 169–174.

Burd, E.M., 2003. Human papillomavirus and cervical cancer. Clin. Microbiol. Rev. 16 (1), 1–17.

Cardone, G., Moyer, A.L., Cheng, N., Thompson, C.D., Dvoretzky, I., Lowy, D.R., Schiller, J.T., Steven, A.C., Buck, C.B., Trus, B.L., 2014. Maturation of the human papillomavirus 16 capsid. mBio 5, e01104–01114. https://doi.org/10.2210/pdb3J6R/pdb. EMDataResource: EMD-5932.

Carter, J., Saunders, V., 2007. Virology: Principles and Applications. Wiley, Chichester.

Casini, G.L., Graham, D., Heine, D., Garcea, R.L., Wu, D.T., 2004. In vitro papillomavirus capsid assembly analyzed by light scattering. Virology 325, 320–327.

Caspar, D.L.D., Klug, A., 1962. Physical principles in the construction of regular viruses. Cold Spring Harbor Symp. Quant. Biol. 27, 1–24.

Conway, M.J., Cruz, L., Alam, S., Christensen, N.D., Meyers, C., 2011. Differentiation-dependent interpentameric disulfide bond stabilizes native human. Papillomavirus type 16. PLoS One 6 (7), e22427.

Duda, R.L., 1998. Protein chainmail: catenated protein in viral capsids. Cell 94, 55–60.

Dürst, M., Gissmann, L., Ikenberg, H., zur Hausen, H., 1983. A papillomavirus DNA from a cervical carcinoma and its prevalence in cancer biopsy samples from different geographic regions. Proc. Natl. Acad. Sci. U.S.A. 80, 3812–3815.

Dykeman, E.C., Stockley, P.G., Twarock, R., 2014. Levinthal's paradox for virus assembly. Proc. Natl. Acad. Sci. U.S.A. 111 (14), 5361–5366. https://doi.org/10.1073/pnas.1319479111.

Elrad, O.M., Hagan, M.F., 2008. Mechanisms of size control and polymorphism in viral capsid assembly. Nano Lett. 8 (11), 3850–3857.

Fields, B.N., Knipe, D.M., Howley, P.M., 2007. Fields Virology. Wolters Kluwer, Health/Lippincott Williams & Wilkins, Philadelphia.

Garcea, R.L., DiMaio, D. (Eds.), 2007. The Papillomaviruses. Springer, New York.

Gibson, M.A., Bruck, J., 2000. Efficient exact stochastic simulation of chemical systems with many species and many channels. J. Phys. Chem. A 104 (9), 1876–1889.

Gillespie, D.T., 1976. A general method for numerically simulating the stochastic time evolution of coupled chemical reactions. J. Comput. Phys. 22 (4), 403–434.

Gillespie, D.T., 1977. Exact stochastic simulation of coupled chemical reactions. J. Phys. Chem. 81 (25), 2340–2361.

Gillison, M.L., 2004. Human papillomavirus-associated head and neck cancer is a distinct epidemiologic, clinical, and molecular entity. Semin. Oncol. 31, 744–754.

Gorshkov, V., Zavalov, A., Privman, V., 2009. Shape selection in diffusive growth of colloids and nanoparticles. Langmuir 25 (14), 7940–7953.

Gorshkov, V., Zavalov, O., Atanassov, P.B., Privman, V., 2011. Morphology of nanoclusters and nanopillars formed in nonequilibrium surface growth for catalysis applications. Langmuir 27 (13), 8554–8561.

Hadden, J.A., Perilla, J.R., Schlicksup, C.J., Venkatakrishnan, B., Zlotnick, A., Schulten, K., 2018. All-atom molecular dynamics of the HBV capsid reveals insights into biological function and cryo-EM resolution limits. eLife 7, e32478. https://doi.org/10.7554/eLife.32478.001.

Hagan, M.F., 2013. Modeling Viral Capsid Assembly. Paper in Arxiv: arXiv:1301.1657.arXiv.org e-Print archive. https://arxiv.org/abs/1301.1657, Accessed date: 8 January 2013.

Hagan, M.F., 2014. Modeling viral capsid assembly. Adv. Chem. Phys. 155, 1–68.

Hagan, M.F., Chandler, D., 2006. Dynamic pathways for viral capsid assembly. Biophys. J. 91, 42–54.

Hagan, M.F., Elrad, O.M., 2010. Understanding the concentration dependence of viral capsid assembly kinetics—the origin of the lag time and identifying the critical nucleus size. Biophys. J. 98, 1065–1074.

Hagan, M.F., Elrad, O.M., Jack, R.L., 2011. Mechanisms of kinetic trapping in self-assembly and phase transformation. J. Chem. Phys. 135, 104115.

Hare, D.J., Chan, J.C., 1968. Role of hydrogen and disulfide bonds in polyoma capsid structure. Virology 34, 481–491.

Hemberg, M., Yaliraki, S.N., Barahona, M., 2006. Stochastic kinetics of viral capsid assembly based on detailed protein structures. Biophys. J. 90, 3029–3042.

Ishii, Y., Tanaka, K., Kanda, T., 2003. Mutational analysis of human papillomavirus type 16 major capsidprotein L1: the cysteines affecting the intermolecular bonding and structure of L1-capsids. Virology 308, 128–136.

Ishizu, K.-I., Watanabe, H., Han, S.-I., Kaneshashi, S.-N., Hoque, M., Yajima, H., Kataoka, K., Handa, H., 2001. Roles of disulfide linkage and calcium ion-mediated interactions in assembly and disassembly of virus-like particles composed of simian virus 40 VP1 capsid protein. J. Virol. 75 (1), 61–72.

Keef, T., Micheletti, C., Twarock, R., 2006. Master equation approach to the assembly of viral capsids. J. Theor. Biol. 242 (3), 713.

Kirnbauer, R., Booy, F., Cheng, N., Lowy, D.R., Schiller, J.T., 1992. Papillomavirus L1 major capsid protein self-assembles into virus-like particles that are highly immunogenic. Proc. Nati. Acad. Sci. U.S.A. 89, 12180–12184.

Li, M., Beard, P., Estes, P.A., Lyon, M.K., Garcea, R.L., 1998. Intercapsomeric disulfide bonds in papillomavirus assembly and disassembly. J. Virol. 72, 2160–2167.

Lidmar, J., Mirny, L., Nelson, D.R., 2003. Virus shapes and buckling transitions in spherical shells. Phys. Rev. E 68, 051910.

Lorman, V.L., Rochal, S.B., 2008. Landau theory of crystallization and the capsid structures of small icosahedral viruses. Phys. Rev. B 77, 224109.

Mach, H., Volkin, D.B., Troutman, R.D., Wang, B., Luo, Z., Kathrin, U., Jansen, K.U., Hi, L., 2006. Disassembly and reassembly of yeast-derived recombinant human papillomavirus virus-like particles (HPV VLPs). J. Pharm. Sci. 95 (10), 2195–2206.

Mateu, M.G., 2013. Assembly, stability and dynamics of virus capsids. Arch. Biochem. Biophys. 531, 65–79.

May, E.R., Feng, J., Brooks III, C.L., 2012. Exploring the symmetry and mechanism of virus capsid maturation via an ensemble of pathways. Biophys. J. 1020 (3), 606–612.

McBride, A.A., Münger, K., 2018. Expert views on HPV infection. Viruses 10 (2), 94.

Michaels, T.C.T., Bellaiche, M.M.J., Hagan, M.F., Knowles, T.P.J., 2017. Kinetic constraints on self-assembly into closed supramolecular structures. Sci. Rep. 7, 12295.

Modis, Y., Trus, B.L., Harrison, S.C., 2002. Atomic model of the papillomavirus capsid. EMBO J. 21 (18), 4754–4762.

Moisant, P., Neeman, H., Zlotnick, A., 2010. Exploring the paths of (virus) assembly. Biophys. J. 99, 1350–1357.

Mukherjee, S., Thorsteinsson, M.V., Johnston, L.B., DePhillips, P.A., Zlotnick, A., 2008. A quantitative description of in vitro assembly of human papillomavirus 16 virus-like particles. J. Mol. Biol. 381, 229–237.

Nguyen, H.D., Reddy, V.S., Brooks III, C.L., 2009. Invariant polymorphism in virus capsid assembly. J. Am. Chem. Soc. 131, 2606–2614.

Noad, R., Roy, P., 2003. Virus-like particles as immunogens. Trends Microbiol. 11 (9), 438–444.

Pastrana, D.V., Peretti, A., Welch, N.L., Borgogna, C., Olivero, C., Badolato, R., Notarangelo, L.D., Gariglio, M., FitzGerald, P.C., McIntosh, C.E., Reeves, J., Starrett, G.J., Bliskovsky, V., Velez, D., Brownell, I., Yarchoan, R., Wyvill, K.M., Uldrick, T.S., Maldarelli, F., Lisco, A., Sereti, I., Gonzalez, C.M., Androphy, E.J., McBride, A.A., Van Doorslaer, K., Garcia, F., Dvoretzky, I., Liu, J.S., Han, J., Murphy, P.M., McDermott, D.H., Buck, C.B., 2018. Metagenomic discovery of 83 new human papillomavirus types in patients with immunodeficiency. mSphere 3 (6) e00645-18.

Perilla, J.R., Schulten, K., 2017. Physical properties of the HIV-1 capsid from all-atom molecular dynamics simulations. Nat. Commun. 8, 15959.

Perilla, J.R., Hadden, J.A., Goh, B.C., Mayne, C.G., Schulten, K., 2016. All-atom molecular dynamics of virus capsids as drug targets. J. Phys. Chem. Lett. 7 (10), 1836–1844.

Prasad, B.V.V., Schmid, M.F., 2012. Principles of virus structural organization. Adv. Exp. Med. Biol. 726, 17–47.

Rapaport, D.C., 2004. Self-assembly of polyhedral shells: a molecular dynamics study. Phys. Rev. E. 70, 051905.

Rapaport, D.C., 2018. Molecular dynamics study of T = 3 capsid assembly. J. Biol. Phys. 44 (2), 147–162. https://doi.org/10.1007/s10867-018-9486-7.

Reddy, T., Sansom, M.S.P., 2016. Computational virology: from the inside out. Biochim. Biophys. Acta Biomembr. 1858 (7 B), 1610–1618.

Salunke, D.M., Caspar, D.L.D., Garcea, R.L., 1986. Self-assembly of purified polyomavirus capsid protein VP1. Cell 46, 895–904.

Sapp, M., Bienkowska-Haba, M., 2009. Viral entry mechanisms: human papillomavirus and a long journey from extracellular matrix to the nucleus. FEBS J. 276, 7206–7216.

Sapp, M., Fligge, C., Petzak, I., Harris, J.R., Streeck, R.E., 1998. Papillomavirus assembly requires trimerization of the major capsid protein by disulfides between two highly conserved cysteines. J. Virol. 72, 6186–6189.

Schiller, J.T., Muller, M., 2015. Next generation prophylactic human papillomavirus vaccines. Lancet Oncol. 16 (5), e217–225.

Schwartz, R., Shor, P.W., Prevelige Jr., P.E., Berger, B., 1998. Local rules simulation of the kinetics of virus capsid self-assembly. Biophys. J. 75, 2626–2636.

Serebrinsky, S.A., 2011. Physical time scale in kinetic Monte Carlo simulations of continuous-time Markov chains. Phys. Rev. E. 83, 037701.

Siber, A., 2006. Buckling transition in icosahedral shells subjected to volume conservation constraint and pressure: relations to virus maturation. Phys. Rev. E 73, 061915.

Siegel, R., Naishadham, D., Jemal, A., 2012. Cancer statistics 2012. CA A Cancer J. Clin. 62, 10–29.

Stanley, M., Lowy, D.R., Frazer, I., 2006. Chapter 12: prophylactic HPV vaccines: underlying mechanisms. Vaccine 24S3 S3/106–S3/113.

Stray, S.J., Ceres, P., Zlotnick, A., 2004. Zinc ions trigger conformational change and oligomerization of hepatitis B virus capsid protein. Biochemistry 43, 9989–9998.

Twarock, R., 2004. A tiling approach to virus capsid assembly explaining a structural puzzle in virology. J. Theor. Biol. 226, 477–482.

Twarock, R., 2006. Mathematical virology: a novel approach to the structure and assembly of viruses. Phil. Trans. R. Soc. A 364, 3357–3373.

Twarock, R., Hendrix, R.W., 2006. Crosslinking in viral capsids via tiling theory. J. Theor. Biol. 3, 419–424.

Vega, J.F., Vicente-Alique, E., Núñez-Ramírez, R., Wang, Y., Martínez-Salazar, J., 2016. Evidences of changes in surface electrostatic charge distribution during stabilization of HPV16 virus-like particles. PLoS One 11 (2), e0149009.

Verdier, T., Foret, L., Castelnovo, M., 2016. Modeling the kinetics of open self-assembly. J. Phys. Chem. B 120 (26), 6411–6420. https://doi.org/10.1021/acs.jpcb.6b03899.

Walboomers, J.M., Jacobs, M.V., Manos, M.M., Bosch, F.X., Kummer, J.A., Shah, K.V., Snijders, P.J., Peto, J., Meijer, C.J., Muñoz, N., 1999. Human papillomavirus is a necessary cause of invasive cervical cancer worldwide. J. Pathol. 189, 12–19.

Wang, J.W., Roden, R.B., 2013a. Virus-like particles for the prevention of human papillomavirus-associated malignancies. Expert Rev. Vaccines 12 (2), 129–141.

Wang, J.W., Roden, R.B.S., 2013b. L2, the minor capsid protein of papillomavirus. Virology 445 (1–2), 175–186.

Whitesides, G.M., Boncheva, M., 2002. Beyond molecules: self-assembly of mesoscopic and macroscopic components. Proc. Natl. Acad. Sci. U.S.A. 99 (8), 4769–4774.

Wilber, A.W., Doye, J.P.K., Louis, A.A., Lewis, A.C.F., 2009. Monodisperse self-assembly in a model with protein-like interactions. J. Chem. Phys. 131, 175102.

Wolf, M., Garcea, R.L., Grigorieff, N., Harrison, S.C., 2010. Subunit interactions in bovine papillomavirus. Proc. Natl. Acad. Sci. U.S.A. 107 (14), 6298–6303.

Zandi, R., van der Schoot, P., Reguera, D., Kegel, W., Reiss, H., 2006. Classical nucleation theory of virus capsids. Biophys. J. 90, 1939–1948.

Zhang, L., Tang, R., Bai, S., Connors, N.K., Lua, L.H.L., Chuan, Y.P., Middelberg, A.P.J., Sun, Y., 2013. Molecular energetics in the capsomere of virus-like particle. J. Phys. Chem. B 117, 5411–5421.

Zhao, Q., Modis, Y., High, K., Towne, V., Meng, Y., Wang, Y., Alexandroff, J., Brown, M., Carragher, B., Potter, C.S., Abraham, D., Wohlpart, D., Kosinski, M., Washabaugh, M.W., Sitrin, R.D., 2012. Disassembly and reassembly of human papillomavirus virus-like particles produces more virion-like antibody reactivity. Virol. J. 9, 52.

Zlotnick, A., 2005. Theoretical aspects of virus capsid assembly. J. Mol. Recognit. 18, 479–490.

Zlotnick, A., Mukhopadhyay, S., 2011. Virus assembly, allostery and antivirals. Trends Microbiol. 19 (1), 14–23.

Zlotnick, A., Stray, S.J., 2003. How does your virus grow. Understanding and interfering with virus assembly. Trends Biotechnol. 21 (12), 536–542.