Protein Cargo Encapsulation by Virus-Like Particles: Strategies and Applications

Donna McNeale¹, Noor Dashti², Li Chen Cheah², Frank Sainsbury¹,²,*

1. Centre for Cell Factories and Biopolymers, Griffith Institute for Drug Discovery, Griffith University, Nathan, QLD 4111, Australia
2. Australian Institute for Bioengineering and Nanotechnology, The University of Queensland, St Lucia, QLD 4072, Australia
* Correspondence: f.sainsbury@griffith.edu.au

ABSTRACT

Viruses and the recombinant protein cages assembled from their structural proteins, known as virus-like particles (VLPs), have gained wide interest as tools in biotechnology and nanotechnology. Detailed structural information and their amenability to genetic and chemical modification make them attractive systems for further engineering. This review describes the range of non-enveloped viruses that have been co-opted for heterologous protein cargo encapsulation and the strategies that have been developed to drive encapsulation. Spherical capsids of a range of sizes have been used as platforms for protein cargo encapsulation. Various approaches, based on native and non-native interactions between the cargo proteins and inner surface of VLP capsids, have been devised to drive encapsulation. Here we outline the evolution of these approaches, discussing their benefits and limitations. Like the viruses from which they are derived, VLPs are of interest in both biomedical and materials applications. The encapsulation of protein cargo inside VLPs leads to numerous uses in both fundamental and applied biocatalysis and biomedicine, some of which are discussed herein. The applied science of protein encapsulating VLPs is emerging as a research field with great potential. Developments in loading control, higher order assembly and capsid optimization are poised to realize this potential in the near future.

INTRODUCTION

Virus-like particles (VLPs) mimic the structure of viruses, albeit devoid of native genetic material. They retain many of the inherent properties of virus capsids, including precise
structure and uniformity of size, biocompatibility, stability, immunogenicity and cell entry behaviour (Lua et al., 2014). These features make them attractive starting materials for a range of diverse applications in health and manufacturing. They are produced using scalable and biocompatible processes using various production hosts, which may or may not be the same as the source host for the cognate virus (Glasgow and Tullman-Ercek, 2014; Le and Muller, 2021). Moreover, VLPs are amenable to genetic and chemical modification (Chen et al., 2016; Schoonen and van Hest, 2014; Szyszka et al., 2022), which can be made with molecular precision to impart new functionalities, further diversifying the potential applications of VLPs.

This review focuses on the use of VLPs as protein cages and their capacity for VLPs to encapsulate heterologous protein cargos. The interior cavity of VLPs have been used for the sequestration of many different protein cargos, most prominently fluorescent proteins and various enzymes (Wilkerson et al., 2018). The former have been instrumental in the design and testing of protein-encapsulating VLPs (Dashti et al., 2018; Minten et al., 2009; O’Neil et al., 2011), as well as their use as intracellular delivery vehicles (Brillault et al., 2017; Dashti et al., 2018) and in vivo imaging agents (Das et al., 2020a; Herbert et al., 2020). The latter results in the creation of biocatalytic nanocompartments for application in vitro and in vivo, and their development has revealed information on the fundamental behaviour of enzymes in confined environments (Jordan et al., 2016; Minten et al., 2011; Wang et al., 2020). Here we discuss the diversity of VLP systems developed for protein encapsulation. Numerous other types of protein cages have been developed in nanotechnology, and some of these are indeed virus-like. We point the reader to several useful reviews for further information on these non-viral protein cages (Bhaskar and Lim, 2017; Edwardson and Hilvert, 2019; Edwardson et al., 2022; Stupka and Heddle, 2020).

Viruses have evolved a variety of mechanisms for the encapsulation of their genome, minor coat proteins, and non-structural proteins, and researchers have explored these strategies for the encapsulation of foreign cargos. From adapting native biomolecular interactions to capsid protein translational fusions, we discuss the various strategies for protein cargo encapsulation within non-enveloped VLPs. The source of these VLPs are viruses that infect prokaryotes, plants, and animals, and the VLPs are diverse in terms of size and capsid protein organization (Figure 1). Properties of the particles and their assembly mechanisms dictate the available encapsulation approaches that may be used for each VLP. The advantages and limitations of these approaches are considered, and we highlight several studies that demonstrate advancements in heterologous cargo encapsulation within VLPs. These advancements are accompanied by applications in biomedicine and biocatalysis, and we discuss a range of applications and progress made within these areas.
Figure 1. Examples of non-enveloped viruses used for the encapsulation of heterologous proteins. PDB IDs: MS2, 2MS2; CCMV, 1ZA7; Qβ, 1QBE; PP7, 1DWN; CMV, 1F15; HBcAg, 1QGT; PyV, 1SIE; P22, 2XYY; BTV core, 2BTV minus PCD chains; and RV core, 3KZ4. Murine polyomavirus is shown as a representative of the Polyomaviridae family. Average external diameters were obtained from VIPERdb (https://viperdb.scripps.edu). Figure lists maximum determined cargo numbers, see text for details.
NON-SPECIFIC ENCAPSULATION

VLP assembly in the presence of high bulk concentrations of cargo proteins can result in encapsulation by non-specific interactions. To achieve specificity with this type of encapsulation approach, assembly must be carried out in vitro from purified components. While this does allow a degree of control, in general, loading density of cargo proteins is low.

Statistical Encapsulation

Statistical encapsulation is achieved when capsid protein subunits are assembled in vitro in the presence of a high concentration of free protein cargo (Figure 2), leading to encapsulation of cargo at a density that is dependent on the cargo concentration. This approach led to the first demonstration of protein encapsulation by a capsid-based VLP, which was assembled from the coat protein (CP) of Cowpea chlorotic mottle virus (CCMV) (Comellas-Aragonès et al., 2007). CCMV VLPs exhibit reversible assembly and disassembly, which is pH dependent. While this can limit encapsulation of pH-sensitive cargos, this study reported a novel means of investigating single enzyme kinetics via the encapsulation of horseradish peroxidase (Comellas-Aragonès et al., 2007). This approach has also been taken for green fluorescent protein (GFP) encapsulation by core particles of Hepatitis B virus (HBV) (Lee and Tan, 2008). Hepatitis B core antigen (HBcAg) VLPs dissociate in the presence of denaturants and reassemble upon their removal (Wingfield et al., 1995).

Amenability to in vitro assembly characterizes the VLPs used for statistical encapsulation. This provides specific advantages related to the chemical control over assembly, for example, avoiding the encapsulation of host proteins and potentially some control over the average number of proteins encapsulated per particle. However, statistical encapsulation requires a large excess of cargo that needs to be separately expressed and purified before loading and subsequently removed from reassembled VLPs before quantitative assays on encapsulation can be performed. Furthermore, loading density is limited by the concentration of cargo proteins required to achieve loading by statistical encapsulation alone. The following strategies use protein engineering to shift the equilibrium towards partitioning of cargo proteins within VLPs assembled both in vitro and in vivo.
Electrostatic interactions

Encapsulation by electrostatic interactions relies on charge complementarity between the interior surface of the VLP and the cargo protein (Figure 2). Many capsid-based viruses carry positive charges on their interior surface, which facilitates encapsulation of their viral genome, and this feature can be exploited to encapsulate other negatively charged cargo. For example, a far-red fluorescent protein, small-ultrared fluorescent protein (smURFP), was encapsulated within two structurally similar VLPs from the Leviviridae family, bacteriophage Qβ and pseudomonas phage seven (PP7) (Herbert et al., 2020). Both viral capsids are composed of 180 copies of a coat protein and assembly is normally directed by interactions with RNA. Allophycocyanin α-subunit (APCα), from which smURFP is derived, has an isoelectric point of 4.64, allowing smURFP to mimic the role of RNA during assembly and thus be encapsulated at neutral pH. Herbert et al (2020) achieved packaging of approximately three smURFP per
particle, producing brightly fluorescent VLPs which, when injected subcutaneously and intravenously, were shown to be effective noninvasive imaging agents (Herbert et al., 2020).

Charge-mediated encapsulation can also be achieved by imparting negative charges on the cargo protein through the conjugation of nucleic acids (Brasch et al., 2017; Glasgow et al., 2012; Lu et al., 2012). In the case of CCMV, the N-terminal region of the CP is enriched in arginine residues, and it was found that chemically linking nucleic acid adapters to cargo proteins can trigger assembly around the cargo protein (Brasch et al., 2017). This represented a significant development in the use of CCMV for the encapsulation of proteins sensitive to the changes in pH previously used to drive assembly. Moreover, it provides a mechanism that is potentially tunable by varying the adapter length or salt concentration during assembly. Using this strategy, Brasch et al (2017) demonstrated the confinement of two separate enzyme pathways, achieving an average encapsulation of one glucose oxidase (GOx), or one GOx and one or two gluconokinase (GCK) per particle. Cucumber mosaic virus (CMV) belongs to the same viral family as CCMV and is structurally similar, however unlike CCMV, CMV requires CP-RNA interactions for assembly. This interaction was mimicked by a DNA oligomer conjugated to streptavidin and a fluorophore to induce assembly of CMV, leading to incorporating the protein cargo into VLPs (Lu et al., 2012). The Bacteriophage MS2 virus-like particle was one of the first non-enveloped VLPs used for protein encapsulation and this was mediated by a sequence-specific RNA-coat protein interaction (Wu et al., 1995). However, in the presence of the osmolyte, trimethylamine N-oxide, assembly can proceed efficiently around cargo modified with non-specific polyanions (Glasgow et al., 2012), obviating the need to work with relatively unstable RNA aptamers. The conjugation of a DNA tag to GFP, for example, resulted in the encapsulation of approximately 6.5 GFPs per capsid. The subsequent use of a genetically fused polyanionic tag (16 acidic amino acid residues) also resulted in capsid assembly in the presence of osmolyte, though the number of encapsulated GFPs or alkaline phosphatases was not reported (Glasgow et al., 2012).

The use of complementary charges to direct assembly provides a relatively simple method for cargo encapsulation, albeit with the requirement of separate cargo purification for in vitro assembly and, in some cases, the chemical modification of cargo molecules with polyanions such as nucleic acids. This strategy has been of particular use in VLPs that require a charged cargo for assembly. Although cargo loading efficiencies are lower compared to encapsulation methods that use specific interactions, the approach is more flexible with regard to polyanion conjugate and loading densities can be somewhat tunable through the variation of adapter length or ionic strength.
SPECIFIC NON-COVALENT INTERACTIONS

Encapsulation via specific interactions can be mediated by either protein-protein or protein-nucleic acid interactions. Furthermore, these interactions can be native, borrowing from the assembly of virions around internal components, or engineered non-native interactions. These approaches may be used for encapsulation during *in vitro* assembly, but particularly in the case of native interactions, the specificity of the interactions affords the possibility of *in vivo* self-sorting of protein cargos into VLPs that assemble in their host cells.

**Native**

**Protein**

Bacteriophage P22 particles have been widely investigated for their encapsulation abilities (Esquirol et al., 2022; O'Neil et al., 2012; O'Neil et al., 2011; Patterson et al., 2012; Patterson et al., 2013a; Patterson et al., 2014; Qazi et al., 2016; Wang et al., 2020). The particle is composed of 420 copies of CP forming an icosahedral capsid and is assembled with the support of 100 to 330 copies of scaffold protein (SP). A truncated SP, which retains its ability to template the assembly of P22 VLPs, can be used to encapsulate cargo proteins via genetic fusion to the cargo N or C termini (Figure 3A). O'Neil et al (2011) reported a method for the *in vivo* encapsulation of fluorescent proteins, GFP and mCherry, genetically fused to an N-terminally truncated SP and co-expressed with P22 CP (O'Neil et al., 2011). P22 capsids showed high loading capacity with 281 GFP molecules or 233 mCherry molecules within each capsid. This high loading capacity enables the co-encapsulation of cargo proteins via sequential fusion to SP in a 1:1 ratio (O'Neil et al., 2012). While the *in vivo* assembly and encapsulation in P22 does not require the separate purification of cargo and CP components, it provides little control over cargo stoichiometry and loading density. To overcome these limitations, an *in vitro* assembly approach was developed (Sharma et al., 2017) where packaging stoichiometry and density of cargos can be modified by varying the ratio of cargo-fused SP and wildtype SP (Sharma and Douglas, 2020).
Figure 3. Examples of specific interactions mediating cargo protein interactions in VLPs. A) SP-mediated in vivo encapsulation in P22 adapted from (Patterson et al., 2012). B) VP2 peptide-mediated in vivo cargo capture and in vitro encapsulation in MPyV VLPs adapted from (Dashti et al., 2018). C) RNA adapter-mediated in vivo encapsulation in Qβ VLPs adapted from (Das et al., 2020b). D) Coiled-coil interactions mediating in vitro CP purification followed by in vitro encapsulation in CCMV VLPs adapted from (Minten et al., 2009).

Polyomavirus (PyV) VLPs have also been widely explored for both in vivo (Bouřa et al., 2005; Catrice and Sainsbury, 2015; Cheah et al., 2021) and in vitro (Abbing et al., 2004; Dashti et al., 2018) encapsulation. The capsid of murine polyomavirus (MPyV) is comprised of 360 copies of the major coat protein VP1 arranged as 72 pentamers. Each of these capsomeres can accommodate the binding of one minor coat proteins VP2/3, which share the same C-terminus. Abbing et al (2004) reported a method for in vitro encapsulation by fusion of GFP to a 49 amino
acid C-terminal fragment of VP2 (VP2C) (Abbing et al., 2004), which is sufficient for VP1 capsomere binding (Chen et al., 1998). GFP-loaded VLPs were assembled by mixing GFP-VP2C with VP1 pentamers resulting in an estimated 64 GFPS per VLP. However, due to the hydrophobicity of VP2, a major drawback to this method is the poor solubility of the anchor protein. To address this limitation, an approach combining in vivo cargo capture with in vitro assembly was established (Dashti et al., 2018). In contrast to self-assembly in eukaryotic cells, MPyV VP1 capsomers do not self-assemble into VLPs when expressed in prokaryotic cells. Co-expression of VP1 and VP2C-EGFP resulted in the formation of soluble VP1/VP2C-EGFP complexes which were subsequently assembled in vitro (Figure 3B). The co-encapsulation of two fluorescent proteins, EGFP and mRuby3, was used to demonstrate Förster resonance energy transfer (FRET) between the co-confined cargo molecules (Dashti et al., 2018). This work reduced the length of the VP2 C-terminal fragment used as a capsomere-binding anchor to 31 aa and the in vivo capture approach enabled fusing the anchor to cargo protein N-termini. This orientation destabilizes the fusion protein in vitro (Abbing et al., 2004) and in vivo (Cheah et al., 2021), however, interaction with VP1 capsomeres masks the hydrophobic side chains of the conserved alpha-helix that mediates binding (Chen et al., 1998). Work on the similar polyomavirus, Simian virus 40 (SV40), showed that placing VP2C fragments at the N-terminus of cargo proteins avoids the external presentation of cargo proteins that sometimes results from C-terminal fusion of VP2C (Inoue et al., 2008). Another polyomavirus, JC polyomavirus, is able to self-assemble in prokaryotic cells, and has also been explored in protein encapsulation studies (Ohtake et al., 2010). GFP fused to the N-terminus of VP2 and co-expressed with VP1 in E.coli was encapsulated in vivo. When GFP was replaced with a 6x Histidine tag, small molecules containing nitrilotriacetic acid (NTA) could be encapsulated via diffusion through the capsid pores and binding to the engineered cargo protein. Release of the cargo was shown at low pH and protonation of the histidines, showing potential as a controlled release drug carrier (Ohtake et al., 2010). Human papillomavirus (HPV) is structurally similar to PyV VLPs and has also been used for in vivo encapsulation in insect cells by an orthologous approach, with fusion of GFP to the minor capsid protein (Windram et al., 2008).

Nucleic acid

Sequence or structure-specific protein-RNA interactions are necessary for the capsid assembly of many ssRNA viruses (Twarock et al., 2018). These secondary structure elements, known as packaging signals (PSs), have been used to trigger the assembly and encapsulation of foreign cargo within VLPs (Ashley et al., 2011; Das et al., 2020a; Fiedler et al., 2010; Rhee et al., 2011; Wu et al., 1995). Genome encapsulation and assembly of the Bacteriophage MS2 capsid is mediated by a 19-nucleotide RNA stem-loop (pac site) (Pickett and Peabody, 1993). The
encapsulation of foreign protein cargo has been achieved through the conjugation of the pacman site RNA to the cargo (Ashley et al., 2011; Wu et al., 1995). Wu et al. (1995) first demonstrated the in vitro encapsulation of a glycoprotein toxin, ricin toxin A-chain (RTA) into MS2 VLPs using this approach, which were successfully delivered into mammalian cells.

Similar to MS2, bacteriophage Qβ assembly uses an RNA hairpin structure that interacts with the interior of the CP to form viral capsids (Witherell and Uhlenbeck, 1989). To test the in vivo encapsulation of a protein cargo within Qβ VLPs, a bifunctional RNA molecule containing an α-Rev RNA aptamer and the genome packaging hairpin flanking the CP RNA sequence was designed, while an N-terminal Rev peptide tag was fused to the cargo enzyme (Fiedler et al., 2010). Co-expression of this dual-plasmid system led to the binding of the Rev-tagged enzyme to the α-Rev aptamer, which was subsequently tethered to the interior of the Qβ capsid via the RNA hairpin (Figure 3C). Three enzymes were tested with this system, and it was shown that the average number of encapsulated cargos could be somewhat controlled by modifying expression conditions, achieving a variable loading of PepE, a 24 kDa enzyme, between 2 and 18 copies per particle. This encapsulation strategy was also used to package fluorescent proteins (FPs) separately into Qβ VLPs, with up to 15 proteins per particle (Rhee et al., 2011), and near-infrared FPs with either 5 or 9 copies per particle (Das et al., 2020a).

Non-native (engineered)

Protein

Non-covalent encapsulation strategies also include the use of non-native binding domains to generate specific interactions between the capsid inner surface and cargo, directing encapsulation of cargo proteins upon assembly. WW domains are small protein modules that mediate protein-protein interactions through the binding of proline-rich ligands (Ingham et al., 2005). Schmidt et al (2001) used the first WW domain of the mouse formin binding protein 11 (FBP11) and the proline-rich motif PPLP to direct the in vitro encapsulation of PPLP-tagged proteins within polyomavirus (MPyV) VLPs (Schmidt et al., 2001). The FBP11 WW domain was fused to the inward-facing N-terminus of VP1, while GFP was modified with the PPLP tag. Increase in cargo-to-capsomere ratio during assembly led to an increase in cargo loading, reaching a maximum of 260 GFP molecules per particle. Although only demonstrated for in vitro loading, the capacity is greater than encapsulation via the minor coat protein VP2, which has a theoretical loading maximum of 72 monomeric cargo molecules per particle.

Coiled-coils are protein structural motifs composed of two to six alpha helices arranged either parallel or antiparallel to each other (Truebsestein and Leonard, 2016). They can be homo- or
hetero-oligomers and contain repeating hydrophobic residues which drives the association of the helices resulting in the coiled-coil formation. Heterodimeric E and K coil peptides were chosen for the directed non-covalent protein encapsulation inside CCMV particles (Minten et al., 2011; Minten et al., 2009; Rurup et al., 2014). Through genetic modification, the K-coil was fused to the interior-facing N-terminus of the CCMV CP and the E-coil to the C-terminus of EGFP. Purification of EGFP-capsid protein complexes was achieved through immobilization of the GFP and capture of the CP directly from lysates using K/E-coil dimerization (Figure 3D)(Minten et al., 2009). *In vitro* assembly of the EGFP-capsid protein complexes required mixing with unmodified capsid protein in various ratios, enabling the encapsulation of up to 15 EGFP molecules per capsid. This approach allowed for higher loading with more control over loading density than by non-specific encapsulation in CCMV and was further explored for the encapsulation of *Pseudozyma antartica* lipase B (PalB) at various loading densities to enable investigation of reaction rate as a function of PalB loading density (Minten et al., 2011). The use of non-native binding domains provides an alternative to natural targeting peptides, such as scaffolding proteins, although it has only been applied to *in vitro* assembly and encapsulation.

**COVALENT INTERACTIONS**

Encapsulation *via* covalent interactions with capsid proteins results from both translational or post-translational, enzymatic, fusion of cargo proteins. Due to the considerable changes to capsid protein properties that result from cargo fusion, this approach is generally used for *in vivo* encapsulation. Nearly all non-enveloped VLPs that have been used for heterologous protein encapsulation can be assembled *in vivo* (Table 1), however, success of this strategy depends on the organization of capsid proteins and size of the capsid. In larger VLPs that can accommodate the full complement of fusion proteins while maintaining high-fidelity assembly, cargo proteins can be loaded with predictable stoichiometries. In smaller VLPs, encapsulation is achieved in mosaic particles containing both fused and unfused capsid proteins.

**Translational CP fusion**

Genetic fusion directly to coat proteins is a conceptually simple method for the encapsulation of foreign protein cargo (Brillault et al., 2017; Charpilienne et al., 2001). In cases where encapsulation is efficient, this approach can allow high loading of cargo proteins at known stoichiometries. However, it is limited by steric constraints on cargo protein size and assembly.
of capsid proteins (Belval et al., 2016; Beterams et al., 2000; Cortes-Perez et al., 2010; Gilbert et al., 2004; Kar et al., 2005; Pascual et al., 2015). Infectious bursal disease virus (IBDV) capsids are formed by 260 trimers of VP2 protein (Luque et al., 2009). In the absence of the scaffolding protein VP3, precursor VP2 (pVP2) with an N-terminal Histidine tag can assemble into the T=13 VLPs (Pascual et al., 2015). Both the N and C termini of pVP2 reside within the interior of the capsid and while direct fusion of EGFP to the N-terminus of pVP2 via the Histidine-tag did not result in VLPs, this was overcome by co-expression of cargo-CP and unmodified CP in insect cells. Stable particles containing approximately 240 copies of EGFP (~1 per pVP2 trimer) were formed when the EGFP fusion was co-expressed with unfused pVP2, indicating that steric hindrance limited capsid formation (Pascual et al., 2015).

The Reoviruses, Rotavirus (RV) and Bluetongue virus (BTV) have also been used for encapsulation by translational fusion (Brillault et al., 2017; Charpilienne et al., 2001; Kar et al., 2005; Thuenemann et al., 2021). The VLP is a triple layer particle composed of 4 different proteins, although recombinant double layer core particles composed of two proteins and single layer sub-core particles composed of only one protein, are also possible. The different particles exhibit a range of physical properties, with the core-like particles the most stable under physiological conditions (Jiménez-Zaragoza et al., 2018; Thuenemann et al., 2013), although the VLP possesses the most effective cell uptake ability (Thuenemann et al., 2021).

Encapsulation of protein cargo is achieved by fusion to each of the 120 copies of the scaffolding protein that makes up the sub-core particle, and fusion to the N-terminus has enabled the encapsulation of 120 GFP molecules in both RV (Charpilienne et al., 2001) and BTV (Brillault et al., 2017).

Although this approach best suits larger VLPs with higher cargo capacity, there are examples of smaller capsids that have been used for cargo encapsulation by direct fusion. Canine parvovirus (CPV) is one such example (Gilbert et al., 2004). Sixty copies of VP2 are sufficient for VLP assembly and GFP fusion to the internal facing N-terminus yielded fluorescent VLPs, although a protein product similar in apparent molecular weight to unfused VP2 was observed. This suggested the formation of mosaic particles resulting from the proteolytic cleavage of the fusion product, which may be necessary for the assembly of VLPs. Another smaller VLP used for encapsulation by translational fusion is Grapevine fanleaf virus (GFLV). Fluorescent protein fusion showed that the N-terminus of GFLV CP is also internal, with recombinant expression in plants leading to red fluorescent protein (RFP) encapsulation in VLPs (Belval et al., 2016). Cargo-loaded VLPs appeared to be comprised of a majority intact CP-cargo fusion. Encapsulation by CP fusion within another smaller VLP, HBV VLPs, has also been demonstrated (Beterams et al., 2000). Two-hundred and forty copies of a 17 kDa
**Staphylococcus aureus** nuclease were encapsulated via fusion to the C-terminus of the CP. CP-cargo fusions remained intact, however, the frequency of aberrant particles was higher than that of VLPs formed by the co-expression of cargo-fused and wildtype CP.

**Post-translational fusions**

**Sortase-mediated**

Surface proteins are anchored to the bacterial cell wall of **S. aureus** via the action of the membrane protein Sortase A (SrtA) (Mazmanian et al., 2001). Proteins harboring the C-terminal sorting signal, an LPXTG motif, are recognized by SrtA, which cleaves between the threonine and glycine within the LPXTG motif, liberating the carboxyl of the threonine, which is subsequently coupled to an N-terminal glycine residue of peptidoglycan cross-bridges. Schoonen et al (2015) developed a method for the site-specific covalent modification of CCMV capsids catalyzed by SrtA (Schoonen et al., 2015). The N-terminus of the CP, located within the interior of the capsid, was engineered to display a glycine residue while GFP was modified with a LPETG-tag at its C-terminus. *In vitro* assembly was conducted following the Sortase-mediated coupling, resulting in VLPs encapsulating 16-18 GFP molecules (Schoonen et al., 2015). The peptide tags required for this SrtA-based strategy are relatively short and therefore modification of capsid and cargo proteins are more likely to be tolerated, making this a potentially broadly applicable strategy for protein cargo encapsulation, and could theoretically be used to mediate *in vivo* encapsulation.

**SpyTag/SpyCatcher-mediated**

The SpyTag/SpyCatcher system is another example of a bacterial-derived protein linkage system from *Streptococcus pyogenes* (Zakeri et al., 2012). The SpyTag (13 amino acids), harbouring a reactive aspartate, and SpyCatcher (12.3 kDa) which contains a reactive lysine, rapidly form a covalent isopeptide bond upon interaction. This system has been used for the *in vivo* encapsulation of two enzymes within bacteriophage MS2 capsids (Giessen and Silver, 2016). The SpyTag was inserted within an internal loop of the MS2 CP, while the SpyCatcher was genetically fused to two cargo proteins. SpyCatcher-tagged enzymes and SpyTag-CP were co-expressed, resulting in the covalent attachment of 2 to 4 molecules of each enzyme within the MS2 VLPs. In contrast to Sortase-mediated cargo capture, SpyCatcher fusions are considerably larger, possibly leading to steric constraints on cargo loading density, however, this method does not require the addition or co-expression of a ligating enzyme, making it relatively simpler than Sortase-mediated encapsulation, and easier to implement for *in vivo* encapsulation.
Table 1. Non-enveloped virus capsids that have been used to encapsulate heterologous proteins.

<table>
<thead>
<tr>
<th>Host kingdom</th>
<th>Family</th>
<th>Virus</th>
<th>Abbr.</th>
<th>Minimal protein composition (no. CP)</th>
<th>Internal diameter</th>
<th>Assembly</th>
<th>Encapsulation strategy</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bacteria</td>
<td>Leviviridae</td>
<td>Enterobacteria phage MS2</td>
<td>MS2</td>
<td>180 CP</td>
<td>21 nm</td>
<td>In vivo or in vitro</td>
<td>DNA aptamer  RNA aptamer Post-translational CP fusion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Enterobacteria phage Obeta</td>
<td>Qβ</td>
<td>180 CP</td>
<td>21.4 nm</td>
<td>In vivo or in vitro</td>
<td>Electrostatic RNA aptamer</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Pseudomonas phage seven</td>
<td>PP7</td>
<td>180 CP</td>
<td>21.6 nm</td>
<td>In vitro</td>
<td>Electrostatic</td>
</tr>
<tr>
<td>Podoviridae</td>
<td></td>
<td>Enterobacteria phage P22</td>
<td>P22</td>
<td>420 CP, 100-330 SP</td>
<td>46.4 nm --&gt; 53.4 nm(^a)</td>
<td>In vivo or in vitro</td>
<td>Native protein tag</td>
</tr>
<tr>
<td>Animalia</td>
<td>Hepadnaviridae</td>
<td>Hepatitis B virus (core antigen)</td>
<td>HBcAg</td>
<td>180 CP, or 240 CP</td>
<td>25.4 nm</td>
<td>In vivo or in vitro</td>
<td>Statistical</td>
</tr>
<tr>
<td></td>
<td>Polyomaviridae</td>
<td>JC Polyomavirus</td>
<td>JCPyV</td>
<td>360 CP</td>
<td>35 nm</td>
<td>In vivo</td>
<td>Native protein tag</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Murine Polyomavirus</td>
<td>MPyV</td>
<td>360 CP</td>
<td>35 nm</td>
<td>In vivo or in vitro</td>
<td>Native and non-native protein tags</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Simian virus 40</td>
<td>SV40</td>
<td>360 CP</td>
<td>35.8 nm</td>
<td>In vivo</td>
<td>Native protein tag</td>
</tr>
<tr>
<td>Reoviridae</td>
<td>Bovavirus</td>
<td>Bluetongue virus</td>
<td>BTV</td>
<td>600 VP7(^b), 120 VP3</td>
<td>46.2 nm</td>
<td>In vivo</td>
<td>Translational CP fusion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Rotavirus</td>
<td>RV</td>
<td>780 VP6, 120 VP2</td>
<td>45.4 nm</td>
<td>In vivo</td>
<td>Translational CP fusion</td>
</tr>
<tr>
<td>Birnaviridae</td>
<td></td>
<td>Infectious bursal disease virus</td>
<td>IBDV</td>
<td>780 CP(^c)</td>
<td>50.4 nm</td>
<td>In vivo</td>
<td>Translational CP fusion</td>
</tr>
<tr>
<td>Parvoviridae</td>
<td></td>
<td>Canine parvovirus</td>
<td>CPV</td>
<td>60 CP</td>
<td>16 nm</td>
<td>In vivo</td>
<td>Translational CP fusion</td>
</tr>
<tr>
<td>Plantae</td>
<td>Bromoviridae</td>
<td>Cowpea chlorotic mottle virus</td>
<td>CCMV</td>
<td>180 CP</td>
<td>18.6 nm</td>
<td>In vivo or in vitro</td>
<td>Statistical  Electrostatic DNA aptamer Non-native protein tag Post-translational CP fusion</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Cucumber mosaic virus</td>
<td>CMV</td>
<td>180 CP</td>
<td>17 nm</td>
<td>In vitro</td>
<td>DNA aptamer</td>
</tr>
<tr>
<td></td>
<td>Secoviridae</td>
<td>Grapevine fanleaf virus</td>
<td>GFLV</td>
<td>60 CP</td>
<td>22 nm</td>
<td>In vivo</td>
<td>Translational CP fusion</td>
</tr>
</tbody>
</table>

\(^a\) P22 VLPs can vary in size with thermal or chemically inducible maturation from the procapsid to expanded and wiffle ball forms (Selivanovitch et al., 2019).

\(^b\) BTV core particles derived from the virion possess a full complement of 780 VP7 proteins making up the T=13 second layer. However, it has been found that recombinant cores lack VP7 trimers at the 5-fold axis of symmetry (Brillault et al., 2017). See Figure 1.

\(^c\) Wild type VP2 assembles into a T=1 capsid whereas the truncated form that enables protein encapsulation assembles into a T=13 capsid (Pascual et al., 2015).
BIOMEDICAL APPLICATIONS

Engineered viruses and VLPs have had clinical success as vaccines and gene delivery vectors. As nanoparticles, both modified viruses and recombinant VLPs are showing preclinical promise as useful scaffolds and delivery vehicles for various inorganic, organic and biomolecular therapeutics (Chung et al., 2020). Size, biocompatibility and engineerability are key features of VLPs that make them attractive as delivery vehicles in biomedical applications. The ability to chemically or genetically modify the exterior of VLPs carrying heterologous cargo proteins has been used to enhance or modify cell uptake, tune cell tropism, or present antigens (Steinmetz et al., 2020). Specific uses for protein loaded VLPs include the delivery of prodrug-converting enzymes, diagnostic imaging, and the encapsulation of T cell epitopes.

Enzyme prodrug therapy is a strategy to increase the specificity and efficacy of a therapeutic by delivering a pro-drug converting enzyme to a specific site. Both CCMV and P22 have been used for the encapsulation of cytochrome P450 BM3 (CYPBM3) for the activation of anticancer drugs (Sanchez-Sanchez et al., 2014; Sanchez-Sanchez et al., 2015; Tapia-Moreno et al., 2017). P22 VLPs loaded with an average of 109 enzymes were transfected into human cervical carcinoma cells using lipofectamine, where they conferred 10-fold higher CYP activity compared to endogenous CYP activity alone (Sanchez-Sanchez et al., 2015). Chemical modification of the exterior of P22 VLPs carrying CYPBM3 to present folic acid increased uptake into a breast cancer cell line and considerably increased their sensitivity to tamoxifen (Tapia-Moreno et al., 2017). In another study, multifunctional P22 VLPs containing CYP were decorated with photosensitizer and targeting moiety on the surface for to allow the combination of enzyme prodrug therapy and photodynamic therapy (Chauhan et al., 2018). BTV VLPs and CLPs have also been developed for this application to take advantage of their natural ability for cell entry (Thuenemann et al., 2021). Here both the triple layer VLP and double layer core-like particles were loaded with Herpes simplex virus 1 thymidine kinase, which is widely used in the conversion of prodrugs to cytotoxic agents. While both particles are capable of efficient cell entry, the VLP is able to mediate endosomal escape and localization in the cytosol, resulting in the death of human glioblastoma cells in the presence of ganciclovir (Thuenemann et al., 2021).

Near-infrared fluorescent proteins (NIR-FPs) are receiving growing attention for their use in non-invasive deep-tissue and whole-body imaging. Encapsulation within VLPs has the added benefits of increasing the stability of the cargo and potential for tunable cell and tissue tropism by exterior modification. To demonstrate the capacity for encapsulated NIR-FPs in in vivo imaging, monomeric mIFP and a red-shifted dimeric iRFP720 variant were separately packaged within Qβ VLPs (Das et al., 2020a). Particles produced similar photochemical properties to
unencapsulated NIR-FP but exhibited increased stabilization towards denaturation and proteolytic digestion. Systemic administration resulted in localization of the particles, and NIR fluorescence, in the liver. Similarly, small-ultrared fluorescent protein (smURFP) was encapsulated within Qβ and PP7 VLPs (Herbert et al., 2020). Encapsulated smURFP showed different tissue and organ localization and remained visible for longer than free smURFP. Although VLPs based on bacteriophages have not yet been shown to have a natural tropism for human cell receptors, they are readily engineered to confer cell binding. Cargo encapsulation within Qβ VLPs via bifunctional RNA was combined with the external display of targeting ligands through CP-fusion, showing that external modification did not adversely affect cargo packaging efficiency (Fiedler et al., 2018). GFP-loaded Qβ VLPs have also been decorated with a CD22 receptor ligand via Cu-catalyzed azide-alkyne cycloaddition and showed strong and selective affinity to CD22+ cells with internalization (Rhee et al., 2011). Together this shows the potential of NIR-FP-loaded VLPs for diagnostic imaging.

VLPs are well established as effective vaccines against both the cognate virus and as presentation platforms for heterologous antigens (Lee et al., 2016; Lua et al., 2014). As a scaffold for the design of vaccines, the repetitive organization of the coat protein is particularly advantageous for external presentation of immunogenic moieties to elicit antibody responses. However, the encapsulation of immunogens can also be effective. Immunization of mice with IBDV VLPs loaded with one or both of the influenza virus antigens HA2 and M2 raised HA-specific or M2-specific antibodies and subsequent mice challenge studies resulted in protection against mortality (Pascual et al., 2015). This approach could serve to protect antigens as well as mediate their delivery to appropriate immune cells. The delivery to professional antigen presenting cells to stimulate humoral immunity is a promising use case for protein encapsulating VLPs. P22 VLPs loaded with the conserved influenza nucleoprotein elicited a robust T cell response that provided broad protection in a multi-strain challenge assay in mice (Patterson et al., 2013b). Similarly respiratory syncytial virus (RSV) M and M2 proteins induced a humoral response that reduced RSV titres following challenge (Schwarz et al., 2016). Combination with external modifications has the potential to enhance or modify the immune response. For example, external presentation of CD40L trimers led to association of FP-loaded P22 VLPs with B lymphocytes (Schwarz et al., 2015).

In addition to infectious disease applications, the encapsulation of immunogens can also be used to stimulate therapeutic immune responses against chronic diseases. Several studies have demonstrated the use of antigen loaded MPyV VLPs for cancer immune therapy in mouse model systems (Eriksson et al., 2011; Tegerstedt et al., 2007; Tegerstedt et al., 2005). Protection against the development of Her2 expressing tumor in mice was shown following
immunization with MPyV VLPs containing the extracellular and transmembrane domains of human Her2 (Tegerstedt et al., 2007). The anti-tumor effect was more efficient when VLPs were loaded into murine dendritic cells (DCs) ex vivo, whereas the unencapsulated antigen, directly administered or ex vivo-loaded on DCs, was unable to protect mice against a lethal dose of D2F2/E2 cells. Similarly, MPyV VLPs loaded with human prostate specific antigen (PSA) were loaded onto DCs and PSA-specific CD4+ and CD8+ cells, which were sufficient to protect mice from PSA-expressing tumor outgrowth (Eriksson et al., 2011). While these examples, clearly demonstrate the utility of VLPs in stimulating the immune system, this inherent property of VLPs remains a barrier to their use as in vivo diagnostics or therapeutic delivery vehicles.

BIOCATALYTIC APPLICATIONS

Over the past decade, VLPs have attracted increasing interest as containers for in vitro and in vivo biocatalysis. Pioneering work on the development of CCMV-based and then P22-based nanoreactors discussed in this review, have demonstrated the potential for fundamental studies of the effect of confinement on enzyme activity kinetics. Despite the fact that the impact of encapsulation on the activity of a given enzyme is not predictable (Esquirol et al., 2022), VLPs have been particularly useful tools for studying the effect of macromolecular crowding on individual enzyme activity due to their structural homogeneity and potential for precise control of the cargo packing density. Control over the number of encapsulated lipase molecules per CCMV VLP showed specific activity decreased at higher packing densities despite higher activity of the encapsulated enzyme compared to free enzyme (Minten et al., 2011). The effect of macromolecular crowding was studied in P22 VLPs by co-encapsulation of alcohol dehydrogenase with varying ratios of wild type SP, which can be selectively removed by treatment with a mild chaotrope (Sharma and Douglas, 2020). The specific activity of alcohol dehydrogenase was inversely related to the concentration of SP, suggesting that high macromolecular crowding could restrict enzyme conformational dynamics. Similarly, expansion of a P22 nanoreactor by heat treatment was found to increase the activity of encapsulated alcohol dehydrogenase, an effect attributed to reduction in macromolecular crowding (Patterson et al., 2012).

Protein cages are often seen to stabilise encapsulated cargo against proteases, high temperatures, and/or chemical denaturants in vitro. This has been demonstrated with VLPs of P22 (Jordan et al., 2016; Wang et al., 2020), MS2 (Giessen and Silver, 2016), and Qβ (Das et al., 2020b; Fiedler et al., 2018). Moreover, in vivo loading into P22 VLPs has been shown to be
a promising strategy to mitigate enzyme misfolding and aggregation during high-level prokaryote expression (Patterson et al., 2013a). Applied to cellular biocatalysis in baker's yeast, *Saccharomyces cerevisiae*, the use of MPyV VLPs to stabilise myoinositol oxygenase (MIOX) led to increased flux through a recombinant metabolic pathway (Cheah et al., 2021).

VLPs can function as a scaffold to spatially organise successive enzymes in a biocatalytic cascade, generating artificial metabolons. The goal is to improve pathway efficiency and minimise intermediate loss to competing reactions, mimicking the function of natural catalytic compartments such as bacterial microcompartments and eukaryotic organelles. Up to three enzymes in a glycolytic cascade have been co-encapsulated in P22 VLPs by fusing them in series with SP (Patterson et al., 2014). In CCMV VLPs, complementary ssDNA tags served to control GOX and GCK stoichiometry or to act at a secondary biocatalyst with GOX when the ssDNA sequence encoded a DNAzyme (Brasch et al., 2017). Post-translational fusion was used to co-encapsulate two enzymes in the indigo biosynthetic pathway within MS2 VLPs *in vivo*. Even with minimal control over loading stoichiometry, indigo bioproduction in *E. coli* increased by ~60% (Giessen and Silver, 2016). Going further to mimic natural compartmentalization, a nested protein cage system has been constructed with P22 VLPs, (Waghwani et al., 2020). Ferritin (Fn) protein cages and the cellobiose-hydrolyzing enzyme (CelB) were co-encapsulated within P22 via SP fusion. This work demonstrated the co-encapsulation of two distinct macromolecules, with Fn-cages themselves serving as subcompartments within P22. These *in vitro* studies and others demonstrate that controlled co-encapsulation of multiple enzymes is central to organising biocatalytic cascades. The ability to achieve this *in vivo*, to organise metabolic pathways with defined stoichiometry in cellular biocatalysis, presents an ongoing challenge in the development of VLP-based biomimetic biocatalysts.

A promising direction for VLP-based nanoreactors is the potential to generate higher-order assemblies such as sheets, gels, and ordered 3D superlattices. Viruses, and VLPs in particular, have long been explored to develop novel nanomaterials due to their natural propensity to form ordered arrays (Sun et al., 2007; Yoshimura et al., 2016). VLPs can be assembled into 3D composites using combinations of covalent and non-covalent interactions mediated by additional proteins, ions, or organic compounds. Examples of hybrid composites include CCMV with photosensitive dendrons (Kostiainen et al., 2010) or avidin (Liljestrom et al., 2014), P22 VLPs with Dec proteins (Uchida et al., 2015) or PAMAM dendrimers (Uchida et al., 2018), and DNA-modified Qβ VLPs (Cigler et al., 2010). For *in vitro* applications, such assemblies could facilitate efficient recycling and recovery of encapsulated enzymes. However, this use case is
dependent on reaching a catalytic advantage that justifies the purification and processing of in vitro biocatalysts. The emergence of more interesting properties in the literature point to this possibility. Higher order assemblies of P22 VLPs have been tuned to massively enhance partitioning of substrates, increasing the catalytic rates of encapsulated alcohol dehydrogenase (Selivanovitch et al., 2021).

OUTLOOK

Recent results from heterologous protein encapsulation in VLPs cover an impressive array of applications, particularly in biomedicine and biocatalysis. On one hand, protein cargo loading has enabled fundamental studies on enzyme activity and confinement. On the other hand, these are synthetic systems developed as biocompatible and biomimicking vehicles with aspirations in biotechnology and nanotechnology. In this regard, significant barriers to clinical or commercial success remain to be addressed. While there is great potential for enhanced in vivo biocatalysis using VLPs as scaffolds for metabolic pathways, challenges remain for the translation of in vitro VLP engineering progress to in vivo settings. These include the implementation of stoichiometrically controlled cargo loading and directed assembly of higher-order structures. As intracellular delivery vehicles, VLPs are capable of long circulation and highly efficient uptake, though the specific pathway of uptake and inherent immunogenicity are important considerations that require more attention. Chemical modification is one approach to addressing some of these concerns. A more elegant approach might be to take advantage of the inherent mutability of virus capsids to improve their physical or chemical properties for a given application. Directed evolution has been shown to generate MS2 capsids with different morphology (Asensio et al., 2016), pH sensitivity (Hartman et al., 2018) as well as permissibility to site-specific modification (Brauer et al., 2019). Viral vectors for use in gene delivery have been evolved into immune-orthologous variants to evade immune system recognition during repeat delivery (Moreno et al., 2019). In an non-viral protein cage, both the capsid protein and the cargo have been co-evolved to improve capsid stability and encapsulation of a nucleic acid cargo (Tetter et al., 2021). These examples provide inspiration for refining protein encapsulating VLPs tailored to specific conditions or applications and point to a general approach applicable to the various platforms discussed in this review.
ACKNOWLEDGEMENTS

This work was carried out on the traditional lands of the Yugarabul, Yuggera, Jagera and Turrbal peoples. F.S acknowledges support from the Commonwealth Scientific and Industrial Research Organisation (CSIRO) in the form of a Synthetic Biology Future Science Platform Fellowship. D.M. was supported by a CSIRO Synbio PhD Top-up Scholarship.

DATA AVAILABILITY STATEMENT

Data sharing not applicable to this article as no datasets were generated or analysed during the current study.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest.

REFERENCES


