Heliyon 8 (2022) e10588

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Identification and characterization of coiled-coil motifs across Autographa californica multiple nucleopolyhedrovirus genome

Jianxiang Qiu [a], Zumei Liu [a], Zhixin Fang [b,**], Wenjiao Wu [a,*]

[a] Medical Research Center, Guangdong Second Provincial General Hospital, Guangzhou 510317, China
[b] Biosafety Laboratory, Guangdong Second Provincial General Hospital, Guangzhou 510317, China

ARTICLE INFO

Keywords:
Baculovirus
AcMNPV
Coiled-coil motif

ABSTRACT

Coiled coils (CCs) are protein structural motifs universally found in proteins and mediate a plethora of biological interactions, and thus their reliable annotation is crucial for studies of protein structure and function. Autographa californica multiple nucleopolyhedrovirus (AcMNPV) is a large double-stranded DNA (dsDNA) virus and encodes 154 proteins. In this study, genome-wide scans of previously uncharacterized CC motifs throughout AcMNPV was conducted using CC prediction software. In total, 24 CC motifs in 19 CC proteins with high confidence were identified. The characteristic of viral CC motifs were analyzed. The CC proteins could be divided into 12 viral structural proteins and 7 non-structural proteins, including viral membrane fusion proteins, enzymes, and transcription factors. Moreover, CC motifs are conserved in the baculoviral orthologs of 14 of the 19 proteins. It is noted that five CC proteins, including Ac51, Ac66, Exon0, Ac13, and GP16, were previously identified to function in the nuclear egress of nucleocapsids, and Ac66 contains multiple CC motifs, the longest of which comprises 252 amino acids, suggesting a role of CC motifs in this process. Taken together, the CC motifs identified in this study are valuable resource for studying protein function and protein interaction networks during virus replication.

1. Introduction

Baculoviruses are enveloped, circular, double-stranded DNA viruses, specifically pathogenic to insects (Rohrmann, 2019b). The baculovirus-insect cell expression system is widely used in exogenous protein expression and vaccine production (Cox, 2012; Mishra, 2020). Baculovirus can also transduce mammalian cells with an efficiency of up to 95% (Sung et al., 2014), and is therefore a prospective gene delivery tool (Ono et al., 2018; Schaly et al., 2021).

The molecular biology of baculoviruses has been widely elucidated. There are two types of morphologically distinct but genetically identical progeny virions in a baculovirus life cycle: budded virions (BVs) and occlusion-derived virions (ODVs). BVs are responsible for virus cell-to-cell spread, while ODVs are responsible for virus spread among insects. Baculoviral nucleocapsids assemble in the virogenic stroma in the nucleus and are then transported across the nuclear membrane, through cytoplasm and finally bud from the plasma membrane to form BVs. In the late phase of infection cycle, nascent nucleocapsids are retained in the nucleus and enveloped by intranuclear microvesicles in the ring zone to form ODVs, which are finally embedded within a proteinaceous crystal matrix to form occlusion bodies (Rohrmann, 2019b). The nuclear egress of nucleocapsids is an important and complex process for BV formation and involves many host proteins, such as cellular actin cytoskeletons (Ohkawa and Welch, 2018), the soluble N-ethylmale-imide-sensitive factor (NSF) (Guo et al., 2017), and the endosomal sorting complexes required for transport (ESCRT) III complex (Yue et al., 2018), and many viral proteins, such as Ac51, Ac66, Exon0, Ac13, Ac11, Ac76, Ac78, GP41, Ac93, Ac142, Ac146, and so on (Fang et al., 2007; Ke et al., 2008; Qiu et al., 2019). How these proteins interact together to promote this process is still not clear enough.

The archetype member of *Baculoviridae* is Autographa californica multiple nucleopolyhedrovirus (AcMNPV), which is also the most extensively studied baculovirus. AcMNPV contains 155 open reading frames and encodes 154 proteins (Rohrmann, 2019a). Although the function of many viral proteins in virus replication has been determined, viral protein interaction networks are far less known.

The coiled coil (CC) is a common structural motif found in proteins throughout eukaryotes, bacteria, and archaea. Generally, CC motifs are mostly found in cytoskeletal motor proteins, transcription factors, the soluble NSF-attachment protein receptors (SNAREs), viral envelop

* Corresponding author.
** Corresponding author.
E-mail addresses: fangzhx@gd2h.org.cn (Z. Fang), wj910103@126.com (W. Wu).

https://doi.org/10.1016/j.heliyon.2022.e10588
Received 25 January 2022; Received in revised form 15 May 2022; Accepted 5 September 2022

proteins, and so on (Burkhard et al., 2001). CC motifs are structurally conserved but can vary in sequence and length. The usual length is 20 or so amino acid residues, but they can be shorter or can span many hundreds of amino acids (Truebestein and Leonard, 2016). Typically, it consists of two or more amphipathic α-helices twisting around each other into a left-handed helix to form a supercoil. Sequences of parallel left-handed CCs are characterized by a seven-residue periodicity (heptad repeat), whose positions are denoted as a-b-c-d-e-f-g. The apolar residues are preferentially positioned in the first (a) and fourth (d) positions of the 'heptad', and a and d positions are found at the interface of the two helices (Burkhard et al., 2001; Mason and Arndt, 2004). The heptad repeat is the basis for the canonical CC structures (Lupas and Bassler, 2017; Lupas et al., 2017). Meanwhile, there are repeats longer than seven residues, such as 11-residue (hendecad repeats) and 15-residue (pentadecad repeats) repeats, which can be described as combinations of three- and four-residue segments and are the basis for non-canonical CC structures (Lupas and Bassler, 2017; Lupas et al., 2017).

In order to identify CC structures, several computational programs have been developed to predict CC regions, the likelihood that helices form CC motifs, and the oligomeric state of CC motifs. These programs include COILS, Paircoil, Multicoil, MARCOIL, Paircoil2, and Deepcoil (Berger et al., 1995; Delorenzi and Speed, 2002; Ludwiczak et al., 2019; Lupas et al., 1991; McDonnell et al., 2006; Wolf et al., 1997). Paircoil2 is an improved version of Paircoil, and it is demonstrated to outperform other common methods (McDonnell et al., 2006). Deepcoil is a recently developed prediction tool of canonical and non-canonical CC motifs with high sensitivity, outperforms other methods, such as COILS, Multicoil, and MARCOIL, and is thought to be a method of choice for accurate genome-wide detection of CC domains (Ludwiczak et al., 2019). Such programs can enable the recognition of CC motifs from protein sequences.

CC motifs have been found in some viral proteins, especially viral envelope proteins, where they play crucial roles. LearnCoil-VFM is a program developed by a previous study which detects CC-like regions in many viral membrane fusion proteins in many retroviruses, paramyxoviruses, filoviruses, and also baculovirus (Singh et al., 1999). Generally, enveloped viruses such as influenza virus, Ebola virus, and human immunodeficiency virus (HIV) depend on the fusion of their membranes with cellular membranes to import their nucleocapsids into cells at the beginning of the infection cycle. The fusion process is mediated by viral surface glycoproteins that contain CC motifs (Burkhard et al., 2001; Mo et al., 2004; Sauter et al., 1992; Watanabe et al., 2000), and combined with the structural studies of membrane fusion process, it is suggested that diverse enveloped DNA and RNA viruses share similar mechanisms in membrane fusion (Singh et al., 1999). So far, some other viral proteins have also been identified to contain CC motifs. For example, Ebola virus VP35 protein contains a CC motif, which mediates VP35 homo-oligomerization and is indispensable for binding to the viral polymerase (Ramaswamy et al., 2018; Reid et al., 2005). A similar motif is found in the protein in the closely related Marburg virus (Moller et al., 2005). Moreover, crystal structure analysis reveals that the CC motif is crucial for VP35 oligomerization (Bruhn et al., 2017; Zinzula et al., 2019). In addition, the CC motif in the spike protein of Rotavirus VP4 cooperates with the actin binding domain for actin remodeling (Condemine et al., 2019). These studies indicate that CC motifs are also common in viruses.

So far, there are some studies concerning about the CC motif analysis in the genome-wide scale across eukaryotes, bacteria, and archaea (Liu and Rost, 2001; Rackham et al., 2010; Wang et al., 2012). A study analyzed 29 proteomes for representatives from eukaryotes, prokaryotes, and archaebacteria, and revealed that there are twice as many CC proteins in eukaryotes (10%) as in prokaryotes and archaes (4%–5%) (Liu and Rost, 2001). However, characterization of CC motifs across large viral genome has not been reported.

In this study, we conducted a survey of CC proteins across the AcMNPV genome. In total, 19 CC-containing proteins were identified with high confidence, accounting for approximately 12.3% of all proteins in AcMNPV. These proteins include viral membrane fusion proteins (GP64 and F protein), transcription factors (PE38 and IE2), enzymes (chitinase and HE65), other viral structural proteins (BRO, BV/ODV-E26, Ac51, Ac66, Ac73, Exon0, P24, CG30, P10, and Pp34), and non-structural ones (Ac13, Ac29, Ac47, and GP16). Interestingly, among these proteins, Exon0, Ac66, Ac51, Ac13, and GP16 are identified to function in the nuclear egress of nucleocapsids, suggesting a role of CC motifs in this process. Further studies on these proteins will reveal the function of viral CC motifs and protein-protein interaction networks mediated by this motif.

2. Materials and methods

2.1. Sequences data

The proteome sequence of AcMNPV was downloaded from National Center for Biotechnology Information (NCBI) (Reference Sequence: NC_001623.1).

2.2. Prediction, identification and selection of CC proteins

Among several CC prediction programs mentioned in the introduction, Paircoil2 (http://cb.csail.mit.edu/cb/paircoil2/paircoil2.html) and Deepcoil (https://toolkit.tuebingen.mpg.de/#/tools/deepcoil) were the latest developed with good performance and were used in this study to predict CC motifs. The default settings of the p-score of Paircoil2 and the probability score of Deepcoil were 0.025 and 0.5, respectively. In our study, all of the 154 AcMNPV protein sequences were run through Paircoil2 using a cutoff p-score of 0.06 (the lower the p-score is, the more likelihood a CC motif will form) and window size of 21 residues. At the same time, protein sequences were also run through Deepcoil using a cutoff probability score of 0.7 (the higher the score is, the more likelihood a CC motif will form). As the maximum allowed sequence length of Deepcoil is 500 aa, those proteins longer than 500 aa were cut apart into two parts and submitted to Deepcoil. Deepcoil2 can predict the probability of "a" and "d" residues in protein sequences and is used in this study to further confirm the prediction results from Paircoil2 and Deepcoil. CC motifs were defined in this study when the protein sequences reached the threshold of both Paircoil2 and Deepcoil and were confirmed by Deepcoil2. The oligomerization of CC proteins were predicted with Mulitcoil (http://cb.csail.mit.edu/cb/multicoil/cgi-bin/multicoil.cgi) (a good tool to predicted the oligomeric state of CC proteins) with window size of 21 residues and a cutoff probability score of 0.5.

2.3. Clustering of the predicted CC proteins

To analyze the possible correlation among these proteins, protein clustering was conducted. The classification of structural or non-structural proteins was determined according to the mass spectrometry of purified baculoviral virions in previous studies (Braunagel et al., 2003; Hou et al., 2013; Wang et al., 2010). The protein function and their roles in virus replication were determined according *Baculovirus Molecular Biology* (Rohrmann, 2019b). Meanwhile, function of some proteins were confirmed according to the latest reports.

2.4. Identification of long CC motifs in AcMNPV proteins

According to a previous study concerning genome-wide identification of CC proteins, small gaps (less than 25 amino acids (aa)) between CC motifs were ignored and CC motifs with length more than 70 aa were defined as long CC motifs (Rose et al., 2007). In the present study, CC motifs were defined as long CCs with more than 70 aa in length and ignoring small gaps (less than 10 aa). Conserved domains of long CC-containing proteins were then predicted by the Conserved Domain Search Service tool in NCBI to predict possible function.

3. Results

3.1. Characterization of predicted CC proteins and motifs in AcMNPV

Initially, a total of 12 (7.8%) and 21 (13.6%) viral proteins were predicted to contain at least one CC motif by Paircoil2 (with a p-score lower than the cutoff score of 0.03) and by Deepcoil (with a probability score higher than the cutoff score of 0.7), respectively (Figure 1A). However, as Ac51 and Exon0 were reported to contain CC motifs previously (Biswas et al., 2018; Qiu et al., 2019), their scores predicted by Paircoil2 were 0.0505 and 0.0589, respectively. Therefore, for a high sensitivity of CC protein prediction, the cutoff score of Paircoil2 was adjusted to 0.06. The adjusted prediction identified 19 proteins that contain CC motifs, which all have scores of more than 0.7 by Deepcoil (Table 1). These proteins were present in AcMNPV genome with a rate of 12.3% (Figure 1A), which is comparable to the average rate of 10% in eukaryotic proteome. These CC proteins are BRO, Ac13, BV/ODV-E26, F protein, Ac29, Ac47, Ac51, Ac66, CG30, HE65, Chitinase, GP64, P24, GP16, Pp34, P10, Exon0, IE2, and PE38 (Table 1). Notably, when using a strict cutoff score of 0.025 in Paircoil2, 9 proteins were identified (Ac13, BV/ODV-E26, F protein, Ac29, Ac66, CG30, P10, IE2, and PE38) and the rate was 5.8%. Therefore, these 9 proteins have higher probability to be true CC proteins. Furthermore, the 24 CC motifs were also submitted to Alphafold to predict their ab initio 3D structures, and the results showed that almost all CC motifs were predicted to form α-helix with high possibility (as shown by the value of pLDDT) (Figure S1 in supplementary data), which also suggests the accurate prediction of CC motifs. In addition, the oligomerization state of the 19 proteins were predicted using Multicoil with a cutoff probability score of 0.5, and five proteins (Ac29, Ac66, CG80, IE2, and PE38) were predicted to form dimer with high possibility. The detailed information of the prediction results are listed in Table 1.

Among the 19 proteins, 16 of them (84.2%) contained only one CC motif, while 2 contained two CC motifs (F protein and PE38) and 1 contained four CC motifs (Ac66) (Figure 1B). In total, 24 CC motifs were identified. The relative length (number of amino acids) of a CC motif within its parental protein (where the motif is derived from) were mostly under 30% (Figure 1C), suggesting that most CC proteins may contain other domains which function together with CC motifs. 84.2% of the predicted CC motifs (16) were 30–50 aa in length (Figure 1D). There were 7 CC motifs with length of more than 70 aa (Figure 1D), which are distributed in BRO, Ac13, Ac66, CG30, IE2, and PE38.

3.2. Clustering of viral CC proteins

To investigate the correlation among these CC proteins, protein clustering was conducted according to their annotated function in a virus life cycle. The 19 proteins identified in the prediction could be classified into 12 viral structural proteins and 7 non-structural proteins (Figure 2A). Proteins in different clusters were displayed in Figure 2B. Among the 19 proteins, 6 proteins play critical roles in virus replication and essential for high virus titer (including Ac51, Ac66, Exon0, GP64, F protein, and PE38) while others are not crucial for virus production in host cells (Figure 2A, B). On the other hand, further analysis showed that some of the predicted CC motifs exhibit conserved distribution in AcMNPV, including transcription factors (IE2 and PE38), enzymes (chitinase and HE65), and viral membrane fusion proteins (GP64 and F protein) (Figure 2C). Except for the six proteins listed above, other proteins were classified as viral structural proteins, including BRO, BV/ODV-E26, Ac51, Ac66, Exon0, Ac73, P24, CG30, P10, and Pp34, and non-structural proteins, including Ac13, Ac29, Ac47, and GP16 (Figure 2C).

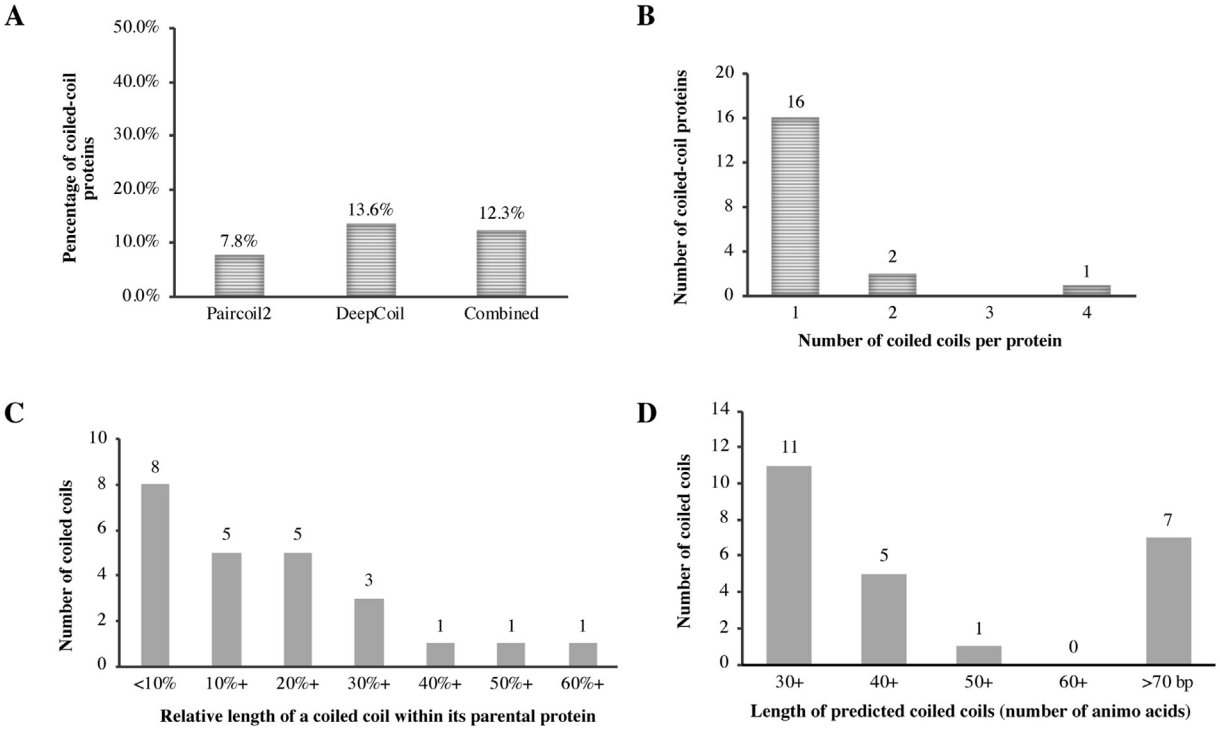

Figure 1. Characterization of predicted CC motifs in AcMNPV. (A) Percentage of CC proteins in AcMNPV predicted by Paircoil2 and Deepcoil. 12 (7.8%) and 21 (13.6%) proteins were predicted containing at least one CC motif by Paircoil2 (with a p-score lower than 0.03) and by Deepcoil (with a probability score higher than 0.7), respectively. After adjusting the cutoff score of Paircoil2 to 0.06, 19 proteins (12.3%) were predicted, which were all included in the 21 proteins predicted by Deepcoil. (B) Number of CC motifs in each of predicted CC proteins. The number of proteins containing 1, 2, 3, or 4 CC motifs are indicated. (C) Relative length (number of amino acids) of a CC motif within its parental proteins (where a CC motif is derived from). The number of CC motifs with different relative length to its parental proteins are indicated. (D) Distribution of the length of predicted CC motifs. The number of CC motifs with different length are indicated.

Table 1. Summary of predicted coiled-coil motifs in AcMNPV

No.	Protein ID	Gene name	ORF No.	CC region (p-score by Paircoil2)	CC prob. by DeepCoil	Oligomerization
1	NP_054031.1	BRO	Orf2[a]	133–173 (0.0468); 184–242 (0.0259)	0.8	–[d]
2	NP_054042.1		Orf13[ac]	142–178 (0.0184); 179–213 (0.0310)	0.8	–
3	NP_054045.1	BV/ODV-E26	Orf16[c]	70–114 (0.0083)	0.9	–
4	NP_054052.1	F protein	Orf23[c]	75–110 (0.0459); 219–268 (0.0216)	0.8	–
5	NP_054058.1		Orf29	1–47 (0.0127)	0.8	Dimer
6	NP_054076.1		Orf47	13–50 (0.0289)	0.8	–
7	NP_054080.1		Orf51[c]	289–318 (0.0505)	0.8	–
8	NP_054096.1		Orf66[b]	14–46 (0.0546); 301–552 (0.0013–0.0319); 589–625 (0.0162); 638–785 (0.0057–0.0529)	0.9	Dimer
9	NP_054118.1	CG30	Orf88	116–248 (0.0043)	0.9	Dimer
10	NP_054134.1	HE65	Orf105	167–199 (0.0494)	0.8	–
11	NP_054156.1	Chitinase	Orf126	315–359 (0.0320)	0.9	–
12	NP_054158.1	GP64	Orf128[c]	307–348 (0.0320)	0.9	–
13	NP_054159.1	P24	Orf129	157–186 (0.0520)	0.7	–
14	NP_054160.1	GP16	Orf130	24–53 (0.0520)	0.7	–
15	NP_054161.1	Pp34	Orf131	176–209 (0.0298)	0.9	–
16	NP_054167.1	P10	Orf137	5–41 (0.0167)	0.8	–
17	NP_054172.1	Exon0	Orf141[c]	169–198 (0.0589)	0.7	–
18	NP_054182.1	IE2	Orf151[ac]	295–377 (0.0037); 378–408 (0.0056)	0.9	Dimer
19	NP_054184.1	PE38	Orf153[a]	18–48 (0.016); 49–66 (0.0371); 185–254 (0.0013)	0.9	Dimer

CC, coiled-coil motifs. prob., probability.
[a] two coiled-coil motifs with small gaps (1 aa) are considered as one coiled-coil motif by ignoring the gap in later analysis.
[b] the regions of 301–552 and 638–785 in Ac66 consist of several short coiled-coil motifs with gaps of 1 aa and the range of p-score is indicated.
[c] these proteins are reported to contain coiled-coil motifs previously.
[d] means that no dimer or trimer is predicted by Multicoil.

3.3. Long CC motifs in AcMNPV genome

The lengths of CC motifs are highly variable. CC motifs with lengths longer than 70 aa were defined, by ignoring small gaps, as long CCs according to a previous report (Rose et al., 2007). Among the CC motifs predicted in AcMNPV genome, 7 long CCs were identified, which are 110, 133, 114, 70, 252 and 197 aa in length and distributed in BRO, CG30, Ac66, IE2, PE38, and Ac66, respectively (Figure 3). Notably, Ac66 contains two long CC motifs with length of 252 aa and 197 aa. Other conserved domains in these six proteins were predicted by Conserved Domain Search Service tool in NCBI. These conserved domains are BRO-N superfamily domain in BRO, DUF3627 superfamily domain in Ac13, Desmo N and ATPase domains in Ac66, Zinc Finger and Myosin_tail_1 superfamily domains in CG30, Zinc Finger domain in PE38, and Ring and SMC superfamily domains in IE2 (Figure 3). The location of CC motifs and conserved domains in each protein were displayed and some of them were overlapping (Figure 3). The secondary structures of the six proteins were also predicted using PSIPRED (http://bioinf.cs.ucl.ac.uk/psipred) and the prediction results were shown in Figure S2 in supplementary data. These results indicated that AcMNPV genome also encodes for long CC proteins with a rate of nearly 3.9%, and these long CC motifs is supposed to function together with other domains to support protein function.

3.4. Conservation of CC motifs in baculoviruses

There is a high occurrence of a heptad repeat pattern in a protein sequence by chance (Singh et al., 1999), making the conservation analysis of a CC motif significant. To further analyze the importance of CC motifs, the conservation of this motif in the orthologs of each protein was analyzed using the ProBLAST/PSI-BLAST tool in MPI Bioinformatics Toolkit (https://toolkit.tuebingen.mpg.de/tools/psiblast). The results indicated that CC motifs could be predicted in the orthologs of 14 of the 19 proteins (74%), including Ac13, BV/ODV-E26, Ac47, HE65, Ac51, Ac66, CG30, GP64, F protein, P24, Pp34, chitinase, PE38, and IE2, indicating a conserved and possibly important function of CC motifs in these proteins. CC motifs seemed not to be conserved in the orthologs of BRO, Ac29, GP16, P10, and Exon0 (Figure 4).

4. Discussion

CC motifs are highly abundant in all organisms, and usually mediate crucial protein–protein interactions in the cell. In this study, we present a comprehensive survey of CC proteins in AcMNPV by combining the results of several CC prediction software, which revealed distribution, characteristic, conservation, and versatile roles of CC motifs in baculovirus.

In total, 19 proteins were identified to contain CC motifs. Among them, some (Ac13, BV/ODV-E26, IE2, GP64, F protein, Exon0, and Ac51) are reported to contain CC motifs previously (Biswas et al., 2018; Qiu et al., 2019; Rohrmann, 2019a; Singh et al., 1999). However, FP25K, which was also predicted to contain a CC motif (Garretson et al., 2016), was not identified in this study, as its scores in two prediction software could not reach the cutoff score. 12.3% of AcMNPV proteins were predicted to contain CC structures (Figure 1A), comparable with the average 10% in eukaryotes (Liu and Rost, 2001), instead of 4–5% in prokaryotes

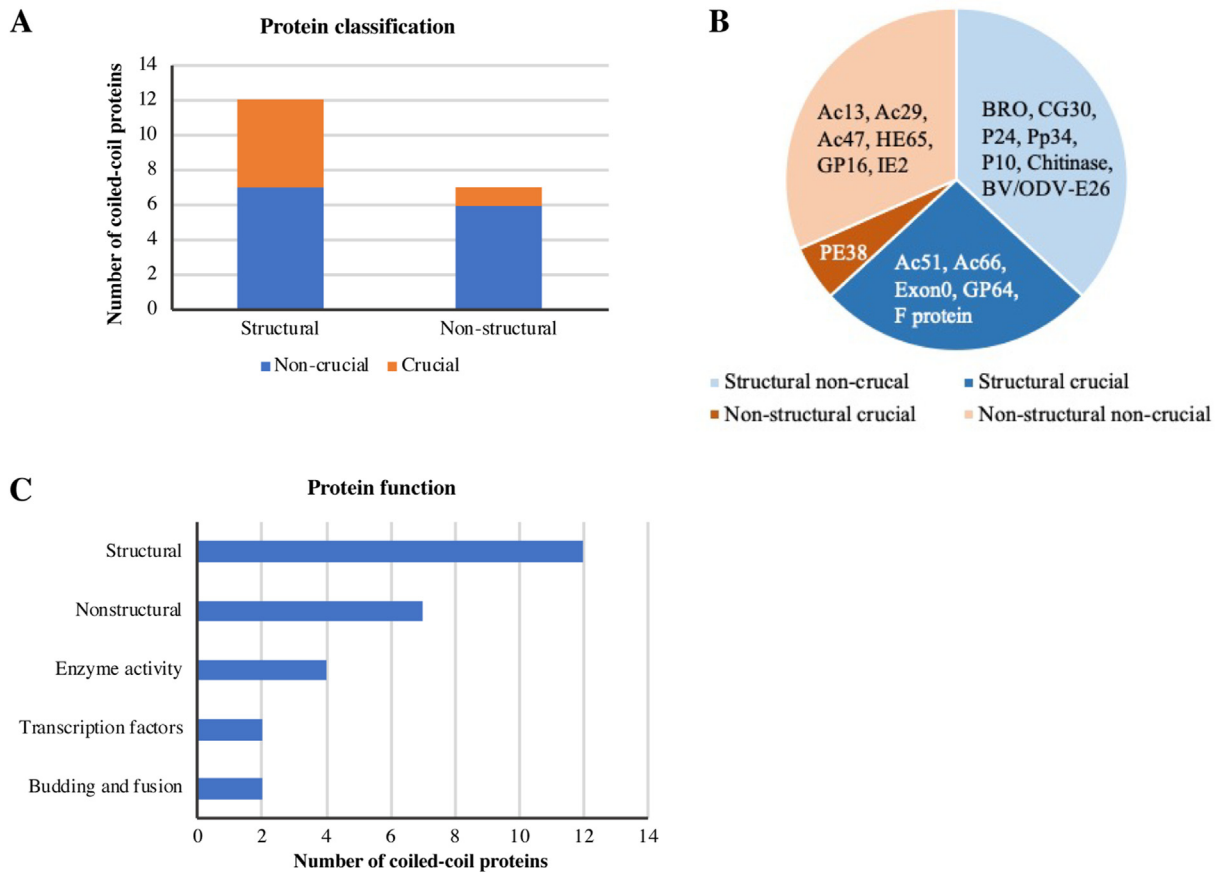

Figure 2. Clustering of viral CC proteins. (A) Viral CC proteins were divided into structural and non-structural proteins. These proteins were further classifies as crucial and non-crucial proteins. Crucial proteins (indicated in orange) are essential for efficient virus replication and high virus titer in host cells whereas non-crucial proteins are not. (B) Proteins indicated in (A) are displayed. Structural proteins are in dark and light blue and non-structural proteins are in dark and light orange. (C) CC proteins were classified into five clusters according to their function during virus replication.

and archaea (Liu and Rost, 2001), suggesting that large eukaryotic DNA viruses may be more similar to eukaryotes in the evolution of CC structures. Conservation of CC structures was not predicted in the orthologs of 5 proteins (BRO, Ac29, GP16, P10, and Exon0) (Figure 4). However, the CC motifs might still play important roles in these five proteins, as the motif in Exon0 was demonstrated to be crucial for its dimerization and virus replication (Biswas et al., 2018; Fang et al., 2008).

4.1. Conserved roles of CC motifs in AcMNPV

CC motifs generally play key roles in directing and controlling protein–protein associations involved in cytoskeleton, transcription, intracellular trafficking, viral infection, dynamic assembly of protein complexes, and many other events (Ford and Fioriti, 2020; Hartmann, 2017; Truebestein and Leonard, 2016; Wang et al., 2012). Interestingly, this property is conserved in baculovirus. Notably, both baculovirus membrane fusion proteins GP64 and F protein contain CC motifs, consistent with the prediction results of a previous study (Singh et al., 1999), which has been mentioned in the introduction. Besides, two transcription factors PE38 and IE2 in AcMNPV were predicted to contain long CC motifs with length of 70 and 114 aa, respectively. Transcription of DNA involves many specific interactions between DNA and DNA-binding proteins. Many transcription factors contain CC motifs that are responsible for specific recognition between molecules (Burkhard et al., 2001). Both PE38 and IE2 are viral immediate-early genes. PE38 is crucial for viral DNA replication and virus replication (Milks et al., 2003), while IE2 deletion delayed infection in Sf21 cells but showed no detrimental effect in Tn-5B1-4 cells (Prikhod'ko et al., 1999). In addition, Bombyx mori nucleopolyhedrovirus IE2 interacts with itself through the CC motif (Imai et al., 2000). The function of CC motif of PE38 remains to be further investigated.

Two viral enzymes were also predicted to contain CC motifs, including chitinase and HE65. It is supposed that CC motifs in enzymes function as molecular rulers, positioning catalytic activities at fixed distances (Truebestein and Leonard, 2016). Chitinase is essential for liquification of insects by degrading chitin in the skeleton and is not required for BV production (Rohrmann, 2019a). HE65 is a putative RNA editing ligase, and a study showed that deletion of HE65 did not affect BV production (Gandhi et al., 2012). The function of CC motifs of chitinase and HE65 in virus replication also remains to be investigated.

4.2. CC motifs in viral structural proteins

Other than proteins described above, the remaining CC-containing proteins can be divided into two categories, namely viral structural proteins (BRO, BV/ODV-E26, Ac51, Ac66, Exon0, Ac73, P24, CG30, P10, and Pp34) and non-structural ones (Ac13, Ac29, Ac47, and GP16) (Figure 2C). The function of these proteins has not been fully determined. Among the structural proteins, Exon0, Ac66, and Ac51 are crucial for virus replication and BV production. They are supposed to function in the pathway of nuclear egress of nucleocapsids to cytoplasmic membrane in the formation of mature BVs (Fang et al., 2007; Ke et al., 2008; Qiu et al., 2019). The CC region of Exon0 has been demonstrated to be crucial for protein dimerization and virus replication (Fang et al., 2008), and Exon0 interacts with Ac66 (Biswas et al., 2018). The function of CC motifs of Ac66 and Ac51 are remained to be determined. It is proposed that CC motifs may play a role to mediate protein interactions during the anterograde transport of nucleocapsids.

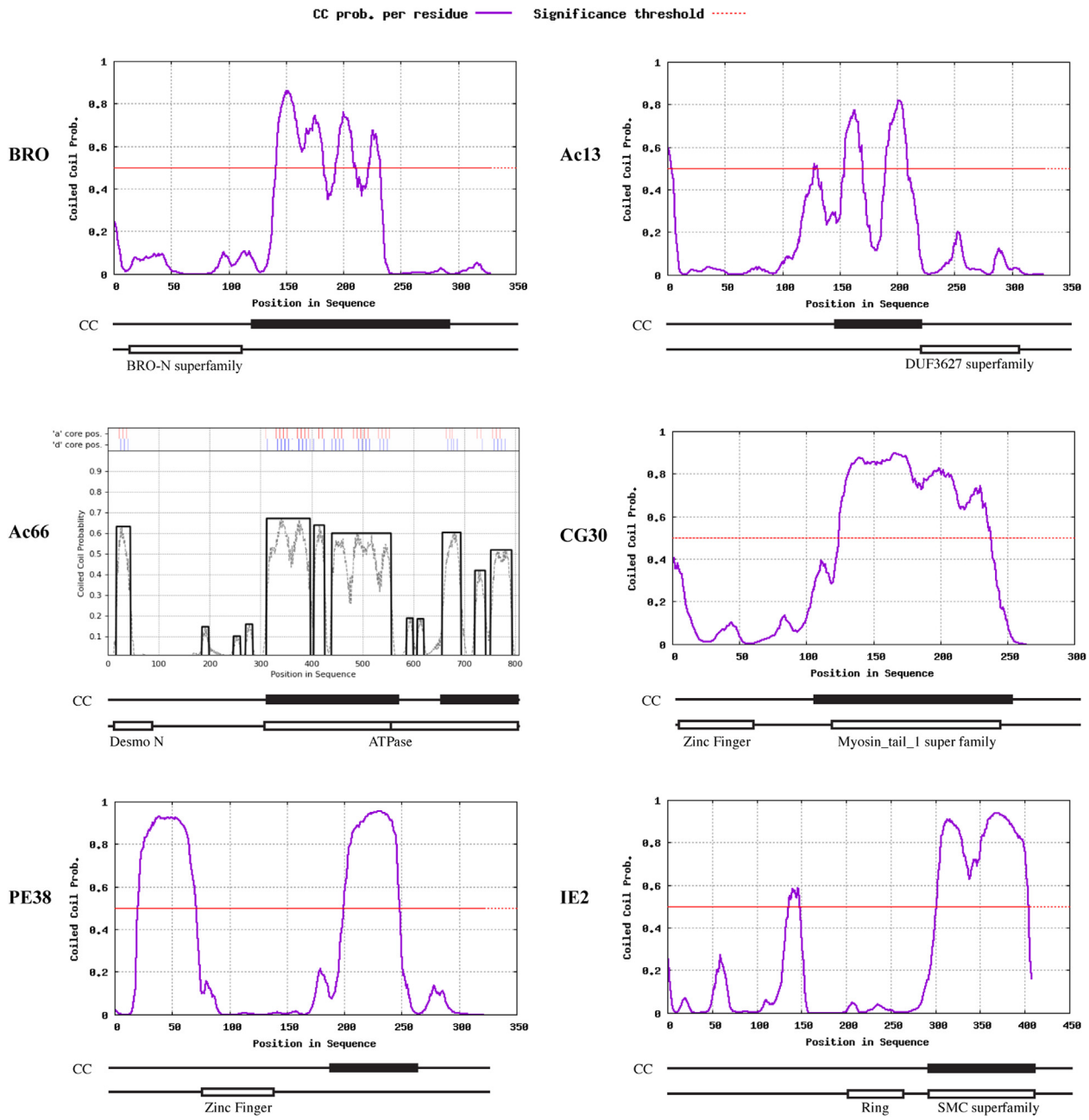

Figure 3. Distribution of long CC motifs in six viral proteins. Pictures of BRO, Ac13, CG30, PE38, and IE2 were the prediction results from Deepcoil. As the maximum allowed sequence length of Deepcoil is 500 aa, picture of Ac66 was the prediction result from Deepcoil2. The location of long CCs and other conserved domains in each protein was indicated.

Functional characterization of BV/ODV-E26, P24, Ac73, BRO, and CG30 revealed that deletion of any of these genes showed no striking effect on phenotype in cell lines (Rohrmann, 2019a; Shao et al., 2019), and function of their CC motifs in virus replication remains to be determined. Pp34 is the polyhedron envelope and is associated with P10 fibrillar structures. Deletion of *pp34* or *p10* gene both resulted in polyhedra with rough surface showing cavities where virions had apparently been dislodged (Rohrmann, 2019a). CC motifs may mediate the association between Pp34 and P10.

4.3. CC motifs in baculovirus non-structural proteins

Among the non-structural proteins, deletion of Ac13 resulted in lower titers of BV (Chen et al., 2021), while deletion of Ac29, Ac47, or GP16 (Ac130) had no striking effect on virus replication and BV production in cell lines (Rohrmann, 2019a; Yang et al., 2014). Ac130 localizes near the nuclear membrane in the cytoplasm, and is supposed to function in the envelopment/de-envelopment of BVs when they cross over the nuclear membrane and pass through cytoplasm (Yang et al., 2014). Ac13 is present in both the inner- and outer nuclear membranes, and a recent study indicates that it is required for efficient nuclear egress of nucleocapsids during virion budding (Chen et al., 2021). Ac29 showed over 80% probability of being related to the amphiphysin BAR domain from *Drosophila* spp., which are sensors of membrane curvature (Rohrmann, 2019a). It may be worthwhile to investigate whether CC motifs in Ac13, Ac29, and GP16 play roles in the egress of nucleocapsids across nuclear membrane.

4.4. Function of CC motifs in the nuclear egress of baculoviral nucleocapsids

The nuclear egress of viral nucleocapsids is a complex process involving many cellular and viral proteins. In the present study, several

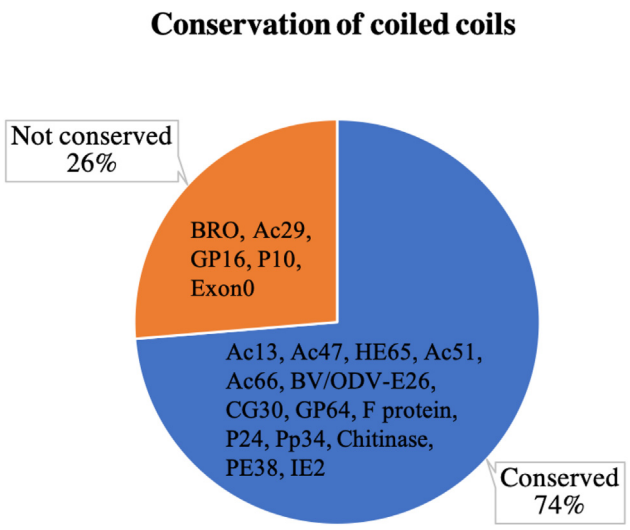

Figure 4. Conservation of CC motifs in the orthologs of each protein. The existence of CC motifs in the orthologs of the 19 proteins were predicted by ProBLAST/PSI-BLAST tool in MPI Bioinformatics Toolkit. CC motifs are conserved in the orthologs of 74% of the predicted proteins, but not in those of 26% of the proteins.

viral proteins identified to function in the nuclear egress of nucleocapsids were predicted to contain CC motifs, including Ac51, Ac66, Exon0, Ac13, and Ac130. Meanwhile, studies indicate that the cellular SNARE system and ESCRT III complex are involved in AcMNPV infection and are both required for the nuclear egress of progeny nucleocapsids of baculoviruses (Guo et al., 2017; Yue et al., 2018). SNARE system triggers fusion of transport vesicles with their target membranes while ESCRT III functions in the membrane scission at the bud neck. SNARE proteins are known to use CC motifs to promote function and form a core complex comprising a heterotetrameric CC (Burkhard et al., 2001; Parry et al., 2008). Moreover, some components of ESCRT III complex are predicted to contain CC motifs (Shim et al., 2006; Sweeney et al., 2006). In addition, some baculoviral proteins are demonstrated to interact with components of SNARE system and ESCRT complex (Guo et al., 2017; Yue et al., 2018). Whether Ac51, Ac66, Exon0, Ac13, and Ac130 are associated with SNARE system and ESCRT complex during the nuclear egress of nucleocapsids remains to be further investigated. A study identified genome-wide CC interaction (CCI) network in Saccharomyces cerevisiae by combining the yeast two-hybrid assay results and computational analysis and revealed that CCI network is specifically and functionally organized; the study showed that CCI played a critical role in the assembly of the kinetochore, and disruption of the CCI network led to defects in kinetochore assembly and cell division (Wang et al., 2012). Collectively, it is speculated that CC motifs may play important roles in the nuclear egress of baculoviral nucleocapsids by mediating the interaction between SNARE system/ESCRT III complex and other viral proteins.

5. Conclusion

The present study conducted a survey of CC motifs in the AcMNPV genome and identified 24 CC motifs and 19 CC-containing proteins with high confidence. These proteins include viral structural proteins, nonstructural proteins, membrane fusion proteins, transcription factors, and enzymes. The CC motifs in membrane fusion proteins highlight a conserved membrane fusion mechanism among enveloped DNA and RNA viruses. An interesting finding is that several CC proteins are involved in the nuclear egress of baculoviral nucleocapsids, suggesting the important function of CC motifs in this process and possible interaction among these proteins. Further functional studies of CC proteins will supplement our knowledge to baculovirus and CC motifs.

Declarations

Author contribution statement

Jianxiang Qiu: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Wenjiao Wu: Conceived and designed the experiments; Wrote the paper.

Zhixin Fang: Performed the experiments; Analyzed and interpreted the data; Wrote the paper.

Zumei Liu: Analyzed and interpreted the data; Wrote the paper.

Funding statement

Jianxiang Qiu was supported by National Natural Science Foundation of China [32000110], Doctoral workstation foundation of Guangdong Second Provincial General Hospital [2020BSGZ035].

Zhixin Fang was supported by National Natural Science Foundation of China [81902064], Science and Technology Planning Project of Guangdong Province of China [2021B1212030008].

Dr Wenjiao Wu was supported by Doctoral workstation foundation of Guangdong Second Provincial General Hospital [2019BSGZ012].

Data availability statement

Data will be made available on request.

Declaration of interests statement

The authors declare no conflict of interest.

Additional information

Supplementary content related to this article has been published online at https://doi.org/10.1016/j.heliyon.2022.e10588.

References

Biswas, S., Willis, L.G., Fang, M., Nie, Y., Theilmann, D.A., 2018. Autographa californica nucleopolyhedrovirus AC141 (Exon0), a potential E3 Ubiquitin ligase, interacts with viral Ubiquitin and AC66 to facilitate nucleocapsid egress. J. Virol. 92 (3).

Berger, B., Wilson, D.B., Wolf, E., Tonchev, T., Milla, M., Kim, P.S., 1995. Predicting coiled coils by use of pairwise residue correlations. Proc. Natl. Acad. Sci. U. S. A. 92 (18), 8259–8263.

Braunagel, S.C., Russell, W.K., Rosas-Acosta, G., Russell, D.H., Summers, M.D., 2003. Determination of the protein composition of the occlusion-derived virus of Autographa californica nucleopolyhedrovirus. Proc. Natl. Acad. Sci. U. S. A. 100 (17), 9797–9802.

Bruhn, J.F., Kirchdoerfer, R.N., Urata, S.M., Li, S., Tickle, I.J., Bricogne, G., Saphire, E.O., 2017. Crystal structure of the Marburg virus VP35 oligomerization domain. J. Virol. 91 (2).

Burkhard, P., Stetefeld, J., Strelkov, S.V., 2001. Coiled coils: a highly versatile protein folding motif. Trends Cell Biol. 11 (2), 82–88.

Chen, X., Yang, X., Lei, C., Qin, F., Sun, X., Hu, J., 2021. Autographa californica multiple nucleopolyhedrovirus orf13 is required for efficient nuclear egress of nucleocapsids. Virol. Sin. 36 (5), 968–980.

Condemine, W., Eguether, T., Courousse, N., Etchebest, C., Gardet, A., Trugnan, G., Chwetzoff, S., 2019. The C terminus of Rotavirus VP4 protein contains an actin binding domain which requires cooperation with the coiled-coil domain for actin remodeling. J. Virol. 93 (1).

Cox, M.M., 2012. Recombinant protein vaccines produced in insect cells. Vaccine 30 (10), 1759–1766.

Delorenzi, M., Speed, T., 2002. An HMM model for coiled-coil domains and a comparison with PSSM-based predictions. Bioinformatics 18 (4), 617–625.

Fang, M., Dai, X., Theilmann, D.A., 2007. Autographa californica multiple nucleopolyhedrovirus EXON0 (ORF141) is required for efficient egress of nucleocapsids from the nucleus. J. Virol. 81 (18), 9859–9869.

Fang, M., Nie, Y., Dai, X., Theilmann, D.A., 2008. Identification of AcMNPV EXON0 (ac141) domains required for efficient production of budded virus, dimerization and association with BV/ODV-C42 and FP25. Virology 375 (1), 265–276.

Ford, L.K., Fioriti, L., 2020. Coiled-coil motifs of RNA-binding proteins: dynamicity in RNA regulation. Front. Cell Dev. Biol. 8.

Gandhi, K.M., Ohkawa, T., Welch, M.D., Volkman, L.E., 2012. Nuclear localization of actin requires AC102 in Autographa californica multiple nucleopolyhedrovirus-infected cells. J. Gen. Virol. 93 (Pt 8), 1795–1803.
Garretson, T.A., McCoy, J.C., Cheng, X.-W., McFadden, G., 2016. Baculovirus FP25K localization: role of the coiled-coil domain. J. Virol. 90 (21), 9582–9597.
Guo, Y., Yue, Q., Gao, J., Wang, Z., Chen, Y.R., Blissard, G.W., Liu, T.X., Li, Z., 2017. Roles of cellular NSF protein in entry and nuclear egress of budded virions of Autographa californica multiple nucleopolyhedrovirus. J. Virol. 91 (20).
Hartmann, M.D., 2017. Functional and structural roles of coiled coils. Subcell. Biochem. 82, 63–93.
Hou, D., Zhang, L., Deng, F., Fang, W., Wang, R., Liu, X., Guo, L., Rayner, S., Chen, X., Wang, H., Hu, Z., 2013. Comparative proteomics reveal fundamental structural and functional differences between the two progeny phenotypes of a baculovirus. J. Virol. 87 (2), 829–839.
Imai, N., Kang, W., Iwabuchi, K., Sato, K., Maeda, S., 2000. Analysis of interaction between molecules of Bombyx mori nucleopolyhedrovirus IE-2 using a yeast two-hybrid system. Acta Virol. 44 (3), 199–202.
Ke, J., Wang, J., Deng, R., Wang, X., 2008. Autographa californica multiple nucleopolyhedrovirus ac66 is required for the efficient egress of nucleocapsids from the nucleus, general synthesis of preoccluded virions and occlusion body formation. Virology 374 (2), 421–431.
Liu, J., Rost, B., 2001. Comparing function and structure between entire proteomes. Protein Sci. 10 (10), 1970–1979.
Ludwiczak, J., Winski, A., Szczepaniak, K., Alva, V., Dunin-Horkawicz, S., 2019. DeepCoil-a fast and accurate prediction of coiled-coil domains in protein sequences. Bioinformatics 35 (16), 2790–2795.
Lupas, A.N., Bassler, J., 2017. Coiled coils – a model system for the 21st century. Trends Biochem. Sci. 42 (2), 130–140.
Lupas, A., Van Dyke, M., Stock, J., 1991. Predicting coiled coils from protein sequences. Science 252 (5009), 1162–1164.
Lupas, A.N., Bassler, J., Dunin-Horkawicz, S., 2017. The Structure and Topology of α-Helical Coiled Coils, Fibrous Proteins: Structures and Mechanisms, pp. 95–129.
Mason, J.M., Arndt, K.M., 2004. Coiled coil domains: stability, specificity, and biological implications. Chembiochem 5 (2), 170–176.
McDonnell, A.V., Jiang, T., Keating, A.E., Berger, B., 2006. Paircoil2: improved prediction of coiled coils from sequence. Bioinformatics 22 (3), 356–358.
Milks, M.L., Washburn, J.O., Willis, L.G., Volkman, L.E., Theilmann, D.A., 2003. Deletion of pe38 attenuates AcMNPV genome replication, budded virus production, and virulence in heliothis virescens. Virology 310 (2), 224–234.
Mishra, V., 2020. A comprehensive guide to the commercial baculovirus expression vector systems for recombinant protein production. Protein Pept. Lett. 27 (6), 529–537.
Mo, H., Konstantinidis, A.K., Stewart, K.D., Dekhtyar, T., Ng, T., Swift, K., Matayoshi, E.D., Kati, W., Kohlbrenner, W., Molla, A., 2004. Conserved residues in the coiled-coil pocket of human immunodeficiency virus type 1 gp41 are essential for viral replication and interhelical interaction. Virology 329 (2), 319–327.
Moller, P., Pariente, N., Klenk, H.D., Becker, S., 2005. Homo-oligomerization of Marburgvirus VP35 is essential for its function in replication and transcription. J. Virol. 79 (23), 14876–14886.
Ohkawa, T., Welch, M.D., 2018. Baculovirus actin-based motility drives nuclear envelope disruption and nuclear egress. Curr. Biol. 28 (13), 2153–2159.
Ono, C., Okamoto, T., Abe, T., Matsuura, Y., 2018. Baculovirus as a tool for gene delivery and gene therapy. Viruses 10 (9).
Parry, D.A., Fraser, R.D., Squire, J.M., 2008. Fifty years of coiled-coils and alpha-helical bundles: a close relationship between sequence and structure. J. Struct. Biol. 163 (3), 258–269.
Prikhod'ko, E.A., Lu, A., Wilson, J.A., Miller, L.K., 1999. In vivo and in vitro analysis of baculovirus ie-2 mutants. J. Virol. 73 (3), 2460–2468.
Qiu, J., Tang, Z., Cai, Y., Wu, W., Yuan, M., Yang, K., 2019. The Autographa californica multiple nucleopolyhedrovirus ac51 gene is required for efficient nuclear egress of nucleocapsids and is essential for in vivo virulence. J. Virol. 93 (3).
Rackham, O.J.L., Madera, M., Armstrong, C.T., Vincent, T.L., Woolfson, D.N., Gough, J., 2010. The evolution and structure prediction of coiled coils across all genomes. J. Mol. Biol. 403 (3), 480–493.
Ramaswamy, V.K., Di Palma, F., Vargiu, A.V., Corona, A., Piano, D., Ruggerone, P., Zinzula, L., Tramontano, E., 2018. Insights into the homo-oligomerization properties of N-terminal coiled-coil domain of Ebola virus VP35 protein. Virus Res. 247, 61–70.
Reid, S.P., Cardenas, W.B., Basler, C.F., 2005. Homo-oligomerization facilitates the interferon-antagonist activity of the ebolavirus VP35 protein. Virology 341 (2), 179–189.
Rohrmann, G.F., 2019a. The AcMNPV genome: gene content, conservation, and function. In: Baculovirus Molecular Biology, fourth ed. National Center for Biotechnology Information, Bethesda MD, pp. 201–275.
Rohrmann, G.F., 2019b. Baculovirus Molecular Biology, fourth ed. National Center for Biotechnology Information, Bethesda, MD.
Rose, A., Stahlberg, E.A., Meier, I., 2007. Genome-wide identification and comparative analysis of coiled-coil proteins. Scalable Comput. Pract. Exp. 8 (2), 167–171.
Sauter, N.K., Hanson, J.E., Glick, G.D., Brown, J.H., Crowther, R.L., Park, S.J., Skehel, J.J., Wiley, D.C., 1992. Binding of influenza virus hemagglutinin to analogs of its cell-surface receptor, sialic acid: analysis by proton nuclear magnetic resonance spectroscopy and X-ray crystallography. Biochemistry 31 (40), 9609–9621.
Schaly, S., Ghebretatios, M., Prakash, S., 2021. Baculoviruses in gene therapy and personalized medicine. Biologics 15, 115–132.
Shao, W., He, L., Chen, Q., Li, J., Deng, F., Wang, H., Hu, Z., Wang, M., 2019. Functional characterization of the group I alphabaculovirus specific gene ac73. Virol. Sin. 34 (6), 701–711.
Shim, J.-H., Xiao, C., Hayden, M.S., Lee, K.-Y., Trombetta, E.S., Pypaert, M., Nara, A., Yoshimori, T., Wilm, B., Erdjument-Bromage, H., Tempst, P., Hogan, B.L.M., Mellman, I., Ghosh, S., 2006. CHMP5 is essential for late endosome function and down-regulation of receptor signaling during mouse embryogenesis. J. Cell Biol. 172 (7), 1045–1056.
Singh, M., Berger, B., Kim, P.S., 1999. LearnCoil-VMF: computational evidence for coiled-coil-like motifs in many viral membrane-fusion proteins. J. Mol. Biol. 290 (5), 1031–1041.
Sung, L.Y., Chen, C.L., Lin, S.Y., Li, K.C., Yeh, C.L., Chen, G.Y., Lin, C.Y., Hu, Y.C., 2014. Efficient gene delivery into cell lines and stem cells using baculovirus. Nat. Protoc. 9 (8), 1882–1899.
Sweeney, N.T., Brenman, J.E., Jan, Y.N., Gao, F.B., 2006. The coiled-coil protein shrub controls neuronal morphogenesis in Drosophila. Curr. Biol. 16 (10), 1006–1011.
Truebestein, L., Leonard, T.A., 2016. Coiled-coils: the long and short of it. Bioessays 38 (9), 903–916.
Wang, R., Deng, F., Hou, D., Zhao, Y., Guo, L., Wang, H., Hu, Z., 2010. Proteomics of the Autographa californica nucleopolyhedrovirus budded virions. J. Virol. 84 (14), 7233–7242.
Wang, Y., Zhang, X., Zhang, H., Lu, Y., Huang, H., Dong, X., Chen, J., Dong, J., Yang, X., Hang, H., Jiang, T., 2012. Coiled-coil networking shapes cell molecular machinery. Mol. Biol. Cell 23 (19), 3911–3922.
Watanabe, S., Takada, A., Watanabe, T., Ito, H., Kida, H., Kawaoka, Y., 2000. Functional importance of the coiled-coil of the Ebola virus glycoprotein. J. Virol. 74 (21), 10194–10201.
Wolf, E., Kim, P.S., Berger, B., 1997. MultiCoil: a program for predicting two- and three-stranded coiled coils. Protein Sci. 6 (6), 1179–1189.
Yang, M., Huang, C., Qian, D.D., Li, L.L., 2014. Functional characterization of Autographa californica multiple nucleopolyhedrovirus gp16 (ac130). Virology 464–465, 341–352.
Yue, Q., Yu, Q., Yang, Q., Xu, Y., Guo, Y., Blissard, G.W., Li, Z., 2018. Distinct roles of cellular ESCRT-I and ESCRT-III proteins in efficient entry and egress of budded virions of Autographa californica multiple nucleopolyhedrovirus. J. Virol. 92 (1).
Zinzula, L., Nagy, I., Orsini, M., Weyher-Stingl, E., Bracher, A., Baumeister, W., 2019. Structures of Ebola and reston virus VP35 oligomerization domains and comparative biophysical characterization in all ebolavirus species. Structure 27 (1), 39–54 e36.