The chromatin binding domain, including the QPQRYG motif, of feline foamy virus Gag is required for viral DNA integration and nuclear accumulation of Gag and the viral genome

Guochao Wei[a,1], Timo Kehl[a], Qiuying Bao[a], Axel Benner[b], Janet Lei[a,2], Martin Löchelt[a,*]

[a] *Division of Molecular Diagnostics of Oncogenic Infections, Research Focus Infection, Inflammation and Cancer, German Cancer Research Center (Deutsches Krebsforschungszentrum, DKFZ), Heidelberg, Germany*
[b] *Division of Biostatistics, German Cancer Research Center (Deutsches Krebsforschungszentrum, DKFZ), Heidelberg, Germany*

ABSTRACT

The retroviral Gag protein, the major component of released particles, plays different roles in particle assembly, maturation or infection of new host cells. Here, we characterize the Gag chromatin binding site including the highly conserved QPQRYG motif of feline foamy virus, a member of the *Spumaretrovirinae*. Mutagenesis of critical residues in the chromatin binding site/QPQRYG motif almost completely abrogates viral DNA integration and reduces nuclear accumulation of Gag and viral DNA. Genome packaging, reverse transcription, particle release and uptake into new target cells are not affected. The integrity of the QPQRYG motif appears to be important for processes after cytosolic entry, likely influencing incoming virus capsids or disassembly intermediates but not Gag synthesized *de novo* in progeny virus-producing cells. According to our data, chromatin binding is a shared feature among foamy viruses but further work is needed to understand the mechanisms involved.

1. Introduction

Virus particles are molecular machines that fulfill several distinct functions during different phases of the virus life cycle. As many viruses have limited genetic coding capacity, the assembly process is highly dependent on the multi-functionality of viral proteins (Freed, 2001; Flint et al., 2009). In addition to the stoichiometric assembly of all relevant proteins into nascent particles, major challenges are genome encapsidation and subsequent trafficking and release in new target cells (Meng and Lever, 2013). The assembly process culminates in the production and release of particles that may undergo further maturation into physically stable and infectious entities capable of surviving the harsh environmental conditions faced when finding new host cells, either within the same organism or upon transmission to a new host (Hutter et al., 2013; Freed, 2015). However, after specific entry or uptake, the virus particle must be able to carefully coordinate particle disassembly, genome release, and the localization of viral proteins present in the particle that may be necessary for the initial steps of the replication cycle (Nisole and Saïb, 2004; Berka et al., 2013; Pöhlmann and Simmons, 2013). Premature genome release or aberrant localization of viral components within the cell could trigger pre-existing antiviral defenses and inevitable restriction of virus replication (Franca et al., 2006; Bhat and Fitzgerald, 2014; Nisole et al., 2005). Viral protein expression and localization are therefore tightly spatio-temporally controlled to perform seemingly mutually exclusive tasks, such as the tight physical stability of released particles or regulated particle disintegration upon entry of a new host cell (Flint et al., 2009; Adamson and Freed, 2010).

In retroviruses, most of these functions are encoded by the major viral structural proteins Gag and Env (Hutter et al., 2013; Freed, 2015). The retroviral Env is mostly involved in cell targeting and entry; only in Spumaviruses is Env required for particle budding (Pietschmann et al., 1999; Gallo et al., 2003; Wilk et al., 2001; Geiselhart et al., 2003, 2004; Lindemann and Goepfert, 2003). In contrast, Gag plays a more diverse role, influencing almost all aspects of the virus replication cycle (Bell and Lever, 2013; Müllers, 2013). The overall functions of Gag and its processing are conserved among retroviruses, though several differences do exist between and within the retroviral subfamilies

Orthoretrovirinae and Spumaretrovirinae (Müllers, 2013). While the functions of orthoretroviral Gag have been studied substantially, comparatively little is known about Gag from spumaretroviruses, or foamy viruses (FV), including simian (SFV), feline (FFV), bovine (BFV), and equine FVs (EFV). For instance, FV Gag proteins encode the matrix-, capsid- and nucleocapsid-like domains (MA, CA and NC, respectively). In contrast to orthoretroviruses, however, they are not proteolytically separated in the mature and infectious FV particle (Freed, 1998; Taylor et al., 2017; Ball et al., 2016; Flügel and Pfrepper, 2003). Only the C-terminus of Gag is processed by the FV protease (PR), releasing a 3–4 kDa large terminal peptide of unknown function (Zemba et al., 1998; Enssle et al., 1997). Additional processing may occur late after infection of new host cells (Flügel and Pfrepper, 2003; Lehmann-Che et al., 2005). Since Pol and Gag expression are independent of each other, the kinetics of PR and reverse transcriptase (RT) activity differ in FVs (Yu et al., 1996). In FVs, reverse transcription occurs prior to particle release, leading to virions that may contain full-length DNA genomes (Yu et al., 1996). Another distinguishing feature of FVs is the Env-dependent release of FV capsids pre-assembled at the microtubule organizing center (MTOC) by budding at the plasma membrane or intracellular membranes (Wilk et al., 2001; Geiselhart et al., 2003, 2004; Lindemann and Rethwilm, 2011; Yu et al., 2006). Enriched FV particle preparations mostly consist of physically intact virions, potentially due to the unique structure of Env, which has a long, N-terminal particle-associated leader peptide (Elp) and forms a dense array on the virus surface (Wilk et al., 2001, 2000; Lindemann et al., 2001).

Though no classical orthoretroviral Cys-His fingers needed for genome binding are present in the C-terminal NC domain of FVs (Müllers, 2013), the NC domain is rich in glycine and arginine residues, forming three so-called glycine-arginine-rich boxes in SFVs, including prototype FV (PFV) (Winkler et al., 1997; Hamann et al., 2014). In PFV, these glycine and arginine residues are involved in RNA binding, genomic RNA packaging and reverse transcription (Hamann et al., 2014; Müllers et al., 2011a; Stenbak and Linial, 2004). In addition, the PFV sequence $\underline{Q}GGYNLRPRTY\underline{Q}P\underline{Q}RYG$ (Gag residues 534 to 550, Fig. 1A) in the central glycine-arginine-rich box contains the chromatin binding site (CBS, underlined), tethering Gag to chromatin via H2A/2B histone binding during mitosis after nuclear membrane breakdown (Tobaly-Tapiero et al., 2008; Müllers et al., 2011b). Recently, it was shown that the Y residue of the overlapping, highly conserved and currently understudied QPQRYG motif is involved in H3 histone binding (Winkler et al., 1997; Lesbats et al., 2017). The CBS-mediated access of Gag to chromosomes is responsible for the nuclear accumulation of Gag in newly infected cells and contributes to the integration site preference of FVs (Müllers et al., 2011b; Lesbats et al., 2017; Hocum et al., 2016). This may be one of the reasons for the requirement of actively cycling cells for FV replication and the general cell cycle-dependency of FV infections (Trobridge and Russell, 2004; Patton et al., 2004). While mutagenesis of the CBS in a cloned PFV genome completely abrogated viral infectivity, titers of PFV-based vectors with a CBS deletion were reduced about five-fold (Tobaly-Tapiero et al., 2008; Müllers et al., 2011b).

The fate of incoming FV particles and all steps preceding integration of the reverse-transcribed viral DNA into the host cell genome are only partially understood. FV capsids or capsid-like assemblies are transported to the MTOC and may persist there or at other cellular sites until the nuclear envelope breaks down during host cell mitosis (Tobaly-Tapiero et al., 2008; Müllers et al., 2011b; Lehmann-Che et al., 2007; Petit et al., 2003). The pre-integration complex (PIC) is thought to be tethered via the CBS in Gag to chromatin before integrase (IN)-mediated integration. Chromatin binding has also been detected in cells expressing only Gag and was shown to mediate nuclear accumulation of Gag (Müllers et al., 2011b).

Here, we show that CBS-mediated PIC binding to chromatin occurs in FFV-infected cells and is therefore a shared feature among FVs. We found that the highly conserved downstream QPQRYG motif that is part of the FV CBS plays an important role in chromatin binding, predominantly in newly infected cells, though it also influences chromatin binding of newly synthesized Gag. Mutations of Y416, R419 and RYG residues of the FFV CBS strongly impair viral infectivity, nuclear accumulation of Gag and, in particular, viral genome integration. These mutations do not affect particle assembly, budding, reverse transcription or cell entry. Alanine mutagenesis of QPQ also influences infectivity and DNA integration but not nuclear accumulation of Gag or the DNA genome. The data characterize a novel component of the Gag chromatin binding machinery that mainly acts in newly-infected target cells after entry before and, importantly, during FV DNA integration. Furthermore, nuclear targeting of Gag in producer cells appears to be dispensable for proper assembly and release.

2. Results

2.1. CBS residues Y416, R419 and the QPQRYG motif are not required for capsid assembly and particle release but crucial for viral infectivity

In previous studies, a CBS was identified in PFV Gag (Müllers, 2013; Tobaly-Tapiero et al., 2008; Müllers et al., 2011b; Lesbats et al., 2017). The homologous region of the distantly related FFV Gag (see Fig. 1A and Section 1) contains, as in PFV and most other FVs, a short and highly conserved six-amino acid sequence QPQRYG (underlined) overlapping the putative FFV CBS GGGYNFRRNPQQPQRYG (FFV Gag amino acids 413–429) (Winkler et al., 1997; Tobaly-Tapiero et al., 2008). However, it is unknown whether the QPQRYG motif in FFV Gag also functions as a CBS or whether it has a distinct function during chromatin binding. The functions of the CBS, including the QPQRYG motif, in FFV Gag were explored in the pCF-7 plasmid genome by alanine mutagenesis. To inactivate the CBS previously described for PFV (Tobaly-Tapiero et al., 2008), the two conserved residues Y416 and R419 in the corresponding sequence of FFV Gag (GGG$\underline{Y}$NF$\underline{R}$RNPQQP-QRYG, underlined, see also Fig. 1A and B), were replaced by alanines in mutant M1. Alanine mutations of residues QPQ (M2) and RYG (M3) of the CBS QPQRYG motif, as well as the residues Q and G immediately downstream of the QPQRYG motif (M4), were also introduced (Fig. 1B).

Protein expression and particle release were characterized upon transfection into HEK293T cells. The mutant Gag proteins had similar steady state levels in transfected cells compared to wild-type (wt) Gag. Particle release, measured by the presence of processed and unprocessed Gag and packaging of mature p65^{PR-RT}, was comparable between wt and mutants (Fig. 1C) with a certain degree of variation between samples and experiments also observed for PFV Gag and full-length Pol and p65^{PR-RT} proteins (for review see Müllers, 2013 and Fig. 2C). Mutation of Y416/R419 (M1) significantly decreased infectious viral titers by 1000-fold (Fig. 1D). The reduction in infectivity of the FFV CBS mutant is similar to a corresponding non-infectious PFV mutant but different from a recombinant and replication-deficient subgenomic PFV eGFP vector (Tobaly-Tapiero et al., 2008; Müllers et al., 2011b), which was much less affected for unknown reasons. Replacement of the first three residues of the QPQRYG motif (QPQ, M2) significantly decreased titers by more than 10-fold, while replacement of the next three residues (RYG residues, M3) reduced titers by even more than 100-fold. Mutagenesis of the two residues downstream of the QPQRYG motif did not impair viral infectivity (M4). This mutant was included in subsequent studies as an additional control.

To determine the contribution of individual residues in M1 and M3, both of which led to high reductions in viral titer, single amino acid replacements of the M1 and M3 sites were introduced into gag in the viral genome pCF-7 (Fig. 1B). Independent replacements of Y416 and R419 in the CBS significantly decreased titers by 125- and 625-fold, respectively (Fig. 1E). A corresponding mutation of R540Q in a PFV vector system resulted only in a 50% decrease of transduction efficacy, probably due to species-, context- and mutation-specific differences (Lesbats et al., 2017). Individual replacements of RYG residues were