The FEBS Journal

FEBS PRESS science publishing by scientists

VIEWPOINT

# Cas9 beyond CRISPR – SUMOylation, effector-like potential and pathogenic adaptation

Umut Sahin

Faculty of Engineering and Natural Sciences, Sabancı University, Istanbul, Turkey

**Keywords**
Cas9; effector proteins; host–pathogen interactions; post-translational modifications; SUMOylation

**Correspondence**
U. Sahin, Faculty of Engineering and Natural Sciences, Sabancı University, Istanbul 34956, Turkey
Tel: +90 2164839538
E-mail: umut.sahin@sabanciuniv.edu

(Received 24 May 2025, revised 17 July 2025, accepted 27 August 2025)

doi:10.1111/febs.70256

The CRISPR/Cas9 system has revolutionized molecular biology and gene editing, yet key aspects of its regulation, especially within eukaryotic environments, remain enigmatic. In this Viewpoint article, I will speculate on and explore the provocative hypothesis that Cas9 may possess previously unrecognized effector-like functions when expressed in host cells, potentially shaped by host-mediated post-translational modifications (PTMs). Of particular interest is SUMOylation at lysine 848, a key residue for DNA binding within the catalytic site, raising the possibility that this modification is not incidental, but functionally significant and precisely regulated. SUMOylation, a eukaryotic PTM, is increasingly recognized as a mechanism that also targets bacterial and viral effector proteins and virulence factors during infection, exerting context-dependent effects that may either enhance or hinder pathogen replication. Could Cas9, beyond its canonical role in bacterial CRISPR immunity, act as a host-modulating effector during infection, akin to known bacterial nucleomodulins such as transcription activator-like (TAL) effectors? If so, this would imply that certain pathogenic bacteria may have evolved Cas9 variants capable of exploiting host PTM machinery and targeting the host genome—an adaptation with potential implications for microbial virulence, host–pathogen interactions, and co-evolutionary dynamics. This perspective underscores the importance of systematically mapping Cas9 PTMs and examining their evolutionary conservation, functional significance, and pharmacological tunability, not only for basic biological insight and to deepen our understanding of microbial strategies, but also to refine the precision and safety of Cas9-based therapeutic platforms.

## Introduction

CRISPR arrays are specialized regions found on prokaryotic and archaeal chromosomes, composed of short repetitive sequences intercalated with short, unique, and variable spacer sequences, initially discovered in *Escherichia coli* [1–4]. Together with CRISPR-associated proteins (Cas proteins), these arrays comprise an intriguing adaptive immune mechanism that defends against foreign invaders, such as phages and plasmids [5]. This fascinating defense mechanism operates through three distinct stages (Fig. 1). In the

**Abbreviations**
Cas9, CRISPR-associated protein 9; CRISPR, clustered regularly interspaced short palindromic repeats; crRNA, CRISPR RNA; HIV-1, human immunodeficiency virus 1; HTLV-1, human T-cell lymphotrophic virus 1; NLS, nuclear localization signal; PTM, post-translational modification; Pup, prokaryotic ubiquitin-like protein; SaCas9, *Staphylococcus aureus Cas9*; sgRNA, single guide RNA; SLO, Streptolysin O; SpCas9, *Streptococcus pyogenes Cas9*; STUbL, SUMO-targeted ubiquitin ligase; *SUMO*, small ubiquitin-like modifier; TAL, transcription activator-like; TALEN, TAL effector nuclease; tracrRNA, trans-activating CRISPR RNA.

adaptation stage, sequences from foreign genetic elements are incorporated (or memorized) into the CRISPR array which then become the abovementioned spacer regions, resulting in acquired immunity that kicks in during subsequent infections. During this time (expression stage), the entire CRISPR array gets transcribed into a precursor CRISPR RNA transcript (pre-crRNA) from which mature crRNA's are excised through processing by Cas endonucleases. Finally, during the interference stage, crRNAs form a duplex with trans-activating CRISPR RNA (tracrRNA), creating a platform for the binding of Cas nucleases, which are subsequently guided to complementary sequences on invading nucleic acids, resulting in their cleavage and neutralization [5–7]. As a result, the system confers immunity and protection during infection in these unicellular organisms, mirroring the adaptive immune response seen in higher eukaryotes.

This remarkable evolutionary innovation, which enables bacteria to defend themselves against invading genetic elements, was soon recognized as a powerful tool for genome editing in higher eukaryotes, achieved by programming bacterial Cas9 enzymes with matching guide RNA sequences to precisely target and cleave eukaryotic DNA [8–10]. In recent years, the CRISPR-Cas9 system has become the gene editing platform of choice in laboratories across the globe, thanks to its versatility, simplicity, and remarkable efficiency. It has been widely applied across various systems, including cell lines, primary cells, and induced pluripotent stem cells, and has facilitated the development of new animal disease models [11–13]. Importantly, it has opened new avenues for therapeutic interventions in the clinical management of various diseases, as recently demonstrated in the treatment of sickle cell disease and beta-thalassemia [14].

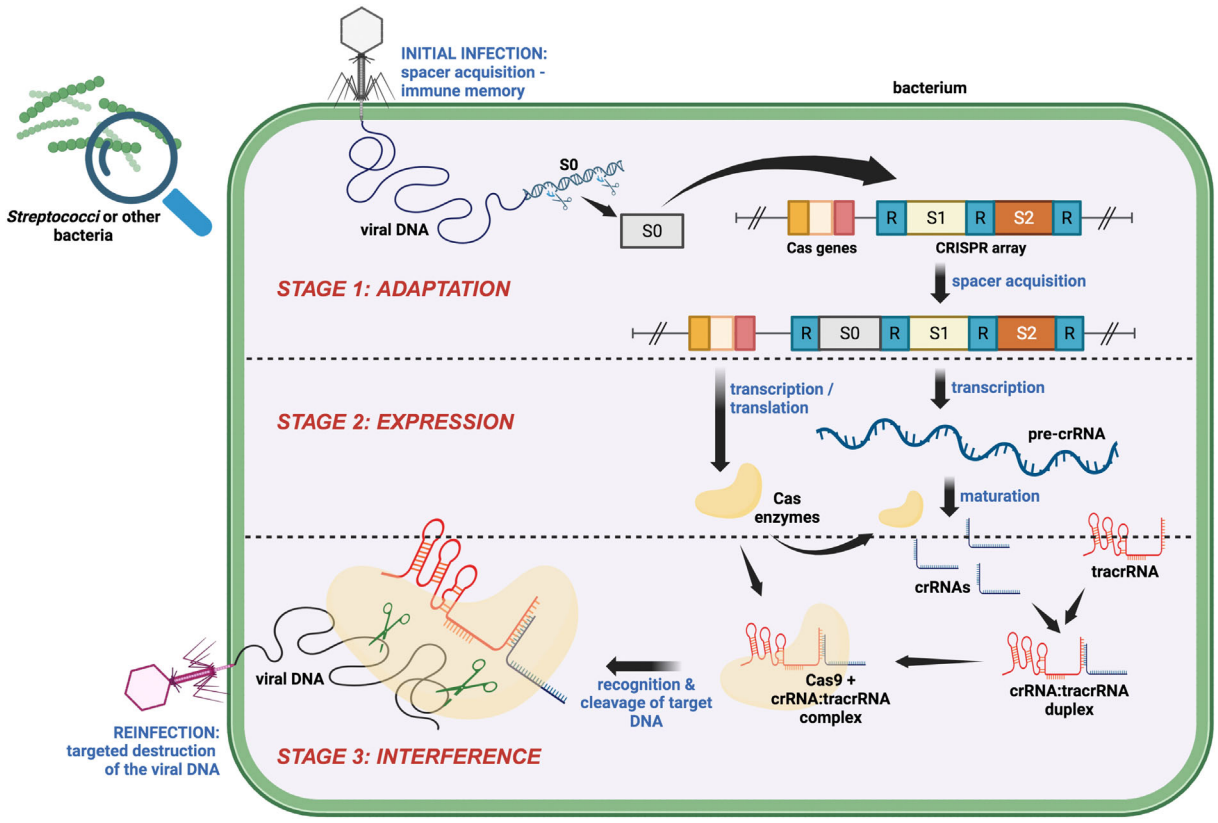


**Fig. 1.** Mechanism of CRISPR-mediated immunity in bacteria. The CRISPR system provides bacteria with adaptive immunity against phage infections and functions through three distinct stages. During the adaptation stage, a fragment of foreign DNA (S0) is incorporated as a spacer into the CRISPR locus, becoming the most recent entry in the array—effectively 'memorizing' the invader for future defense. The locus also contains genes encoding Cas proteins and trans-activating CRISPR RNA (tracrRNA). In the expression stage, the CRISPR array is transcribed into a long pre-CRISPR RNA (pre-crRNA), which is then processed into individual mature crRNAs. Finally, in the interference stage, if the bacterium is reinfected, the Cas9:crRNA:tracrRNA complex identifies and disables the matching foreign DNA.

The widespread adoption of the CRISPR/Cas9 system as a technological platform, along with its promising clinical applications, and its remarkable and elegant molecular design—from both scientific and evolutionary perspectives—highlights the need for a deeper understanding of the mechanisms that potentially regulate it. These regulatory processes could operate in natural environments, such as within bacteria, or during bacterial infections of eukaryotic cells. They could also have important implications in research and clinical contexts, particularly when Cas9 is artificially introduced into eukaryotic cells for gene editing applications. While the development of CRISPR-based gene editing technology has advanced rapidly, our understanding of the regulatory networks that govern this system remains limited. Importantly, these mechanisms may significantly influence both on-target and off-target cleavage activities. Understanding these regulatory pathways could also provide valuable insights into host–pathogen interactions and the co-evolutionary dynamics between them, as outlined below.

## The Cas9 enzyme

Cas enzymes are the catalytic engines that carry out CRISPR-associated functions. While some Cas enzymes, such as Cas1 and Cas2, function as part of the adaptation module, aiding spacer sequence acquisition, others (i.e., Cas9) serve as effector proteins of the interference module and cleave the target DNA.

While some bacteria such as *Escherichia coli* have sophisticated CRISPR/Cas toolkit (classified as class I CRISPR/Cas systems) containing multi-subunit effector complexes among other players, others have a rather simple, single multi-domain endonuclease (i.e., Cas9, Cas12a) that takes part in the interference stage [15,16]. These class II CRISPR/Cas systems, defined by a single endonuclease protein, can then be further engineered to provide modularity where the target specificity is determined by a nuclease-independent single guide RNA sequence (sgRNA), in which crRNA and tracrRNA are synthetically linked [10,17]. In this way, the system provides both simplicity and flexibility for researchers, and when combined with its efficiency, it has emerged as a preferred platform for genome-editing strategies. Among the various CRISPR-associated proteins, Cas9 from *Streptococcus pyogenes* (SpCas9), and to a lesser extent from *Staphylococcus aureus* (SaCas9), have gained particular popularity.

Cas9 has a bilobular shape, a feature also shared by Cas12a, containing a DNA recognition lobe (REC) and a nuclease lobe (NUC) (Fig. 2) [17,18]. The recognition lobe is formed by majority of the N-terminal residues and responsible for recognizing the target nucleotide sequences. The C-terminal nuclease lobe contains two catalytic cores with distinct endonuclease activities. The HNH core (also referred to as the HNH domain) cleaves the target DNA strand with which the sgRNA initially base pairs. On the other hand, the RuvC core, formed by 3 distinct RuvC domains, cleaves the complementary nontarget DNA strand. The combined actions of the HNH and RuvC catalytic cores ultimately create a double-strand break at the target site [18–21]. In addition to these 3 key structures (HNH core and RuvC core, together forming the NUC lobe; as well as the REC lobe), Cas9 also contains additional regions critical to its function. A helical region rich in positively charged residues that bridges the first RuvC domain to the REC lobe (thus named as bridge helix) helps stabilize the negatively charged sgRNA:target DNA duplex which is formed upon binding to target sequence. In parallel, the displaced nontarget DNA strand is also stabilized by the positively charged residues in the linker regions (L1 and L2) flanking the HNH domain [17].

The elucidation of the three-dimensional structure of *Streptococcus pyogenes* Cas9 at atomic resolution has provided critical insights into its conformational dynamics and the complex interactions it forms with both sgRNA and target DNA throughout different stages of the CRISPR pathway. These structural insights have contributed to our understanding of target recognition, binding, and catalysis [17–22]. While certain residues are directly involved in catalysis and endonucleolytic cleavage, many others, originating from various domains and linker regions, participate in sequence-specific or sequence-independent DNA binding. This structural complexity, coupled with its relatively large size, renders Cas9 highly susceptible to regulation by post-translational modifications (PTMs), which may influence its binding affinity to nucleic acids, enzymatic activity, stability, subcellular localization, and, critically, its on-target and off-target cleavage efficiencies.

## Post-translational modifications of Cas9

Post-translational modifications play a crucial role in rapidly, dynamically, and often reversibly regulating protein function. They have evolved to expand the functional diversity of the proteome and meet the increasing complexity of organisms over time. While PTMs are commonly linked to eukaryotic

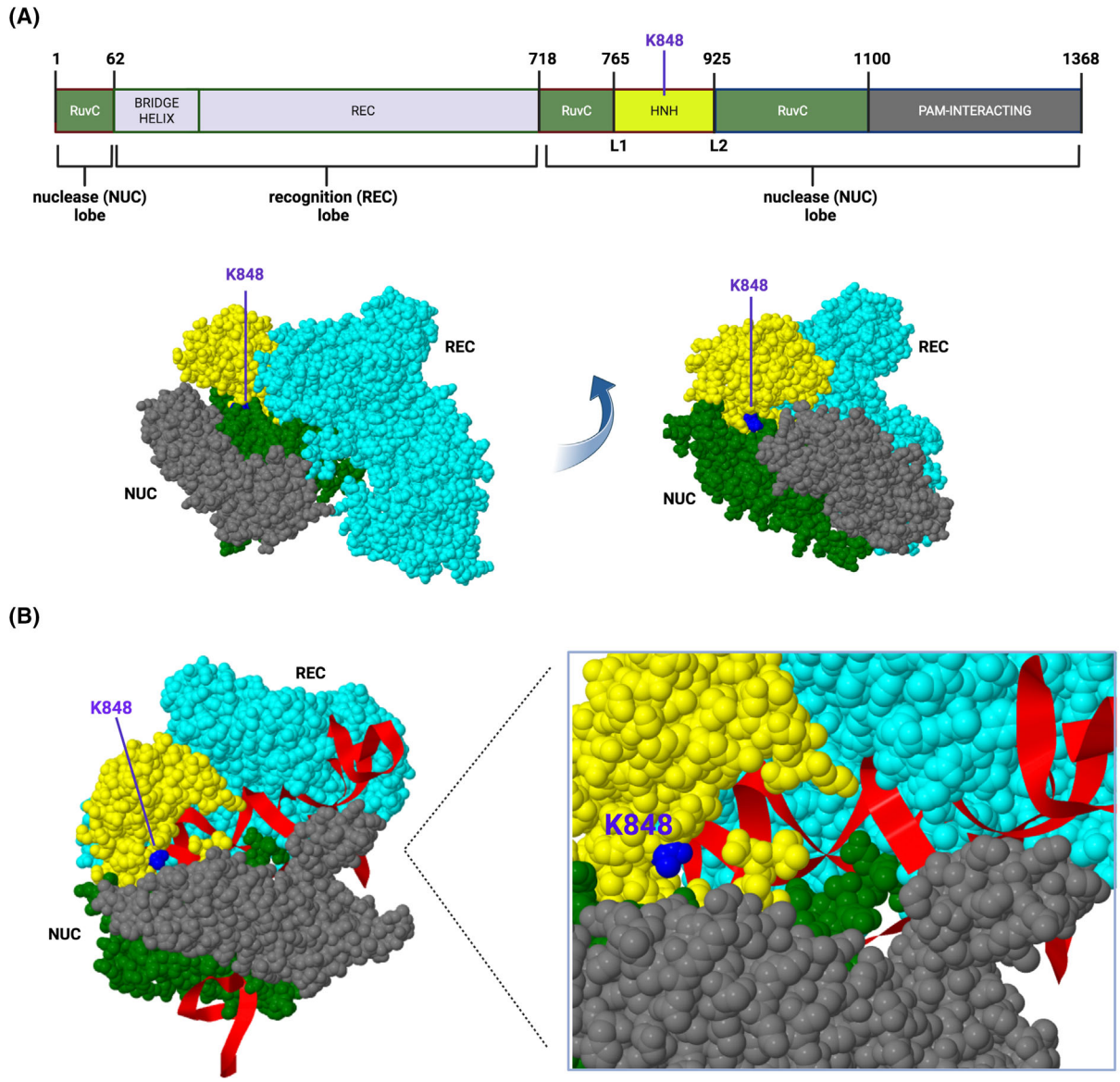


**Fig. 2.** Structural features of Cas9 highlighting lysine 848, the primary SUMO2/3 conjugation site. (A) The domain architecture of the *Streptococcus pyogenes* Cas9 enzyme is illustrated, highlighting the N-terminal recognition (REC) lobe, which includes the bridge helix, and the C-terminal nuclease (NUC) lobe, which houses the catalytic HNH and RuvC domains. The NUC lobe also contains the PAM-interacting domain, responsible for recognizing the protospacer adjacent motif (PAM), a short DNA sequence located just downstream of the target DNA cleavage site, a crucial step for DNA cutting. Lysine 848, the residue targeted for SUMOylation within the HNH catalytic domain, is marked on the 3D structure shown from two different perspectives (PDB: 4CMP). (B) A close-up view of lysine 848 is provided, showing its position relative to the DNA:RNA hybrid (red ribbon) within the catalytic core. Its solvent accessibility suggests it is readily accessible to SUMOylation enzymes (PDB: 4OO8). All structural images were generated using JSMOL.

systems, many bacterial species also employ a diverse array of modifications, including phosphorylation, acetylation, glycosylation, and even ubiquitin-like modifications such as pupylation, to control essential cellular processes such as cell division and virulence, and to adapt to environmental stresses [23].

Given its substantial size, comprising 1368 amino acids, it would be surprising if SpCas9 were not subject to a diverse array of PTMs, whether in its native bacterial context or upon introduction into eukaryotic systems, either through natural infection or artificial expression. Remarkably, despite this enzyme's

biological significance and the pressing need to understand its regulatory landscape, current knowledge on Cas9-associated PTMs remains limited.

As briefly touched upon earlier, PTMs may modulate Cas9 function under three distinct scenarios: (a) within its native bacterial context; (b) upon translocation into a eukaryotic host during infection, where it may be subject to host-mediated modifications; and (c) during heterologous expression in eukaryotic systems for genome-editing purposes (Fig. 3). In both eukaryotic settings, Cas9 has the potential to interface with host PTM pathways, including the SUMOylation machinery, which will be discussed further below.

Initial evidence suggesting that the CRISPR/Cas9 system may be subject to regulation by PTMs emerged from studies in which ubiquitin was artificially fused to the N terminus of Cas9 [24]. This modification led to enhanced proteasomal degradation and a reduced half-life of the enzyme when expressed in HEK293 cells or introduced into nonhuman primate embryos via mRNA injection. Notably, the authors demonstrated that this controlled destabilization of Cas9 effectively minimized the emergence of mosaic mutations, a frequent outcome of prolonged Cas9 expression and activity. This study not only underscored the critical role of temporal control in enhancing the specificity and clinical applicability of CRISPR-based genome editing, but also implicated a PTM as a modifiable regulatory mechanism with potential to fine-tune Cas9 activity for desired outcomes.

Building on this earlier work involving artificial ubiquitin fusion to Cas9, my group presented the first direct biochemical evidence of endogenous Cas9 ubiquitylation in 2022 [25]. In this study, we demonstrated that Cas9 undergoes extensive ubiquitin modification at multiple lysine residues upon either transient overexpression or stable expression in diverse eukaryotic cell systems. These ubiquitylated forms

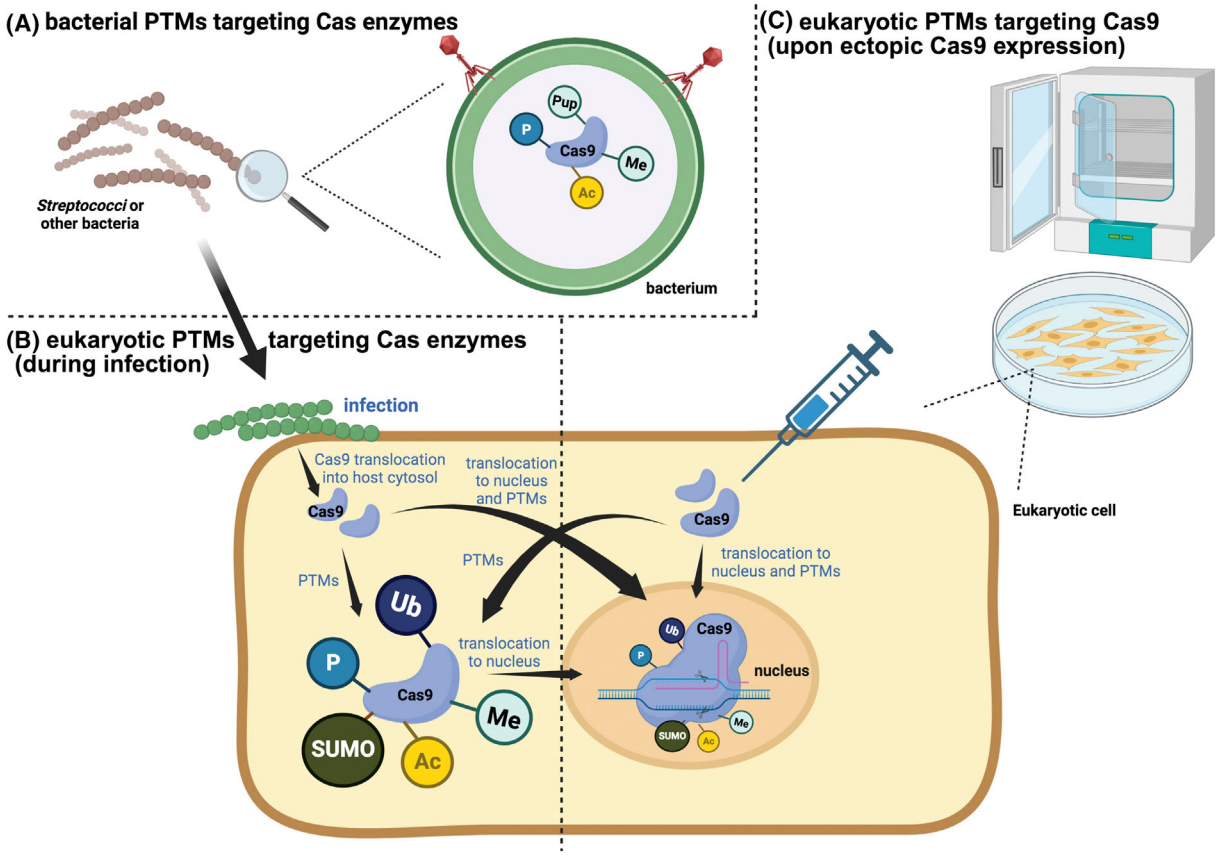


**Fig. 3.** Three distinct scenarios in which post-translational modifications (PTMs) may target Cas9 and influence its activity include (A) Cas9 undergoing prokaryotic PTMs within its native bacterial environment; (B) Cas9 being transferred to a eukaryotic host during infection, where it is modulated by host-derived PTMs such as SUMOylation; (C) Cas9 being subject to eukaryotic PTMs when heterologously expressed for CRISPR-based genome-editing applications. Ac, acetyl; Me, methyl; P, phosphate; Pup, prokaryotic ubiquitin-like protein; Ub, ubiquitin. With the exception of ubiquitination and SUMOylation, the post-translational modifications depicted on Cas9 are currently hypothetical.

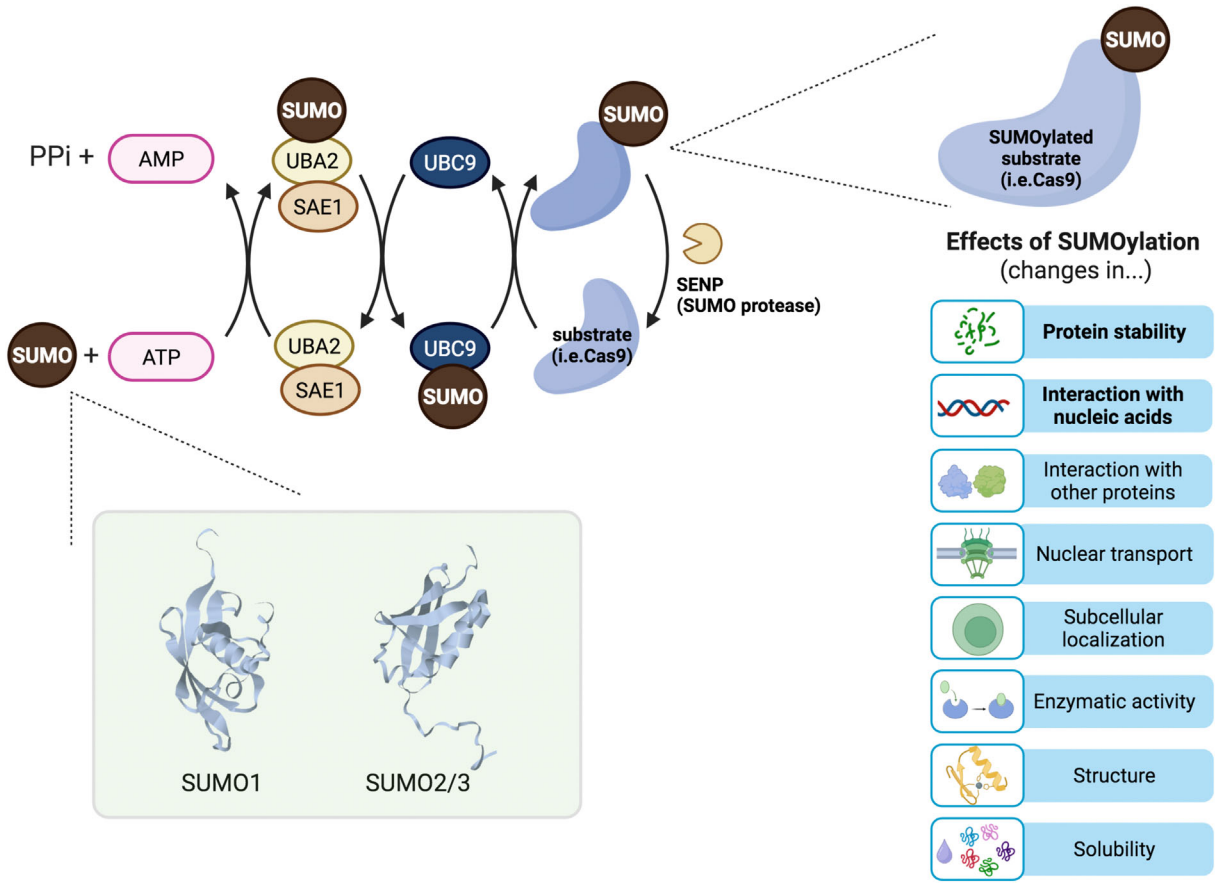


**Fig. 4.** The SUMO conjugation pathway and biochemical effects of SUMOylation. The SUMO protein is first activated by a heterodimeric E1 enzyme complex (SAE1/UBA2) in an energy-dependent manner. It is then transferred to the central E2 conjugating enzyme, UBC9. UBC9 can directly mediate the covalent attachment of SUMO to substrate proteins, such as Cas9, a process that can be further enhanced by SUMO E3 ligases (not depicted here). SUMOylation is a highly dynamic and reversible modification, regulated by SUMO-specific proteases (also known as SENPs), which remove SUMO from target proteins. The right side of the figure illustrates the biochemical and molecular consequences of SUMO modification. Cas9 features impacted by SUMOylation as supported by experimental evidence are highlighted in bold. Both SUMO1 (PDB: 2N1V) and SUMO2 (PDB: 2N1W) proteins are conjugated to lysine residues on substrate proteins via their C-terminal glycine residues, which are located on the upper side of the 3D structures shown in the inset. The structures were generated using the JSMOL molecular viewer.

accumulated in response to proteasome inhibition, and we further showed that Cas9 physically associated with proteasomal complexes. Collectively, these findings established ubiquitylation as a key PTM, negatively regulating Cas9 stability via targeted proteasomal degradation, a mechanism that had previously been overlooked in the context of heterologous Cas9 expression in eukaryotic systems.

Among the many lysine residues identified as ubiquitylation sites on spCas9 via mass spectrometry, one in particular, lysine 848, stands out, as it is also targeted by another PTM: the small ubiquitin-like modifier (SUMO) [25]. Like ubiquitin, SUMO is covalently conjugated to lysine residues on substrate proteins through a dedicated enzymatic cascade in a process known as SUMOylation (Fig. 4) [26,27]. SUMOylation is a highly versatile PTM, capable of modulating a wide range of protein properties, including enzymatic activity, conformation, stability, solubility, subcellular localization, and interaction capacity with proteins or nucleic acids, depending highly on the biological context (Fig. 4) [26]. While some bacteria employ a ubiquitin-like system involving Pup (prokaryotic ubiquitin-like protein) to tag substrates for proteasomal degradation via pupylation, SUMOylation appears to be a strictly eukaryotic modification, with no known functional or evolutionary equivalents in prokaryotes [23]. Consequently, although the ubiquitylation of Cas9 can be conceptually aligned with bacterial systems such as pupylation, the observation that

Cas9 is also subject to a uniquely eukaryotic PTM, SUMOylation, raises compelling and complex questions about relevance, contexts and consequences of this modification—especially during natural infections or in engineered eukaryotic environments.

The human genome encodes five SUMO paralogs, SUMO1 through SUMO5, of which SUMO1, SUMO2, and SUMO3 are ubiquitously expressed and have been the subject of extensive functional characterization over the past three decades [26,28]. Due to their high-sequence identity (~ 95%), SUMO2 and SUMO3 are often collectively referred to as SUMO2/3 (Fig. 4). Notably, SUMO2/3 is highly responsive to cellular stressors such as heat shock, oxidative stress, infection, and DNA damage, all of which typically trigger its conjugation to target proteins, thereby modulating their functions and stability. Unlike SUMO1, SUMO2/3 harbors an internal SUMOylation consensus site, enabling the formation of poly-SUMO2/3 chains on substrates, a feature analogous to poly-ubiquitin chains. These poly-SUMO modifications are instrumental in orchestrating diverse downstream processes, including recruitment of SUMO-targeted ubiquitin ligases (STUbLs) in some cases, resulting in proteasomal degradation [26,27,29–31].

Several observations indicate that spCas9 undergoes SUMOylation with remarkable specificity and precision, making it unlikely that this modification arises as a random or artifactual consequence of ectopic expression in eukaryotic cells, but may rather be deliberate and functionally meaningful. Despite the presence of numerous lysine residues within the enzyme, 10 of which are located within a canonical SUMOylation consensus motif, only a single residue, Lys 848, is modified by the specific SUMO paralog, SUMO2/3, when Cas9 is expressed in various eukaryotic systems, whether transiently or stably. Furthermore, Lys 848 is particularly interesting because it resides within the catalytic site of the HNH nuclease core and stands out as one of the key positively charged residues that facilitates Cas9's sequence-independent interaction with the DNA molecule on both target and off-target sites (Fig. 2) [22,25]. Indeed, disruption of this nonspecific interaction between Cas9 and DNA through neutralization of the positive charge on Lys 848 effectively reduces the rate of off-target cleavage events [32].

The highly specific SUMOylation of spCas9 at Lys 848 by SUMO2/3 introduces a previously unrecognized regulatory layer for this widely used genome-editing enzyme. This observation also raises intriguing questions about the evolutionary significance of this modification, particularly if it occurs naturally during infections in eukaryotic cells by Cas9-harboring bacteria.

To date, it remains unclear whether Cas9 undergoes additional post-translational modifications beyond ubiquitylation and SUMOylation. However, given its large size and central role in Class II CRISPR systems in bacteria, it is likely that Cas9 is regulated by a variety of PTMs, such as phosphorylation, acetylation, or pupylation, which have yet to be identified. For the purpose of this review, I will focus on the potential implications and significance of SUMOylation.

## Why is Cas9 subject to SUMOylation? An unexpected modification with big implications

Studies have confirmed the biochemical impact of Cas9 SUMOylation. Disruption of SUMO2/3 conjugation at Lys 848, either by substituting this residue with arginine, thereby preserving the positive charge while preventing SUMOylation, or by eliminating the adjacent aspartic acid residue at position 850, which is essential for efficient SUMOylation of Lys 848, results in reduced enzyme stability and impaired sequence-specific DNA binding when guided by sgRNAs [25]. A comparable phenotype was also observed upon pharmacological inhibition of Cas9 SUMOylation, despite Lys 848 remaining intact [25]. While it remains to be established whether SUMOylation also influences Cas9's enzymatic activity, this modification, which is strikingly situated at the catalytic site of the HNH domain clearly exerts notable biochemical effects, suggesting that it is unlikely to be a random occurrence. However, further studies are required to determine whether this represents an evolutionarily selected regulatory mechanism.

As discussed above, in nature, Cas9 is a bacterial protein and is part of the prokaryotic adaptive immune system. Whether it is introduced into eukaryotic cells during infection in the native biological context is unknown. There are many examples of bacterial effectors being secreted into eukaryotic host cells via secretion systems, such as Type III or Type IV secretion systems, though this has not been shown for Cas9 [33]. Although *Streptococcus pyogenes* does not have the classical secretion systems like Type III or Type IV that are common in many Gram-negative pathogens, as a Gram-positive bacterium, it uses various other mechanisms to secrete its effector proteins into the host, such as Streptolysin O (SLO)-Dependent Translocation where SLO forms pores in host cell membranes through which bacterial effectors are translocated into host [34–38]. *Streptococcus pyogenes* indeed employs multiple secretion mechanisms for its effectors, such as the general secretion pathway (Sec

Pathway), Sortase-mediated anchoring, as well as the nonclassical secretion pathway that involves vesicles [39]; however, the presence of the SLO-dependent translocation system particularly makes it worth speculating whether Cas9 could be co-opted into translocation into the host via such a system, as this process does not require specific recognition motifs or signal peptides being present on the effector.

If during infection Cas9 were ever delivered to a host cell, even at low levels or via membrane disruption, retained long enough in the cytoplasm or nucleus, and structured appropriately, as it clearly is, it is also highly plausible that it would become a target of the eukaryotic SUMOylation machinery. Naturally, this raises the intriguing question of the physiological significance of a bacterial protein being subjected to a eukaryotic PTM. Countless studies have so far shown that noneukaryotic proteins, once introduced into eukaryotic cells, can be modified by host post-translational modification systems, including SUMOylation, ubiquitination, and phosphorylation [40–42]. In some instances, this benefits the invading pathogen, which hijacks the host's PTM machinery to enhance the function of its effector proteins. In other cases, however, such modifications may act as a host defense strategy, neutralizing the pathogen's virulence factors through specific PTMs. Both of these scenarios are relevant in the context of SUMOylation during host–pathogen interactions. SUMO peptides play a pivotal role in eukaryotic innate immunity. Beyond acting as key downstream effectors of interferon-mediated antiviral responses that limit pathogen replication, they can also directly hinder bacterial and viral propagation by modifying essential virulence factors [26,43–45]. For instance, SUMOylation of the HTLV-1 Tax oncoprotein decreases its stability by facilitating its ubiquitylation [46]. On the other hand, as an example of SUMOylation benefiting the pathogen, the SUMOylation of HIV-1 integrase, an enzyme essential for the integration of viral DNA into the host genome and for the progression of the viral life cycle, is necessary for the efficient execution of early stages of HIV-1 replication and plays a regulatory role in viral infectivity [47]. It is now widely recognized that SUMOylation of either the host's own proteins or bacterial/viral effector proteins during infection can have context-dependent, either agonistic or antagonistic, biochemical effects.

## Is Cas9 a bacterial effector?

The precise targeting of Lys 848, which plays a critical role in DNA binding, by a specific SUMO paralog (SUMO2/3), makes it plausible to propose that Cas9 could also function as a bacterial effector with previously uncharacterized roles in eukaryotic hosts.

Cas9, while originally identified for its role in the bacterial adaptive immune system, has several characteristics that could align with those of bacterial effector proteins. Firstly, the enzyme has cross-species functional potency, meaning that it is capable of interacting with and cleaving eukaryotic DNA. This makes it functionally similar to other bacterial effector proteins that are capable of manipulating host cellular mechanisms. Secondly, as discussed extensively above, it is subjected to a eukaryotic PTM, as many other bacterial effectors, and this modification seems to occur in a very precise and specific manner. Thirdly, given its biochemical function, Cas9 has the potential to contribute to bacterial virulence by affecting a plethora of host genes involved in immune response or other key cellular processes: Cas9 could potentially target key immune-related genes, or it can interfere with host cell apoptosis pathways allowing the pathogen to evade host cell death to ensure its own survival. Alternatively, by introducing double-strand breaks, it can overwhelm the host DNA repair pathways. Cas9's ability to modify DNA certainly makes it an interesting candidate for effector-like activity in the context of infection, and this could be fine-tuned by SUMOylation.

Within this speculative context of Cas9's potential role as a bacterial effector in eukaryotic hosts, one might envision multiple ways by which the enzyme targets the host DNA. Bacteria might release guide RNA molecules in vesicles or through pores, which could then associate with Cas9 once inside the cell. Alternatively, preformed guide RNA-Cas9 complexes may directly be released into the host through the secretion systems. RNA molecules that mimic guide RNA structures may also be synthesized by the host genome. Cas9 could also target genomic DNA directly without the need for a guide RNA, in this case acting as a sequence-independent effector or a nuclease that induces double-strand breaks at random locations. Cas9's endonucleolytic activity, especially if untethered from guide RNA specificity, raises parallels to DNA-targeting toxins or nucleases such as colicins [48,49]. Here, the breakages could lead to genomic instability, which might be leveraged by the pathogen to disrupt host cell functions or evade immune responses. In light of the recent findings that SUMO interacts with DNA in a sequence-independent manner, these nonspecific interactions between Cas9 and host DNA may indeed be mediated by the enzyme's SUMOylation [50].

In CRISPR/Cas9 genome-editing platforms, spCas9 is routinely engineered with one or more nuclear localization signals (NLS) at the N- or C terminus to

ensure its efficient nuclear localization. Native Cas9 proteins typically lack canonical NLSs; however, during infection, it is plausible that bacteria may hijack host cellular pathways to promote Cas9 nuclear translocation [51]. Indeed, bacterial pathogens can exploit host cellular mechanisms to transport certain effector proteins directly into the nucleus, employing strategies such as endosomal-lysosomal trafficking or retrograde transport. Certain bacterial effectors are also known to target proteins of the nuclear pore complex (e.g., nucleoporins) to modulate nuclear import/export. One can also easily imagine that SUMOylation (or other PTMs) could unmask cryptic NLSs or generate novel interfaces for interaction with importins, thereby enhancing the nuclear entry of Cas9 [51].

Notably, Cas9 expression is not exclusive to *Streptococcus pyogenes*, an extracellular pathogen. A wide range of other bacteria, including various Gram-positive and Gram-negative pathogenic species with distinct secretion systems, also encode Cas9 enzymes. Taken together, considering the various potential routes through which Cas9 could be secreted into eukaryotic host cells and its capacity to interact with host DNA in both guide RNA-dependent and RNA-independent manners, it is plausible that Cas9's function extends beyond its canonical role in bacterial CRISPR immunity. Instead, it may be co-opted during infection to influence host immune responses, compromise cellular integrity, or disrupt DNA repair pathways. In this context, the broad and seemingly nonspecific ubiquitylation of Cas9 on multiple lysine residues, leading to its degradation by the proteasome, could reflect an evolved host surveillance mechanism aimed at limiting the persistence or neutralizing the toxicity of this bacterial protein [25]. This observation, however, also underscores a potential limitation of the Cas9 effector hypothesis. The SUMOylation of Cas9 at Lys 848 could represent a host defense mechanism—a form of quality control targeting foreign proteins for degradation or sequestration, similar to how ubiquitylation operates. The apparent specificity for Lys 848 might result from structural accessibility or a coincidental match to the SUMOylation consensus motif. However, this seems unlikely: if it were merely due to motif availability, one would expect Lys 848 to be modified not only by SUMO2/3 but also by SUMO1. Furthermore, Cas9 contains nine other lysines that fall within canonical SUMOylation consensus motifs, most of which are solvent-exposed, yet none of these undergo SUMOylation. This suggests a level of selectivity that is unlikely to be incidental.

To rigorously address this limitation, a comparative analysis of SUMOylation at Lys 848—or an analogous residue with similar structural and functional properties—across Cas9 orthologs from both pathogenic and nonpathogenic bacteria will be critical, as discussed in the next section.

More broadly, several key pieces of biological evidence that would strengthen the effector hypothesis are currently lacking. These include the following: (a) demonstration of Cas9 secretion or delivery into host cells, then to nucleus, during natural infection; (b) identification of bacterial or host-derived guide RNAs within host cells; and (c) evidence of Cas9-dependent virulence or modulation of host processes. While their absence does not invalidate the effector hypothesis, they do represent some limitations. Nonetheless, each of these aspects can be systematically explored and addressed, both conceptually and experimentally, as outlined and proposed throughout this article.

## Conclusion: evolutionary insights and future prospects

While not all bacteria harbor class II CRISPR systems and Cas9 proteins, among those that do possess them, a compelling question arises as to whether Cas9 SUMOylation is a conserved phenomenon across phylogenetically distinct species, particularly in pathogenic strains that infect eukaryotic cells and therefore have the opportunity to interact with the host SUMOylation machinery. If Cas9 SUMOylation, in particular, on a residue homologous to *Streptococcus pyogenes*' Lys 848 in the catalytic site is found to be an evolutionarily conserved process, it would have far-reaching implications in microbiology, host–pathogen interactions, evolutionary biology, and also biotechnology, resulting in a paradigm shift in our understanding of Cas9 functions.

As mentioned above, traditionally, Cas9 is viewed as a bacterial defense protein, central to CRISPR-based adaptive immunity. Conservation of SUMOylation would imply selective pressure to retain host-interfacing capability and suggests that Cas9 interacts broadly with eukaryotic hosts, potentially redefining this protein as a host-modulating effector during infection. Nucleomodulins are a class of bacterial effectors that target and modulate host nuclear processes, including transcription, mRNA splicing, and chromatin remodeling, and are secreted by pathogenic bacteria such as *Listeria monocytogenes*, *Shigella flexneri*, and *Legionella pneumophila* [52,53]. Notably, some of these pathogens actively manipulate the host SUMOylation machinery to enhance their survival or replication, underscoring a functional interface between pathogen-derived effectors and host post-translational

modifications [42,54,55]. In this context, it will be important to investigate whether Cas9 qualifies as a nucleomodulin, with SUMOylation acting as a regulatory switch that modulates or repurposes its activity within the host nucleus.

It is also noteworthy that a specific class of bacterial nucleomodulins, known as Transcription activator-like (TAL) effectors, is secreted by pathogens into plant cells, where they induce the expression of host genes that facilitate infection. These proteins were later engineered into TAL effector nucleases (TALENs), which served as early tools for genome editing prior to the emergence of the CRISPR/Cas9 platform [56,57].

In conclusion, additional phylogenetic and biochemical studies will be crucial to determine whether Cas9 SUMOylation is an evolutionarily conserved feature and if this modification occurs under physiological conditions during natural host–pathogen interactions. Conserved Cas9 SUMOylation across species, particularly at sites homologous to SpCas9 Lys 848, mediating interactions with the DNA in the HNH catalytic core, would imply functional relevance, rather than random modification or artifact, and may redefine Cas9 more than just a genome editor and potentially as a virulence factor, and finally link bacterial immunity and host immunity in very unexpected ways. Furthermore, it will be vital to systematically investigate and document other PTMs that might act on Cas9, both in bacteria and during ectopic expression in eukaryotic systems. This will not only provide insights into the regulatory mechanisms influencing this essential enzyme but could also have therapeutic implications for optimizing CRISPR/Cas9-based precision genome editing. Engineered Cas9 variants could unintentionally retain or even amplify effector-like properties, contributing to off-target effects or immune responses in clinical settings. Mapping and eliminating PTM sites or motifs involved in host interactions may be essential for refining the safety of Cas9-based therapies and may offer means to fine-tune its activity, thereby improving control over both on-target and off-target effects.

## Acknowledgements

Research in U.S.'s laboratory, including past and ongoing work on Cas9, has been supported by EMBO (EMBO YIN IG3336, several EMBO Small Grants, and a Scientific Exchange Grant) and by TÜBİTAK (Grant Numbers 223Z048 and 121C230).

## Conflict of interest

The author declares no conflict of interest.

## References

1 Bolotin A, Quinquis B, Sorokin A & Ehrlich SD (2005) Clustered regularly interspaced short palindrome repeats (CRISPRs) have spacers of extrachromosomal origin. *Microbiology (Reading)* **151**, 2551–2561.
2 Mojica FJ, Diez-Villasenor C, Garcia-Martinez J & Soria E (2005) Intervening sequences of regularly spaced prokaryotic repeats derive from foreign genetic elements. *J Mol Evol* **60**, 174–182.
3 Pourcel C, Salvignol G & Vergnaud G (2005) CRISPR elements in *Yersinia pestis* acquire new repeats by preferential uptake of bacteriophage DNA, and provide additional tools for evolutionary studies. *Microbiology (Reading)* **151**, 653–663.
4 Jansen R, Embden JD, Gaastra W & Schouls LM (2002) Identification of genes that are associated with DNA repeats in prokaryotes. *Mol Microbiol* **43**, 1565–1575.
5 Barrangou R, Fremaux C, Deveau H, Richards M, Boyaval P, Moineau S, Romero DA & Horvath P (2007) CRISPR provides acquired resistance against viruses in prokaryotes. *Science* **315**, 1709–1712.
6 Brouns SJ, Jore MM, Lundgren M, Westra ER, Slijkhuis RJ, Snijders AP, Dickman MJ, Makarova KS, Koonin EV & van der Oost J (2008) Small CRISPR RNAs guide antiviral defense in prokaryotes. *Science* **321**, 960–964.
7 Garneau JE, Dupuis ME, Villion M, Romero DA, Barrangou R, Boyaval P, Fremaux C, Horvath P, Magadan AH & Moineau S (2010) The CRISPR/Cas bacterial immune system cleaves bacteriophage and plasmid DNA. *Nature* **468**, 67–71.
8 Cong L, Ran FA, Cox D, Lin S, Barretto R, Habib N, Hsu PD, Wu X, Jiang W, Marraffini LA *et al.* (2013) Multiplex genome engineering using CRISPR/Cas systems. *Science* **339**, 819–823.
9 Mali P, Yang L, Esvelt KM, Aach J, Guell M, DiCarlo JE, Norville JE & Church GM (2013) RNA-guided human genome engineering via Cas9. *Science* **339**, 823–826.
10 Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA & Charpentier E (2012) A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. *Science* **337**, 816–821.
11 Adli M (2018) The CRISPR tool kit for genome editing and beyond. *Nat Commun* **9**, 1911.
12 Doudna JA & Charpentier E (2014) Genome editing. The new frontier of genome engineering with CRISPR-Cas9. *Science* **346**, 1258096.
13 Hsu PD, Lander ES & Zhang F (2014) Development and applications of CRISPR-Cas9 for genome engineering. *Cell* **157**, 1262–1278.
14 Frangoul H, Altshuler D, Cappellini MD, Chen YS, Domm J, Eustace BK, Foell J, de la Fuente J, Grupp

S, Handgretinger R *et al.* (2021) CRISPR-Cas9 gene editing for sickle cell disease and beta-thalassemia. *N Engl J Med* **384**, 252–260.
15 Makarova KS, Zhang F & Koonin EV (2017) SnapShot: class 2 CRISPR-Cas systems. *Cell* **168**, 328.e1.
16 Makarova KS, Wolf YI, Alkhnbashi OS, Costa F, Shah SA, Saunders SJ, Barrangou R, Brouns SJ, Charpentier E, Haft DH *et al.* (2015) An updated evolutionary classification of CRISPR-Cas systems. *Nat Rev Microbiol* **13**, 722–736.
17 Bhattacharya S & Satpati P (2023) Insights into the mechanism of CRISPR/Cas9-based genome editing from molecular dynamics simulations. *ACS Omega* **8**, 1817–1837.
18 Anders C, Niewoehner O, Duerst A & Jinek M (2014) Structural basis of PAM-dependent target DNA recognition by the Cas9 endonuclease. *Nature* **513**, 569–573.
19 Jinek M, Jiang F, Taylor DW, Sternberg SH, Kaya E, Ma E, Anders C, Hauer M, Zhou K, Lin S *et al.* (2014) Structures of Cas9 endonucleases reveal RNA-mediated conformational activation. *Science* **343**, 1247997.
20 Jiang F, Zhou K, Ma L, Gressel S & Doudna JA (2015) Structural biology. A Cas9-guide RNA complex preorganized for target DNA recognition. *Science* **348**, 1477–1481.
21 Nishimasu H, Ran FA, Hsu PD, Konermann S, Shehata SI, Dohmae N, Ishitani R, Zhang F & Nureki O (2014) Crystal structure of Cas9 in complex with guide RNA and target DNA. *Cell* **156**, 935–949.
22 Palermo G, Chen JS, Ricci CG, Rivalta I, Jinek M, Batista VS, Doudna JA & McCammon JA (2018) Key role of the REC lobe during CRISPR-Cas9 activation by 'sensing', 'regulating', and 'locking' the catalytic HNH domain. *Q Rev Biophys* **51**, e91.
23 Macek B, Forchhammer K, Hardouin J, Weber-Ban E, Grangeasse C & Mijakovic I (2019) Protein post-translational modifications in bacteria. *Nat Rev Microbiol* **17**, 651–664.
24 Tu Z, Yang W, Yan S, Yin A, Gao J, Liu X, Zheng Y, Zheng J, Li Z, Yang S *et al.* (2017) Promoting Cas9 degradation reduces mosaic mutations in non-human primate embryos. *Sci Rep* **7**, 42081.
25 Ergunay T, Ayhan O, Celen AB, Georgiadou P, Pekbilir E, Abaci YT, Yesildag D, Rettel M, Sobhiafshar U, Ogmen A *et al.* (2022) Sumoylation of Cas9 at lysine 848 regulates protein stability and DNA binding. *Life Sci Alliance* **5**, e202101078.
26 Celen AB & Sahin U (2020) Sumoylation on its 25th anniversary: mechanisms, pathology, and emerging concepts. *FEBS J* **287**, 3110–3140.
27 Vertegaal ACO (2022) Signalling mechanisms and cellular functions of SUMO. *Nat Rev Mol Cell Biol* **23**, 715–731.
28 Pichler A, Fatouros C, Lee H & Eisenhardt N (2017) SUMO conjugation – a mechanistic view. *Biomol Concepts* **8**, 13–36.
29 Hay RT (2013) Decoding the SUMO signal. *Biochem Soc Trans* **41**, 463–473.
30 Sahin U, de The H & Lallemand-Breitenbach V (2022) Sumoylation in physiology, pathology and therapy. *Cells* **11**, 814.
31 Gareau JR & Lima CD (2010) The SUMO pathway: emerging mechanisms that shape specificity, conjugation and recognition. *Nat Rev Mol Cell Biol* **11**, 861–871.
32 Slaymaker IM, Gao L, Zetsche B, Scott DA, Yan WX & Zhang F (2016) Rationally engineered Cas9 nucleases with improved specificity. *Science* **351**, 84–88.
33 Costa TR, Felisberto-Rodrigues C, Meir A, Prevost MS, Redzej A, Trokter M & Waksman G (2015) Secretion systems in gram-negative bacteria: structural and mechanistic insights. *Nat Rev Microbiol* **13**, 343–359.
34 Meehl MA & Caparon MG (2004) Specificity of streptolysin O in cytolysin-mediated translocation. *Mol Microbiol* **52**, 1665–1676.
35 Gauthier A & Finlay BB (2001) Bacterial pathogenesis: the answer to virulence is in the pore. *Curr Biol* **11**, R264–R267.
36 Magassa N, Chandrasekaran S & Caparon MG (2010) *Streptococcus pyogenes* cytolysin-mediated translocation does not require pore formation by streptolysin O. *EMBO Rep* **11**, 400–405.
37 Bricker AL, Cywes C, Ashbaugh CD & Wessels MR (2002) NAD+-glycohydrolase acts as an intracellular toxin to enhance the extracellular survival of group a streptococci. *Mol Microbiol* **44**, 257–269.
38 Madden JC, Ruiz N & Caparon M (2001) Cytolysin-mediated translocation (CMT): a functional equivalent of type III secretion in gram-positive bacteria. *Cell* **104**, 143–152.
39 Schneewind O & Missiakas D (2014) Sec-secretion and sortase-mediated anchoring of proteins in gram-positive bacteria. *Biochim Biophys Acta* **1843**, 1687–1697.
40 Ribet D & Cossart P (2010) Pathogen-mediated posttranslational modifications: a re-emerging field. *Cell* **143**, 694–702.
41 Ribet D & Cossart P (2010) Post-translational modifications in host cells during bacterial infection. *FEBS Lett* **584**, 2748–2758.
42 Ribet D & Cossart P (2018) Ubiquitin, SUMO, and NEDD8: key targets of bacterial pathogens. *Trends Cell Biol* **28**, 926–940.
43 Sahin U, Ferhi O, Carnec X, Zamborlini A, Peres L, Jollivet F, Vitaliano-Prunier A, de The H & Lallemand-Breitenbach V (2014) Interferon controls SUMO availability via the Lin28 and let-7 axis to impede virus replication. *Nat Commun* **5**, 4187.

44 Everett RD, Boutell C & Hale BG (2013) Interplay between viruses and host sumoylation pathways. *Nat Rev Microbiol* **11**, 400–411.
45 Fan Y, Li X, Zhang L, Zong Z, Wang F, Huang J, Zeng L, Zhang C, Yan H, Zhang L *et al*. (2022) SUMOylation in viral replication and antiviral defense. *Adv Sci (Weinh)* **9**, e2104126.
46 Dassouki Z, Sahin U, El Hajj H, Jollivet F, Kfoury Y, Lallemand-Breitenbach V, Hermine O, de The H & Bazarbachi A (2015) ATL response to arsenic/interferon therapy is triggered by SUMO/PML/RNF4-dependent tax degradation. *Blood* **125**, 474–482.
47 Zamborlini A, Coiffic A, Beauclair G, Delelis O, Paris J, Koh Y, Magne F, Giron ML, Tobaly-Tapiero J, Deprez E *et al*. (2011) Impairment of human immunodeficiency virus type-1 integrase SUMOylation correlates with an early replication defect. *J Biol Chem* **286**, 21013–21022.
48 Cascales E, Buchanan SK, Duche D, Kleanthous C, Lloubes R, Postle K, Riley M, Slatin S & Cavard D (2007) Colicin biology. *Microbiol Mol Biol Rev* **71**, 158–229.
49 Klein A, Wojdyla JA, Joshi A, Josts I, McCaughey LC, Housden NG, Kaminska R, Byron O, Walker D & Kleanthous C (2016) Structural and biophysical analysis of nuclease protein antibiotics. *Biochem J* **473**, 2799–2812.
50 Eilebrecht S, Smet-Nocca C, Wieruszeski JM & Benecke A (2010) SUMO-1 possesses DNA binding activity. *BMC Res Notes* **3**, 146.
51 Le LHM, Ying L & Ferrero RL (2021) Nuclear trafficking of bacterial effector proteins. *Cell Microbiol* **23**, e13320.
52 Bierne H & Cossart P (2012) When bacteria target the nucleus: the emerging family of nucleomodulins. *Cell Microbiol* **14**, 622–633.
53 Hanford HE, Von Dwingelo J & Abu Kwaik Y (2021) Bacterial nucleomodulins: a coevolutionary adaptation to the eukaryotic command center. *PLoS Pathog* **17**, e1009184.
54 Ribet D, Hamon M, Gouin E, Nahori MA, Impens F, Neyret-Kahn H, Gevaert K, Vandekerckhove J, Dejean A & Cossart P (2010) *Listeria monocytogenes* impairs SUMOylation for efficient infection. *Nature* **464**, 1192–1195.
55 Sidik SM, Salsman J, Dellaire G & Rohde JR (2015) Shigella infection interferes with SUMOylation and increases PML-NB number. *PLoS One* **10**, e0122585.
56 Mak AN, Bradley P, Bogdanove AJ & Stoddard BL (2013) TAL effectors: function, structure, engineering and applications. *Curr Opin Struct Biol* **23**, 93–99.
57 Cermak T, Doyle EL, Christian M, Wang L, Zhang Y, Schmidt C, Baller JA, Somia NV, Bogdanove AJ & Voytas DF (2011) Efficient design and assembly of custom TALEN and other TAL effector-based constructs for DNA targeting. *Nucleic Acids Res* **39**, e82.