The Journal of Chemical Physics ARTICLE scitation.org/journal/jcp

Disruption of electrostatic contacts in the HNH nuclease from a thermophilic Cas9 rewires allosteric motions and enhances high-temperature DNA cleavage

Cite as: J. Chem. Phys. 157, 225103 (2022); doi: 10.1063/5.0128815
Submitted: 30 September 2022 • Accepted: 29 November 2022 •
Published Online: 14 December 2022

Helen B. Belato,[1] Carmelissa Norbrun,[1] Jinping Luo,[2] Chinmai Pindi,[3] Souvik Sinha,[3] Alexandra M. D'Ordine,[1] Gerwald Jogl,[1] Giulia Palermo,[3,a)] and George P. Lisi[1,a)]

AFFILIATIONS

[1] Department of Molecular Biology, Cell Biology and Biochemistry, Brown University, Providence, Rhode Island 02912, USA
[2] Brown University Transgenic Mouse and Gene Targeting Facility, Providence, Rhode Island 02903, USA
[3] Departments of Bioengineering and Chemistry, University of California Riverside, Riverside, California 92521, USA

Note: This paper is part of the JCP Special Topic on New Views of Allostery.
[a)] **Authors to whom correspondence should be addressed:** gpalermo@engr.ucr.edu and george_lisi@brown.edu

ABSTRACT

Allosteric signaling within multidomain proteins is a driver of communication between spatially distant functional sites. Understanding the mechanism of allosteric coupling in large multidomain proteins is the most promising route to achieving spatial and temporal control of the system. The recent explosion of CRISPR-Cas9 applications in molecular biology and medicine has created a need to understand how the atomic level protein dynamics of Cas9, which are the driving force of its allosteric crosstalk, influence its biophysical characteristics. In this study, we used a synergistic approach of nuclear magnetic resonance (NMR) and computation to pinpoint an allosteric hotspot in the HNH domain of the thermostable *Geo*Cas9. We show that mutation of K597 to alanine disrupts a salt-bridge network, which in turn alters the structure, the timescale of allosteric motions, and the thermostability of the *Geo*HNH domain. This homologous lysine-to-alanine mutation in the extensively studied mesophilic *S. pyogenes* Cas9 similarly alters the dynamics of the *Sp*HNH domain. We have previously demonstrated that the alteration of allostery via mutations is a source for the specificity enhancement of *Sp*Cas9 (e*Sp*Cas9). Hence, this may also be true in *Geo*Cas9.

INTRODUCTION

The use of CRISPR-Cas9 technology has rapidly expanded the possibility for genome editing in a variety of organisms.[1] As a form of bacterial immune response, Cas9 proteins vary from species-to-species in their size and perhaps the molecular factors that dictate cleavage and specificity.[2,3] A thermostable Cas9 from the bacterium *Geobacillus stearothermophilus* (*Geo*Cas9) maintains a domain architecture similar to canonical *Streptococcus pyogenes* Cas9, utilizing a single-guide RNA (sgRNA) to recognize, unwind, and cleave a double-stranded DNA target after recognition of its protospacer adjacent motif (PAM).[4] *GeoCas9* contains a truncated recognition (REC) lobe that docks the RNA:DNA hybrid and similarly sized PAM interacting (PI), HNH, and RuvC domains that work concertedly to cleave double-stranded DNA. Each of the nuclease active sites is spatially distinct from the PAM recognition site of *Geo*Cas9, necessitating structural and dynamic changes that propagate DNA binding information to allosterically couple the subdomains.

A significant body of work has investigated the mechanisms of allostery and specificity in attempts to enhance spatial and temporal control of the canonical mesophilic *S. pyogenes* Cas9 (*Sp*Cas9).

Experimental[5,6] and computational[7–9] studies of *Sp*Cas9 suggested the catalytic HNH domain to be the core of an allosteric relay and a driver of specificity.[5,6] Expanding mechanistic studies to a highly stable thermophilic Cas9 with noted activity in human plasma have the potential to facilitate genome editing *in vivo*, especially considering that the architecture of the *Geo*Cas9 HNH domain (*Geo*HNH) is very similar to that of *Sp*HNH.[10] It has also been demonstrated that single lysine-to-alanine point mutations in *Sp*HNH reduced off-target effects while maintaining on-target cleavage *in vitro* and *in vivo*.[11] One lysine, K855, is conserved in *Geo*HNH and many other Cas9 HNH domains, including *Nme*HNH (*N. meningitidis*), *Sa*HNH (*S. aureus*), *Cj*HNH (*C. jejuni*), and *St*HNH (*S. thermophilus*). It was hypothesized that a specificity-enhancing K855A mutation in *Sp*Cas9 weakened the interaction between HNH and the target DNA, thereby exerting an electrostatic control favoring the re-hybridization of DNA.[11] Recent biophysical studies of this mutant showed that K855A alters *Sp*HNH dynamics and allosteric signal propagation to the rest of the protein, demonstrating that the combination of allosteric and electrostatic effects is critical to optimizing the specificity of *Sp*Cas9.[8] This analogous residue in *Geo*HNH, K597, might influence the dynamics of the domain, being the center of a highly flexible site with the lowest nuclear magnetic resonance (NMR) order parameter (S^2) of any residue in *Geo*HNH.[10] Despite the different intrinsic dynamic profiles of *Sp*HNH (μs–ms) and *Geo*HNH (ps–ns) that may be related to chemical signal propagation,[10] we wondered what effect mutation of this conserved lysine would have on *Geo*Cas9 function and *Geo*HNH structure and dynamics. Here, we examine the *Geo*HNH K597A variant and show the importance of this residue in the maintenance of wild-type (wt) structure, dynamics, thermostability, and DNA cleavage function. Our findings reveal that the K597A mutation modulates the protein motions and thermostability of *Geo*HNH, disrupting a network of critical salt-bridge interactions in its core. We also show that K597A alters the dynamics and structure of *Geo*Cas9, analogously to the specificity-enhancing K855A mutation in *Sp*Cas9,[8] suggesting that lysine-to-alanine mutations might be critical for improving the function and specificity of *Geo*HNH and other Cas9 HNH domains.

RESULTS

Mutation of K597 induces structural changes in *Geo*HNH

We introduced a K597A mutation into an isolated *Geo*HNH construct and used solution NMR to determine any associated structural changes via chemical shift perturbations (CSP) in $^{1}H–^{15}N$ HSQC NMR spectra (Fig. 1). Resonances with significant CSP or those broadened beyond detection are plotted on the *Geo*HNH structure in Fig. 1. These resonances are highly localized to a loop spanning residues 583–595 and the antiparallel beta sheet opposite the K597A mutation that is important for the interaction with the target DNA and contains the catalytic histidine, H582. NMR line broadening in the 583–595 loop is indicative of a dynamic change consistent with ms timescale motions that are not found in wild-type (wt)-*Geo*HNH.

Further structural changes derived from per-residue C_α, C_β, and C_o CSP between wt- and K597A *Geo*HNH (Fig. S1) localize to regions of the domain that interface with RuvC and REC (Fig. S2). Nonetheless, the predominantly α-helical structure of *Geo*HNH is maintained, as confirmed by circular dichroism (CD) (Fig. S3) and x-ray crystallography (Fig. S1). NMR-derived secondary structure analysis from C_α and C_β chemical shifts show that K597A *Geo*HNH is very similar to wt-, with a minimal change in helical content due perhaps to the high helical propensity of the alanine residue introduced by mutation.[12] The global structural changes in *Geo*HNH due to the mutation are modest and do not appear to define the K597A variant differently than wt-*Geo*HNH. Since protein dynamics are the driver of allosteric commination in several Cas9s,[6,8,13–16] we wondered if instead, the flexibility of this variant was substantially altered.

The timescale of intrinsic GeoHNH dynamics is altered by the K597A mutation

To assess the dynamics of K597A *Geo*HNH relative to those of wt-*Geo*HNH, we measured longitudinal (R_1), transverse (R_2), and heteronuclear $^{1}H–[^{15}N]$ NOE relaxation parameters (Fig. S4).

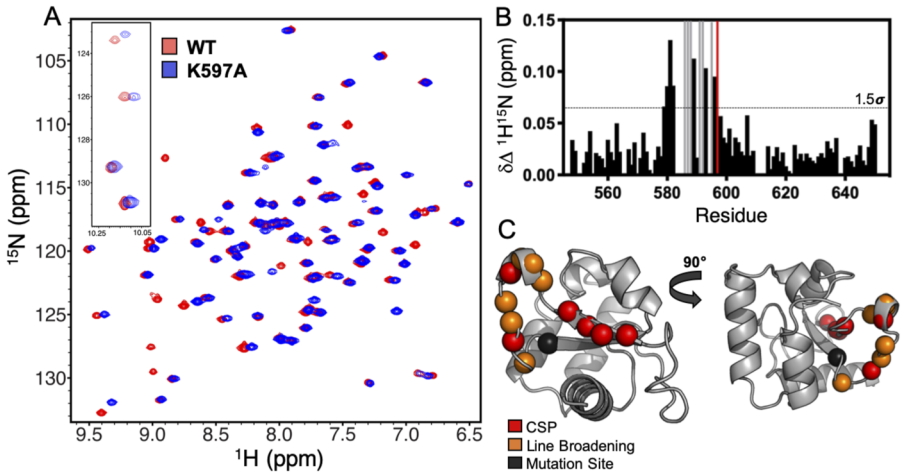

FIG. 1. (a) 600 MHz $^{1}H–^{15}N$ TROSY HSQC NMR spectra of wt-GeoHNH (red) and K597A GeoHNH (blue). (b) NMR chemical shift perturbations (CSP) caused by K597A mutation are plotted for each residue. Residues that have broadened beyond detection are marked with a gray bar. The mutation site is marked with a red bar. (c) Residues with CSP above 1.5 standard deviations of the 10% trimmed mean are shown as red spheres on the *Geo*HNH structure. Residues that have broadened beyond detection are shown as orange spheres. The mutation site is shown as a black sphere.

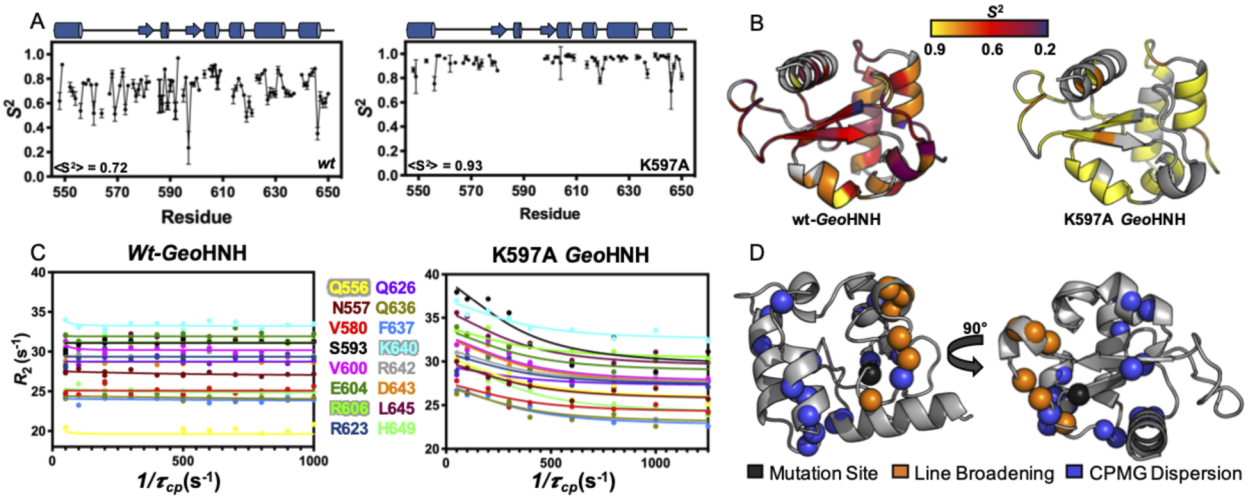

FIG. 2. Mutation-induced shift in the dynamic profile of *Geo*HNH (a) Order parameters (S^2) for wt- and K597A *Geo*HNH determined from Model-free analysis of T_1, T_2, and ^{1}H–[^{15}N] NOE measurements. The average S^2 for the entire domain is also indicated. Cartoons of the *Geo*HNH secondary structures are shown above the plot in blue. (b) S^2 is mapped onto the wt- and K597A *Geo*HNH structures, where the heat map reports the magnitude of S^2 according to the legend. (c) CPMG relaxation dispersion profiles for residues in wt- and K597A *Geo*HNH collected at 25 °C at 600 MHz. Curved profiles for K597A *Geo*HNH indicate the presence of detectable ms motions at these sites. (d) Residues with CPMG dispersion are mapped onto the *Geo*HNH structure as blue spheres. The mutation site is shown as a black sphere. Line broadened residues in ^{1}H–^{15}N TROSY HSQC spectra of K597A *Geo*HNH are shown as orange spheres, which also qualitatively indicate ms motions.

We used these parameters to conduct a Model-free analysis and determine per-residue order parameters (S^2) for K597A *Geo*HNH. We previously measured very low S^2 across wt-*Geo*HNH, suggestive of substantial flexibility on the ps-ns timescale.[10] Interestingly, values of S^2 for K597A *Geo*HNH are notably higher than those of wt-*Geo*HNH (Fig. 2), affecting the entire structure and indicating that its fast motion arises primarily from global tumbling of the protein in solution. We, therefore, carried out Carr-Purcell-Meiboom-Gill (CPMG) relaxation dispersion NMR experiments to assess the influence of μs–ms timescale motions of K597A *Geo*HNH, which were not observed in our previous study of wt-*Geo*HNH. We now find 16 residues with CPMG relaxation dispersion profiles in K597A *Geo*HNH (Fig. 2, Table S2), indicating a mutation-induced shift in the dynamic profile of *Geo*HNH. Conformational exchange rates, k_{ex}, derived from dual-field CPMG experiments can be globally fit to $k_{ex} = 1101 \pm 208\ s^{-1}$ (Fig. 2, Table S2). Data fitting indicates a two-site exchange process, with the population of the minor state being ~13%. The concerted millisecond timescale motion of these sites appears to be due to a disruption of stabilizing salt bridges by the lysine-to-alanine mutation.

Disruption of GeoHNH salt bridge network modulates its thermostability

K597 is at the center of a salt bridge network that stabilizes the 583–589 loop and neighboring helices in the wt-*Geo*Cas9 [Fig. 3(a), left]. The polar contact network of K597 includes (1) D592, which forms a salt bridge with the catalytic H582, (2) Y594, which forms salt bridges with K641 and E555, and (3) K641, which forms salt bridges with E555, S638, and L645. We speculated that disruption of this salt bridge network could play a role in altering the dynamics and thermostability of K597A *Geo*HNH. This hypothesis is strengthened by the critical role that salt-bridge interactions exert during the conformational activation of the HNH domain in *Sp*Cas9.[17] A recent evolutionary study also reported a substantial increase in the salt bridge networking of thermostable adenylate kinases (Adk) relative to their mesophilic homologs.[18] In fact, addition of salt bridges to mesophilic Adks, or their removal from thermophilic Adks, modulated the T_m of the enzyme substantially. Consistent with this hypothesis, our x-ray structure of the K597A mutant reveals a disruption of the salt bridge network involving the 583–589 loop [Fig. 3(a), right]. As a result of these structural perturbations, we detect altered dynamics and a much lower T_m (~42 °C) for K597A *Geo*HNH relative to that of wt-*Geo*HNH [~62 °C, Fig. 3(b)].

The structural dynamics and the thermostability of wt- and K597A *Geo*HNH were further assessed through molecular dynamics (MD) simulations at different temperatures. MD simulations were carried out starting from the x-ray structure of the wt-*Geo*HNH, also introducing the K597A mutation (see Methods). Upon mutation, we observed increased backbone fluctuations (i.e., the root-mean-square fluctuations—RMSF—of the Cα atoms) in the mutant compared with the wt-*Geo*HNH at 25 °C (Fig. S5). Detailed analysis of the intra-domain contacts was performed for wt- and K597A *Geo*HNH at 25 °C (i.e., wt-*Geo*HNH – K597A *Geo*HNH), which reveals a significant reduction of stable contacts in the K597A mutant, while the wt-*Geo*HNH gains stable contacts [Figs. 3(c) and 3(d)]. Specifically, we observed a reduction of contacts in the core region of the salt bridge network formed by K597, in line with the x-ray structure of the K597A mutant [Fig. 3(a), right] and with the hypothesis derived by NMR analysis (*vide supra*). To further investigate the effect of K597A on the dynamics and salt bridge

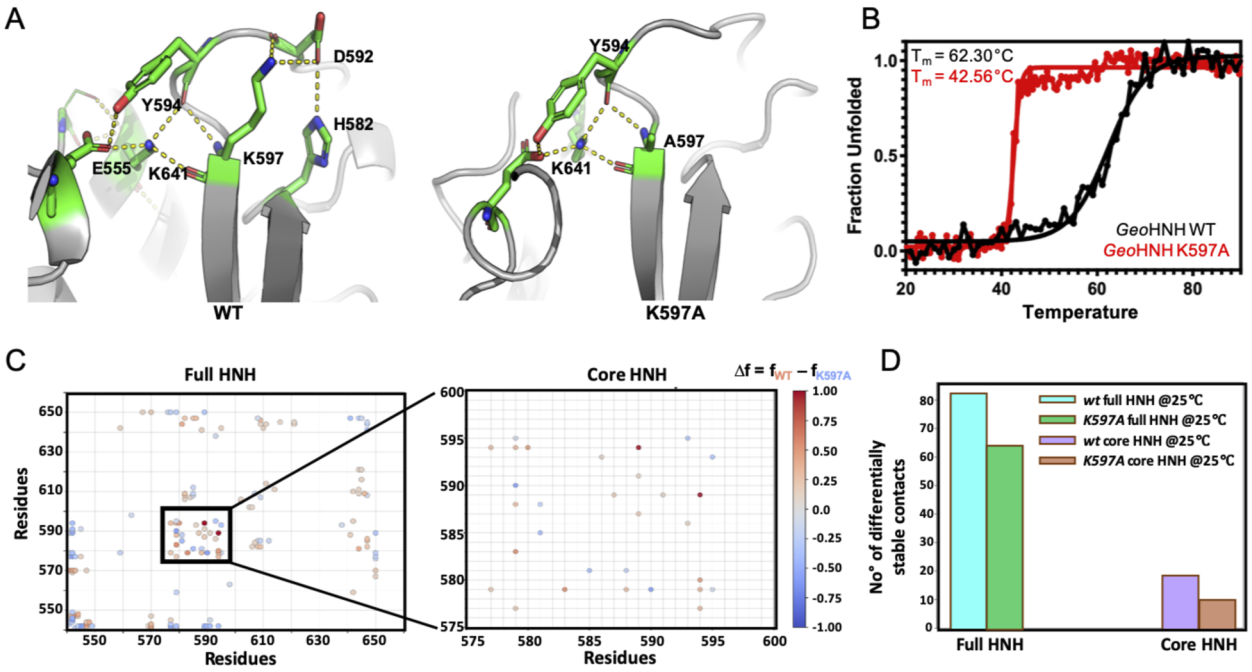

FIG. 3. K597A is a control switch for thermostability in *Geo*HNH. (a) X-ray crystal structures of wt-*Geo*HNH and K597A *Geo*HNH highlighting K597 (left) at the center of a salt bridge network in *Geo*HNH. These critical contacts are lost with A597 (right). (b) Temperature-dependent CD spectra reveal that K597A *Geo*HNH unfolds at a much lower temperature (~42 °C) than wt-*Geo*HNH (~62 °C). (c) Differential contact analysis map between the wild-type *Geo*HNH and its K597A mutant, showing reduction of stable contacts in K597A *Geo*HNH. Values ranging from 0 to 1 (red) describe contacts that are more stable in wt-*Geo*HNH, while values from 0 to −1 (blue) describe contacts that are more stable in the mutant. A close-up view highlights the K597A proximal core domain (d) Total number of differentially stable contacts in the full *Geo*HNH domain and in the K597A proximal core domain at 25 °C.

network of *Geo*HNH, we performed MD simulations of the wt- and mutant at their respective melting temperatures (i.e., T_m = 65 °C for the wt, and T_m = 45 °C for K597A). The differential RMSF of the Cα atoms between the room and melting temperatures of the wt- and K597A *Geo*HNH is shown in Fig. S6. Both the wt- and mutant *Geo*HNH show increased fluctuations at melting temperatures, in agreement with NMR experiments. However, the extent of the fluctuations in wt-*Geo*HNH are comparatively less substantial at the reported T_m than what is observed in the K597A mutant, consistent with a gradual transition from folded to unfolded state [see Fig. 3(a)].

To further investigate the effect of temperature on K597A *Geo*HNH, we compared the temperature dependence of NMR chemical shifts ($\delta^{15}N/\Delta T$ and $\delta^{1}H/\Delta T$) between wt- and K597A *Geo*HNH. Using these slopes, we identified residues with greater than expected temperature fluctuations (i.e., outside 95% confidence boundaries) and mapped them onto the *Geo*HNH structure [Fig. 4(c)]. Interestingly, the majority of these residues are clustered on the 640–645 helix and adjacent loop that was determined to exhibit a global k_{ex} of 1101 ± 208 s^{-1} in CPMG experiments. This region also appears most heavily influenced by the disruption of salt bridges. Collectively, these data suggest that the *Geo*HNH structure and the timescale of its intrinsic dynamics are important determinants of thermal stability. K597 appears to be a control switch for thermostability in *Geo*HNH and in *Geo*Cas9, which remarkably displays an unfolding transition that is ~15 °C *higher* than that of wt-*Geo*Cas9 [Fig. 4(a)]. This observation may be explained by the unique motions of K597A *Geo*HNH that strengthen contacts with the adjacent RuvC and REC domains. A recent study of *Sp*HNH revealed that allosteric signal transmission between HNH and the adjacent RuvC and REC domains required intrinsic dynamics of HNH occurring on the microsecond-to-millisecond timescale.[7] In our current study of K597A *Geo*HNH, we observe a similar dynamic trend and, therefore, speculate that communities of residues with micro-millisecond motions located at the *Geo*HNH domain interfaces facilitate stronger communication with neighboring domains. An enhancement of domain–domain contacts may negate the dissolution of salt bridge interactions observed in the isolated *Geo*HNH, resulting in a net stabilization of full-length K597A *Geo*Cas9 that is reflected in the increased thermostability. Such a thermostability switch is rare, and the molecular factors governing these opposing results require extremely careful analysis in follow-up biophysical studies.

To test the effect of the K597A mutation on the function of full-length *Geo*Cas9, we assayed the *in vitro* cleavage ability of wt- and K597A *Geo*Cas9 against a target DNA strand after incubation at four temperatures (37, 60, 75, and 85 °C). Figure 4(a) shows that both proteins are able to cleave target DNA at 37 and 60 °C, suggesting that the K597A mutation does not hamper full-length *Geo*Cas9 cleavage of on-target DNA. Both wild-type and mutant

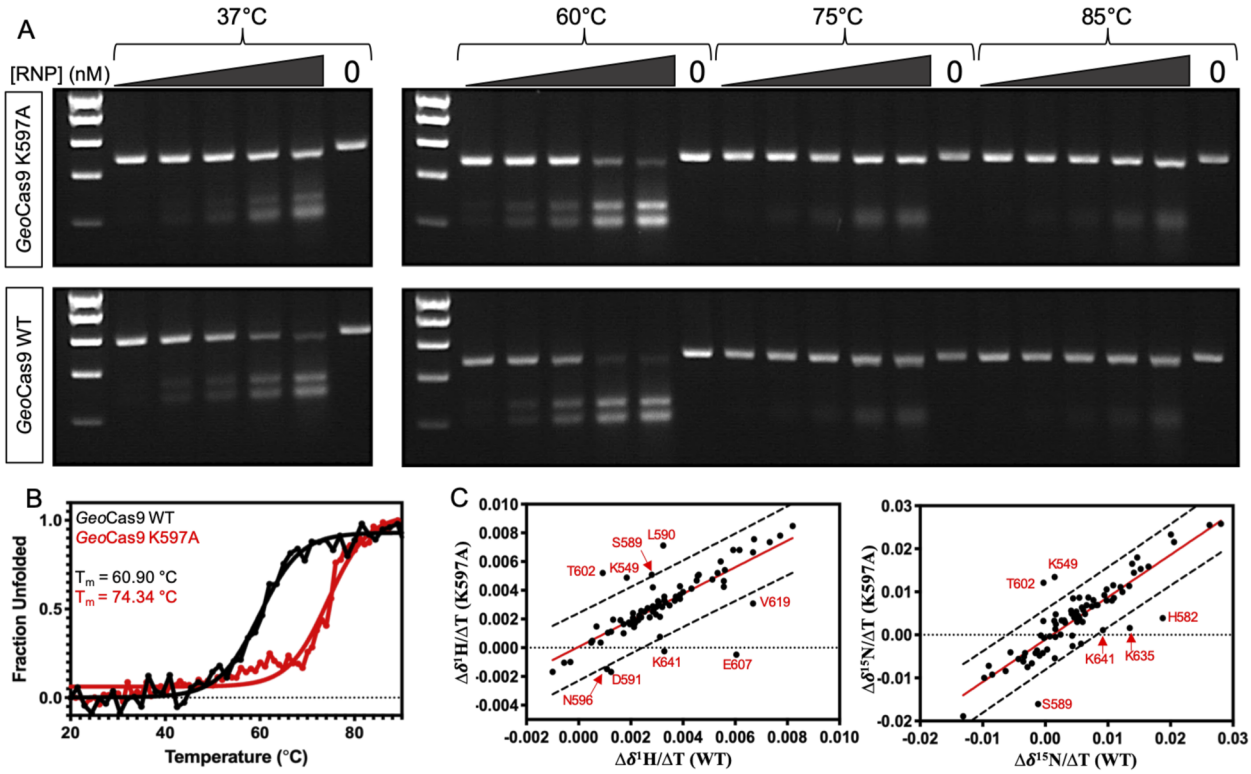

FIG. 4. K597A is a control switch of thermostability in *Geo*Cas9. (a) RNPs of wt- or K597A *Geo*Cas9 and sgRNA were incubated at 37, 60, 75, or 85 °C for 10 min, after which the RNPs were used for individual cleavage reactions at 37 °C for 30 min. *in vitro* cleavage assays indicate that the K597A mutation does not effect on-target DNA cleavage at 37 and 60 °C. K597A *Geo*Cas9 is also able to cleave target DNA at 75 and 85 °C more efficiently than wt. RNP concentration in each lane (left-to-right) is 100, 200, 300, 600, 900, and 0 nM at each temperature tested. Molecular weight markers on agarose gels (top-to-bottom) are 1000, 800, 600, 400, and 200 base pairs. (b) Temperature-dependent CD spectra reveal that a K597A *Geo*Cas9 (i.e., full-length) mutant is more thermostable than wt-*Geo*Cas9 by ~15 °C. (c) Correlations between temperature-dependent proton and nitrogen chemical shifts in ^{1}H–^{15}N TROSY HSQC spectra of wt- and K597A *Geo*HNH are shown for the nitrogen (right) and proton (left) dimensions. Temperature-dependent shifts outside 95% confidence boundaries are highlighted in red.

*Geo*Cas9 cleave target DNA more efficiently closer to its physiological temperature (60 °C). At low RNP concentrations, no cleavage occurs at elevated temperatures (75 and 85 °C) consistent with the report of Dounda and co-workers,[4] but at higher RNP concentrations we note modest cleavage activity by K597A *Geo*Cas9 and only trace amounts of cleavage by wt-*Geo*Cas9. The observation of even this small amount of cleavage for the wt-*Geo*Cas9 was unexpected but could be rationalized by the possibility that its guide RNA may be highly stabilizing to the RNP complex, enabling activity at temperatures higher than the T_m of the apo protein [Fig. 4(b)].

Interestingly, the homologous lysine-to-alanine mutation in *Sp*Cas9 was found to modestly reduce the pK_a of the catalytic histidine, reducing activity by making the histidine a weaker base. We investigated the chemistry of the catalytic histidine in *Geo*HNH and noted no change in pK_a of the active site His or other histidines in the domain (Figure S7). This suggests that any change in cleavage activity between wt-*Geo*Cas9 and K597A *Geo*Cas9 is not due to the chemistry of the active site, but rather due to changes in *Geo*HNH dynamics and intradomain contacts. We should note, however, that substrate-induced effects on the His pK_a were not investigated for the variants in question and cannot be ruled out. Overall, the NMR, computational, and biophysical data reported above support the hypothesis that the identity of the residue at position 597 controls chemical contacts affecting the thermostability of both *Geo*HNH and *Geo*Cas9 and the dynamics of *Geo*HNH.

DISCUSSION

A single-point mutation of a dynamic residue detected by NMR generates a highly thermostable *Geo*Cas9 (T_m ~ 74 °C). The binding of sgRNA likely increases its T_m further, allowing for target DNA cleavage at 75 and 85 °C. Paradoxically, the K597A mutation within the isolated *Geo*HNH domain has a clear destabilizing effect that is evident not only by CD spectroscopy but also in the way *Geo*HNH is purified for NMR studies. Unlike wt-*Geo*HNH, the K597A variant is insoluble after cell lysis and must be purified from the pellet with a more complex unfolding-and-refolding method. We observed little-to-no protein in the soluble lysate. The instability of the isolated *Geo*HNH domain is likely due to the disruption of the lysine-centered salt bridge network that is no longer promot-

ing tight electrostatic contacts within the protein. This is supported by MD simulations that highlight a loss of stable contacts in the mutant.

However, enhanced conformational sampling of K597A *Geo*HNH resulting from the loss of this salt bridge network, evidenced by a measurable 13% excited state population from CPMG experiments (no excited state is observed in wt-*Geo*HNH) may facilitate a more energetically favorable full-length *Geo*Cas9 structure through strong dynamic interdomain correlations of *Geo*HNH with adjacent *Geo*Cas9 domains, namely the Rec lobe and RuvC nuclease (Fig. S8). A similar scenario has been proposed in studies of the allosteric crosstalk between domains in canonical mesophilic Cas9s.[8]

Traditional models of allostery involving molecular pathways are characterized by higher energy, large-scale conformational changes occurring on slower (i.e. millisecond) timescales.[19] Although the influence of faster timescale protein dynamics on allostery has been noted,[20–22] including in wt-*Geo*HNH, there are comparatively fewer examples. Here, we revealed how the internal motions of the HNH domain and its relationship to the rest of the *Geo*Cas9 structure can modulate its thermostability. By manipulating an allosteric residue in *Geo*HNH without altering its active site, we modulated its allosteric motions, which are dominated by fast internal dynamics (in the wt enzyme) and are intimately linked to conformational entropy and revealed an opportunity for expanding *Geo*Cas9 nuclease activity and thermostability. Coupled communication between wt-*Geo*HNH and the rest of the *Geo*Cas9 domains is likely achieved by global redistribution of the fast internal motions and conformational entropy. Mutation of K597, which had the lowest NMR order parameter in a domain with already globally low S^2 on average, disrupted the intrinsic dynamics of *Geo*HNH, resulting in an overall shift to slower dynamics and the appearance of a lowly populated excited state. The K597A mutation in *Geo*Cas9 alters the dynamics of HNH similar to what was observed with the homologous K855A mutation in *Sp*Cas9, where the altered allosteric signaling was shown to be a signature for specificity enhancement. At present, we have not observed a similar specificity enhancement in K597A *Geo*Cas9. A battery of further specificity tests against target DNAs is being conducted *in vitro* and *in vivo* with wt-*Geo*Cas9, the K597A variant, and a library of other variants for a follow-up study. Taken together, our findings suggest that this lysine-to-alanine mutation might be critical for improving the function of *Geo*HNH and other Cas9 HNH domains. Further biophysical characterization of adjacent domains *Geo*Rec3 and *Geo*RuvC will increase our understanding of how the altered structure and dynamics of *Geo*HNH propagates to these regions to modulate the structural, dynamics, and functional properties of *Geo*Cas9, including its specificity. Such insight could lead to further engineering of *Geo*Cas9 that shifts its stability and activity profile into hyperthermophilic ranges.

MATERIALS AND METHODS

K597A *Geo*HNH purification

The K597A mutation was introduced into a previously reported construct of the *Geo*HNH domain (residues 511–662).[10] The plasmid was transformed into BL21 (DE3) cells (New England BioLabs). Isotopically labeled samples were grown in M9 minimal media containing $CaCl_2$, $MgSO_4$, and MEM vitamins, and supplemented with[15] N ammonium chloride and ^{13}C glucose (1.0 and 2.0 g/l, respectively; Cambridge Isotope Laboratories). Cells were induced with 1 mM IPTG after reaching an OD_{600} of 0.8–1.0 and grown for 4 h at 37 °C post induction. The cells were harvested by centrifugation, resuspended in a denaturing lysis buffer (10 mM Tris–HCl, 100 mM sodium phosphate, 6M guanidine hydrochloride (GuHCl), 10 mM reduced glutathione, and 1 mM PMSF, pH 8.0), lysed by ultrasonication, and cell debris was removed by centrifugation.

The resulting supernatant was poured into a gravity column packed with 10 ml of Ni-NTA agarose beads. The column was washed with the initial lysis buffer, followed by a gradient of the same buffer without GuHCl and reduced glutathione over 100 ml. Elution of K597A *Geo*HNH in its denatured form was performed with 1 column volume of a buffer containing 10 mM Tris–HCl, 100 mM sodium phosphate, and 250 mM imidazole (pH 8.0). K597A *Geo*HNH was refolded by dilution via dropwise addition of the 10 ml eluent into 100 ml of a refolding buffer containing 10 mM Tris–HCl, 100 mM sodium phosphate, and 750 mM arginine (pH 8.0). The refolded protein was dialyzed exhaustively against a buffer containing 20 mM HEPES, 150 mM KCl, 1 mM DTT, and 1 mM EDTA (pH 7.4). The sample was concentrated to 10 ml, loaded on a Ni-NTA column and pure K597A *Geo*HNH was collected from the flow-through and dialyzed into NMR buffer (20 mM HEPES, 80 mM KCl, 1 mM DTT, 1 mM EDTA, 5% glycerol, and 10% (v/v) D_2O at pH 7.4).

Full-length K597A *Geo*Cas9 purification

The K597A mutation was introduced into a *Geo*Cas9 plasmid available on Addgene (No. 87700). wt-*Geo*Cas9 and the K597A variant used for CD spectroscopy and *in vitro* cleavage assays were purified as previously described.[4]

NMR spectroscopy

Two-dimensional ^1H–^{15}N HSQC NMR spectra were collected at 600 MHz and NMR spin relaxation experiments were carried out at 600 and 850 MHz on Bruker Avance NEO and Avance III HD spectrometers, respectively. All NMR spectra were processed with NMRPipe[23] and analyzed in NMRFAM-SPARKY.[24] CPMG experiments were adapted from the report of Palmer and co-workers[25] with a constant relaxation period of 20 ms and ν_{CPMG} values of 50 (×2), 100, 200, 300 (×2), 400, 600, 800 (×2), 1000, and 1250 Hz. Relaxation dispersion curves were generated, and exchange parameters were obtained from global fits of the data carried out with RELAX[26] using the R2eff, NoRex, and CR72 models, as well as in-house fitting in GraphPad Prism via

$$R_{2,eff} = R_2^0 + R_{ex} * \left(1 - \left(\frac{2 * x * \tan h\left(k_{ex} / \tau_{cp}{}^2\right)}{k_{ex}}\right)\right),$$

where R_{ex} is

$$R_{ex} = \frac{p_a p_b \Delta\omega^2}{k_{ex}}.$$

An error was determined with replicate measurements. Longitudinal and transverse relaxation rates were measured with relaxation times of 20 (×2), 60 (×2), 100, 200, 400, 600(×2), 800, and 1200 ms for T_1 and 16.96, 33.92(×), 50.88(×2), 67.84(×2), 84.8, and 101.76(×2) *ms* for T_2. Peak intensities were quantified in Sparky and the resulting decay profiles were analyzed in Mathematica with errors determined from the fitted parameters. Steady-state ^{1}H–[^{15}N] NOE was measured with a 6 s relaxation delay followed by a 3 s saturation (delay) for the saturated (unsaturated) experiments. All relaxation experiments were carried out in a temperature-compensated interleaved manner. Model-free analysis using the Lipari-Szabo formalism was carried out on dual-field NMR data in RELAX[26] with fully automated protocols.

Histidine ε^1 sidechain carbon chemical shifts were measured via HMQC (*hmqcphpr* pulse sequence) using NMR samples that were successively buffer exchanged into NMR buffer at pH values ranging from pH 6–8. The ε_1 carbon sidechain of the catalytic H592 in *Geo*HNH was assigned by mutagenesis to alanine. As previously reported for *Sp*HNH,[27] the H840 pK_a was determined by fitting the ^{1}H chemical shift trajectories to a modified Henderson–Hasselbalch equation, which can be found in Fig. S7.

Circular dichroism (CD) spectroscopy

*Geo*Cas9 and *Geo*HNH proteins were buffers exchanged into CD buffer (20 mM sodium phosphate, pH 7.5), and samples were diluted to 1 μM and loaded into a 1 mm quartz cuvette (JASCO instruments). A CD spectrum was first measured between 200 and 250 nm, after which the sample was progressively heated from 20 to 90 °C in 1.5 °C increments while ellipticity was monitored at 222 and 208 nm. Measurements made on a CD buffer baseline were subtracted from the sample measurements.

X-ray crystallography

K597A *Geo*HNH protein purified as described above was crystallized by sitting drop vapor diffusion at room temperature. K597A *Geo*HNH at 11 mg/ml in 10 mM HEPES, 100 mM KCl, pH 7.5 was mixed with an equal volume of crystallizing condition from PACT Premier screen well C11: 0.2M calcium chloride hexahydrate, 0.1M HEPES pH 7.0, 20% polyethylene glycol 6000. Crystals were cryoprotected in crystallizing conditions containing 30% ethylene glycol. Diffraction images were collected at the NSLS-II AMX beamline at Brookhaven National Laboratory under cryogenic conditions. Images were processed using XDS[28] and Aimless[29] in CCP4 before the structure was solved by molecular replacement with Phaser in Phenix.[30] The wt-*Geo*HNH domain structure[10] (residues 33–140, PDB ID: 7MPZ) was used as the search model. The K597A *Geo*HNH structure was finished by iterative rounds of manual building in Coot[31] and refinement in Phenix.

In vitro DNA cleavage assay

*Geo*Cas9 sgRNA templates containing 21nt spacers targeting the mouse *Tnnt2* gene locus were introduced into EcoRI and BamHI sites in pUC57 (Genscript). The plasmid was transformed into BL21(DE3) cells (New England BioLabs) and subsequent restriction digest of the plasmid DNA was done using the BamHI restriction enzyme from New England BioLabs according to the manufacturer's instructions. Linearized plasmid DNA was immediately purified using the DNA Clean and Concentrator-5 kit from Zymo Research according to the manufacturer's instructions. RNA transcription was performed *in vitro* with the HiScribeTM T7 High Yield RNA Synthesis Kit from New England BioLabs. DNA substrates containing the target cleavage site (479 base pairs) were made by polymerase chain reaction (PCR) using mouse genomic DNA as a template and primer pairs (5′CAAAGAGCTCCTCGTCCAGT3′) and (5′ATGGACTCCAGGACCCAAGA3′) followed by a column purification (NucleoSpin® Gel and PCR Clean-up kit, Macherey-Nagel). Guide RNA sequence and DNA target can be found in Table S1. For the *in vitro* activity assay [Fig. 4(a)], RNP formation was achieved by incubating 3 μM *Geo*Cas9 (wild-type or K597A mutant) and 3 μM sgRNA at either 37, 60, 75, or 85 °C for 10 min in the reaction buffer (20 mM Tris, 100 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, 5% glycerol, pH 7.5). The 10 μl cleavage reactions were set up by mixing RNP at varying concentrations with 149 ng of PCR products on ice and then incubated at 37 °C for 30 min. The reaction was quenched by adding 1 μl of proteinase K (20 mg/ml) to each reaction for an incubation at 56 °C for 10 min. 6× DNA loading buffer was added to each reaction and 10 μl reaction mixtures/lane was loaded onto an agarose gel.

Molecular dynamics (MD) simulations

Molecular dynamics (MD) simulations were based on the x-ray crystal structure of *Geo*HNH deposited in the Protein Data Bank (PDB: 7MPZ). This structure was used as a model of the wt-*Geo*HNH, while the K597A mutant was generated by introducing the alanine residue in place of lysine at position 597 in the same structure. Both systems were solvated with explicit water molecules and neutralized by adding Cl^- ions, resulting in a cubic periodic box of ~69.3 Å · ~68.1 Å · ~66.9 Å. The AMBER ff14SB force field[32] was employed to describe the protein, and the three-site TIP3P model[33] was used to describe the water molecules. The LINCS algorithm was used to restrain all bonds involving hydrogens.[34] A particle mesh Ewald method with a 10 Å cutoff was used to calculate electrostatics. The equations of motion were integrated with the leapfrog Verlet algorithm with a time step of 2 fs. Both wt-*Geo*HNH and its K597A mutant were simulated at room temperature and at their melting temperatures (i.e., 65 °C for the wt-*Geo*HNH and 45 °C for the K597A mutant). These systems were subjected to energy minimization and equilibrated in the canonical *NVT* ensemble for ~500 ps. Further equilibration was performed in the isothermal-isobaric *NPT* ensemble for ~10 ns allowing the density of the systems to stabilize around 1.01 g/cm^{-3}. Production runs were carried out reaching ~3 μs at each temperature and in three replicates. This resulted in ~9 μs for each simulated system and a total simulation time of ~36 μs. All simulations were performed using the GPU-empowered version of the AMBER20 code.[35]

To characterize the difference in stability of the HNH intradomain contacts across the systems under investigation, we first computed the contact maps for each system. Contacts between two heavy atoms of two residues were considered formed within a distance cutoff ≤4.5 Å. Then, the difference in stability of the common

contacts, which were present in both systems, was computed by considering the formula: $\Delta f = f_{WT} - f_{K597A}$, where f_{WT} denotes the number of frames in which a contact persists in wt-*Geo*HNH over the total number of frames; while f_{K597A} refers to the same measure for the K597A mutant. The same formula was also used to perform differential contact analysis at different temperatures. The extent of the stability of a contact in either of the systems compared with the other system is represented by $|\Delta f|$.

SUPPLEMENTARY MATERIAL

See the supplementary material for additional NMR spectra, pH titrations, relaxation data and associated parameters, RMSF calculations for wt- and K597A *Geo*HNH, and structural analysis.

ACKNOWLEDGMENTS

This work was supported by NIH R01 GM136815 (to G.P.L. and G.P.) and NSF Grant No. MCB 2143760 (to G.P.L.). This work was also supported by NIH R01 GM141329 (to G.P.) and NSF Grant No. CHE-1905374 (to G.P.). Computer time for Molecular Dynamics simulations was provided by XSEDE (TG-MCB160059) and NERSC (M3807). H.B.B. was supported by NSF GRFP Award No. 2040433. Research reported in this publication was supported by the National Institute of General Medical Sciences of the National Institutes of Health under Award Nos. T32GM077995 (H.B.B.) and T32GM139793 (C.N.). The content was solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

This research used the AMX beamline of the National Synchrotron Light Source II, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Brookhaven National Laboratory under Contract No. DE-SC0012704. The Center for BioMolecular Structure (CBMS) was primarily supported by the National Institutes of Health, National Institute of General Medical Sciences (NIGMS) through a Center Core P30 Grant (No. P30GM133893), and by the DOE Office of Biological and Environmental Research (Grant No. KP1607011).

AUTHOR DECLARATIONS

Conflict of Interest

The authors have no conflicts to disclose.

Author Contributions

H.B.B.: produced *Geo*HNH and *Geo*Cas9 proteins and sgRNA, conducted the NMR experiments. C.N.: conducted NMR experiments. C.P.: conducted MD simulations. S.S.: conducted MD simulations. A.M.D.: Solved x-ray crystal structure of K597A *Geo*HNH. G.J.: Supervision – x-ray crystallography. J.L.: *in vitro* cleavage assay. G.P.: Supervision – MD simulations. G.P.L.: supervision – NMR spectroscopy.

Helen B. Belato: Conceptualization (lead); Data curation (lead); Formal analysis (lead); Funding acquisition (equal); Project administration (lead); Supervision (lead); Writing – original draft (lead); Writing – review & editing (lead). **Carmelissa Norbrun**: Data curation (lead); Formal analysis (equal); Writing – original draft (lead). **Jinping Luo**: Data curation (equal); Formal analysis (equal). **Chinmai Pindi**: Data curation (equal); Formal analysis (equal); Writing – review & editing (equal). **Souvik Sinha**: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Funding acquisition (equal); Project administration (equal); Supervision (equal); Writing – review & editing (equal). **Alexandra M. D'Ordine**: Data curation (equal); Formal analysis (equal); Writing – review & editing (equal). **Gerwald Jogl**: Data curation (equal); Formal analysis (equal); Supervision (equal); Writing – review & editing (equal). **Giulia Palermo**: Formal analysis (equal); Funding acquisition (equal); Project administration (equal); Supervision (equal); Writing – review & editing (equal). **George P. Lisi**: Conceptualization (equal); Data curation (equal); Formal analysis (equal); Funding acquisition (equal); Project administration (equal); Supervision (equal); Writing – review & editing (equal).

DATA AVAILABILITY

The data that support the findings of this study are available from the corresponding author upon reasonable request. The x-ray crystal structure of K597A *Geo*HNH is available in the Protein Data Bank under accession ID 8F43.

REFERENCES

[1] J. A. Doudna, "The promise and challenge of therapeutic genome editing," Nature **578**, 229–236 (2020).
[2] A. Mir, A. Edraki, J. Lee, and E. J. Sontheimer, "Type II-C CRISPR-Cas9 biology, mechanism, and application," ACS Chem. Biol. **13**, 357–365 (2018).
[3] F. Jiang and J. A. Doudna, "The structural biology of CRISPR-Cas systems," Curr. Opin. Struct. Biol. **30**, 100–111 (2015).
[4] L. B. Harrington *et al.*, "A thermostable Cas9 with increased lifetime in human plasma," Nat. Commun. **8**, 1424 (2017).
[5] S. H. Sternberg, B. LaFrance, M. Kaplan, and J. A. Doudna, "Conformational control of DNA target cleavage by CRISPR–Cas9," Nature **527**, 110–113 (2015).
[6] J. S. Chen *et al.*, "Enhanced proofreading governs CRISPR–Cas9 targeting accuracy," Nature **550**, 407–410 (2017).
[7] K. W. East *et al.*, "Allosteric motions of the CRISPR-Cas9 HNH nuclease probed by NMR and molecular dynamics," J. Am. Chem. Soc. **142**, 1348 (2020).
[8] L. Nierzwicki *et al.*, "Enhanced specificity mutations perturb allosteric signaling in CRISPR-Cas9," Elife **10**, e73601 (2021).
[9] Ł. Nierzwicki, P. R. Arantes, A. Saha, and G. Palermo, "Establishing the allosteric mechanism in CRISPR-Cas9," Wiley Interdiscip. Rev. Comput. Mol. Sci. **11**, e1503 (2021).
[10] H. B. Belato *et al.*, "Structural and dynamic insights into the HNH nuclease of divergent Cas9 species," J. Struct. Biol. **214**, 107814 (2022).
[11] I. M. Slaymaker *et al.*, "Rationally engineered Cas9 nucleases with improved specificity," Science **351**, 84–88 (2016).
[12] C. N. Pace and J. M. Scholtz, "A helix propensity scale based on experimental studies of peptides and proteins," Biophys. J. **75**, 422–427 (1998).
[13] M. Jinek *et al.*, "Structures of Cas9 endonucleases reveal RNA-mediated conformational activation," Science **343**, 1247997 (2014).
[14] G. Palermo, Y. Miao, R. C. Walker, M. Jinek, and J. A. McCammon, "Striking plasticity of CRISPR-Cas9 and key role of non-target DNA, as revealed by molecular simulations," ACS Cent. Sci. **2**, 756–763 (2016).
[15] C. Huai *et al.*, "Structural insights into DNA cleavage activation of CRISPR-Cas9 system," Nat. Commun. **8**, 1375 (2017).
[16] G. Palermo *et al.*, "Protospacer adjacent motif-induced allostery activates CRISPR-Cas9," J. Am. Chem. Soc. **139**, 16028–16031 (2017).

[17]G. Palermo *et al.*, "Key role of the REC lobe during CRISPR-Cas9 activation by 'sensing,' 'regulating,' and 'locking' the catalytic HNH domain," Q. Rev. Biophys. **51**, e91 (2018).
[18]V. Nguyen *et al.*, "Evolutionary drivers of thermoadaptation in enzyme catalysis," Science **355**, 289–294 (2017).
[19]H. N. Motlagh, J. O. Wrabl, J. Li, and V. J. Hilser, "The ensemble nature of allostery," Nature **508**, 331–339 (2014).
[20]C. M. Petit, J. Zhang, P. J. Sapienza, E. J. Fuentes, and A. L. Lee, "Hidden dynamic allostery in a PDZ domain," Proc. Natl. Acad. Sci. U. S. A. **106**, 18249–18254 (2009).
[21]A. J. Wand, "The dark energy of proteins comes to light: Conformational entropy and its role in protein function revealed by NMR relaxation," Curr. Opin. Struct. Biol. **23**, 75–81 (2013).
[22]S.-R. Tzeng and C. G. Kalodimos, "Protein activity regulation by conformational entropy," Nature **488**, 236–240 (2012).
[23]F. Delaglio *et al.*, "NMRPipe: A multidimensional spectral processing system based on UNIX pipes," J. Biomol. NMR **6**, 277–293 (1995).
[24]W. Lee, M. Tonelli, and J. L. Markley, "NMRFAM-SPARKY: Enhanced software for biomolecular NMR spectroscopy," Bioinformatics **31**, 1325–1327 (2015).
[25]J. P. Loria, M. Rance, and A. G. Palmer, "A relaxation-compensated Carr–Purcell–Meiboom–Gill sequence for characterizing chemical exchange by NMR spectroscopy," J. Am. Chem. Soc. **121**, 2331–2332 (1999).
[26]M. Bieri, E. J. d'Auvergne, and P. R. Gooley, "relaxGUI: A new software for fast and simple NMR relaxation data analysis and calculation of ps-ns and μs motion of proteins," J. Biomol. NMR **50**, 147–155 (2011).
[27]Ł. Nierzwicki *et al.*, "Principles of target DNA cleavage and the role of Mg^{2+} in the catalysis of CRISPR–Cas9," Nat. Catal. **5**, 912–922 (2022).
[28]W. Kabsch, "XDS," Acta Crystallogr., Sect. D: Biol. Crystallogr. **66**, 125–132 (2010).
[29]M. D. Winn *et al.*, "Overview of the CCP4 suite and current developments," Acta Crystallogr., Sect. D: Biol. Crystallogr. **67**, 235–242 (2011).
[30]D. Liebschner *et al.*, "Macromolecular structure determination using X-rays, neutrons and electrons: Recent developments in Phenix," Acta Crystallogr., Sect. D: Biol. Crystallogr. **75**, 861–877 (2019).
[31]P. Emsley, B. Lohkamp, W. G. Scott, and K. Cowtan, "Features and development of *Coot*," Acta Crystallogr., Sect. D: Biol. Crystallogr. **66**, 486–501 (2010).
[32]J. A. Maier *et al.*, "ff14SB: Improving the accuracy of protein side chain and backbone parameters from ff99SB," J. Chem. Theory Comput. **11**, 3696–3713 (2015).
[33]W. L. Jorgensen, J. Chandrasekhar, J. D. Madura, R. W. Impey, and M. L. Klein, "Comparison of simple potential functions for simulating liquid water," J. Chem. Phys. **79**, 926–935 (1983).
[34]B. Hess, H. Bekker, H. J. C. Berendsen, and J. G. E. M. Fraaije, "LINCS: A linear constraint solver for molecular simulations," J. Comput. Chem. **18**, 1463–1472 (1997).
[35]al, C. D. e, AMBER 20, University of California, San Francisco, 2020.