



Deposited via The University of Sheffield.

White Rose Research Online URL for this paper:

<https://eprints.whiterose.ac.uk/id/eprint/241724/>

Version: Published Version

Article:

Thompson, M.J., Shahari, N.N.B.M., Ouanounou, R.J. et al. (2026) Functional roles for FEN1 phosphate steering residues in multi-step substrate verification prior to reaction. *Journal of Biological Chemistry*. 113225. ISSN: 0021-9258

<https://doi.org/10.1016/j.jbc.2026.113225>

Reuse

This article is distributed under the terms of the Creative Commons Attribution (CC BY) licence. This licence allows you to distribute, remix, tweak, and build upon the work, even commercially, as long as you credit the authors for the original work. More information and the full terms of the licence here:

<https://creativecommons.org/licenses/>

Takedown

If you consider content in White Rose Research Online to be in breach of UK law, please notify us by emailing eprints@whiterose.ac.uk including the URL of the record and the reason for the withdrawal request.

Journal Pre-proof

Functional roles for FEN1 phosphate steering residues in multi-step substrate verification prior to reaction

Mark J. Thompson, Nur Nazihah B. Md Shahari, Reuben J. Ouanounou, Nathan Gittens, Barbara Ciani, L. David Finger, Jane A. Grasby

PII: S0021-9258(26)02097-1

DOI: <https://doi.org/10.1016/j.jbc.2026.113225>

Reference: JBC 113225

To appear in: *Journal of Biological Chemistry*

Received Date: 12 March 2026

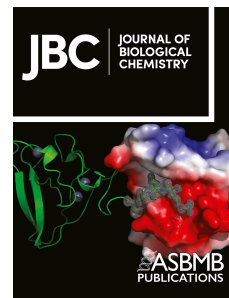
Revised Date: 13 May 2026

Accepted Date: 26 May 2026

Please cite this article as: Thompson MJ, Shahari NNBM, Ouanounou RJ, Gittens N, Ciani B, Finger LD, Grasby JA, Functional roles for FEN1 phosphate steering residues in multi-step substrate verification prior to reaction, *Journal of Biological Chemistry* (2026), doi: <https://doi.org/10.1016/j.jbc.2026.113225>.

This is a PDF of an article that has undergone enhancements after acceptance, such as the addition of a cover page and metadata, and formatting for readability. This version will undergo additional copyediting, typesetting and review before it is published in its final form. As such, this version is no longer the Accepted Manuscript, but it is not yet the definitive Version of Record; we are providing this early version to give early visibility of the article. Please note that Elsevier's sharing policy for the Published Journal Article applies to this version, see: <https://www.elsevier.com/about/policies-and-standards/sharing#4-published-journal-article>. Please also note that, during the production process, errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

© 2026 THE AUTHORS. Published by Elsevier Inc on behalf of American Society for Biochemistry and Molecular Biology.



Functional roles for FEN1 phosphate steering residues in multi-step substrate verification prior to reaction

Mark J. Thompson, Nur Nazihah B. Md Shahari, Reuben J. Ouanounou, Nathan Gittens[†], Barbara Ciani, L. David Finger[‡], and Jane A. Grasby^{*}

From the School of Mathematical and Physical Sciences, University of Sheffield, Sheffield, UK

* For correspondence: Jane A. Grasby, j.a.grasby@sheffield.ac.uk.

[†] Present address for Nathan Gittens: Structural & Biophysical Sciences, Glaxo SmithKline, Stevenage, UK.

[‡] Present address for L. David Finger: FabricNano, 4th Floor, 184–192 Drummond Street, London, NW1 3HP, UK.

Running title: Phosphate steering in the FEN1 reaction.

Keywords: phosphate steering, flap endonuclease 1, enzyme mechanism, enzyme kinetics, DNA repair, DNA replication, DNA–protein interaction, circular dichroism spectroscopy.

Abstract

Flap endonuclease 1 (FEN1) removes 5'-flaps from double-flap DNA junction intermediates during replication and repair. Substrate recognition and reaction site selection depend on two intrinsically disordered regions: the α 4– α 5 helical arch, through which the 5'-flap threads prior to catalysis, and the adjacent 3'-flap binding pocket, which allosterically signals disorder-to-order transition of the arch upon sensing a 3'-flap, committing the enzyme–DNA complex towards reaction. The phosphate steering hypothesis proposes that conserved, positively charged amino acids in α 4/ α 5 facilitate passage of 5'-flap DNA through the arch during threading and position the target phosphate diester for hydrolysis; however, supporting evidence is limited and mechanistic details are currently lacking. We investigated functional roles for these residues using kinetic and spectroscopic methods, finding that alanine substitutions of Arg103, Arg104, Arg129 and Lys132 modestly reduce the catalytic rate and the stability of 5'-flap threading. Following arch ordering, distortion of the reacting DNA duplex is necessary for active site transfer of the target cut site, and we identified key roles in this process for two substrate-facing residues from α 5, Lys125 and Arg129. Concurrently, 'back-of-arch' residues Arg104 and Lys132 contact the +1 phosphate to precisely position the target phosphodiester for hydrolysis. Helicity-disrupting mutations in α 4/ α 5, designed to impair ordering, reduced the catalytic rate and severely inhibited allosteric signalling of 3'-flap recognition to the active site. These findings define critical functional roles for phosphate steering residues in the FEN1 mechanism, and inform a deeper understanding of how coordinated substrate verification optimises targeting specificity to preserve genome integrity.

Introduction

As an essential component of the DNA replication and repair apparatus, flap endonuclease 1 (FEN1) catalyses the removal of single-stranded DNA (ssDNA) overhangs, known as 5'-flaps, from DNA junction structures produced during strand displacement synthesis (Fig. 1A). It is described as a structure-sensing nuclease due to recognising target substrates independent of nucleotide sequence, requiring the two DNA helices of the substrate to bend by approximately 100° to create most of the protein–DNA interface (1). This feature ensures selection of DNA junction structures over continuous genomic double-stranded DNA (dsDNA), which must not be indiscriminately hydrolysed (Fig. 1B).

Early work established the optimal FEN1 substrate as a double-flap conformer of the two-way junction (as shown in Fig. 1A), with a single nucleotide 3'-flap and any length of 5'-flap, including zero (2, 3). The hydrolytic reaction occurs one nucleotide into the reacting duplex, which when combined with the requirement for a single nucleotide 3'-flap ensures that FEN1 action produces nicked dsDNA suitable for direct ligation. This high specificity ensures efficient and accurate completion of DNA replication or repair, guarding against genomic instability (4).

Whereas the ideal FEN1 substrate has long been known, the molecular basis for this selectivity was elucidated only relatively recently. Crystal structures of human FEN1 (hFEN1) and paralogues in complex with DNA provided a wealth of information to support more detailed mechanistic understanding (Fig. 1B,C) (1, 5). The 5'- and 3'-flap binding sites were clearly defined, and ordering of the $\alpha 4$ – $\alpha 5$ flexible loop region was seen in complexes with DNA, to form a structural feature referred to as the 'helical arch' that straddles the enzyme's active site (Fig. 1B). Through NMR studies, we previously confirmed earlier assumptions that this region is an unstructured flexible loop in substrate-free FEN1, and while parts of $\alpha 4$ were seen to transiently sample α -helical space, it was concluded that $\alpha 5$ is always disordered (6). Similarly, in X-ray crystal structures of hFEN1 (without DNA substrate), $\alpha 4$ is disordered or partially ordered but $\alpha 5$ is always disordered (7), consistent with the NMR data. In the presence of DNA substrate (or product), published structures indicate a fully ordered helical arch with the 5'-flap (if present) threaded through this protein structural feature (1, 5).

Accordingly, biochemical experiments supported a 'disorder–thread–order' model of substrate binding wherein the 5'-flap threads through the initially disordered flexible loop, followed by ordering of $\alpha 4$ – $\alpha 5$ to initiate catalysis (8). This requirement for threading means the substrate must have a free 5'-terminus, preventing aberrant processing of continuous ssDNA–dsDNA junctions that arise during normal replication.

A key step in control of FEN1 catalysis was recognised as transfer of the scissile phosphodiester bond onto active site metal ions, or 'active site transfer' (9, 10), which was assumed to involve conformational changes in the substrate after initial protein–DNA binding (5–7). This was seemingly confirmed through a complex between the catalytically impaired hFEN1-D86N mutant and a double-flap substrate, positioned for reaction, where significant deviation of the reacting duplex region from a regular B-DNA structure was observed (5). As implicated in prior functional studies (10), recognition of the +1 phosphate (yellow dot in Fig. 1A) seemed to play a particularly important role in active site transfer.

It was also suggested that several positively charged and conserved protein side chains in the helical arch—near to but distinct from the FEN1 active site—help ‘steer’ the substrate into position for reaction in a twisting or rolling type of motion, corresponding to the aforementioned DNA conformational change. These residues (Arg103, Arg104, Arg129 and Lys132; Fig. 1C) were termed ‘phosphate steering’ residues, and in addition to a role in moving the scissile phosphodiester to within catalytic distance of the active site metal ions (i.e., active site transfer), they were proposed to help orient the 5'-flap away from the active site during threading to preserve the integrity of flap ssDNA. Thus, phosphate steering was defined as “electrostatic interactions that can dynamically position the phosphodiester backbone” (5).

Distinguishing between these roles in 5'-flap threading and active site transfer is crucial because the helical arch domain must be disordered during the former process but properly structured to enable the latter, since ordering controls assembly of the active site by bringing two essential conserved catalytic residues, Lys93 and Arg100, into position for reaction. This is explained by the ‘disorder–thread–order’ model that implies threading and active site transfer are distinct processes. Subsequent work offered strong evidence that the key event driving disorder-to-order transition of the arch is recognition of the 3'-flap, which is bound over 20 Å from the active site in a mechanism that effectively makes it an allosteric regulator of active site transfer and, by implication, of catalysis (7).

Despite the intuitive appeal of the phosphate steering hypothesis, and a supportive molecular dynamics study (11), a direct role for the proposed steering residues in either 5'-flap threading or active site transfer has not been conclusively demonstrated. In this paper, we aim to answer important outstanding questions regarding the specific functional role(s) of phosphate steering residues, considering how and where these roles integrate with known molecular details of the FEN1 reaction pathway.

Results

Phosphate steering mutants show reduced threading stability

As noted above, the hFEN1 phosphate steering residues were defined as Arg103 and Arg104 in helix $\alpha 4$, and Arg129 and Lys132 in $\alpha 5$ (Fig. 1C). Conservation of positive charge at these positions was seen widely among archaea and eukaryotes (5). Further sequence analysis (Fig. 2A), using a set of 179 curated FEN1 sequences reported previously (7), confirmed this earlier conclusion but also revealed strong conservation of Lys125, which contacts the template strand of the substrate along with other $\alpha 5$ residues (Lys128 and Arg129; Fig. 2B). We therefore prepared alanine mutants of all five residues in the present study.

These mutations noticeably impacted the reactivity of hFEN1 towards a model double-flap substrate, albeit to different degrees (Fig. 2C). Specifically, single turnover kinetics experiments were performed using an idealised static double-flap substrate S1 that exists in a single conformer of the DNA junction, and contains a 5 nucleotide 5'-flap and the canonical single nucleotide 3'-flap (see Fig. 2D for a schematic, and Table S1 and Fig. S1 for sequence and structural information). To elucidate the maximal single turnover rate (k_{STmax}), which directly reflects the enzyme's ability to perform its hydrolytic reaction, substrate (5 nM) and enzyme (400 nM) concentrations were respectively set at approximately an order of magnitude below

and above the K_M value of hFEN1 (~35 nM) (3, 7, 12) to ensure measured values reflected the maximal single turnover rate under saturating conditions. Our previous work has demonstrated very similar rates at 400 nM and 1,000 nM [E] for hFEN1 and a range of phosphate steering and 3'-flap binding mutants (5, 7), establishing that saturating conditions are adequately achieved at 400 nM enzyme. A 6-carboxyfluorescein (FAM) label was attached to the 5'-flap terminus to enable quantitative analysis of quenched reaction aliquots by HPLC, as described previously (5, 7). Earlier results indicate that tethering this fluorophore does not impair threading (7, 8), consistent with a report that hFEN1 accommodates 5'-flap lengths of up to 30 nucleotides with no significant impact on single turnover rate (13).

A very modest impact on k_{STmax} was seen for R103A and K132A (within fivefold of wt hFEN1, consistent with previous results for K132A) (5), whereas R129A and K125A were both approximately 10-fold slower (Figs. 2C and S2). The other mutant, R104A, exhibited a more substantial rate reduction of 40-fold and an additive effect when combined with K132A, as documented in the original phosphate steering study (5). Although single turnover rate defects were already known for most of these mutations, their impact on 5'-flap threading has not yet been evaluated.

To address this, we utilised an established assay to compare the proportion of enzyme–substrate complex where the substrate is bound in a threaded state at equilibrium (7, 8). This is achieved by incorporating an additional biotin label at the 5'-flap terminus (substrate S2; see Table S1 and Fig. S1), and exploiting the fact that a streptavidin–biotin complex is far too large to fit through the hFEN1 helical arch, even in a disordered state. Given that threading requires the presence of divalent metal ions, catalytically inert Ca^{2+} is used during equilibration (i.e., enzyme–substrate binding) to prevent reaction. Streptavidin is either added prior to equilibration to prevent threading ('blocked' substrate), or introduced after equilibrating the enzyme and substrate, to render threading irreversible ('trapped' substrate). Single turnover reaction profiles are then followed after introducing excess Mg^{2+} to initiate catalysis.

For wt hFEN1, all of the substrate is known to be threaded when bound to the enzyme (7), revealed by the fact that the initial ('fast') phase of the reaction reaches equivalent endpoints for either the trapped or control (no streptavidin) complexes (i.e., $y_1^{trap} = y_1^{ctrl}$, as defined and shown in Fig. 3A). The blocked substrate reacts 100-fold slower. With FEN1 mutants or substrates lacking a 3'-flap, threading defects are often observed (i.e., $y_1^{trap} < y_1^{ctrl}$), as seen for phosphate steering mutants R104A, R125A and R129A, which all exhibited modest but discernible threading deficiencies (Figs. 3B and S3). The double mutant R104A/K132A showed a more marked compromise in threading ability as might be expected given its 200-fold rate reduction (Fig. 2C), further noting that these 'back-of-arch' residues both make important contacts to the +1 phosphate.

Since this is an equilibrium assay and k_{on} is diffusion-limited for hFEN1 (14), we have previously interpreted threading deficiencies as representing reduced population and/or lifetime of the threaded (and ordered) complex, which is in turn assumed to be stabilised by the additional protein–DNA contacts formed upon helical arch structuring (6, 7). In an earlier report, based on multiple lines of evidence from kinetic and spectroscopic experiments, it was concluded that helical arch ordering is a key step in the FEN1 reaction pathway that 'locks' the

5'-flap in position, stabilising the complex to be longer-lived and proceed towards active site transfer and catalysis (7); otherwise, the FEN1–DNA complex is more prone to dissociation. The data implied that ordering of the helical arch is a strict requirement for the FEN1 reaction and is controlled by 3'-flap binding. If phosphate steering mutations do not affect this disorder-to-order transition, the observed threading deficiencies likely reflect a destabilising influence on the FEN1–DNA complex (i.e., increased k_{off}). As such, it was unsurprising to observe a meaningful threading defect for R104A and R104A/K132A (hereafter 104/132) given these residues make contacts with the +1 phosphate group that sits directly beneath the helical arch in the reactive configuration (Fig. 1C). Certain observations suggest the threading and rate defects for 104/132 may reflect a decreased binding lifetime: the K_{d} of wt hFEN1 has been reported as 3.3–4.8 nM (depending on 5'-flap length from 6–50 nucleotides) (15), based on single-molecule FRET experiments, whereas we previously determined the K_{d} of 104/132 as 44.9 nM via bulk FRET measurements (with a 5'-flap length of 6 nucleotides) (5). Assuming k_{on} is unaffected, the 200-fold rate deficiency for 104/132 is then at least partly explained by an order of magnitude increase in k_{off} (arising from the loss of +1 phosphate contacts), yet there also seems to be a role for functional defects in the threaded-and-ordered FEN1–DNA complex, prompting us to consider potential deficiencies in active site transfer as well.

Interestingly and somewhat unexpectedly, the trend in threading defects differed to that seen in kinetics results, with the K125A and R129A mutants showing greater threading deficiency than initially anticipated based on single turnover rates alone (10-fold slower than wt hFEN1). This suggested that $\alpha 5$ contacts to the continuous DNA template strand, and not just the 5'-flap, might be more important in the FEN1 mechanism than has previously been recognised, prompting further investigation of their role in active site transfer, alongside Arg104 and Lys132.

Resolving active site transfer measurements with FEN1 substrates

The phosphate steering hypothesis proposed that as well as promoting threading, certain positively charged protein side chains play a role in active site transfer (i.e., guiding the reacting DNA duplex into the FEN1 active site), with the above results also suggesting a role for Lys125. We considered how best to interrogate this process in the present context, noting that previous efforts to measure active site transfer, by ourselves and others, have exploited modified substrates incorporating 2-aminopurine (2AP) nucleobases near to the reaction site.

Since 2AP fluorescence is strongly quenched by neighbouring nucleobases in the absence of enzyme (16), measurements of 2AP fluorescence have been used to infer active site transfer. One such approach (17) relied on static double-flap substrate S3, incorporating adjacent 2AP bases at the +1 and –1 positions (see Table S1 and Fig. S1 for sequence and structural information). A threefold increase in fluorescence was reported over the first 100 s of a Mg^{2+} -supported reaction in the presence of slow-reacting hFEN1 mutant D34A. No product formation was detected over this timescale, so the fluorescence change was interpreted as observation of active site transfer prior to catalysis (17). Since this result contradicted our previous findings, we explored 2AP fluorescence changes further by investigating S3 with wt hFEN1 and a similar mutant, D34N, which exhibits ~80,000-fold slower reaction with substrate S1 ($k_{\text{STmax}} = 0.013$ vs $1,020 \text{ min}^{-1}$) (7) and evidence for off-target hydrolysis (Fig. S4).

We have previously demonstrated that addition of catalytically inert Ca^{2+} ions to hFEN1 complexes with 2AP-containing substrates produces a threaded enzyme–substrate complex and alters the time-resolved 2AP fluorescence parameters of the substrate (9, 18). In our hands, there was no effect of Ca^{2+} on steady-state fluorescence of the wt hFEN1–S3 complex over the first 100 s (Fig. S5A, cyan plot), whereas addition of Mg^{2+} ions produced a rapid and significant increase in steady-state fluorescence (Fig. S5A, purple plot), in agreement with the aforementioned report (17) and readily attributable to the disruption of quenching interactions upon hydrolytic reaction. Under these conditions, the wt hFEN1 reaction is rapid (half-time ~ 40 ms) and is essentially complete at the first data point (5 s), producing an approximately 10-fold increase in fluorescence. A slower reaction phase (assumed to reflect unbinding/rebinding of unproductively bound substrate, in line with previous studies) (3, 7) explains the additional increase seen from 5–100 s. Thus, fluorescence changes for wt hFEN1–S3 are fully explicable in terms of substrate hydrolysis, without invoking a contribution from active site transfer.

We proceeded to study the D34N–S3 complex. As with wt hFEN1, no fluorescence changes were observed in the presence of Ca^{2+} ions over 100 s (Fig. S5, beige plots). In this case, Mg^{2+} also failed to induce a change on the same timescale (Fig. S5, red plots), consistent with the essentially undetectable product formation over this period with S1 (Fig. S4). However, we did observe a gradual increase in fluorescence over 4 h (data not shown), after which time the slow hydrolysis of S1 by hFEN1-D34N is mostly complete (half-time = 55 min). We were therefore unable to detect possible active site transfer via 2AP fluorescence increases during the early stages of mutant FEN1 reactions, and steady-state fluorescence changes with wt enzyme were attributable solely to product formation.

In contrast, DNA conformational changes corresponding to active site transfer can be observed spectroscopically in near-UV circular dichroism (CD) spectra using the exciton-coupled CD (ECCD) assay, which similarly relies on substrates with tandem 2AP pairs (7, 9, 10, 12, 18, 19). Neighbouring (i.e., stacked) 2AP bases in dsDNA form an exciton pair, whose CD signal intensity maximum at 326 nm is highly sensitive to the relative orientation of the two bases (17, 20). When 2AP residues are placed at the -1 and -2 positions of an idealised double-flap substrate such as S4 (Table S1 and Fig. S1), the initial ECCD signal in the presence of Ca^{2+} ions is reduced to near-zero upon addition of wt hFEN1 (Fig. 4A). This spectral shift is assumed to reflect significant conformational change in the bound substrate, namely active site transfer, based on prior functional and crystallographic observations (5, 9, 10).

Phosphate steering residues have distinct roles in active site transfer

The ECCD assay was applied to interrogate the role of phosphate steering residues in active site transfer. Among the present set of point mutants, only K125A and R129A showed defects in this conformational change with double-flap substrate S4 (Fig. 4B), suggesting a role for $\alpha 5$ contacts in promoting the FEN1 reaction. Stated differently, conserved electrostatic contacts made from $\alpha 5$ of hFEN1 to the template strand of the substrate may be important in attaining and/or stabilising the catalytically competent conformation.

Mutation of ‘front-of-arch’ phosphate steering interactions (R103A/R129A, hereafter 103/129) or the corresponding back-of-arch contacts (104/132; see Fig. 1C) also allowed a role in active site transfer to be discerned for the former, by utilising both S4 and an analogous

exonucleolytic substrate (S5, where the 5'-flap length is zero; see Table S1 and Fig. S1 for sequence and structural information). Only the 103/129 double mutant displayed significantly compromised active site transfer with these substrates (Fig. 4C), indicating an important role in active site transfer for Arg103 and Arg129, but not Arg104 and Lys132. Together with the data for K125A, these results suggest that front-of-arch contacts from $\alpha 4$ and $\alpha 5$ can promote active site transfer, possibly in concert with short-range 'rocking' motions of the structured arch that were proposed based on NMR and single-molecule FRET data (6), and may help 'roll' the reacting duplex region into the enzyme active site. Other results with a hyperthermophilic FEN1 (from *Archaeoglobus fulgidus*) showed that threading efficiency and active site transfer ability were strongly temperature-dependent, improving as the temperature was raised and the degree of disorder in the arch region of the substrate-free protein increased (7).

Intriguingly, the above findings (Figs. 3B and 4C) indicate that while the back-of-arch residues (Arg104 and Lys132) contribute towards threading stability, presumably by contacting the critical +1 phosphate, they make little (if any) contribution to active site transfer. Their role was further probed using substrates S6 and S7, which are analogous to S4 and S5 but incorporate 2AP modifications at the +1 and -1 positions (as opposed to the -1 and -2 positions; see Table S1 and Fig. S1 for details). In earlier work we concluded that spectral changes with +1/-1 modified substrates provide a readout of proper substrate positioning for catalysis in the FEN1 active site, distinct from active site transfer assessed by substrates with -1/-2 modifications (7, 10). When interrogated with substrate S6 in the ECCD assay, both phosphate steering double mutants (103/129 and 104/132) showed comparable defects in substrate positioning (Fig. 4D). However, with exonucleolytic substrate S7 (where the +1 phosphate corresponds to the 5'-terminus of the reacting strand), the 104/132 double mutant was significantly compromised in substrate positioning compared to wt hFEN1, whereas 103/129 was not. These results are consistent with the established role for back-of-arch phosphate steering residues Arg104 and Lys132 in positioning the +1 phosphate to correctly define the site of reaction (5).

An alternative approach to investigate these interactions was to neutralise the negative charge of the +1 phosphate group by methylphosphonate substitution, as in double-flap substrate S8 (see Table S1 and Fig. S1). With wt hFEN1, this substrate was processed around 100-fold slower than unmodified counterpart S1 under both multiple and single turnover conditions (Fig. 4E, blue bars, and Fig. S6A). Significantly, the impact on rate was much less for the 104/132 double mutant, with only around a 10-fold rate reduction under the same conditions (Fig. 4E, beige bars, and Fig. S6B); the ratio of average rates for all four protein/substrate combinations is summarised in Fig. 4F. A conservative approach was adopted in these experiments, using 1,000 nM [E] to accommodate potential weaker binding and ensure saturating conditions were maintained with both substrates, especially considering the slow rates observed with S8 (Fig. S6). These methylphosphonate substitution results reinforce a functional role for electrostatic interactions between back-of-arch phosphate steering side chains (i.e., Arg104 and Lys132) and the substrate +1 phosphate (5), and imply involvement of these contacts in the rate-determining step of the FEN1 reaction. Since earlier work established that this step corresponds to conformational change of protein and/or substrate (10), it seems likely that single turnover rate defects (Fig. 2C) may be explained in terms of compromised active site transfer ability and/or substrate positioning, after threading has occurred.

Helix $\alpha 5$ is a key mediator of the FEN1 reaction pathway

Taken together, the above findings confirmed a role for phosphate steering residues in the hFEN1 reaction mechanism. However, since mutation of these residues maps to functional impacts on either 5'-flap threading, active site transfer or substrate positioning, all of which are essential for efficient reaction, it becomes apparent that detrimental effects arise through loss of specific contacts. Our results therefore imply that positively charged phosphate steering side chains have discernible functional roles, rather than merely providing a generalised region of positive charge to non-specifically promote threading and/or positioning of substrate DNA.

We have particularly uncovered an important role for contacts made from hFEN1 to the template strand via Lys125 and Arg129, both situated in helix $\alpha 5$ (Fig. 2B). The functional significance of these contacts has not been fully appreciated to date, despite the fact they show strong conservation at the residue level across archaea and eukaryotes (Fig. 2A). Earlier evidence for their role is, in fact, apparent from results obtained with hFEN1 mutant L130P, for which structuring of $\alpha 5$ was presumed to be either inhibited or prevented completely (18). With an idealised double-flap substrate, this mutant reacted 3,000-fold slower than wt hFEN1 under single turnover conditions, and significantly, the ECCD assay revealed severely impaired active site transfer ability.

Although $\alpha 5$ is disordered in the absence of substrate (6), the multiple contacts it makes once structured (to $\alpha 4$, the $\alpha 2$ – $\alpha 3$ loop region and the substrate template strand; Fig. 5A) position it as a key structural feature in the allosteric model of FEN1 function that links 3'-flap recognition to catalysis (1, 7, 14). In this model, 'sensing' of a 3'-flap by the $\alpha 2$ – $\alpha 3$ loop (over 20 Å from the active site) initiates the essential disorder-to-order transition of the helical arch ($\alpha 4$ and $\alpha 5$), leading to active site transfer and catalysis. Consistent with the disorder–thread–order model of FEN1 function, these events all appear to occur after threading, but $\alpha 5$ has strikingly low α -helical propensity in the free protein based on predictions from the AGADIR algorithm (21) (Fig. 5B). Nevertheless, the present results indicate that not only must $\alpha 5$ order properly to enable the aforementioned processes, specific contacts with the substrate are necessary for efficient reaction. We hence proceeded to investigate the role of 3'-flap recognition in helical arch ordering.

Ordering of $\alpha 4$ and $\alpha 5$ is driven by 3'-flap recognition

To pursue further evidence for disorder-to-order transition of the helical arch and its dependence upon 3'-flap recognition, glycine mutations were introduced in $\alpha 4$ and $\alpha 5$ at sites predicted to substantially disrupt helix formation according to AGADIR (Figs. 5B and 5C). The $\alpha 4$ sequence contains two distinct maxima in helical character such that two variants were selected for analysis: the A107G mutation was expected to sharply inhibit structuring of the top half of $\alpha 4$ (Fig. 5B), whereas the double mutation A98G/A107G was predicted to impair structuring of the full sequence. For $\alpha 5$, a double mutation (V123G/T127G) was found necessary to appreciably reduce its low inherent helical propensity.

Kinetic evaluation of these new mutants, again under single turnover conditions with substrate S1, showed a 25-fold rate reduction for hFEN1-A107G and 100-fold slower reaction for the two double mutants, relative to wt hFEN1 (Fig. 5D, purple bars, and Fig. S2). To maintain

saturating conditions, these reactions were performed at 1,000 nM [E] as a precautionary measure due to the substantial impact of the mutation(s) on reaction rate. While the observed kinetic effects are consistent with impaired arch ordering in the glycine mutants, it is apparently not prevented completely because k_{STmax} remains appreciably faster than seen with the analogous substrate lacking a 3'-flap, which reacts almost 1,000-fold slower (7). Hence, the next question to ask was how mutation of the 3'-flap interaction would impact activity of the glycine mutants. Recognition of the 3'-flap was shown to rely principally upon a conserved 'wedge residue' in the middle of the $\alpha 2$ – $\alpha 3$ loop region, corresponding to Leu53 in hFEN1 (7); mutating this residue to alanine slowed the single turnover reaction with S1 by a factor of ~1,000, similar to removal of the 3'-flap. In the same study, characterisation of hFEN1-L53A and other 3'-flap binding variants indicated a statistically significant correlation between single turnover rate and active site transfer ability (measured via the ECCD assay with substrate S4), suggesting 3'-flap recognition as the key event that initiates arch ordering.

Each of the three glycine mutants was therefore prepared with an additional L53A mutation, then single turnover rates were measured (again at 1,000 nM [E]) to assess the role of 3'-flap binding in arch ordering (Fig. 5D, orange bars). The observed activity of these combination mutants was drastically reduced in every case, to a similar extent to that seen for hFEN1-L53A. Overall, the kinetic results with $\alpha 4$ – $\alpha 5$ glycine mutants seem consistent with the proposal that the arch must order fully to promote reaction, and that this ordering is critically dependent upon wedge residue Leu53 binding a 3'-flap.

Interestingly, the ECCD assay with S4 (Figs. 5E and 5F) revealed a greater impact on active site transfer for mutation of $\alpha 5$ (V123G/T127G) compared with $\alpha 4$ (A98G/A107G; Fig. 5F, purple bars), despite the fact $\alpha 5$ is more distant from the active site, and that as noted earlier, two invariant catalytic residues (Lys93 and Arg100) lie at the base of $\alpha 4$ and rely upon its ordering to correctly position them for reaction (18). Meanwhile, the L53A combination mutants (Fig. 5F, orange bars) gave essentially identical ECCD results and were similarly compromised to the hFEN1-L53A single mutant, suggesting that structuring of $\alpha 4$ may first rely upon that of $\alpha 5$ in a mechanism dependent on the wedge residue.

The $\alpha 5$ glycine mutations also compromised the previously demonstrated ability of hFEN1 to induce a 3'-flap in equilibrating 'upstream' substrate S9 (see Table S1 and Fig. S1), which can exist as either a single- or double-flap conformer; earlier work established that the single-flap form predominates in free solution, whereas addition of hFEN1 strongly shifts the equilibrium towards the double-flap conformer, based on the ECCD signal at 330 nm (Fig. 5G) (7). Here, the V123G/T127G double mutant showed reduced capacity to induce the 3'-flap conformer, and the combination mutant (L53A/V123G/T127G) proved essentially unable to reconfigure the substrate, similarly to the hFEN1-L53A point mutant (Fig. 5H). If structuring of $\alpha 4$ first requires that of $\alpha 5$ as inferred above, the foregoing results are entirely consistent with our earlier assumption that failure to properly order the arch must result in shorter-lived binding of DNA, and that interaction of wedge residue Leu53 with the 3'-flap is a crucial contact in mediating this structural transition that commits the hFEN1–DNA complex towards reaction (7).

Discussion

The phosphate steering hypothesis proposed that FEN1 exploits a group of conserved or semi-conserved, positively charged residues to dynamically guide the phosphodiester backbone of the reacting DNA strand and thereby control reactivity (5). These residues are thought to orient the 5'-flap DNA away from the active site during threading (as the flap passes through the disordered helical arch), and to play a direct role in positioning the scissile phosphodiester group for reaction once the helical arch has ordered.

This paper offers new evidence supporting defined roles for phosphate steering residues in the substrate conformational changes necessary for FEN1 reaction. Specifically, front-of-arch residues Arg103 and Arg129 promote transfer of the reacting portion of the substrate duplex into the enzyme's active site, after 5'-flap threading. While active site transfer was originally perceived as local DNA unpairing (9), later findings informed a refined understanding described more generally as distortion or 'rolling' of the reacting duplex to access the active site via an irregular DNA structure (5). In this study we identified $\alpha 5$ as a mediator of this substrate motion, including an unexpected role for Lys125 (another conserved front-of-arch residue), suggesting it should be added to the phosphate steering repertoire. Major contributions to the active site transfer process appear to come from both Lys125 and Arg129 (in helix $\alpha 5$), with a potential minor role for Arg103 (in $\alpha 4$), consistent with the level of conservation of each of these residues across eukaryotic and archaeal FEN1s. As mentioned above, our results suggest that front-of-arch contacts, in particular those to the template strand from $\alpha 5$, may be directly involved in a motion involving the ordered helical arch that 'rolls' the reacting duplex into the active site, as also proposed elsewhere (5, 6).

Furthermore, precise positioning of the scissile and +1 phosphates (relative to the active site metal ions) is necessary to achieve the catalytically competent conformation and define the correct site of reaction, in a subtly related but distinct process to active site transfer. Final substrate positioning relies on interactions of the +1 phosphate with back-of-arch phosphate steering residues Arg104 and Lys132, as proposed in crystallographic studies (5) and supported here by multiple lines of evidence from kinetic, spectroscopic (i.e., ECCD) and methylphosphonate substitution experiments.

These established functional roles for phosphate steering residues emphasise how the helical arch actively participates in the FEN1 reaction pathway, rather than merely serving as a passive structural feature as originally thought (i.e., to select for 5'-flaps as opposed to continuous ssDNA) (1, 8, 13). The arch could, in fact, be described as a facilitator of communication due to the number of important contacts made in the catalytically competent conformation. These contacts recognise both strands of the reacting DNA duplex and 'bridge' the 3'-flap (in the upstream duplex) to the active site in what has been characterised as an allosteric process (7).

In the current model, recognition of the 3'-flap initiates an allosteric cascade wherein ordering of the $\alpha 2$ - $\alpha 3$ loop (which forms the major contacts in the 3'-flap binding pocket) and then the adjacent helical arch ($\alpha 5$ followed by $\alpha 4$) leads to active site transfer in a fully structured hFEN1-DNA complex. Once the scissile phosphate group is transferred onto the active site metal ions, the hydrolysis reaction will be rapid. On top of the general requirement for

FEN1 to bind at a DNA junction, the allosteric cascade mechanism could be thought of as sequentially validating different substrate features, since our results show that a loss of contacts at any stage of this mechanism will compromise the efficiency of the process.

Although 3'-flap contacts are the most influential in this context, phosphate steering interactions with both the template and reacting strands must be present to enable efficient active site transfer and substrate positioning. As such, these residues have important roles in ensuring substrate specificity and defining the site of reaction. The molecular details of the overall allosteric cascade process might therefore be thought of as multi-step substrate verification—providing an elaborate example of kinetic proofreading, which is widely exploited by nucleic acid binding proteins to achieve high specificity (22, 23) and often relies on intrinsically disordered regions (24–26)—since efficient FEN1 reaction relies not only upon 3'-flap recognition but also requires transmission of this structural information to the active site in a series of conformational changes of both protein and DNA substrate, including disorder-to-order transitions of the $\alpha 2$ – $\alpha 3$ loop and helical arch. In turn, these conformational changes depend upon specific, conserved contacts with the substrate from multiple protein residues. The present results illustrate how some of these contacts are provided by phosphate steering residues, affirming their important functional roles in mediating active site transfer and proper substrate positioning for reaction. Although technically challenging, it may be informative to further dissect the relative influence of phosphate steering contacts on binding lifetime and conformational changes leading to reaction, by measuring k_{off} for hFEN1 and relevant mutants in future work.

The question of whether phosphate steering residues are involved in 5'-flap threading is harder to resolve, however, since their exact positioning during this process is uncertain due to the disordered state of the arch at this stage (6, 8). Considering that the effects of mutating these residues can be satisfactorily explained in functional terms (with reference to the fully threaded-and-ordered complex), and because shorter ssDNA products arising from reaction in the 5'-flap are not observed with these mutants, it is difficult to support any direct role for these residues in promoting threading itself.

Nevertheless, they may indirectly facilitate threading by influencing the structural dynamics of the disordered helical arch. As is apparent from sequence analysis (Fig. 2A), similar numbers of positively and negatively charged residues are distributed across the $\alpha 4$ – $\alpha 5$ region. Assessment using the Classification of Intrinsically Disordered Ensemble Regions (CIDER) tool (27) concluded that the charge distribution in this domain serves to maintain its disordered state in an 'open' rather than 'collapsed' conformation, as would be ideal for threading (6). Positively charged phosphate steering residues must necessarily contribute to this structural characteristic, inferring an indirect but significant effect in maintaining the arch in an open form conducive to 5'-flap threading.

In summary, we conclude that phosphate steering residues do have significant roles in orienting the FEN1 substrate for reaction. However, a complex interplay between these contacts, and the consequences of 3'-flap binding transmitted through conformational changes of the protein and the DNA substrate (including disorder-to-order transitions in the former), is ultimately responsible for the exquisite specificity of the structure-sensing FEN1 nuclease.

Experimental procedures

Plasmid constructs

All new hFEN1 mutants were generated from the previously reported pET-28b-hFEN1-(His)₆ construct (3) by site-directed mutagenesis, using the QuikChange protocol (Agilent Technologies, Inc.) as described (7). Mutagenic primers were purchased from Fisher Scientific UK (desalted grade), reconstituted in ultrapure water and used as supplied. The sequences were: D34N – 5'-ctcatagaggcattaatggccaccttacggccaaa-3', 5'-tttggccgtaaggtggccattaatgcctctatgag-3'; A98G – 5'-ctcactgcgttgccagctcgctga-3', 5'-tcaggcgagctgggcaaacgcagtgag-3'; A107G – 5'-agctgcttctcctcagcccgc-3', 5'-ggcgggctgagggagagaagcagct-3'; A123G/A127G – 5'-caccagccgcttaccgaattttccccctctgctcggc-3', 5'-gccgagcaggaggggaaaaattcggaagcggctggtg-3'; K125A – 5'-ccagccgcttagtaatgctccacctctgctcgg-3', 5'-ccgagcaggaggtggaagcattcactaagcggctgg-3';

Protein expression

Proteins were expressed in Rosetta (DE3)pLysS competent cells cultured in 2× YT medium, following reported procedures for cell growth, induction, collection and lysis (7).

Protein purification

Protein purification was accomplished using an ÄKTA Pure system at 4 °C as previously described (7). Briefly, proteins were purified via Ni-NTA immobilised metal affinity chromatography (IMAC) followed by anion exchange and heparin affinity chromatography. Stock solutions were prepared in storage buffer (50 mM HEPES pH 7.5, 100 mM KCl, 1 mM CaCl₂, 5 mM DTT, 0.02% NaN₃, 50% v/v glycerol) and kept at -20 °C until needed. Protein-containing fractions were identified and final sample purity was assessed via SDS-PAGE.

Oligonucleotide synthesis and DNA substrates

All oligonucleotides listed in Table S1 were purchased from LGC Biosearch Technologies or Kaneka Eurogentec SA, with HPLC purification. Samples were dissolved in ultrapure water and stock concentrations measured by A₂₆₀ using calculated extinction coefficients. Prior to use, DNA constructs (Table S2 and Fig. S1) were annealed in folding buffer (50 mM HEPES pH 7.5, 100 mM KCl) by heating at 95 °C for 5 min, then allowing samples to stand at room temperature for 30 min.

Single turnover enzyme kinetics

To determine k_{STmax} values, reactions carried out at 37 °C with 400 nM or 1,000 nM enzyme (as specified in Fig. S2) and 5 nM substrate S1 were followed by quenching aliquots at appropriate time points, as documented previously (7). For faster reactions, the rapid quench-flow technique was employed using an RQF-63 instrument (TgK Scientific Ltd, UK), in line with prior reports (3, 5, 7, 18). Briefly, solutions of enzyme and substrate were prepared at 2× final reaction concentration in buffer RB (50 mM HEPES pH 7.5, 100 mM KCl, 8 mM MgCl₂, 2.5 mM DTT, 0.1 mg mL⁻¹ BSA). For each time point, 80 µL aliquots of enzyme and substrate were mixed and

quenched in 1.5 M NaOH containing 20 mM EDTA, after a controlled time delay of 0.0045–91.041 s. For slower reactions, manual sampling was performed as described with 250 mM EDTA (pH 8.0) used as quenching agent (7).

For all quenched aliquots, the amount of product was quantified via denaturing HPLC using a Transgenomic WAVE system equipped with an L-7485 fluorescence detector and a 4.6 × 50 mm OligoSep Cartridge (ADS Biotec, Inc). Product and substrate were separated using previously reported buffers and gradient conditions (7). Specifically, an injected sample with ≥300 fmol DNA was eluted using buffer A (0.1% MeCN, 2.5 mM tetrabutylammonium bromide, 1 mM EDTA) and buffer B (80% MeCN, 2.5 mM tetrabutylammonium bromide, 1 mM EDTA), using the following gradient: 5–30% B over 1 min, 30–55% B over 4.5 min, 55–100% B over 1.6 min, 100% B for 1.4 min, 100–5% B over 0.1 min, 5% B for 2.4 min. Integration of the resultant chromatograms was used to determine the ratio and hence the concentrations of product and substrate with respect to time. If multiple product peaks were observed, the combined total was considered for rate measurements. The first-order rate constant (k_{STmax}) was derived by global fitting to the experimental data using nonlinear least squares regression in GraphPad Prism 7 or later as previously described. One- and two-phase exponential models were tested for each dataset, and model selection was informed by Akaike's information criterion (AIC) (7).

Multiple turnover enzyme kinetics

Normalised rates of reaction ($v/[E]$) were determined at 37 °C with 50 nM substrate S1 and enzyme (10 pM hFEN1 or 200 pM R104AK132A), or with 50 nM methylphosphonate-containing substrate S8 and enzyme (250 pM hFEN1 or 2,000 pM R104AK132A), under identical buffer conditions to the single turnover reactions, as reported previously (3, 7). Triplicate reactions were initiated in 180 µL total volume, then 20 µL aliquots were quenched into 250 mM EDTA pH 8.0 at 2, 4, 6, 8, 10, 12 and 20 min (S1) or 8, 12, 16, 20, 25, 32 and 40 min (S8). Rate values were derived by linear regression in Microsoft Excel.

Trapping and blocking experiments

Trapping/blocking experiments were conducted with substrate S2, bearing both FAM and biotin labels at the 5'-terminus of the 5'-flap, under saturating single turnover conditions using rapid quench-flow and manual sampling techniques as above. In each experimental run, 'trapped' reactions were conducted by adding streptavidin after incubation of substrate with the requisite hFEN1 protein, whereas streptavidin was mixed with the substrate before incubation with the hFEN1 protein for 'blocked' experiments, as detailed in earlier reports (7, 8). Control reactions were conducted as for the trapped reactions, except blank buffer was added instead of streptavidin. For all reactions, 8 M urea containing 300 mM EDTA (pH 8.0) was used as the quench solution. Quenched reaction aliquots were analysed by HPLC as described above, then data were processed in GraphPad Prism 7 or later as documented previously (7).

Fluorescence spectroscopy

Fluorescence measurements were conducted using substrate S3, which contains adjacent 2AP nucleobases at the +1 and -1 positions relative to the scissile phosphate of the reacting DNA strand (see Fig. 1A). Assays used 2.5 μM of wt hFEN1 or the D34N mutant and 0.5 μM of S3, mixed with 1 mM EDTA to produce a pre-formed complex as described (17). To initiate the experiment, 8 mM CaCl_2 or MgCl_2 was added as appropriate, then reactions were monitored at various time points from 5–100 s using a FluoroMax-3 instrument (Horiba Ltd, Japan) at 37 °C with a slit length of 5 nm. Excitation and emission wavelengths were 315 nm and 375 nm, respectively. Raw fluorescence values were normalised relative to reference values (i.e., readings for samples with EDTA only, where no metal ions were added).

ECCD assay

Measurements of the ECCD signal produced by two tandem 2AP residues in the indicated substrates (S4–S7 and S9; Fig. S1) were conducted as described (7). Briefly, spectra were acquired from samples containing 10 μM DNA and 12.5 μM protein in buffer CaRB (50 mM HEPES pH 7.5, 100 mM KCl, 10 mM CaCl_2 ; 500 μL total sample volume) at 20 °C, using a Chirascan Plus CD spectrometer (Applied Photophysics Ltd, UK). Processing of raw data relied on the Pro-Data Viewer 4.4.2.0 software supplied with the instrument. Spectra were baseline-corrected by subtracting a blank (recorded with CaRB) then subjected to smoothing (window size = 10), then the data were transformed to units of $\Delta\epsilon$ per mol 2AP residue. Peak heights at 326 nm (downstream substrates) or 330 nm (upstream substrate) were used to quantify the extent to which different proteins are able to induce DNA conformational change. Due to variation in signal amplitude between substrate batches, values for each replicate were calculated relative to a substrate-only sample recorded on the same experimental run. Representative plots (Figs. 4A, 5E and 5G) were prepared in GraphPad Prism 10.

Sequence analysis

Sequence analysis was conducted as documented previously (7). To summarise, reviewed FEN1 sequences (archaea, $n = 51$; eukaryotes, $n = 128$) from the UniProt database were analysed via multiple sequence alignment with ClustalX 2.1 (28), then the region corresponding to hFEN1 residues 86–133 was extracted and realigned. The output was used to generate sequence logos (Fig. 2A) with WebLogo 3 (29).

Statistical analysis

Pairwise comparisons in Figs. 5 and 6 used two-tailed, two-sample, unequal variance Student's t-tests in Microsoft Excel. Statistical significance is denoted by asterisk(s) according to the following limits: $p < 0.05$ (*), < 0.01 (**), < 0.001 (***), < 0.0001 (****). All datasets in Figures 4B–4D, 5F and 5H were analysed via Shapiro–Wilk Test in GraphPad Prism 11, and the assumption of normality was met in every case.

Data availability

All data are contained within the manuscript and supporting information. Raw data supporting the results of this study will be provided by the corresponding authors upon request.

Supporting information

This article contains supporting information.

Acknowledgements

The authors thank Emma Jardine for expression and purification of the hFEN1-D34N mutant.

Author contributions

M. J. T., B. C., L. D. F. and J. A. G. conceptualisation; M. J. T., L. D. F. and J. A. G. methodology; M. J. T., N. N. B. Md S., R. J. O. and N. G. investigation; M. J. T., N. N. B. Md S. and R. J. O. formal analysis; M. J. T. and J. A. G. visualisation; M. J. T., R. J. O. and L. D. F. writing – original draft; M. J. T. and J. A. G. writing – review & editing; M. J. T., L. D. F. and J. A. G. supervision; J. A. G. project administration; L. D. F. and J. A. G. funding acquisition.

Funding and additional information

This work was funded by BBSRC grants BB/M00404X/1 (to J. A. G.) and BB/R018251/1 (to J. A. G. and L. D. F.). Additionally, R. J. O. and N. N. B. Md S. thank the BBSRC (BB/M011151/1) and the Malaysian Government, respectively, for funding their scholarships.

Conflict of interest

The authors have no conflicts of interest to declare regarding the content of this article.

References

1. Tsutakawa, S. E., Classen, S., Chapados, B. R., Arvai, A. S., Finger, L. D., Guenther, G., *et al.* (2011) Human flap endonuclease structures, DNA double-base flipping, and a unified understanding of the FEN1 superfamily. *Cell* **145**, 198–211.
2. Liu, Y., Kao, H.-I., and Bambara, R. A. (2004) Flap endonuclease 1: a central component of DNA metabolism. *Annu. Rev. Biochem.* **73**, 589–615.
3. Finger, L. D., Blanchard, M. S., Theimer, C. A., Sengerová, B., Singh, P., Chavez, V., *et al.* (2009) The 3'-flap pocket of human flap endonuclease 1 is critical for substrate binding and catalysis. *J. Biol. Chem.* **284**, 22184–22194.
4. Balakrishnan, L., and Bambara, R. A. (2013) Flap endonuclease 1. *Annu. Rev. Biochem.* **82**, 119–138.
5. Tsutakawa, S. E., Thompson, M. J., Arvai, A. S., Neil, A. J., Shaw, S. J., Algasier, S. I., *et al.* (2017) Phosphate steering by flap endonuclease 1 promotes 5'-flap specificity and incision to prevent genome instability. *Nat. Commun.* **8**, 15855.

6. Bennet, I. A., Finger, L. D., Baxter, N. J., Ambrose, B., Hounslow, A. M., Thompson, M. J., *et al.* (2018) Regional conformational flexibility couples substrate specificity and scissile phosphate diester selectivity in human flap endonuclease 1. *Nucleic Acids Res.* **46**, 5618–5633.
7. Thompson, M. J., Gotham, V. J. B., Ciani, B., and Grasby, J. A. (2018) A conserved loop–wedge motif moderates reaction site search and recognition by FEN1. *Nucleic Acids Res.* **46**, 7858–7872.
8. Patel, N., Atack, J. M., Finger, L. D., Exell, J. C., Thompson, P., Tsutakawa, S., *et al.* (2012) Flap endonucleases pass 5'-flaps through a flexible arch using a disorder-thread-order mechanism to confer specificity for free 5'-ends. *Nucleic Acids Res.* **40**, 4507–4519.
9. Finger, L. D., Patel, N., Beddows, A., Ma, L., Exell, J. C., Jardine, E., *et al.* (2013) Observation of unpaired substrate DNA in the flap endonuclease-1 active site. *Nucleic Acids Res.* **41**, 9839–9847.
10. Algasaier, S. I., Exell, J. C., Bennet, I. A., Thompson, M. J., Gotham, V. J. B., Shaw, S. J., *et al.* (2016) DNA and protein requirements for substrate conformational changes necessary for human flap endonuclease-1-catalyzed reaction. *J. Biol. Chem.* **291**, 8258–8268.
11. Donati, E., Vidossich, P., and De Vivo, M. (2021) Molecular mechanism of phosphate steering for DNA binding, cleavage localization, and substrate release in nucleases. *ACS Catal.* **11**, 13244–13254.
12. Algasaier, S. I., Finger, L. D., Bennet, I. A., and Grasby, J. A. (2018) Flap endonuclease 1 mutations A159V and E160D cause genomic instability by slowing reaction on double-flap substrates. *Biochemistry* **57**, 6838–6847.
13. Tarantino, M. E., Bilotti, K., Huang, J., and Delaney, S. (2015) Rate-determining step of flap endonuclease 1 (FEN1) reflects a kinetic bias against long flaps and trinucleotide repeat sequences. *J. Biol. Chem.* **290**, 21154–21162.
14. Rashid, F., Harris, P. D., Zaher, M. S., Sobhy, M. A., Joudeh, L. I., Yan, C., *et al.* (2017) Single-molecule FRET unveils induced-fit mechanism for substrate selectivity in flap endonuclease. *Elife* **6**, e21884.
15. Zaher, M. S., Rashid, F., Song, B., Joudeh, L. I., Sobhy, M. A., Tehseen, M. *et al.* (2018) Missed cleavage opportunities by FEN1 lead to Okazaki fragment maturation via the long-flap pathway. *Nucleic Acids Res.* **46**, 2956–2974.
16. Jones, A. C., and Neely, R. K. (2015) 2-Aminopurine as a fluorescent probe of DNA conformation and the DNA–enzyme interface. *Q. Rev. Biophys.* **48**, 244–279.
17. Song, B., Hamdan, S. M., and Hingorani, M. M. (2018) Positioning the 5'-flap junction in the active site controls the rate of flap endonuclease-1-catalyzed DNA cleavage. *J. Biol. Chem.* **293**, 4792–4804.

18. Patel, N., Exell, J. C., Jardine, E., Ombler, B., Finger, L. D., Ciani, B., *et al.* (2013) Proline scanning mutagenesis reveals a role for the flap endonuclease-1 helical cap in substrate unpairing. *J. Biol. Chem.* **288**, 34239–34248.
19. Exell, J. C., Thompson, M. J., Finger, L. D., Shaw, S. J., Debreczeni, J., Ward, T. A., *et al.* (2016) Cellularly active *N*-hydroxyurea FEN1 inhibitors block substrate entry to the active site. *Nat. Chem. Biol.* **12**, 815–821.
20. Johnson, N. P., Baase, W. A., and Von Hippel, P. H. (2004) Low-energy circular dichroism of 2-aminopurine dinucleotide as a probe of local conformation of DNA and RNA. *Proc. Natl. Acad. Sci. U. S. A.* **101**, 3426–3431.
21. Lacroix, E., Viguera, A. R., and Serrano, L. (1998) Elucidating the folding problem of α -helices: local motifs, long-range electrostatics, ionic-strength dependence and prediction of NMR parameters. *J. Mol. Biol.* **284**, 173–191.
22. Zhou, H.-X. (2012) Intrinsic disorder: signaling via highly specific but short-lived association. *Trends Biochem. Sci.* **37**, 43–48.
23. Boeger, H. (2022) Kinetic proofreading. *Annu. Rev. Biochem.* **91**, 423–447.
24. Dyson, H. J. (2012) Roles of intrinsic disorder in protein–nucleic acid interactions. *Mol. Biosyst.* **8**, 97–104.
25. Babu, M. M. (2016) The contribution of intrinsically disordered regions to protein function, cellular complexity, and human disease. *Biochem. Soc. Trans.* **44**, 1185–1200.
26. Madhurima, K., Nandi, B., Munshi, S., Naganathan, A. N., and Sekhar, A. (2023) Functional regulation of an intrinsically disordered protein via a conformationally excited state. *Sci. Adv.* **9**, eadh4591.
27. Holehouse, A. S., Das, R. K., Ahad, J. N., Richardson, M. O. G., and Pappu, R. V. (2017) CIDER: resources to analyze sequence-ensemble relationships of intrinsically disordered proteins. *Biophys. J.* **112**, 16–21.
28. Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., *et al.* (2007) Clustal W and Clustal X version 2.0. *Bioinformatics* **23**, 2947–2948.
29. Crooks, G. E., Hon, G., Chandonia, J.-M., and Brenner, S. E. (2004) WebLogo: a sequence logo generator. *Genome Res.* **14**, 1188–1190.

Figure legends

Figure 1. FEN1 function, structure and phosphate steering residues, and their relation to mechanism. *A*, FEN1 processes the products of strand displacement synthesis arising from DNA replication and repair into a nicked dsDNA structure (P_2) that can be directly ligated, releasing the 5'-flap as an ssDNA by-product (P_1). The red dot indicates the scissile phosphate diester (i.e., the site of hydrolytic reaction in the substrate), while the yellow dot denotes the +1 phosphate. Neighbouring nucleobases (-2, -1 and +1) are numbered relative to the scissile phosphate. *B*, Structure of hFEN1-D86N in complex with substrate DNA (PDB code 5UM9) (5), with key protein features annotated and showing the 5'-flap threaded through the ordered helical arch ($\alpha 4$ – $\alpha 5$). In this structure, the substrate is positioned to react except that the D86N

mutation precludes binding of the second active site metal ion that would activate the attacking water molecule. *C*, Top-down view of the same complex, highlighting the positions of phosphate steering side chains (front-of-arch residues Arg103 and Arg129, and back-of-arch residues Arg104 and Lys132; blue sticks), as well as the +1 and +2 phosphate non-bridging oxygen atoms (orange).

Figure 2. Phosphate steering contacts are conserved in FEN1s and contribute to reactivity. *A*, Sequence logos based on alignment of 179 archaeal and eukaryotic FEN1s illustrate varying degrees of conservation of phosphate steering residues (yellow highlights), as noted previously (5). Strong conservation of hFEN1-Lys125 (cyan highlight) is also observed. Residue numbering corresponds to the hFEN1 sequence. *B*, Three positively charged side chains from $\alpha 5$ make electrostatic contacts to the continuous template strand (yellow) of the DNA substrate. Conserved residues Lys125 (orange) and Arg129 (red) are in close proximity for electrostatic interactions, as is semi-conserved residue Lys128 (teal). Distances are labelled in Å and the 5'-flap strand is shown in pink. Image prepared using PDB code 5UM9 (5). *C* and *D*, Mutation of phosphate steering residues and Lys125 to alanine has a varying impact on FEN1 single turnover activity (*C*) measured using an idealised double-flap substrate S1 (*D*), noting 104/132 indicates the back-of-arch R104A/K132A double mutant. For R103A, K125A and R129A, rates are reported as mean and standard error from global fitting of time course data via nonlinear regression (Fig. S2); replicates were as follows: R103A, $n = 4$ from two independent experiments performed in duplicate; K125A, $n = 3$ technical replicates; R129A, $n = 4$ from two independent experiments performed in duplicate. Results for wt hFEN1, R104A, K132A and 104/132 were obtained similarly and are reproduced from a previous report (5).

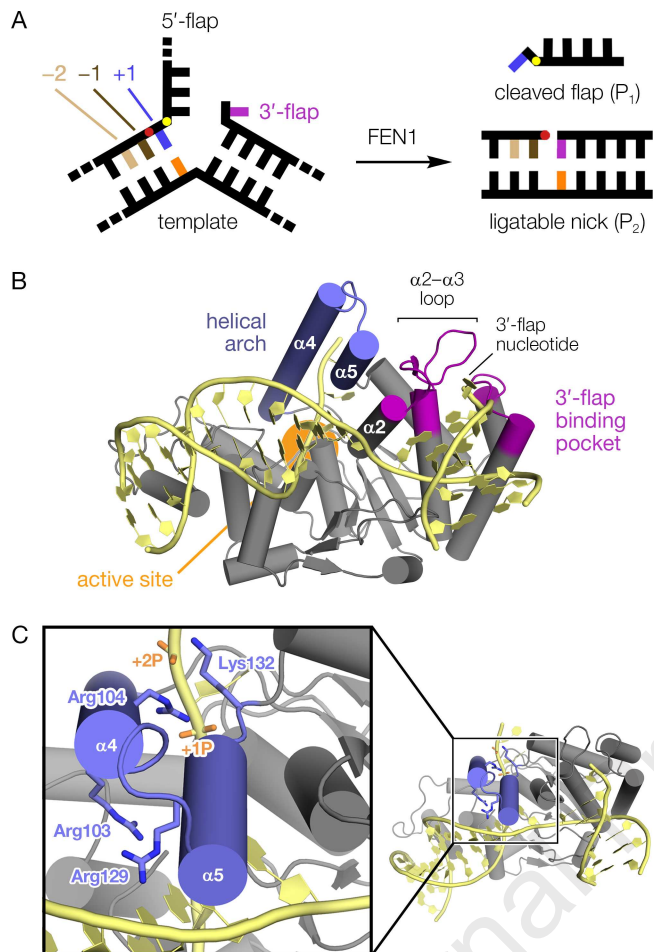
Figure 3. Phosphate steering mutants exhibit 5'-flap threading defects. *A*, Kinetic profiles from threading assays with wt hFEN1 (left) and double phosphate steering mutant 104/132 (right), using biotinylated substrate S2. With the wt enzyme, trapped substrate decays to the same initial endpoint (defined by the modelled first phase from two-phase exponential curve fitting) as the free substrate with no streptavidin added, implying that the 5'-flap is fully threaded after pre-equilibration of enzyme and substrate. With the double mutant these reactions reach different endpoints, indicating some unbound/unthreaded substrate at the point of introducing streptavidin. Error bars denote SD of $n = 4$ replicates per condition from two independent experiments performed in duplicate. *B*, Threading efficiencies, calculated based on the results in (*A*) and analogous experiments with R103A, R104A, K125A and R129A (Fig. S3). Except for R103A, all mutants exhibit a threading defect. Graphed values represent the ratio of endpoint values for the first reaction phase from global two-phase model fitting.

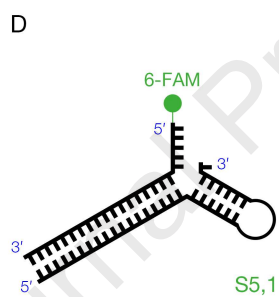
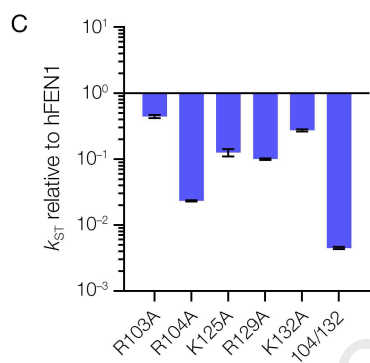
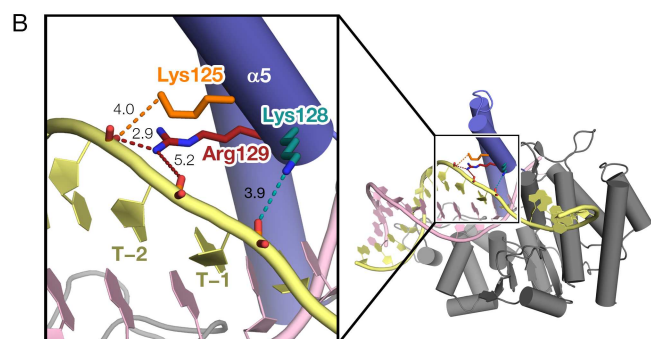
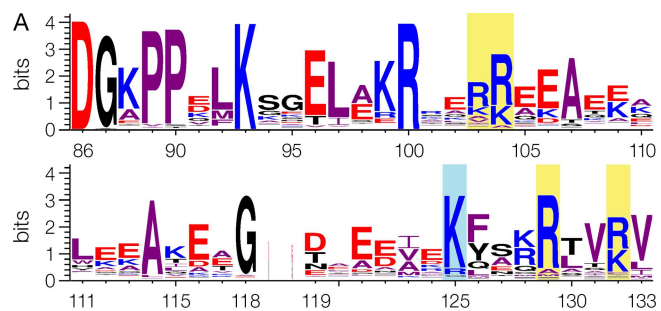
Figure 4. Phosphate steering residues play distinct roles in the FEN1 reaction mechanism. *A*, Representative ECCD spectra of substrate S4 alone (purple plot) or after mixing with hFEN1 (beige plot). The initial maximum at 326 nm, due to exciton coupling between adjacent 2APs (20), is reduced to near-zero in the presence of enzyme, indicating a significant conformational change of the DNA upon binding (interpreted as active site transfer). A schematic of S4 is shown in the purple circle, with tandem 2AP residues at the -1 and -2 positions highlighted in red. *B*, Among the phosphate steering single mutants, K125A and R129A are severely compromised in their ability to bring about active site transfer with S4. Signal intensities at 326 nm are expressed as a percentage of that for substrate alone on the

same experimental run. Statistical analysis via pairwise two-tailed, two-sample, unequal variance Student's t-tests revealed significant differences versus wt hFEN1 for K125A ($p = 0.0015$) and R129A ($p = 0.0002$). *C*, ECCD results for front-of-arch mutant 103/129 and back-of-arch mutant 104/132, with S4 (purple bars) and related exonucleolytic substrate S5 (orange bars; see orange circle for schematic). Statistical analysis as above showed significantly compromised active site transfer for 103/129 versus 104/132 with both S4 ($p = 0.0031$) and S5 ($p = 0.0098$); 103/129 also exhibited impaired active site transfer compared with hFEN1 (S4, $p = 0.0001$; S5, $p = 0.0012$). *D*, Parallel results to (*C*) with analogous substrates S6 (purple bars and circle) and S7 (orange bars and circle), containing 2AP residues at the +1 and -1 positions (indicated in red). In these substrates, reduced ECCD signal intensity with wt hFEN1 is interpreted as proper active site positioning of the substrate for reaction. A difference between mutants is seen with S7, where the +1 phosphate is the 5'-terminus of the reacting strand, and statistical analysis, again as defined above, demonstrated significantly impaired substrate positioning for 104/132 versus 103/129 ($p = 0.0138$) and wt hFEN1 ($p = 0.0065$). With S6, no significant difference between the two mutants was observed ($p = 0.1961$). *E*, Rates of reaction under multiple turnover (left) or single turnover (right) conditions for model double-flap substrate S1 and S8, containing a methylphosphonate substitution at the +1 position. Both wt hFEN1 (blue bars) and 104/132 (olive bars) show reduced reactivity with S8. *F*, Ratio of average rates for S1/S8 under multiple turnover (MT, beige bars) and single turnover (ST, brown bars) conditions shows a much lower impact of +1 methylphosphonate substitution for the back-of-arch mutant 104/132, supporting the proposal that Arg104 and Lys132 interact with the +1 phosphate. Note: Graphs in (*B*–*D*) include overlaid scatter plots of replicate ECCD measurements, which were each performed independently on different experimental runs (*B*, $n = 6$ for hFEN1, $n = 4$ for R104A, $n = 3$ otherwise; *C*, $n = 6$ for hFEN1, $n = 3$ otherwise; *D*, $n = 3$ for all conditions); in panel (*E*), multiple turnover, $n = 12$ for hFEN1 with both substrates, $n = 8$ for S1 with 104/132, and $n = 6$ for S8 with 104/132; in all cases, error bars denote SD. Single turnover data in (*E*) were derived via global model fitting to kinetic data, as with previous examples, for reactions performed in triplicate (Fig. S6).

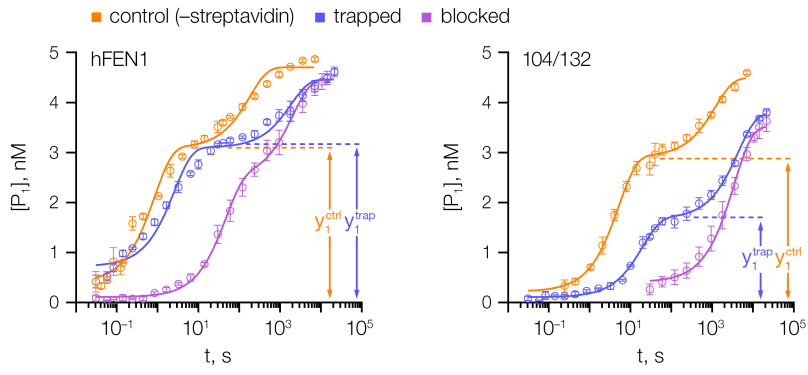
Figure 5. Arch ordering is important in the FEN1 mechanism and relies upon 3'-flap interaction. *A*, Partial surface renders (transparency) emphasise the multiple interactions made by helix $\alpha 5$ once ordered, based on PDB structure 5UM9 (5). *B*, AGADIR predictions (at 20 °C) of α -helicity for $\alpha 4$ (left) and $\alpha 5$ (right) sequences, in wt hFEN1 (purple) or with the indicated mutations. *C*, Location of residues mutated to glycine in the hFEN1 structure. *D*, Single turnover reaction rates with S1 for the indicated proteins (x-axis), without (purple bars) or with (orange bars) an additional L53A wedge residue mutation. *E*, Representative ECCD spectra for S4 without (purple plot) or with (beige plot) addition of hFEN1, superimposing results for $\alpha 5$ mutant V123G/T127G (123/127; cyan plot) and triple mutant L53A/V123G/T127G (53/123/127; red plot), where active site transfer is essentially abolished. *F*, ECCD results for proteins in (*D*) with S4. Note that mutating $\alpha 5$ (123/127) affects active site transfer more than mutation of $\alpha 4$ in A107G and A98G/A107G (98/107). Statistical analysis via a pairwise two-tailed, two-sample, unequal variance Student's t-test revealed active site transfer is significantly worse for 123/127 versus 98/107 ($p = 0.0004$). Combination mutants with L53A behaved similarly to L53A alone, displaying severely impaired active site transfer. *G*, Representative ECCD spectra for S9, which exists mostly in single-flap form when alone in solution (blue plot and schematic in blue circle, with 2AP bases in red), but displays a much-reduced signal in the presence of wt hFEN1 (beige

plot), consistent with conformational shift to the double-flap form with a single nucleotide 3'-flap (schematic in beige circle). This ability to induce a 3'-flap is partially inhibited in $\alpha 5$ mutant 123/127 (cyan plot), and substantially impaired for combination wedge mutant 53/123/127 (red plot). *H*, ECCD results for S9 with the specified proteins (x-axis). Statistical analysis as above demonstrated significantly compromised 3'-flap formation by 123/127 versus wt hFEN1 ($p = 0.0002$), and also for combination mutant 53/123/127 versus 123/127 ($p = 0.0013$). As for S4, the combination mutant acts similarly to L53A alone ($p = 0.4694$). Note: Panel (D) reports single turnover rates and standard error from global model fitting to kinetic data, as in previous cases, for replicate reactions (Fig. S2); replicates were as follows: 123/127, $n = 4$ (using rapid quench flow technique, QF) from two independent experiments performed in duplicate plus $n = 3$ technical replicates (manual sampling for longer time points); 53/123/127, $n = 6$ from two independent experiments performed in triplicate (manual sampling); A107G, $n = 4$ (QF) from two independent experiments performed in duplicate plus $n = 3$ technical replicates (manual sampling for longer time points); 53/107, $n = 3$ technical replicates (manual sampling); 98/107, $n = 4$ (QF) from two independent experiments performed in duplicate plus $n = 3$ technical replicates (manual sampling for longer time points); 53/98/107, $n = 6$ from two independent experiments performed in triplicate (manual sampling). Previously reported values are used for wt hFEN1 and L53A (5, 7). Graphs in (F) and (H) include overlaid scatter plots of replicate ECCD measurements, which were each performed independently on different experimental runs (*F*, $n = 4$ for hFEN1/Leu53 and $n = 3$ otherwise; *H*, $n = 5$ for hFEN1, $n = 6$ for L53A, $n = 3$ otherwise); error bars denote SD.

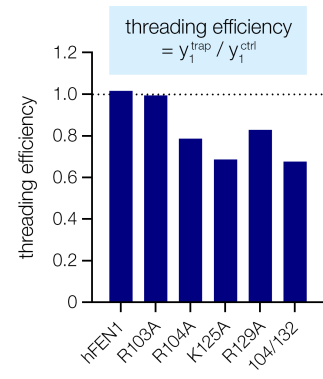


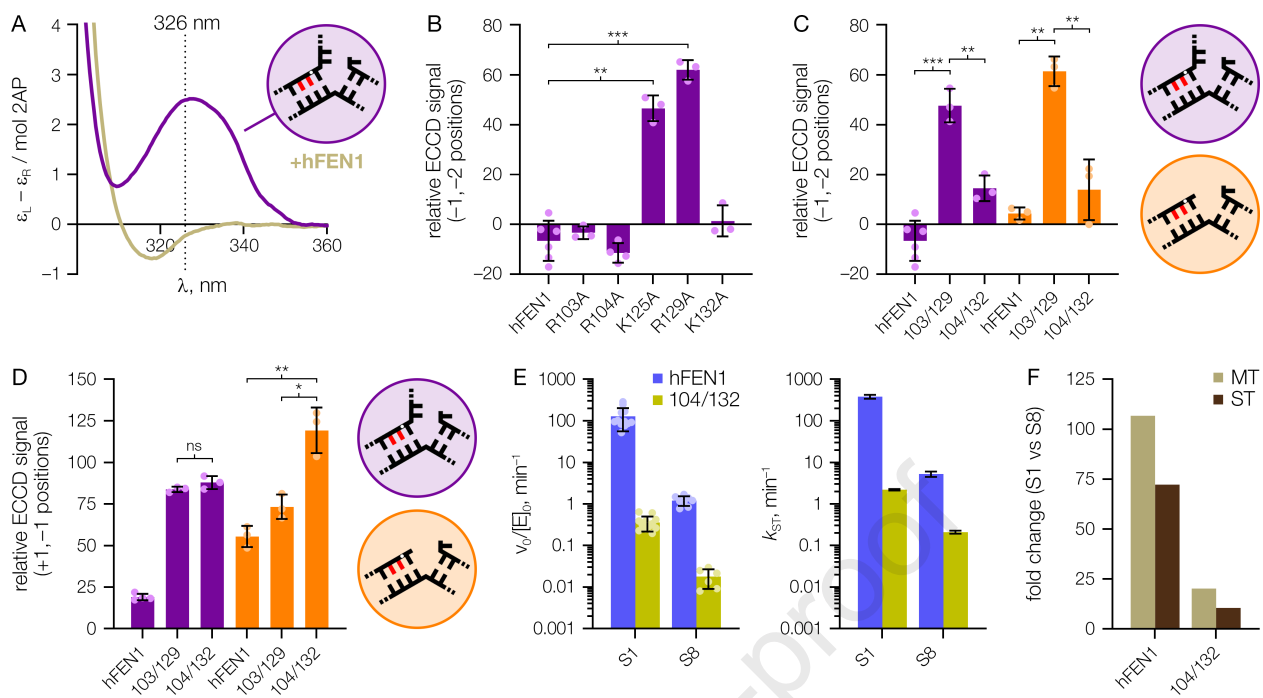


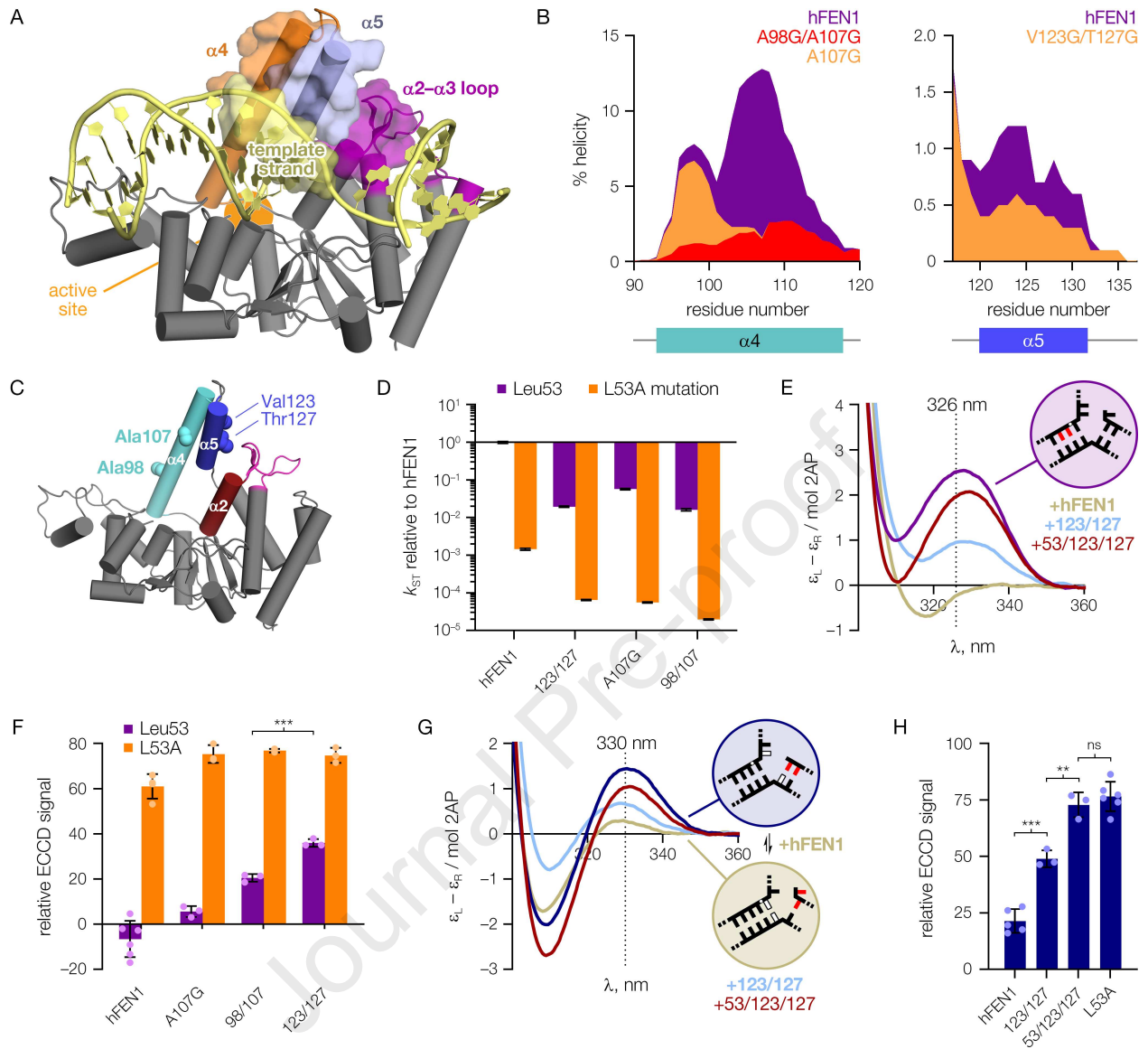
A



B







Mark J. Thompson: Conceptualisation, Methodology, Investigation, Formal analysis, Visualisation, Writing – original draft, Writing – review & editing, Supervision. **Nur Nazihah B. Md Shahari:** Investigation, Formal analysis. **Reuben J. Ouanounou:** Investigation, Formal analysis, Writing – original draft. **Nathan Gittens:** Investigation. **Barbara Ciani:** Conceptualisation. **L. David Finger:** Conceptualisation, Methodology, Writing – original draft, Supervision, Funding acquisition. **Jane A. Grasby:** Conceptualisation, Methodology, Visualisation, Writing – review & editing, Supervision, Project administration, Funding acquisition.

Journal Pre-proof

Declaration of Interest Statement

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

The author is an Editorial Board Member/Editor-in-Chief/Associate Editor/Guest Editor for this journal and was not involved in the editorial review or the decision to publish this article.

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests: