*Published online 11 August 2009* *Nucleic Acids Research, 2009, Vol. 37, No. 18* ***6161–6173***
*doi:10.1093/nar/gkp597*

# Mechanism of DNA substrate recognition by the mammalian DNA repair enzyme, Polynucleotide Kinase

**N. K. Bernstein[1], M. Hammel[2], R. S. Mani[3], M. Weinfeld[3], M. Pelikan[4], J. A. Tainer[5,6] and J. N. M. Glover[1,*]**

[1]Department of Biochemistry, University of Alberta, Edmonton, Alberta T6G 2H7, Canada, [2]Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA, [3]Department of Experimental Oncology, Cross Cancer Institute, Edmonton, Alberta, Canada, [4]Department of Mathematics and Computer Science, University of Missouri in St. Louis, St. Louis, MO 63121, [5]Department of Molecular Biology, Skaggs Institute of Chemical Biology, The Scripps Research Institute, La Jolla, CA 92037 and [6]Life Sciences Division, Department of Molecular Biology, Lawrence Berkeley National Laboratory, Berkeley, CA 94720, USA

Received April 7, 2009; Revised June 29, 2009; Accepted June 30, 2009

## ABSTRACT

**Mammalian polynucleotide kinase (mPNK) is a critical DNA repair enzyme whose 5′-kinase and 3′-phoshatase activities function with poorly understood but striking specificity to restore 5′-phosphate/3′-hydroxyl termini at sites of DNA damage. Here we integrated site-directed mutagenesis and small-angle X-ray scattering (SAXS) combined with advanced computational approaches to characterize the conformational variability and DNA-binding properties of mPNK. The flexible attachment of the FHA domain to the catalytic segment, elucidated by SAXS, enables the interactions of mPNK with diverse DNA substrates and protein partners required for effective orchestration of DNA end repair. Point mutations surrounding the kinase active site identified two substrate recognition surfaces positioned to contact distinct regions on either side of the phosphorylated 5′-hydroxyl. DNA substrates bind across the kinase active site cleft to position the double-stranded portion upstream of the 5′-hydroxyl on one side, and the 3′-overhang on the opposite side. The bipartite DNA-binding surface of the mPNK kinase domain explains its preference for recessed 5′-termini, structures that would be encountered in the course of DNA strand break repair.**

## INTRODUCTION

Mammalian polynucleotide kinase (mPNK) plays key roles in repairing DNA backbone breaks that commonly result from both damage and repair processes and that often do not contain the 5′-phosphate/3′-hydroxyl termini required for completion of DNA repair. 5′-hydroxyl and 3′-phosphate DNA ends are frequently encountered and can be generated by ionizing radiation, chemical agents and enzymatic action. For instance, endonucleases Neil1 and Neil2 excise abasic sites leaving a 3′-phosphate (1). DNase II cleaves double stranded DNA, producing 3′-phosphate and 5′-hydroxyl termini (2). DNA 3′-phosphate ends also result from exposure to neocarzinostatin or from cleavage of camptothecin-trapped Topoisomerase I - DNA adducts by tyrosyl-DNA phosphodiesterase 1 (Tdp1) (3). Another common non-canonical terminus, 3′-phosphoglycolate, may be converted to 3′-phosphate by Tdp1 (4). Accumulating evidence suggests that mPNK is critical for the repair of these non-ligatable ends. The essential role of PNK in DNA damage repair and radio resistance has been demonstrated both *in vitro* (5–7) and *in vivo* (8–10).

mPNK is involved in the repair of both single strand breaks (SSB), via the base excision repair (BER) pathway, and double strand breaks (DSB), via non-homologous end joining (NHEJ) (11,12). In the course of BER and NHEJ, mPNK physically interacts with the scaffold proteins XRCC1 and XRCC4, respectively, by binding to specific

*To whom correspondence should be addressed. Tel: +1 780 492 2136; Fax: +1 780 492 0886; Email: mark.glover@ualberta.ca

The authors wish it to be known that, in their opinion, the first two authors should be regarded as joint First Authors.

phosphothreonine sites on XRCC1 and XRCC4 via its FHA domain (8,13,14).

The crystal structure of mouse mPNK revealed its overall architecture (13) including three functional domains: 5′ DNA kinase and 3′-phosphatase domains, closely associated within the catalytic segment (PK) and an N-terminal FHA (Forkhead-associated) domain. The kinase and phosphatase possess distinct substrate specificities and have been shown to function independently of one another, with the phosphatase having a faster catalytic rate (15). The phosphatase has a minimal substrate requirement of three nucleotides and will accept either single-stranded, double stranded, nicked or gapped DNA substrates (13). The kinase preferentially phosphorylates 5′-hydroxyl groups within DNA nicks, gaps or at 5′-recessed DSBs, discriminating against single stranded or blunt double-stranded 5′-termini (13,16). This kinase substrate preference is consistent with roles in the repair of nicks, as well as single nucleotide gaps generated by Neil-family endonucleases (7,17). The unique mPNK substrate preference also suggests a specific role for the kinase in DSB repair by NHEJ. In NHEJ, mPNK may regenerate recessed 5′-hydroxyl termini, while single stranded or blunt double stranded 5′-hydroxyl termini are processed by the nuclease Artemis (18).

The mPNK FHA domain is a critical phosphopeptide-binding module found in many DNA repair proteins. In contrast to most currently characterized FHA domains, which in addition to phosphothreonine (pT) have a secondary specificity site at pT + 3, the mPNK FHA domain requires a negatively charged sequence N-terminal to the phosphothreonine, due to electrostatic complementarity between the acidic peptide and a positively charged peptide-binding surface of the FHA (13). The FHA domain is attached to PK by a 33 amino-acid linker, which was undefined in the crystal structure and was proposed to act as a flexible tether between the FHA and PK domains in the solution (13). This flexibility may be important in facilitating interactions between mPNK and its diverse protein partners and DNA substrates.

The effective orchestration of DNA end repair by mPNK critically relies on the recognition of the correct DNA termini for subsequent processing. Here we measure the DNA binding and 5′ kinase activities of a panel of site-directed mutants surrounding the kinase active site. This data, together with shape and structural information for mPNK-DNA complexes from SAXS experiments, characterize mPNK–kinase substrate interactions. The bipartite DNA-binding surface surrounding the mPNK kinase active site revealed by these studies explains the preference of mPNK for the internal 5′-hydroxyl termini, which it must recognize during DNA end repair processes.

## MATERIALS AND METHODS

### Cloning, expression and purification of mPNK constructs

Mutants of mouse PNK were cloned with an N-terminal $His_6$ tag to facilitate purification. Mutations, described in the text, were introduced by PCR mutagenesis or using the QuikChange II kit (Stratagene). Proteins were overexpressed in *E.coli* BL21 Gold cells in a 2× YT medium, induced with 0.1 mM IPTG and grown for 18 h at 22°C. Cells were lysed by treatment at 0°C with chicken egg white lysozyme in a lysis buffer (250 mM NaCl, 50 mM Tris pH 8.0, 1% BME) in the presence of proteinase inhibitors (0.1 mg/mL PMSF, benzamidine and leupeptin), followed by sonication ($8 \times 20''$). Cleared lysate was incubated with NiNTA beads (Qiagen) for 2 h. PNK was eluted by a gradient of imidazole (10–250 mM). The protein was further purified by size exclusion chromatography on a Superdex 75 column (Amersham) in 150 mM KCl, 50 mM Tris pH 8.3, 1 mM DTT. The purified protein was exchanged into 150 mM KCl, 10 mM Tris (pH 8.5), 1 mM DTT.

### Kinase activity assays

DNA oligonucleotides (IDT) were purified by ion exchange on a Source Q column (Amersham) in 10 mM NaOH with a gradient of 1 M NaCl, desalted using Sep-Pak cartridges (Waters) and annealed by slow cooling from 95°C in 150 mM NaCl and 10 mM HEPES (pH 7.5). Kinase activity was measured as previously described (13). Briefly, 10 μL reactions containing 2 μM gapped DNA substrate, 40 μM ATP, 5 μCi $\gamma^{32}$P-ATP, 80 μM succinate pH 5.5, 10 mM $MgCl_2$, 1 mM DTT and PNK (0.5, 1, 2 and 4 μM) were incubated 4 min at 37°C and stopped by addition of a 10 μL urea loading buffer (8 M urea, 1× TBE, 20% glycerol, bromophenol blue) and heating at 95°C. The reactions were run on a 12% denaturing polyacrylamide gel and quantitated by autoradiography.

### DNA-binding assays

DNA binding to mPNK (wild type and mutants) was measured by steady-state fluorescence spectroscopy, as described previously (13).

### SAXS data collection and evaluation

Solution scattering data were collected at the SIBYLS beamline BL 12.3.1 ALS Berkeley, California, and processed as previously described (19). Tunable wavelength (λ) and the sample-to-detector distances were set to 1.0–1.5 Å and 1.5 m, respectively, resulting in scattering vectors (q) ranging from 0.008 $Å^{-1}$ to 0.31 $Å^{-1}$ for $q = 4\pi\sin(\theta)/\lambda$, where $2\theta$ is the scattering angle. SAXS data at short and long time exposures (6 s, 60 s) were merged to define the entire scattering profile. Different protein concentrations were tested for aggregation and examined by Guinier plots (20). The radius of gyration ($R_G$) was derived by the Guinier approximation $I(q) = I(0)\ \exp(-q^2R_G^2/3)$ with the limits $qR_G < 1.3$. The curves measured for different protein concentrations (1.9–7.7 mg/ml) displayed a concentration dependence arising from interparticle interaction and interference-free scattering profiles were estimated by extrapolating the measured scattering curves to infinite dilution. Observed molecular weight for the PNK sample was estimated by calculating the normalized scattering intensity at zero angle [$q = 0$; $I(0)$] relative to four standard

proteins: lysozyme, xylanase, BSA and glucose isomerase (21). GNOM (22) was used to calculate the pair-distance distribution functions, $P(r)$, and define the maximum dimension of the macromolecule, $D_{max}$.

Scattering data were collected for free DNA substrates (H1 and H3, 33 μM each), for free proteins (mPNK 1.9–7 mg/ml and PK 1.9–7.7 mg/ml) and for protein:DNA complexes in 100 mM KCl, 33 mM Tris pH 7.5, 0.7 mM DTT, 0.7 mM $MgSO_4$. The DNA sequences were TATGATAC**GCA**GTATCATA**CCAAT** (H1) and TATGATAC GGCGCCTGGGG**GCA**CCCC AGGCGCCGTATCATA**CCAAT** (H3) (single-stranded loops and 3′ tails are highlighted in bold). The relative amounts of protein and DNA used in preparation of complexes were determined by titration on a 10% native polyacrylamide gel at points where all the free DNA was shifted into the protein:DNA band.

### Envelope modeling

The SAXS envelopes were reconstructed from the experimental data using DAMMIN (23). Twenty bead models obtained for each SAXS experiment were averaged by DAMAVER (24) to construct the average model representing general structural features of each reconstruction. Bead models were converted to volumetric SITUS format with the pdb2vol kernel convolution utility (25). To better define regions of the envelopes associated with the DNA, we calculated difference maps by first aligning the envelope obtained from the free protein with those obtained from the two complexes (PK/H1 or PK/H3). Subtraction of the aligned maps was carried out to define regions of difference density.

### Rigid body modeling of macromolecular assemblies with map objects (EMAP)

The EMAP docking approach was used to generate PK–substrate complex models and to rank these models based on their fit to the experimental data using CHARMM (26) (http://www.charmm.org/html/documentation/c34b1/emap.html). Both the protein and DNA substrates were represented as rigid domains with intact van der Waals and electrostatic properties and were moved as independent objects. The grid-threading Monte Carlo method was used to sample 1200 orientations of H1 over the PK surface, followed by an energy-based docking search. This strategy produced conformationally reasonable docked assemblies with a minimum of steric clashes. The theoretical scattering profile, $R_G$ and the corresponding fit to the experimental scattering curve were calculated using the program CRYSOL (27).

### Applying molecular dynamics on domains as rigid bodies

We developed a rigid body modeling strategy 'BILBOMD', in which molecular dynamics (MD) simulations were used to explore conformational space (28). Typically, we performed the MD simulation on the inter-domain linkers at high temperature, where the additional kinetic energy prevents regions from becoming trapped in a local minimum. The MD simulations provided an ensemble of molecular models from which SAXS curves were calculated by CRYSOL (27) and compared with the experimental curve. Using this method and strategy, different conformations of the FHA in mPNK were produced and validated against the experimental data.

The mPNK crystal structure (PDB ID: 1yj5) was used to construct the initial atomic models by connecting the C-terminus of the FHA domain with the phosphatase domain. This missing linker region (110–140), undefined in the crystal structure, was modeled as a random coil. CHARMM (26) with the all-atom force field version 27.0 (29) was used for molecular dynamics (MD) simulation. The initial atomic model was taken as the starting point of the simulations. Only the modeled linker atoms were allowed to move, while the FHA module at the extremity was treated as a rigid body, with no internal motion. The simulations were performed as previously described (30). A similar rigid body modeling strategy was used in the exhaustive conformational search of PK:DNA complexes with two distance restraints. The DNA 3′-overhang was allowed to be flexible and to be in close proximity (5 Å) to Lys483. The 5′-hydroxyl group of the DNA substrate was restrained to be ~2 Å from OD1 of Asp396.

### Minimal Ensemble Search (MES)

Considering the flexibility of the FHA domain or the presence of uncomplexed PK in PK/DNA samples, the coexistence of different conformations that contribute to the experimental scattering curve was taken into account. Based on the ensemble optimization method described by Bernado *et al.* (31), we developed an algorithm that searches for the minimal ensemble (MES) of the conformations from the pool of all generated conformations in previous MD simulations (28). The multiconformational scattering $I(q)$ from such a minimal ensemble was computed by averaging the individual scattering profiles from the conformers:

$$I(q) = \frac{1}{N(I_1(q) + I_2(q) + \ldots\ldots + I_N(q))},$$

where $I_{1,2,3,\ldots,N}(q)$ were the scattering profiles from the single conformers and the momentum transfer.

A genetic algorithm-based search was used to select an appropriate ensemble from a pool of all generated conformations. The scattering curves from all structures in the pool were pre-computed and the selection was performed using these curves. The final model achieved the best fit to the experimental curve $I_{exp}(q)$ by minimizing the discrepancy $\chi^2$ between the experimental and calculated multi-conformational curve:

$$\chi^2 = \frac{1}{K-1}\sum_{j=1}^{k}\left[\frac{\mu I(q_j) - I_{exp}(q_j)}{\sigma(q_j)}\right]^2,$$

where $K$ was the number of experimental points, $\sigma(q)$ were standard deviations and μ was a scaling factor (31).

Comparison of the structural properties of the selected models in the ensemble subset allowed us to distinguish

the degree of flexibility of the experimental system. As shown in Figure 3D, the spread of structural parameters $R_G$, CαRMSD for the MES-derived ensembles relative to those determined for the entire pool, correlated strongly with the level of disorder. We observed large differences between structural parameters of the selected ensemble models for flexible/disordered systems like the FHA domain in mPNK.

## RESULTS

### PNK interaction with its kinase substrate

The kinase active site of mPNK is located in a deep cleft (Figure 1A). Inspection of the electrostatic potential surface of the active site reveals two positively charged surfaces flanking the catalytic Asp396, which acts as a general base to activate the substrate 5′-hydroxyl group for phosphorylation. 'Surface 1' contains Arg395, Arg403 and Arg432 while 'Surface 2' consists of Arg482 and Lys483 (Figure 1A). Likewise, the mPNK DNA substrate may also be viewed as a bipartite structure. The preferred kinase substrate is a dsDNA with at least eight base pairs, and a 3′ overhang of four or more bases (13). The substrate 5′-hydroxyl is thus flanked on two sides. The 'upstream' side is the double-stranded portion, and the 'downstream' side is the 3′-overhang, which can be either single- or double-stranded (Figure 1C).

### The effects of kinase domain mutations on substrate binding and catalysis

To test the role of the structurally implied DNA binding surfaces, we used site-directed mutagenesis to perturb the two DNA-binding areas. One set of mutations was designed to specifically disrupt interactions with 'Surface 1': R395A, W401A, R403A, R403E, T423A and R432A. Although Thr423 and Trp401 are not basic, they were proposed to have a role in DNA binding (13). Thr423 is conserved in T4 PNK and interacts with a backbone phosphate. Trp401 binds a sulphate ion in the native mPNK structure via the partially positively charged edge of its indole ring. A second set of mutations was constructed to target 'Surface 2': R482A, K483A, K482/R483AA and K482/R483EE. Additionally, the catalytic Asp396 (D396N) was mutated.

To characterize the mutants' affinity for a minimal DNA substrate (H1), we initially employed a fluorescence assay, where intrinsic mPNK tryptophan fluorescence is quenched upon ligand binding (Figure 1B, Table S1). Dissociation constants ($K_d$) were determined for all mPNK mutants binding to H1, the model substrate with a recessed 5′-hydroxyl. Binding of mPNK and two mutants (R432A and K482/R483EE) to DNA hairpin H1, a blunt-ended analog of H1 was also measured. The general trend observed was a decrease in binding affinity for most mutants relative to wild type. The most notable differences were found for R395A, D396N, R403A, R403E, K482/R483AA and K482/R483EE binding to H1. A comparison of binding affinities for recessed versus blunt-ended 5′-hydroxyl termini revealed that both the wild type and R432A binding to the blunt-ended substrate was weaker than to the recessed one. R482/K483EE, however, exhibited the reverse pattern, showing a binding preference for the blunt-ended DNA.

To more precisely analyze the mPNK mutants, we assayed the kinase activity of all mutants on a model substrate with two phosphorylatable 5′-ends: a recessed one within an internal 1-nucleotide gap and a blunt one (Figure 1C–F). mPNK selectively phosphorylates the internal 5′-end much more readily than the blunt 5′-end. Reactions were carried out over a range of enzyme concentrations and phosphorylated products were visualized by denaturing PAGE and autoradiography. Enzyme reactivity levels were analyzed in two ways. First, the efficiency of phosphorylation of the internal and the blunt 5′-end for each mutant was compared to the wild-type control (Figure 1E). Second, we compared the efficiency of phosphorylation of the internal site relative to the blunt site, normalized to the reactivity of the internal 5′-hydroxyl at the highest protein concentration for each mutant (Figure 1F). In this way, we could detect changes in the relative activity on two different substrate ends in a single reaction and avoid the introduction of additional error from the wild-type data.

This analysis revealed that all the mutants except D396N retained activity on the recessed 5′ terminus that was comparable to wild-type PNK. In D396N, the catalytic aspartate responsible for activation of the substrate 5′-hydroxyl is replaced with an isosteric asparagine. This mutant exhibited higher selectivity for the recessed 5′-hydroxyl and a dramatic decrease in the overall kinase activity to 23% of WT (Figure 1E and F). The significant drop in catalytic activity in the D396N mutant is expected due to the loss of the general base, and the residual activity observed may be attributed to proximity and orientation effects in the active site. The contribution of these factors to catalysis varies in members of the adenylate kinase family. For example, in 6-phosphofructo-2-kinase (PDB: 1bif) and in human UMP/CMP kinase (PDB: 1tev), the residue corresponding to D396 is glycine, suggesting that a general base at this position is not required in these two enzymes (32,33). The complete absence of phosphorylation of the blunt 5′-hydroxyl by D396N PNK also supports the idea that Asp396 is required for substrate binding, since the loss of this interaction demands binding of the downstream portion of DNA. Moreover, the binding affinity of D396N for the recessed 5′-hydroxyl substrate is significantly lower than for wild type, confirming Asp396 as a crucial binding determinant (Table S1).

While activity levels on the recessed terminus were similar to wild type, there was a significant variation among mutants in activity on the blunt terminus. In particular, mutation of residues Thr423, Arg432, Arg395 and Trp401 resulted in markedly increased preference for the recessed 5′-hydroxyl relative to wild-type mPNK. Mutation of Arg403 to alanine did not have a noticeable effect, but a more radical substitution, R403E, increased the selectivity for the recessed 5′-hydroxyl. While mutation of Arg482 did not affect the substrate specificity, mutants involving Lys483 displayed enhanced preference for the blunt 5′-hydroxyl.

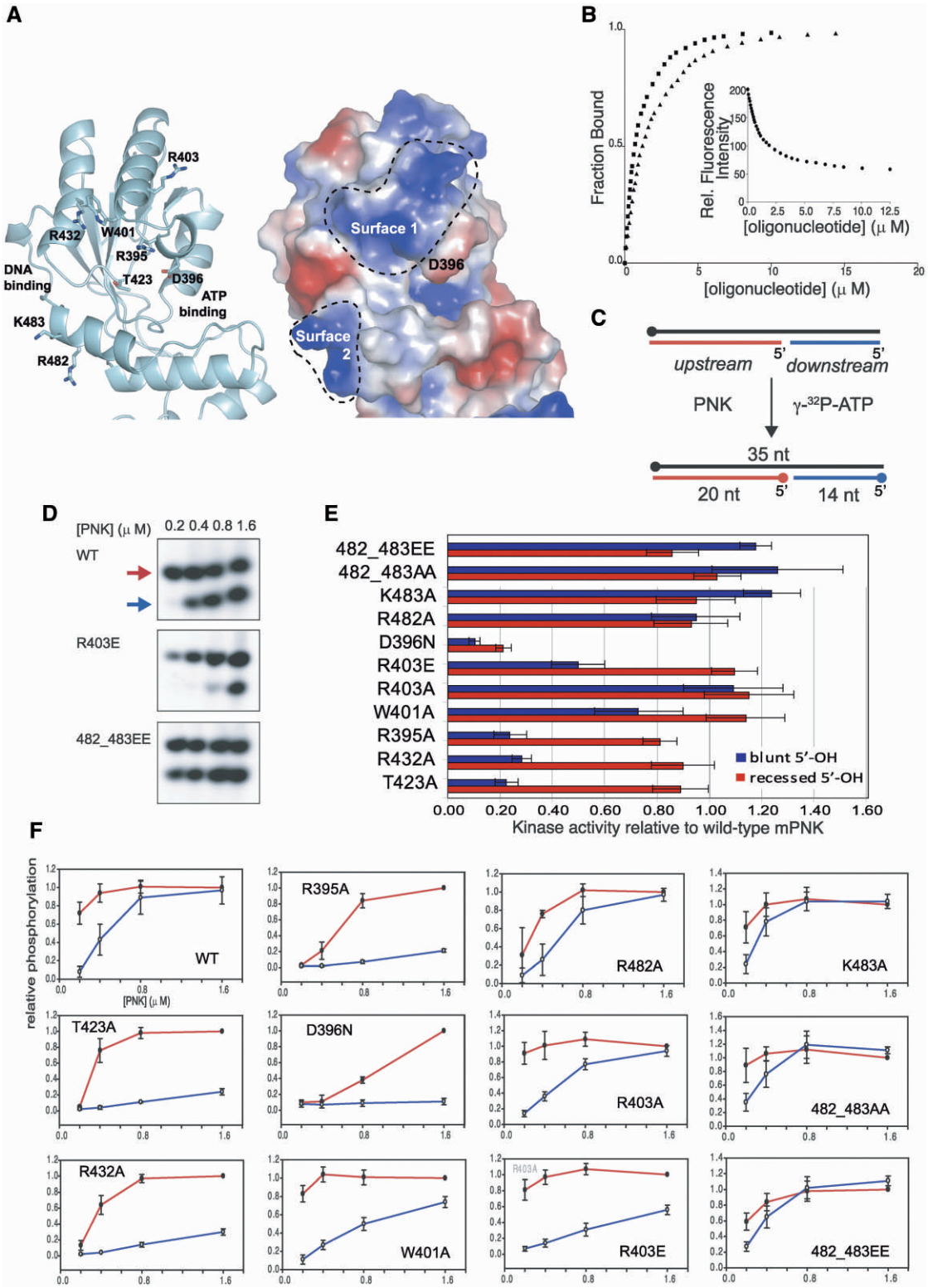


**Figure 1.** Probing mPNK–DNA interactions by site-directed mutagenesis. (**A**) View of the mPNK kinase highlighting the residues mutated in this study (left) and the positively charged surfaces surrounding the active site cleft (right). (**B**) Fluorescence titration of mPNK with hairpin substrates (filled triangle: blunt and filled square: recessed). The inset figure shows a sample raw data plot for the recessed hairpin substrate. (**C**) Schematic representation of the DNA substrate used in the kinase activity and binding assays. (**D**) Representative gels of kinase assays. The red arrow indicates phosphorylated 20-mer oligonucleotide and the blue arrow indicates the 14-mer oligonucleotide of the model substrate (C), corresponding to the recessed and blunt 5′-hydroxyl termini, respectively. (**E**) Comparison of the overall kinase activity of PNK mutants on blunt (blue bars) and recessed (red bars) 5′-hydroxyl termini, relative to wild-type PNK. Overall activity is calculated as an average of ratios of mutant to wild-type activity, weighted by band intensity for mutant activity at each point. (**F**) Results of the kinase activity assay of mPNK mutants on the model substrate. Graphs depict relative phosphorylation level as a function of protein concentration, normalized to the activity of each mutant on the recessed 5′OH at the highest protein concentration. The red and blue curves correspond to the recessed and blunt 5′-hydroxyl termini, respectively.

A general trend was observed for the group of 'upstream' mutations where selectivity for the recessed 5′-hydroxyl increased relative to wild-type mPNK. Whereas wild-type mPNK lost its preference for the internal 5′-hydroxyl at high enzyme concentrations, mutants within 'Surface 1' (R395A, D396N, T423A, R432A and W401A) remained selective for the internal 5′ terminus even at the highest enzyme concentrations tested.

Conversely, mutation of the residues in 'Surface 2', Arg482 and Lys483, resulted in a different pattern of reactivity. Mutation of Arg482 to alanine did not affect substrate specificity, but replacing Lys483 with alanine or glutamate (K483A, R482/K483AA, R482/K483EE) resulted in a loss of specificity for the recessed 5′-hydroxyl. In fact, at the higher protein concentrations, the double mutants showed a slight (~12%) preference for the blunt-ended substrates (Figure 1E and F). Binding affinity for the recessed 5′-hydroxyl to R482A and K483A mutants were only slightly lower than for the wild-type, but the double mutants exhibited significantly weaker binding (Table S1). Additionally, the binding preference of the R482/K483EE mutant for blunt-ended over the 5′-recessed DNA confirms the interaction of this area with the downstream portion of the substrate. Discrimination against the 5′-recessed hydroxyl may be due to electrostatic repulsion between the DNA backbone and the glutamates in R482/K483EE.

We note that the sequences of the two phosphorylated DNA strands differ (Figure S1). However, in control experiments on substrates with the same sequence at both 5′ termini, the trend described above was observed (data not shown), ruling out sequence specificity as the reason for the enzyme's preference for recessed 5′ hydroxyl.

### Structural analysis of mPNK binding to DNA in solution

To experimentally characterize solution structures for mPNK complexes, we employed SAXS to define the shapes and conformations of mPNK complexes including their functional flexibility. The interactions of mPNK complexes with DNA substrates have consistently eluded crystallographic experiments and typify functionally flexible systems that do not form ordered crystals, but are suitable for SAXS analyses (19,34,35). To facilitate the interpretation of the solution scattering data, we used two DNA hairpins (H1 and H3) based on the minimal preferred substrate, but differing in length of the double-stranded portion (8 and 19 base-pairs, respectively). Based on earlier work, the protein was suspected to possess intrinsic domain flexibility (13), and this question was addressed by SAXS analysis prior to determination of the complex structure.

### Domain flexibility in mammalian PNK

To test the domain flexibility of mPNK, we examined X-ray scattering by mPNK constructs. SAXS experiments indicate that the FHA domain of PNK is flexibly attached to the catalytic (PK) segment. mPNK has an extended maximal protein dimension ($D_{max}$) and $R_G$ values (calculated based on Guinier plots, Figure 2B) of ~150 Å and $35.3 \pm 0.2$ Å, respectively. This is consistent with hydrodynamic studies (36), but significantly larger than values from the crystal structure ($D_{max}$ ~105 Å and $R_G = 30.9$ Å: PDB 1yj5) (Table 1). Furthermore, the FHA domain shows large-scale conformational changes, based upon the broadened $P(r)$ function and elongation of the $P(r)$ tail, a characteristic of an extended conformation (Figure 2C). To test if partial dimerization explains the experimentally observed elongation, we performed two independent experiments. First, measurement of intensities at zero angles and four reference proteins to estimate the molecular weight of PNK (see Experimental Procedures) showed mPNK to be a monomer in solution. Second, the $P(r)$ function for PK, derived from solution scattering (Figure 2C), indicates a globular, less extended particle with a smaller $D_{max}$ of ~110 Å (Table 1). As expected, the 14-residue, solvent-exposed N-terminal tail resulted in some residual elongation in the $P(r)$ function, consistent with most of the extension of full-length mPNK coming from the FHA domain.

To further test mPNK flexibility in solution, we used rigid body modeling. In our strategy, a molecular dynamics (MD) simulation was used to probe the conformational distribution of FHA that best matched the solution scattering data (see 'Materials and methods' section). Based on several thousand conformations of mPNK and their calculated SAXS profiles, the SAXS-validated best-fitting model ($\chi^2 = 2.1$) had the FHA linker in an extended conformation (Figure 3A and B). Considering the extended and flexible nature of the FHA linker, conformer combinations were selected using a minimal ensemble search (MES) (28) approach and a mixture of two conformers fit the data better ($\chi^2 = 1.8$) than the single model. Fits improved for mixtures of three, four or five conformers ($\chi^2 = 1.5$ for the assembly of five conformers) (Figure 3A and C). Aside from the variation of the selected conformers (Figure 3C), the broad range of their calculated RMSD and $R_G$ values (Figure 3D) indicated that in solution the FHA domain is not held in a fixed orientation with respect to the rest of the protein. By contrast, the atomic model of PK matched its experimental scattering curve ($\chi^2 = 2.0$) (Figure 2A). For the PK solution structure, the N-terminal extension (Figure 5C) is important to match the data in the entire $q$ range. Removal of this extension significantly diminished the goodness-of-fit to $\chi^2 \sim 4.5$, consistent with the extended character of the FHA linker.

### Visualization of the PK/DNA complex

To test the effect of DNA binding to mPNK, we examined x-ray scattering results with the PK catalytic segment. We reasoned that removal of the mobile FHA would facilitate interpretation and comparison of SAXS parameters to structural models. The properties of the free PK in solution were compared to complexes of PK with the short (H1, 8 bp) and long (H3, 19 bp) DNA hairpins.

We examined differences between the PK and PK/DNA complexes (Figure 2). The significant shoulder in the $P(r)$ function at $r \sim 50$ Å in the DNA-containing complexes indicates binding of DNA to PK, with greater differences

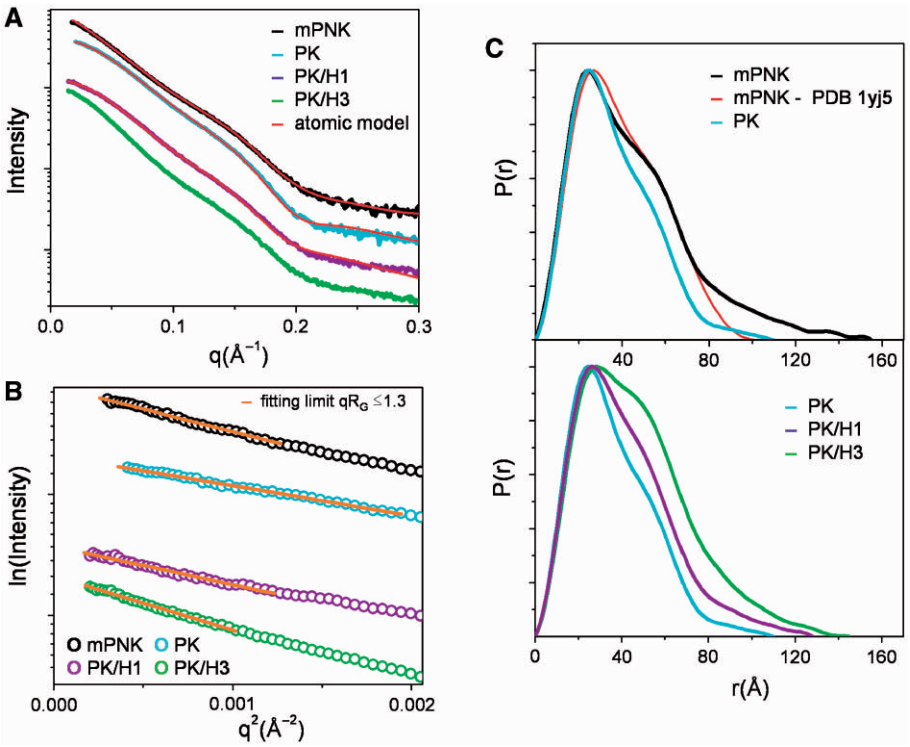


**Figure 2.** Solution structural information from X-ray scattering data. (**A**) Experimental SAXS curves of the mPNK and PK-assemblies (mPNK: black, PK: blue, PK/H1: violet, PK/H3: green). Red line shows the theoretical scattering from the atomic models shown in Figure 3C: mPNK and Figure 5C: PK and PK/DNA complexes. (**B**) Guinier plots for the experimental data with linear fit (orange line) in the limit $qR_G > 1.3$. (**C**) Distance distribution functions of the mPNK, PK and PK/DNA complexes computed from experimental SAXS data shown in the same color scheme as panel A. The $P(r)$ function for crystal structure of mPNK (PDB 1yj5) is shown in red. The $P(r)$ functions are normalized to unity at their maxima.

**Table 1.** Global SAXS parameters for experimental and theoretical data

| Assembly | $R_G$ (Å)[a] | $D_{max}$ (Å)[b] | $\chi^2$ single model[c] |
|---|---|---|---|
| mPNK | 35.3 ± 0.2 | ~150 | 2.1/1.5[d] |
| mPNK (crystal structure) | 30.9 | 105 | 2.6 |
| PK | 28.1 ± 0.1 | ~110 | 2.0 |
| PK/H1 | 31.0 ± 0.3 | ~130 | 2.5 |
| PK/H3 | 35.4 ± 0.2 | ~140 | |

[a]$R_G$ radius of gyration given by the Guinier approximation (20).
[b]$D_{max}$ maximum protein distance estimated from $P(r)$ function.
[c]The goodness of fit $\chi^2$ for the best-fit atomic model.
[d]Multiconformational fit $\chi^2$ for MES models.

for the longer DNA (H3) (Figure 2C and Table 1). To visualize the three-dimensional shapes of the PK/DNA complexes in solution, we calculated their SAXS envelopes and compared them with those calculated for free PK. Single envelopes showed only slight variations between independent runs (Figure 4). The difference between the envelopes occurs since fitting the scattering curve does not suggest a single structure and averaging is needed to construct an accurate model for the observed scattering. The average envelopes for free PK match the shape of the PK domain, with the knob-like protrusion at the top corresponding to the solvent exposed N-terminal tail (Figure 5C). Envelopes constructed for PK/H1 and PK/H3 also displayed this N-terminal extension, with an additional feature visible at the opposite extremity of the model. This region is larger for the longer DNA (Figure 4B), indicating that this part of the model corresponds to the position of DNA. Based on the shape differences in the SAXS envelopes, we conclude that both DNA hairpins are bound to the kinase with one end extending away from the protein into solvent.

To further test the position of DNA predicted from SAXS envelopes, we performed a validation of the exhaustive search of the protein–DNA interfaces by an EMAP docking approach (see 'Materials and Methods' section). In this procedure, the PK position was fixed and the H1 DNA was docked to PK by applying a grid-threading Monte Carlo sampling approach. 1200 orientations of H1 were sampled over the PK surface (Figure 5A), and these were further refined by an energy-based docking search. Theoretical scattering profiles of these models were calculated and the models were scored based on their fit with the experimental scattering profiles. The best-fit model has $\chi^2 = 3.8$ and superimposes well with the SAXS envelopes. The DNA 5′-end is located in close proximity to the kinase domain while the hairpin loop at the opposite end of the DNA points to the solvent (Figures 5C, 6A and B). The two best-fit models show similar positions of the H1 DNA relative to PK, however, the DNAs are bound in opposite orientations (Figure 5C). This comparison suggests that SAXS data alone are not sufficiently sensitive to distinguish between DNA models in either of these two orientations.

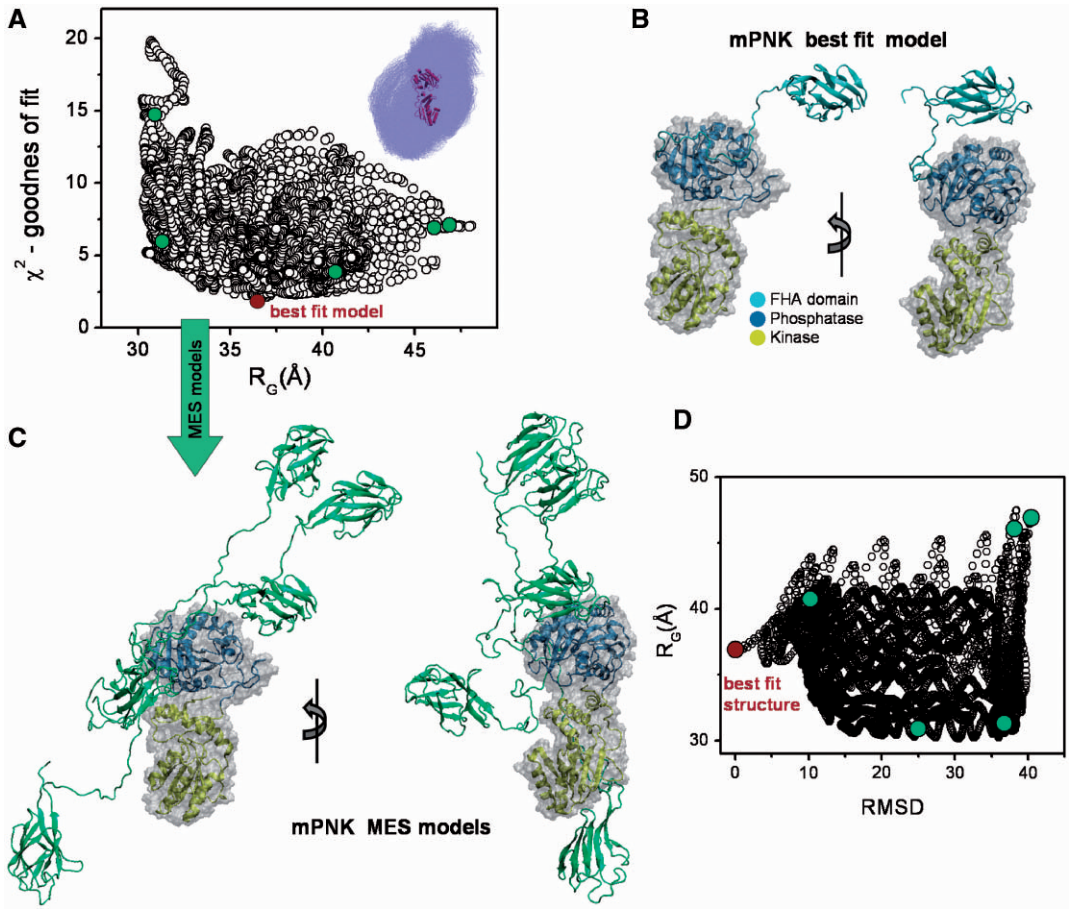


**Figure 3.** FHA domain mobility in solution. (**A**) Comparison of discrepancy with experimental data ($\chi^2$) values for 6000 mPNK models with their $R_G$ values. The best fit model within the region of the experimental $R_G$ value (35 ± 0.2 Å) is indicated by the red circle. Green circles indicate MES-selected models, shown in panel (C). Inset: 6000 models of mPNK conformers colored blue superimposed on PK shown in cartoon representation. (**B**) Best fit structure for mPNK, obtained by conformational sampling, matching experimental data with $\chi^2 = 2.1$. (**C**) Five MES-selected models, representing the conformational space sampled by the FHA domain. The models were superimposed with their PK domains, shown in surface/ribbon representation. In (B) and (C), the kinase domain is in yellow, phosphatase in blue and FHA in green. (**D**) Comparison of discrepancy ($R_G$) values for all 6000 mPNK models with their RMSD values referenced to the best fit structure shown in (B) (red circle). MES-selected models are indicated by green circles.

To test the implications of the biochemical results for mPNK solution conformations, we performed an exhaustive conformational search applying two biochemically based distance restraints. One restraint was applied to keep the 5′-hydroxyl within ~2 Å from OD1 atom of the general base, Asp396. The second restraint was applied to maintain at least one phosphate of the flexible DNA 3′ overhang to within ~5 Å of Lys483. Two best-fit models of PK/H1 complex gave an excellent fit to the experimental data ($\chi^2 = 2.5$, Figure 2A) and also fit well into the *ab initio* SAXS envelope (Figure 5C).

## DISCUSSION

### An accurate solution structural model for mPNK-DNA recognition

Detailed insights into the catalytic mechanism of enzymes such as mPNK that have functionally important flexibility have been hindered by a lack of structural information for enzyme–substrate complexes. For such systems, we developed and herein present and apply an approach that uses structural information from small angle X-ray scattering together with detailed functional data from a panel of site-directed mPNK mutants to test and refine models of mPNK–substrate recognition. Our analysis began with the classification of mutants into two groups based on their specificity for either blunt ended or internal/recessed 5′-ends. The 'Surface 1' group showed enhanced selectivity for the internal 5′-end compared to wild type. We suggest that this enhanced specificity may result from a deficiency in recognition of the 'upstream' portion of the DNA, and thereby conclude that the mutated residues play a role in the recognition of this region. In contrast, mutation within 'Surface 2' led to a reduced selectivity for the recessed 5′-end compared to a blunt-ended substrate, indicating that this region may be critical for recognition of the downstream portion of the substrate.

Our SAXS-based model of an mPNK–substrate complex (Figure 6A) builds upon existing crystal structures to explain the biochemical and mutational results. The DNA-binding surface of the kinase in solution is structurally complementary to the substrate. Thus, D396 interacts

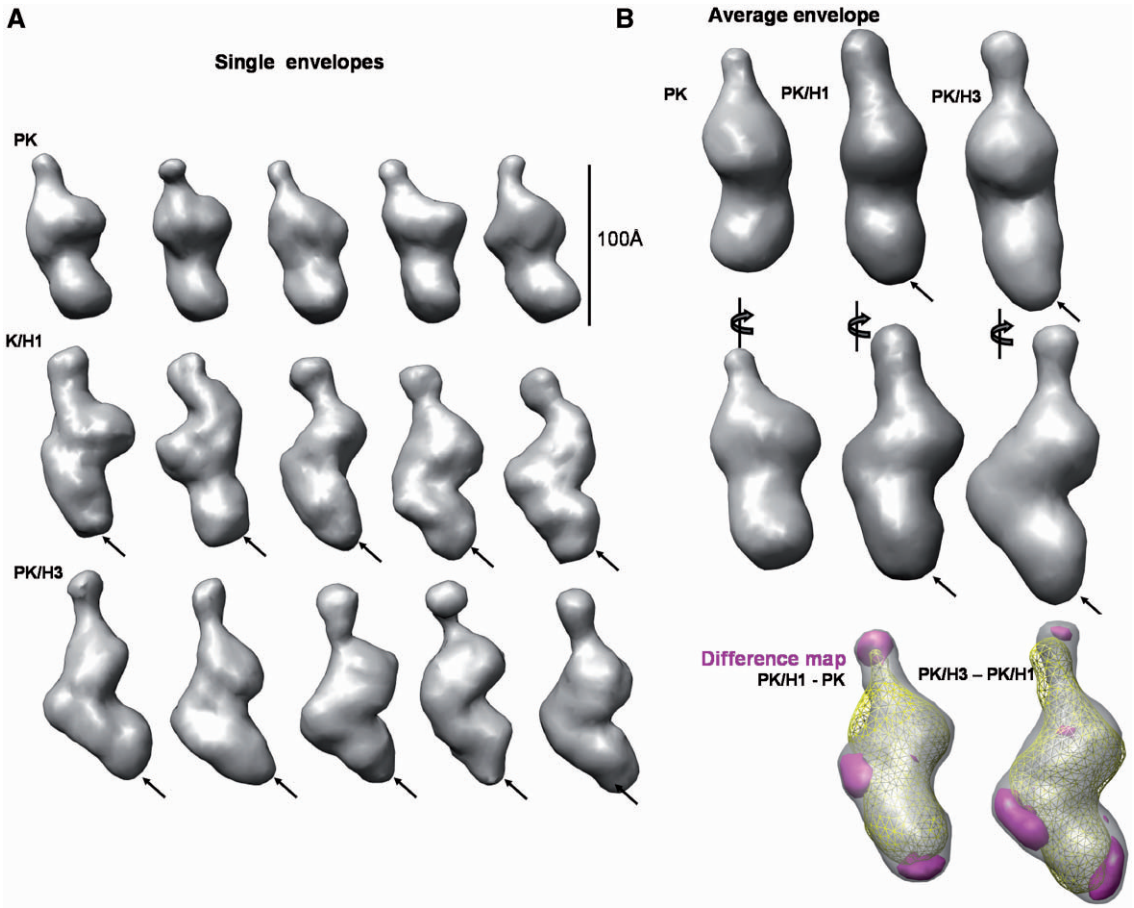


**Figure 4.** DNA binding to the catalytic component of mPNK. (**A**) Five representative DAMMIN models of PK, PK/H1 and PK/H3 assemblies are shown in volumetric SITUS format. (**B**) Two views of the average SAXS envelopes rotated by 90° are shown. Bottom panel shows experimental difference maps for PK/H1–PK or PK/H3–PK/H1, respectively. In the bottom left image, the PK/H1 envelope is shown in grey, while the PK envelope is shown as yellow mesh. In the bottom right image, the PK/H3 envelope is in green while the PK/H1 is in yellow mesh. In both images the difference between these two envelopes is displayed in purple.

with the 5′-OH, the 'upstream' portion of the substrate binds to 'Surface 1', and the downstream structure contacts 'Surface 2' (Figure 1A and 6A). The orientation of the substrate was identified with the use of two DNA hairpins of different lengths.

In the 'Surface 1'–'upstream' interaction, R395 and the partially positively charged edge of the W401 indole ring are positioned to make electrostatic contacts with the phosphodiester backbone of residues 2 nucleotides 3′- to the 5′-hydroxyl terminus. A positively charged surface patch formed by Arg403 and Arg432 is more distant from the site of catalysis, interacting with the double-stranded part of the substrate around base pairs 6–8 (Figure 6A and B). This provides an explanation for the minimal substrate length requirement of eight base pairs (13,16). Mutations of Arg432 to alanine and Arg403 to glutamate resulted in substantially higher selectivity for an internal 5′-hydroxyl compared to a blunt-ended substrate.

Disruption of the second positively charged patch, 'Surface 2' (Figure 1A), results in a significant reduction in specificity for the internal 5′-end compared to the blunt end. The effect is most dramatic for mutation of Lys483. Arg482 appears to be less important, possibly due to being farther from the active site than Lys483.

We used the biochemical information that we derived for mPNK to formulate distance restraints to guide the docking of DNA and mPNK molecular models into the SAXS data. We introduced two conservative restraints, one which restricted the 5′-hydroxyl of the DNA substrate to be near Asp396 and one which restrained the 3′ overhang to be near Lys483. These restraints not only helped to ensure that the substrate was docked into the kinase active site in a biochemically reasonable manner but also gave models that fit the SAXS data significantly better than models derived from unrestrained rigid body docking (Figure 5).

### Structural basis for the differential binding specificities of mammalian and T4 PNK

Mammalian polynucleotide kinase and its bacterial ortholog, T4 polynucleotide kinase, display strikingly different kinase substrate preferences. While the mammalian enzyme prefers recessed and internal 5′-hydroxyl groups within double stranded substrates, the bacterial enzyme strongly prefers single stranded 5′-hydroxyl ends.

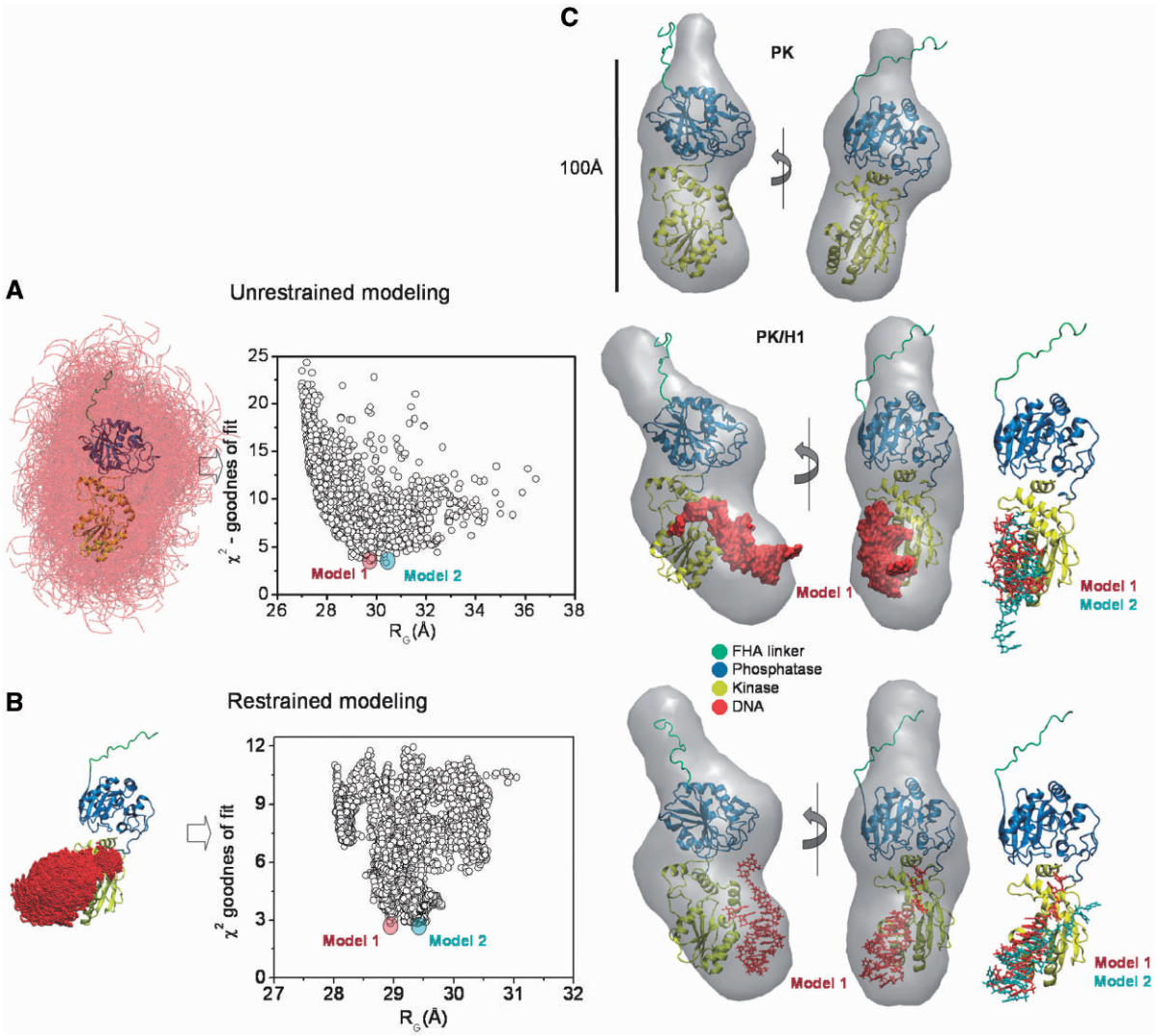


**Figure 5.** Rigid body modeling of PK/DNA complexes. (**A**) Unrestrained rigid body modeling of the PK/H1 DNA complex. The 1200 PK/H1 models generated through the unrestrained EMAP procedure shown on the left with the H1 DNA in transparent red tube representation. PK is shown as ribbons with the FHA linker: cyan, phosphatase: blue, kinase: yellow. A comparison of the $\chi^2$ values for all modeled assemblies with their $R_G$ values is plotted in the right panel. Two best-fit models within the region of the experimental $R_G$ value ($31.0 \pm 0.3$ Å for PK/H1) are indicated by red and cyan circles and shown in the panel (C). Two views (rotated by 90°) of the first best-fit PK/H1 model are shown superimposed on its envelope, with similar views of the free PK structure (C). (**B**) Molecular dynamics modeling of the PK/H1 complex restrained by SAXS data and biochemical restraints. The superimposition of the 6000 PK/H1 complexes modeled through the restrained BILBOMD procedure is shown on the left and the $\chi^2$ vs. $R_G$ values for these complexes are plotted in the right panel. Two best-fit models are indicated by the red and cyan circles and shown in panel (C). (**C**) Best fit models for PK and the PK/H1 complex from unrestrained and restrained refinements, superimposed on corresponding average SAXS envelopes. Left panel shows comparison of two best-fit models from unrestrained and restrained refinements.

Crystallographic studies of the T4 kinase domain bound to short oligonucleotide substrates reveal that three nucleotides at the 5′ terminus of the substrate bind within a deep cleft leading to the active site (37). The mPNK active site, however, is much more open than that of T4, due to a structural difference in the α9-loop-α10 segment. In T4, this region protrudes over the active site and forms a barrier that separates the DNA and ATP binding grooves and precludes binding of a double-stranded substrate. In mPNK, the truncated α9-loop-α10 region can dock against the DNA major groove, allowing access of the 5′-hydroxyl to the kinase active site without distortion of the double-stranded DNA (Figure 6B and C).

## Implications for 5′DNA end repair

The preference of the mPNK kinase domain for internal or recessed 5′-ends matches the DNA substrates that the enzyme will encounter in the repair of both single-strand and double-strand breaks. The repair of certain single-strand breaks is particularly interesting in that both the kinase and phosphatase activities of mPNK may be required. For example, camptothecin induces topoisomerase I-DNA adducts that are cleaved to 3′-phosphate/5′-hydroxyl-containing DNA nicks by the phosphatase TDP1 (3). The arrangement of the two active sites, which are separated by ~40 Å, could not simultaneously

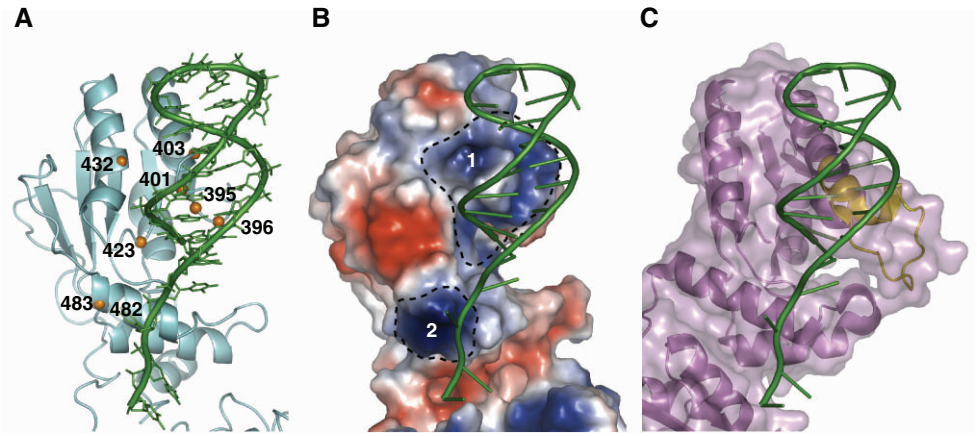


**Figure 6.** DNA binding to the kinase active site. (**A**) Optimized SAXS model of DNA hairpin H1 (green) bound to the kinase domain of mPNK (cyan). The residues mutated in this study are highlighted by orange spheres at Cα. (**B**) DNA H1 bound to mPNK, colored according to surface electrostatic potential. (**C**) T4 PNK (purple), structurally superimposed onto the mPNK:DNA hairpin complex. DNA is in green and mPNK is omitted for clarity. Steric interference between the DNA and the α9-loop-α10 region of T4 PNK (yellow) may account for the requirement of single stranded substrates for the T4 kinase.

interact with 3′-phosphate and a 5′-hydroxyl within a nick or single-nucleotide gap. The phosphatase-active site binds its substrates with much higher affinity than the kinase domain, suggesting that engagement of the phosphatase-active site may be the first step in the repair of 3′-phosphate/5′-hydroxyl single-strand breaks (15). Experiments monitoring the repair of these ends suggest that the 3′-phosphate end must be bound, dephosphorylated and released before the 5′-hydroxyl end can be bound by the kinase domain. Binding of substrate to the phosphatase active site sterically inhibits binding of a second substrate to the kinase active site (15), likely due to the fact that the two active sites are located on the same face of the mPNK catalytic domain. The orientation of the mPNK kinase substrate suggests that the 'downstream' regions of the substrate would protrude over the phosphatase active site, offering a possible explanation for the inhibition of the kinase domain by substrate binding at the phosphatase active site. PNK recruitment to single strand breaks in cells requires interactions with CK2-phosphorylated XRCC1 via the PNK FHA domain (8). Recent work by Ali *et al*. suggests that two molecules of PNK can bind to a single XRCC1, enabling one copy to act as the kinase, and the other as a phosphatase (38). Whether PNK or XRCC1 binds first to the strand break is controversial. Experiments by Parsons *et al*. (39) indicated that binding to damaged DNA substrates precedes binding by XRCC1, i.e. PNK binding to DNA does not require XRCC1. After catalysis, the mPNK kinase domain retains significant affinity for the 5′-phosphorylated product, releasing it only for subsequent processing by DNA ligase III (40). On the other hand, Loizou and coworkers showed that CK2-mediated phosphorylation of XRCC1 enhances recruitment of XRCC1 and PNK to DNA strand breaks following exposure of cells to hydrogen peroxide (8).

While the preference of the mPNK kinase domain for recessed or internal 5′-hydroxyl ends is perfectly suited to its role in SSB repair processes, this specificity presents a potential problem for the repair of double strand breaks, where blunt or single stranded 5′-termini may be encountered. However, these ends could first be processed by the DNA–PK-associated nuclease, Artemis (13). In response to phosphorylation by DNA–PK, Artemis binds and cleaves double strand breaks with 5′ or 3′ overhanging ends (18). Large 3′ overhangs are trimmed to leave an ∼4 nucleotide 3′ overhang, the minimal mPNK kinase substrate. 5′ overhangs are rapidly converted to blunt ends and more slowly processed to remove further 5′ and 3′ nucleotides to yield a short ∼4 nucleotide 3′ overhanging end (41). In this way, Artemis could convert 5′ overhanging or blunt ends to 3′ overhangs containing 5′-phosphate groups. Double strand breaks with 3′ overhangs and 5′-hydroxyl termini would require the additional action of mPNK to produce the 5′-phosphate termini required for ligation.

Taken together, the combined solution results on mPNK interactions presented here provide a detailed understanding of how its structural biology helps to orchestrate DNA end repair. Such an understanding for mPNK is likely important as the specific recognition, processing and control of such DNA ends appear to be a critical repair protein function for base excision repair (42,43), primer removal in DNA replication (44) and double-strand break repair (45). Here the integrated solution structure of mPNK reveals a flexible overall architecture coupled to a bipartite DNA-binding surface at the mPNK kinase active site that is structurally complementary to its substrate. These experimentally characterized features suggest how mPNK may interact with diverse DNA substrates and protein partners while maintaining a preference for internal 5′-hydroxyl termini, which it must recognize during DNA end repair. As our results on mPNK establish that SAXS can be readily combined with biochemical information to define flexible complexes that furthermore can be robustly combined with other structural methods, the approach employed here may be generally applicable to enhancing recent high-throughput SAXS characterizations of pathways involving flexible components (34).

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENT

The authors thank the Berkeley Lab Advanced Light Source and SIBYLS staff at beamline 12.3.1 for providing X-ray facilities.

## FUNDING

Grants from the National Cancer Institute of Canada (NCIC, to J.N.M.G.), the Canadian Institutes of Health Research (CIHR, to M.W.) and the National Institutes of Health (NIH) Structural Cell Biology of DNA Repair Machines P01 grant CA92584 (to J.N.M.G./J.A.T.) and in part by the Office of Science, Office of Biological and Environmental Research, U.S. Department of Energy, under Contract DE-AC02-05CH11231 for SIBLYS beamline efforts. J.N.M.G. acknowledges the support of the Howard Hughes International Scholar program. M.P. was supported by the NSF under grant ECS-0547013 and the USAF, under grant FA9550-06-1-0096. The computational part of this work was supported by the NERSC start up project (m870). Funding for open access charge: National Institutes of Health grant CA92584.

*Conflict of interest statement*. None declared.

## REFERENCES

1. Hazra,T.K., Izumi,T., Boldogh,I., Imhoff,B., Kow,Y.W., Jaruga,P., Dizdaroglu,M. and Mitra,S. (2002) Identification and characterization of a human DNA glycosylase for repair of modified bases in oxidatively damaged DNA. *Proc. Natl Acad. Sci. USA*, **99**, 3523–3528.
2. Evans,C.J. and Aguilera,R.J. (2003) DNase II: genes, enzymes and function. *Gene*, **322**, 1–15.
3. Pouliot,J.J., Yao,K.C., Robertson,C.A. and Nash,H.A. (1999) Yeast gene for a Tyr-DNA phosphodiesterase that repairs topoisomerase I complexes. *Science*, **286**, 552–555.
4. Inamdar,K.V., Pouliot,J.J., Zhou,T., Lees-Miller,S.P., Rasouli-Nia,A. and Povirk,L.F. (2002) Conversion of phosphoglycolate to phosphate termini on 3′ overhangs of DNA double strand breaks by the human tyrosyl-DNA phosphodiesterase hTdp1. *J. Biol. Chem.*, **277**, 27162–27168.
5. Karimi-Busheri,F., Lee,J., Tomkinson,A.E. and Weinfeld,M. (1998) Repair of DNA strand gaps and nicks containing 3′-phosphate and 5′-hydroxyl termini by purified mammalian enzymes. *Nucleic Acids Res.*, **26**, 4395–4400.
6. Whitehouse,C.J., Taylor,R.M., Thistlethwaite,A., Zhang,H., Karimi-Busheri,F., Lasko,D.D., Weinfeld,M. and Caldecott,K.W. (2001) XRCC1 stimulates human polynucleotide kinase activity at damaged DNA termini and accelerates DNA single-strand break repair. *Cell*, **104**, 107–117.
7. Wiederhold,L., Leppard,J.B., Kedar,P., Karimi-Busheri,F., Rasouli-Nia,A., Weinfeld,M., Tomkinson,A.E., Izumi,T., Prasad,R., Wilson,S.H. *et al.* (2004) AP endonuclease-independent DNA base excision repair in human cells. *Mol. Cell*, **15**, 209–220.
8. Loizou,J.I., El-Khamisy,S.F., Zlatanou,A., Moore,D.J., Chan,D.W., Qin,J., Sarno,S., Meggio,F., Pinna,L.A. and Caldecott,K.W. (2004) The protein kinase CK2 facilitates repair of chromosomal DNA single-strand breaks. *Cell*, **117**, 17–28.
9. Meijer,M., Karimi-Busheri,F., Huang,T.Y., Weinfeld,M. and Young,D. (2002) Pnk1, a DNA kinase/phosphatase required for normal response to DNA damage by gamma-radiation or camptothecin in Schizosaccharomyces pombe. *J. Biol. Chem.*, **277**, 4050–4055.
10. Rasouli-Nia,A., Karimi-Busheri,F. and Weinfeld,M. (2004) Stable down-regulation of human polynucleotide kinase enhances spontaneous mutation frequency and sensitizes cells to genotoxic agents. *Proc. Natl Acad. Sci. USA*, **101**, 6905–6910.
11. Hegde,M.L., Hazra,T.K. and Mitra,S. (2008) Early steps in the DNA base excision/single-strand interruption repair pathway in mammalian cells. *Cell Res.*, **18**, 27–47.
12. Lieber,M.R. (2008) The mechanism of human nonhomologous DNA end joining. *J. Biol. Chem.*, **283**, 1–5.
13. Bernstein,N.K., Williams,R.S., Rakovszky,M.L., Cui,D., Green,R., Karimi-Busheri,F., Mani,R.S., Galicia,S., Koch,C.A., Cass,C.E. *et al.* (2005) The molecular architecture of the mammalian DNA repair enzyme, polynucleotide kinase. *Mol. Cell*, **17**, 657–670.
14. Koch,C.A., Agyei,R., Galicia,S., Metalnikov,P., O'Donnell,P., Starostine,A., Weinfeld,M. and Durocher,D. (2004) Xrcc4 physically links DNA end processing by polynucleotide kinase to DNA ligation by DNA ligase IV. *EMBO J.*, **23**, 3874–3885.
15. Dobson,C.J. and Allinson,S.L. (2006) The phosphatase activity of mammalian polynucleotide kinase takes precedence over its kinase activity in repair of single strand breaks. *Nucleic Acids Res.*, **34**, 2230–2237.
16. Karimi-Busheri,F. and Weinfeld,M. (1997) Purification and substrate specificity of polydeoxyribonucleotide kinases isolated from calf thymus and rat liver. *J. Cell Biochem.*, **64**, 258–272.
17. Das,A., Wiederhold,L., Leppard,J.B., Kedar,P., Prasad,R., Wang,H., Boldogh,I., Karimi-Busheri,F., Weinfeld,M., Tomkinson,A.E. *et al.* (2006) NEIL2-initiated, APE-independent repair of oxidized bases in DNA: evidence for a repair complex in human cells. *DNA Repair (Amst)*, **5**, 1439–1448.
18. Ma,Y., Pannicke,U., Schwarz,K. and Lieber,M.R. (2002) Hairpin opening and overhang processing by an Artemis/DNA-dependent protein kinase complex in nonhomologous end joining and V(D)J recombination. *Cell*, **108**, 781–794.
19. Putnam,C.D., Hammel,M., Hura,G.L. and Tainer,J.A. (2007) X-ray solution scattering (SAXS) combined with crystallography and computation: defining accurate macromolecular structures, conformations and assemblies in solution. *Q. Rev. Biophys.*, **40**, 191–285.
20. Guinier,A. and Fournet,F. (1955) *Small Angle Scattering of X-rays*. Wiley, New York.
21. Mylonas,E. and Svergun,D.I. (2007) Accuracy of molecular mass determination of proteins in solution by small-angle X-ray scattering. *J. Appl. Cryst.*, **40**, 245–249.
22. Svergun,D. (1992) Determination of the regularization parameter in indirect-transform methods using perceptual criteria. *J. Appl. Crystal.*, **28**, 768–773.
23. Svergun,D.I. (1999) Restoring low resolution structure of biological macromolecules from solution scattering using simulated annealing. *Biophys. J.*, **76**, 2879–2886.
24. Volkov,V.V. and Svergun,D.I. (2003) Uniqueness of ab initio shape determination in small-angle scattering. *J. Appl. Crystal.*, **36**, 860–864.
25. Wriggers,W., Milligan,R.A. and McCammon,J.A. (1999) Situs: a package for docking crystal structures into low-resolution maps from electron microscopy. *J. Struct. Biol.*, **125**, 185–195.
26. Brooks,B.R., Bruccoleri,R.E., Olafson,B.D., States,D.J., Swaminathan,S. and Karplus,M. (1983) A program for macromolecular energy, minimization, and dynamics calculations. *J. Comp. Chem.*, **4**, 187–217.
27. Svergun,D., Barberato,C. and Koch,M.H.J. (1995) CRYSOL—a program to evaluate x-ray solution scattering of biological macromolecules from atomic coordinates. *J. Appl. Crystallogr.*, **28**, 768–773.
28. Pelikan,M., Hura,G.L. and Hammel,M. (2009) Structure and flexibility within proteins as identified through small angle X-ray scattering. *Gen. Physiol. Biophys.*, **28**, 174–189.
29. MacKerell,J.A.D., Bashford,D., Bellott,M., Dunbrack,R.L.J., Evenseck,J.D., Field,M.J., Ficher,S., Gao,J., Guo,H., Ha,S. *et al.* (1998) All-atom empirical potential for molecular modeling an dynamics studies of proteins. *J. Phys. Chem. B*, **102**, 3586–3616.
30. Hammel,M., Fierobe,H.P., Czjzek,M., Kurkal,V., Smith,J.C., Bayer,E.A., Finet,S. and Receveur-Brechot,V. (2005) Structural

basis of cellulosome efficiency explored by small angle X-ray scattering. *J. Biol. Chem.*, **280**, 38562–38568.
31. Bernado,P., Mylonas,E., Petoukhov,M.V., Blackledge,M. and Svergun,D.I. (2007) Structural characterization of flexible proteins using small-angle X-ray scattering. *J. Am. Chem. Soc.*, **129**, 5656–5664.
32. Rider,M.H., Bertrand,L., Vertommen,D., Michels,P.A., Rousseau,G.G. and Hue,L. (2004) 6-phosphofructo-2-kinase/fructose-2,6-bisphosphatase: head-to-head with a bifunctional enzyme that controls glycolysis. *Biochem. J.*, **381**, 561–579.
33. Segura-Pena,D., Sekulic,N., Ort,S., Konrad,M. and Lavie,A. (2004) Substrate-induced conformational changes in human UMP/CMP kinase. *J. Biol. Chem.*, **279**, 33882–33889.
34. Hura,G.L., Menon,A.L., Hammel,M., Rambo,R.P., Poole,F.L., Tsutakawa,S.E., Jenney,F.E., Frankel,K.A., Hopkins,R.C., Scott,J. *et al.* (2009) Rapid and robust proteomics-scale solution structural analyses determined efficiently by x-ray scattering (SAXS). *Nat. Methods*, in press.
35. Tsutakawa,S.E., Hura,G.L., Frankel,K.A., Cooper,P.K. and Tainer,J.A. (2007) Structural analysis of flexible proteins in solution by small angle X-ray scattering combined with crystallography. *J. Struct. Biol.*, **158**, 214–223.
36. Mani,R.S., Karimi-Busheri,F., Cass,C.E. and Weinfeld,M. (2001) Physical properties of human polynucleotide kinase: hydrodynamic and spectroscopic studies. *Biochemistry*, **40**, 12967–12973.
37. Eastberg,J.H., Pelletier,J. and Stoddard,B.L. (2004) Recognition of DNA substrates by T4 bacteriophage polynucleotide kinase. *Nucleic Acids Res.*, **32**, 653–660.
38. Ali,A.A., Jukes,R.M., Pearl,L.H. and Oliver,A.W. (2009) Specific recognition of a multiply phosphorylated motif in the DNA repair scaffold XRCC1 by the FHA domain of human PNK. *Nucleic Acids Res.*, **37**, 1701–1712.
39. Parsons,J.L., Dianova,II, Boswell,E., Weinfeld,M. and Dianov,G.L. (2005) End-damage-specific proteins facilitate recruitment or stability of X-ray cross-complementing protein 1 at the sites of DNA single-strand break repair. *FEBS J.*, **272**, 5753–5763.
40. Mani,R.S., Fanta,M., Karimi-Busheri,F., Silver,E., Virgen,C.A., Caldecott,K.W., Cass,C.E. and Weinfeld,M. (2007) XRCC1 stimulates polynucleotide kinase by enhancing its damage discrimination and displacement from DNA repair intermediates. *J. Biol. Chem.*, **282**, 28004–28013.
41. Yannone,S.M., Khan,I.S., Zhou,R.Z., Zhou,T., Valerie,K. and Povirk,L.F. (2008) Coordinate 5′ and 3′ endonucleolytic trimming of terminally blocked blunt DNA double-strand break ends by Artemis nuclease and DNA-dependent protein kinase. *Nucleic Acids Res.*, **36**, 3354–3365.
42. Garcin,E.D., Hosfield,D.J., Desai,S.A., Haas,B.J., Bjoras,M., Cunningham,R.P. and Tainer,J.A. (2008) DNA apurinic-apyrimidinic site binding and excision by endonuclease IV. *Nat. Struct. Mol. Biol.*, **15**, 515–522.
43. Mol,C.D., Izumi,T., Mitra,S. and Tainer,J.A. (2000) DNA-bound structures and mutants reveal abasic DNA binding by APE1 and DNA repair coordination [corrected]. *Nature*, **403**, 451–456.
44. Chapados,B.R., Hosfield,D.J., Han,S., Qiu,J., Yelent,B., Shen,B. and Tainer,J.A. (2004) Structural basis for FEN-1 substrate specificity and PCNA-mediated activation in DNA replication and repair. *Cell*, **116**, 39–50.
45. Williams,R.S., Moncalian,G., Williams,J.S., Yamada,Y., Limbo,O., Shin,D.S., Groocock,L.M., Cahill,D., Hitomi,C., Guenther,G. *et al.* (2008) Mre11 dimers coordinate DNA end bridging and nuclease processing in double-strand-break repair. *Cell*, **135**, 97–109.