

Cite this: *Phys. Chem. Chem. Phys.*, 2011, **13**, 11239–11247

www.rsc.org/pccp

PAPER

Correlation between electron localization and metal ion mutagenicity in DNA synthesis from QM/MM calculations†

Robin Chaudret,^a Jean-Philip Piquemal^{ab} and G. Andrés Cisneros^{*c}

Received 16th November 2010, Accepted 14th April 2011

DOI: 10.1039/c0cp02550j

DNA polymerases require two divalent metal ions in the active site for catalysis. Mg^{2+} has been confirmed to be the most probable cation utilized by most polymerases *in vivo*. Other metal ions are either potent mutagens or inhibitors. We used structural and topological analyses based on *ab initio* QM/MM calculations to study human DNA polymerase λ (Pol λ) with different metals in the active site. Our results indicate a slightly longer O3'–P α distance (~ 3.6 Å) for most inhibitor cations compared to the natural and mutagenic metals (~ 3.3 – 3.4 Å). Optimization with a larger basis set for the previously reported transition state (TS) structures (Cisneros *et al.*, *DNA Repair*, 2008, **7**, 1824.) gives barriers of 17.4 kcal mol⁻¹ and 15.1 kcal mol⁻¹ for the Mg^{2+} and Mn^{2+} catalyzed reactions respectively. Relying on the key relation between the topological signature of a metal cation and its selectivity within biological systems (de Courcy *et al.*, *J. Chem. Theor. Comput.*, 2010, **6**, 1048.) we have performed electron localization function (ELF) topological analyses. These analyses show that all inhibitor and mutagenic metals considered, except Na^+ , present a “split” of the outer-shell density of the metal. This “splitting” is not observed for the non-mutagenic Mg^{2+} metal. Population and multipole analyses on the ELF basins reveal that the electronic dipolar and quadrupolar polarization is significantly different with Mg^{2+} compared to all other cations. Our results shed light at the atomic level on the subtle differences between Mg^{2+} , mutagenic, and inhibitor metals in DNA polymerases. These results provide a correlation between the electronic distribution of the cations in the active site and the possible consequences on DNA synthesis.

1. Introduction

Replication and repair of DNA by DNA polymerases are critical processes. Errors in either of these transactions can result in mutations, some of which can lead to disease or even death.^{1,2} DNA polymerases catalyze the addition of an incoming nucleotide on the nascent DNA chain by nucleophilic attack of the primer terminus O3' on the P α of the incoming nucleotide with resulting release of pyrophosphate (PPi). This mechanism has been proposed to be facilitated by two metal ions in the active site.^{3,4} Studies with several polymerases have confirmed that the most probable cation used *in vivo* by DNA polymerases is Mg^{2+} .⁵ Experimental and theoretical studies of this reaction have supported the two metal hypothesis (see ESI† for reaction scheme).^{6–19} These two metals have

been denoted as metal 1 and metal 2. Metal 1 binds to the primer base (Mg1 in Fig. 1A). Metal 2 binds to the incoming nucleotide (Mg2 in Fig. 1A).

Some metal cations, including chromium, nickel, cadmium and cobalt, have been shown to be carcinogenic and/or genotoxic. These effects arise due to the inhibition of the DNA repair process.²⁰ In particular, this inhibition affects several DNA repair pathways including base excision repair (BER), nucleotide excision repair (NER) and mismatch repair (MMR).^{20,21} Another possibility for carcinogenic action is a loss of fidelity of replication of DNA induced by carcinogenic metals.²¹ Replacement of the two Mg^{2+} ions in the active site of DNA polymerases by certain other metals inhibits and/or causes an increase in errors during DNA replication (loss in DNA replication fidelity). Indeed, several carcinogenic or mutagenic metals decrease fidelity, while other non-mutagenic metals inhibit DNA synthesis.²²

This inhibitory or mutagenic effect has been studied on the basis of crystallographic analyses of ternary structures of DNA polymerases. Pelletier *et al.* solved structures of DNA polymerase β (Pol β) with different cations in the active site.²³ Their results show that the cation in the metal 2 site coordinates differently to the phosphate of the incoming

^aUPMC Univ Paris 06, UMR 7616 Laboratoire de Chimie Théorique, case courrier 137, 4 place Jussieu, F-75005, Paris, France

^bCNRS UMR 7616, Laboratoire de Chimie Théorique, case courrier 137, 4 place Jussieu, F-75005, Paris, France

^cDepartment of Chemistry, Wayne State University, Detroit, MI 48202, USA. E-mail: andres@chem.wayne.edu

† Electronic supplementary information (ESI) available: Optimized distances for the calculated structures, ELF attractor positions, and ELF analysis for individual cations. See DOI: 10.1039/c0cp02550j

to +2 for all others) results in less repulsive Coulomb interactions and the enzyme can provide an environment that helps maintain both Na^+ at a distance almost equal to that found in the Mg^{2+} structure.

Crystal structures for Pol λ with Na^+ in the active site have been previously reported.²⁹ These structures, along with kinetic results,⁸ suggest that Na^+ inhibits catalysis by forming a catalytically inactive structure.²⁹ In the experimental structure, the reported nucleophilic attack distance is 4.66 Å. This is due to a C3'-endo conformation of the ribose on the primer-terminal nucleotide, compared to a C2'-endo conformation in the catalytically active structure. However, it is not possible to differentiate between Na^+ and Mg^{2+} crystallographically since they are isoelectronic. For this particular crystal structure, Na^+ was assigned as the ion in the metal 1 site and Mg^{2+} in the metal 2 site based on distances for the 1st coordination shell of the metals and lack of octahedral coordination for the cation in the metal 1 site.²⁹

We performed additional optimizations of a system with Na^+ and Mg^{2+} in the metal 1 and metal 2 sites respectively. When the primer-terminal ribose is in the C2'-endo conformation the nucleophilic attack distance is increased only to 3.6 Å. This is the same as the distance obtained with two Na^+ in the active site. When the conformation of the ribose is changed to C3'-endo the distance increases to 4.6 Å. However, in our hands, this structure reverts to the catalytically active C2'-endo conformation after optimization.

The reversion to C2'-endo could be due to a possible bias imposed by the MM environment. Another possibility is that the C2'-endo conformation is more stable since this results in an octahedral conformation for the cation in the metal 1 site (Na^+ in this case) as opposed to the C3'-endo orientation. The latter structure results in a distorted tetrahedral geometry.²⁹ The higher coordination may thus lead to a more stable structure in solution (as in our simulations), whereas the C3'-endo structure could be thermodynamically lower or equivalent to the C2'-endo conformation in the solid phase.

As mentioned in the introduction, the reaction mechanism of Pol λ has been studied previously for Mg^{2+} and Mn^{2+} .²⁷ The calculated reaction path comprises two stable transition states (TSs). The first TS (TS1) corresponds to the de-protonation of O3' on the primer base. The second TS corresponds to the nucleophilic attack of O3' on the αP of the incoming nucleotide (see ESI†). The resulting energy difference between the barriers was not large enough to determine a cation preference. This was mainly due to the small basis sets used in the previous studies.²⁷

We have re-optimized the transition state (TS) structures with the 6-31G(d) basis for all atoms. The re-optimized structures give energy differences (relative to the reactant) of 17.4 and 12.5 kcal mol⁻¹ for the two TSs in the Mg^{2+} reaction path. For the Mn^{2+} structures the energy differences are 15.1 and 11.7 kcal mol⁻¹. Frequency calculations show that only the structures corresponding to the first TS have one imaginary frequency.

For the TS2 structures, all calculated frequencies are real, that is, these are not true TSs. This is consistent with the observed increase in the nucleophilic attack distance with respect to the original TS structures. The calculated distances

are 2.7 Å and 3.0 Å for Mg^{2+} and Mn^{2+} respectively. This can be contrasted with the previously reported distances of 2.3 Å and 2.4 Å, obtained with the smaller basis set.²⁷

The experimental barrier is estimated to be around 16.6 kcal mol⁻¹.⁸ Moreover, experimentally it is known that the estimated barrier for the Mn^{2+} catalyzed reaction in Pol λ is around 1 kcal mol⁻¹ lower than that for Mg^{2+} , compared to 2.3 kcal mol⁻¹ for our calculations.²⁶ Our results for the re-optimized TSs point to the first TS (proton abstraction from O3' on the primer base) as rate limiting. This is similar to the exonuclease reaction catalyzed by the ϵ subunit of DNA polymerase III, where once the nucleophile is formed the reaction proceeds downhill.³⁰

The potential energy barriers obtained with the larger basis are in good agreement with the experimental reports both for the barrier height and cation preference. However, there are no obvious structural differences between the TS structures for the two cations (see ESI†). That is, the difference in energies for the catalytic step does not appear to arise because of structural changes between the two different cations.

The structural and energetic results from the optimized structures do not provide a clear explanation for the difference in inhibitory or mutagenic activity between the cations compared to the “natural” metal. Based on this, we now turn to ELF topological analyses for a deeper insight and to investigate differences at the electronic level.

2.2 ELF analyses

To gain a deeper understanding of the differences between the Pol λ structures with different cations in the active site we performed ELF calculations based on the wavefunctions obtained from our QM/MM optimizations. Topological analyses of the ELF have been used extensively to analyze chemical bonding and reactivity.^{31,32} In addition, ELF topological analyses have been applied to model systems of biological significance.³³ Here, we went a step further and applied ELF to QM/MM wavefunctions as done recently to study low-barrier hydrogen bonds.³⁴ In this case, we have limited the analyses to a subset of the QM atoms as detailed in the Methods section.

The ELF topological analyses on the studied systems show large differences between the cations. As shown in Fig. 2 the structure with Mg^{2+} shows a single basin around each cation. That is, there is a single basin which corresponds to the core density of the cation. In contrast, all other cations show several attractor positions, which suggest a spatial delocalization around them (see Fig. 2 and ESI†). This has been shown in an ELF study of model systems for proteins from the blood coagulation cascade, where the concept of “subvalence” was introduced³⁵ linking the topological signature of a metal cation to its selectivity within biological systems. de Courcy *et al.* showed that the outer-shell core basins of the cations can be delocalized.³⁵ Indeed, this delocalization was suggested to be correlated to their chemical hardness, η_{A} .^{36,37}

A correlation is also observed between the hardness of the cation, η_{A} , and the subvalence splitting in Pol λ . The hardest cation, Mg^{2+} , has a hardness of 32.55 and shows no splitting of the subvalence. The inhibitor cations Na^+ and Ca^{2+} have

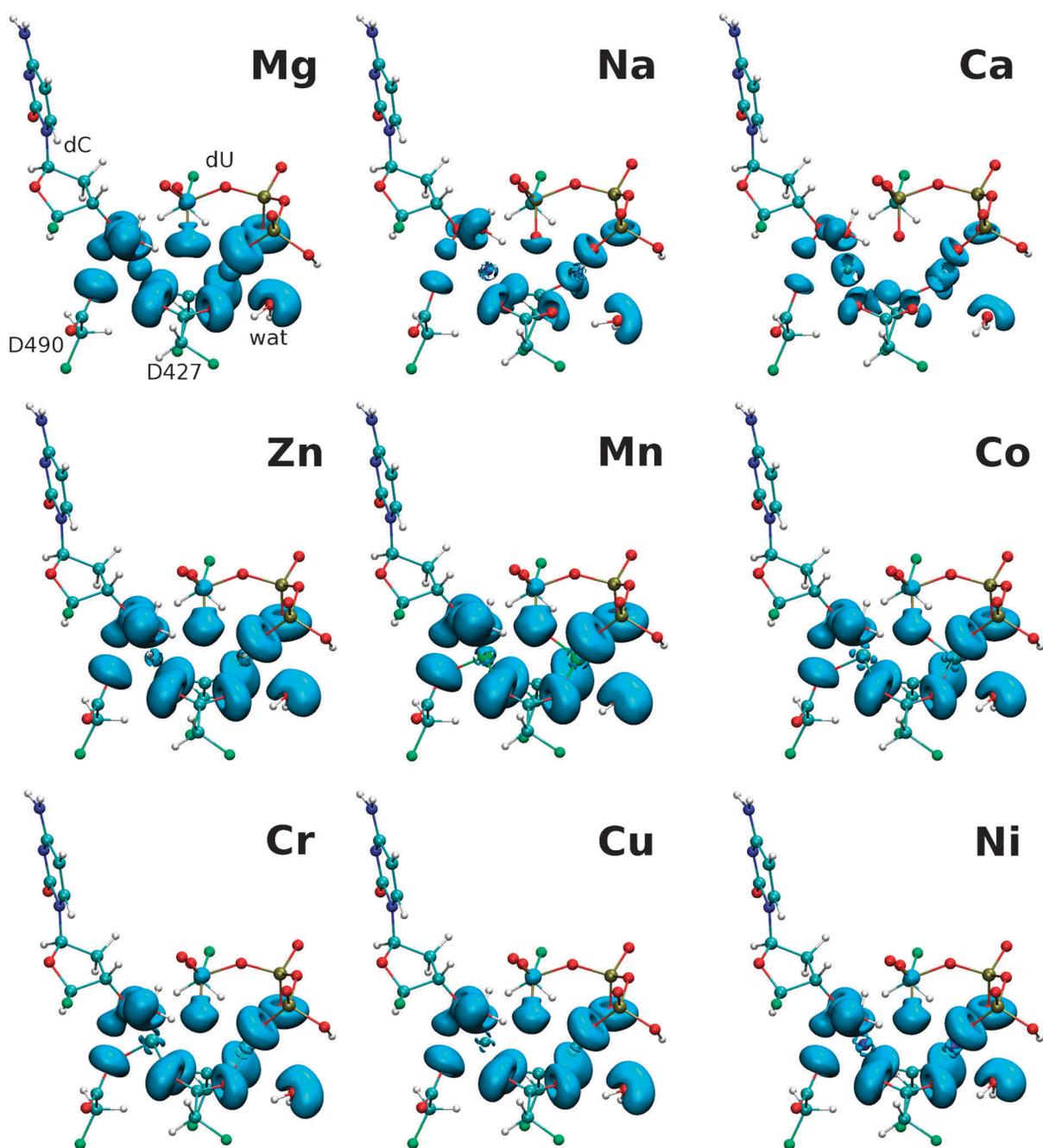


Fig. 2 ELF topological analyses of the Pol λ active site with different cations. Figures show the atoms in the QM region, only part of the incoming dU is included in the QM (see Methods section). The isovalue for the ELF isosurfaces is 0.8 for structures with Mg $^{2+}$ and mutagenic metals except for Mn $^{2+}$ (0.73). For the inhibitor cations the isovalues are 0.865, 0.88 and 0.77 for the Na $^{+}$, Ca $^{2+}$ and Zn $^{2+}$ structures respectively (see ESI \dagger for images of Na $^{+}$ and Ca $^{2+}$ at a smaller isovalue).

η_A values of 21.1 and 19.52 respectively. In this case, no well defined splitting was observed on the Na $^{+}$ cation. Conversely, the splitting of the subvalence is beginning to become apparent on the Ca $^{2+}$. The last inhibitor, Zn $^{2+}$, has a much lower η_A of 10.88 and exhibits considerable splitting. Finally all mutagenic cations have hardness values between 9.2 and 7.23 and exhibit a strong delocalization of the outer-shell core basins.

Differences in the valence regions were also observed. One such difference is for the lone pairs (LP) of O2A (Fig. 2). This basin is different for the structure with the “natural” metal

compared to all other cations (see also ESI \dagger). Moreover three of the five considered mutagens show three basins for the valence region of O2A corresponding to the lone pairs.

We have carried out population and distributed multipole analyses on these basins to understand the physics that arise from these distributions. Table 1 shows the populations and multipoles for the valence of the oxygens from the triphosphate that bind to the metals (see Fig. 1 for labels). There are significant differences in the distributions of the inhibitor and mutagenic metals compared to Mg $^{2+}$.

Table 1 ELF populations (Pop) and multipole analyses (first (M1) and second (M2) moments (in au)) for the lone pair basins, V(O), of the oxygen atoms from the incoming phosphate that bind to the metals (see Fig. 1). The first number corresponds to the basin labeled “LP1” and the second to “LP2”

		Natural		Inhibitors				Mutagens		
		Mg ²⁺	Na ⁺	Ca ²⁺	Zn ²⁺	Co ²⁺	Cr ²⁺	Cu ²⁺	Mn ²⁺	Ni ²⁺
V(O2A)	Pop	3.61, 2.41	4.95, 0.98	3.76, 2.20	1.98, 3.98	1.89, 4.08	2.19, 2.22, 1.60 ^a	0.89, 1.05, 0.45 ^a	2.03, 2.48, 1.53 ^a	1.82, 4.15
	M1	2.11, 0.83	4.23, 0.71	2.48, 0.80	1.36, 3.24	0.85, 3.14	0.81, 0.87, 0.46 ^a	0.89, 1.05, 0.45 ^a	0.57, 0.80, 0.55 ^a	0.65, 3.81
	M2	1.90, 0.56	3.53, 0.61	2.78, 0.50	1.10, 4.18	0.67, 3.77	0.69, 0.75, 0.34 ^a	0.73, 0.93, 0.32 ^a	0.44, 0.62, 0.52 ^a	0.43, 5.13
V(O2B)	Pop	2.48, 4.09	6.45 ^b	2.67, 3.87	1.71, 4.81	1.69, 4.84	1.88, 4.72	1.46, 5.08	1.75, 4.79	1.72, 4.81
	M1	0.93, 3.14	6.85 ^b	1.25, 3.70	0.58, 4.15	0.64, 4.22	0.67, 3.85	0.54, 4.54	0.75, 4.14	0.63, 4.17
	M2	0.81, 4.16	7.09 ^b	1.04, 5.79	0.66, 5.21	0.65, 5.41	0.73, 4.18	0.59, 5.29	0.59, 4.98	0.65, 5.20
V(O3G)	Pop	2.40, 4.12	2.91, 3.41	2.69, 3.78	1.54, 4.93	1.58, 4.91	1.79, 4.77	1.22, 5.28	1.65, 4.85	1.55, 4.93
	M1	0.97, 2.81	1.43, 2.06	1.18, 2.26	0.65, 4.11	0.61, 4.04	0.69, 3.80	0.53, 4.74	0.65, 3.90	0.61, 4.08
	M2	0.94, 1.79	1.09, 1.80	0.89, 1.29	0.62, 2.76	0.617, 2.62	0.79, 2.40	0.55, 3.73	0.64, 2.36	0.61, 2.71

^a The valence density for the lone pairs of these atoms splits into three basins. ^b The valence density for the lone pairs of this atom is found in a single basin.

The number of basins and calculated populations for the lone pairs presents an interesting picture (see Fig. 1B and Table 1). For the lone pairs on O2A in the mutagenic metals there are two main cases for the number of basins. The first corresponds to structures where there are three basins for the lone pairs of O2A (Cr²⁺, Cu²⁺ and Mn²⁺). The remaining mutagenic cations (Co²⁺ and Ni²⁺) have only two basins corresponding to the two lone pairs, LP1 and LP2 (see Table 1 and Fig. 1B). The populations for the lone pairs on O2A for the mutagenic metals with two basins and for Zn²⁺ are observed to be inverted between LP1 and LP2 compared to Mg²⁺. That is, in the case of Mg²⁺ the populations in LP1 and LP2 are 3.61 and 2.41 respectively. On the other hand, the magnitude of the populations is reversed for Co²⁺, Ni²⁺ and Zn²⁺. In the case of Na⁺, the populations are consistent with Mg²⁺, although there is a much larger population on LP1 compared to LP2.

In the case of O2B and O3G, the populations are observed to be smaller for the corresponding LP1 and LP2 for most of the cations except for Ca²⁺ and Na⁺. An interesting feature in the latter structure is that there is only one basin on O2B (see Table 1), implying that there is no splitting of the lone pairs and they occupy the same center for that oxygen. That is, the lone pairs on O2B are found in the same location instead of on two centers when Na⁺ is present in the active site.

The multipolar analyses reveal striking differences in the populations between the “natural” (Mg²⁺) and other cations. In particular mutagenic metals display much larger first moments (M1) on the ELF basins than the ones calculated for the Mg²⁺ structure. For O2A, the M1 follow the population and the distribution is observed to be inverted, with the larger dipole on LP2 for the mutagenic cations compared to LP1 for Mg²⁺. In the case of the inhibitors, Na⁺ and Ca²⁺ show larger M1 for O2A and O2B, albeit to a lesser extent than the mutagenic cations. Zn²⁺ shows a similar behavior to the mutagenic cations on the lone pair basins. However, the calculated M1 on the basins of the split subvalence are on the order of 1 au for the largest one, which is in agreement with the other two inhibitors. Conversely, the largest M1 for the subvalence basins in the mutagenic metals is around 0.5 au. Moreover, the Zn²⁺ structure shows a distorted geometry for the lone pairs of

O2A, somewhat similar to that found on the Na⁺ structure (see ESI†).

For the second moments (M2) the differences are even more evident. In all cases, the calculated M2 for O2A and O2B are markedly different for the Mg²⁺ structure compared to all other structures. For O3G a significant increase in M2 is observed for all other metals except for Na⁺ compared to the Mg²⁺ cation.

Taken together, the ELF results suggest that the major differences between the cations relate to the electronic distribution. This is consistent with an electrostatic contribution to catalysis.³⁸ Our results show that there are significant differences in the dipolar and quadrupolar polarization of the active site when inhibitor or mutagenic metals occupy the active site, compared to the “natural” cation. Moreover, the Mg²⁺ cation does not exhibit any splitting of the subvalence compared to all inhibitor and mutagenic cations (except Na⁺). In the case of Na⁺, the valence distribution shows marked differences with Mg²⁺, in particular, for the lone pairs of O2A. Additionally, the largest changes in population and multipoles are observed on O2A, O2B and O3G. This points to a marked change in dipolar and quadrupolar polarization effects on the incoming nucleotide depending on the cation that occupies the active site. Thus, the inhibitory or mutagenic effect can be ascribed, at least in part, to the polarization and hyper-polarization of the phosphate group where the reaction will take place and on the leaving pyrophosphate.

The ELF analysis on the re-optimized TS structures also corroborates the polarization hypothesis. As can be seen from Table 2, the populations for O2A, O2B and O3G are significantly different between the Mg²⁺ and Mn²⁺ TSs. In particular, VO2B is observed to have 2 and 3 basins, respectively, in the TS1 and TS2 structures with Mn²⁺, compared to only one in the Mn²⁺ TSs. The populations and moments for both TS1 structures are similar to the ones in the reactant structure, with the exception of O2B in the TS2 with Mn²⁺.

The ELF analysis shows a population of 2 electrons on the H–OD2 basin which shows that the proton has been fully transferred to the aspartate residue in TS1 with Mg²⁺. This basin in the Mn²⁺ TS1 structure shows a population of 0.6 electrons. This is also consistent with the structural

Table 2 ELF populations (Pop) and multipole analyses (first (M1) and second (M2) moments (in au) for the lone pair basins, V(O), of the oxygen atoms from the incoming phosphate that bind to the metals, O3', OD2 on D490 (see Fig. 1)

		Mg^{2+}		Mn^{2+}	
		TS1	TS2	TS1	TS2
V(O2A)	Pop	3.42, 2.61	3.06, 2.99	2.04, 2.33, 1.67	1.58, 3.08, 1.38
	M1	1.67, 0.93	1.49, 1.25	0.65, 0.67, 0.49	0.54, 1.21, 0.66
	M2	1.67, 0.69	1.42, 0.95	0.52, 0.56, 0.36	0.35, 0.90, 0.33
V(O2B)	Pop	2.49, 4.10	2.42, 1.72, 1.51	6.58	4.82
	M1	0.94, 3.15	0.93, 1.11, 0.90	7.06	4.01
	M2	0.82, 4.11	0.81, 0.85, 0.74	8.15	4.16
V(O3G)	Pop	2.45, 4.08	2.32, 4.20	1.61, 4.89	1.51, 5.01
	M1	1.02, 2.57	0.91, 2.97	0.70, 3.96	0.74, 4.21
	M2	0.96, 1.71	0.92, 2.01	0.64, 2.38	0.62, 3.03
V(O3')	Pop	3.2, 3.2, 1.48	4.88, 1.97	3.67, 1.50, 1.70	6.90
	M1	1.34, 0.53, 0.74	3.74, 0.70	1.9, 0.88, 1.08	7.65
	M2	0.86, 0.53, 0.59	3.08, 0.57	0.89, 0.72, 0.60	8.87
V(OD2)	Pop	2.09 ^a , 5.10	1.91 ^a , 5.23	0.60 ^a , 4.97, 1.62	1.91 ^a , 5.17
	M1	1.98, 4.66	1.48, 5.05	0.14, 4.36, 0.83	1.53, 4.89
	M2	2.33, 5.36	1.25, 6.17	0.3, 4.86, 0.71	1.41, 6.00

^a This basin corresponds to the H–OD2 bond resulting from the proton abstraction from O3' by OD2 on D490 (see text).

features where the O3'–H and H–OD2 distances are 1.61 Å and 1.04 Å for Mg^{2+} and 1.48 Å and 1.08 Å for Mn^{2+} .

The multipolar analysis also reveals a markedly different dipolar and quadrupolar polarization for all the atoms between the Mg^{2+} and Mn^{2+} structures. In particular, O3' in the TS2 structure with Mn^{2+} shows a single basin with very large M1 and M2 values. Conversely, the TS2 structure with Mg^{2+} shows two basins with significantly smaller M1 and M2. These results are consistent with, and provide a plausible explanation for the experimental observation of increased mutagenicity of Mn^{2+} in the highly homologous Pol β as discussed in ref. 23. These calculations do not take into account the cation binding affinity, which are higher for the transition metals than for Mg^{2+} .

To confirm the differences in polarization we have carried out energy decomposition analysis using the Restricted Variational Space (RVS) method on active site model systems based on the Mg^{2+} and Zn^{2+} structures.^{39,40} These models consist of nine monomers that represent the two metal ions and small fragments for the first coordination shell. The fragments include formate ions for the aspartates, methyl triphosphate for the incoming nucleotide, 1 methyl ribose for the primer terminus and two water molecules. The calculations were done at the RHF/6-31G(d) level.

The total RVS intermolecular interaction (BSSE corrected) for the Mg^{2+} model is $-1562.5 \text{ kcal mol}^{-1}$ and for the Zn^{2+} model is $-1584.3 \text{ kcal mol}^{-1}$. For the Mg^{2+} structure, 86% of the total interaction energy corresponds to the first order components, Coulomb and Exchange repulsion. In contrast, the first order terms contribute only 83% of the total intermolecular interaction in the Zn^{2+} structure. Moreover, the calculated polarization energies for the Mg^{2+} cations are less than $-0.1 \text{ kcal mol}^{-1}$ for both cations. On the other hand the polarization energies for the Zn^{2+} cations are -0.8 and $-1.5 \text{ kcal mol}^{-1}$. These model RVS calculations show that the strongest cation is very weakly polarizable, compared to a softer cation that is more polarizable. This leads to the splitting of the subvalence density similar to that observed in blood coagulation proteins.³⁵

ELF calculations have also been performed on three active site model systems. These correspond to the two above (Mg^{2+} and Zn^{2+}) as well as Mn^{2+} . In all three cases, the subvalence splitting is similar to the one observed in the enzyme (see ESI†). This is consistent with the initial proposal of subvalence splitting, where this phenomenon is observed in formate–cation complexes³⁵ in addition to models of enzyme active sites. Therefore, the splitting of the subvalence is a property of the cation.

The population and moments of the basins in the clusters show a marked difference with respect to the full system, in particular for Zn^{2+} and Mn^{2+} . This reveals the importance of the inclusion of the electrostatic environment from the MM subsystem in the QM/MM calculation. Indeed, the choice of atomic charges plays an important part and ideally it would be preferable to employ a FF that provides an accurate description of the charge density such as the Gaussian Electrostatic Model.⁴¹ This FF has been shown to provide correct polarization of QM/MM wavefunctions in non-bonded model systems.⁴²

These results show that the subvalence splitting in the cations induces a different dipolar and quadrupolar polarization on the ligands. The dipolar and quadrupolar polarization depends on the extent of the splitting. This results in a different environment for the ligands in the active site. The difference in electronic environment affects the catalytic ability of the polymerase and may explain (at least in part) the inhibitory or mutagenic effects of the different cations.

3. Computational methods

3.1 Structure optimizations

QM/MM optimizations were carried out based on our previously optimized reactant structure, see ref. 27 for details. Briefly, the initial structure of Pol λ was taken from the X-ray structure for the pre-catalytic state (pdbid 2PFO).²⁹ This structure was subjected to minimization and MD relaxation using the PMEMD module of AMBER9.⁴³ After equilibration, a single snapshot was selected and the QM

subsystem was chosen to include both active site metals, the side chains of residues D427, D429 and D490, the primer dC nucleotide (excluding C5' and the phosphate group), a part of the incoming nucleotide (triphosphate and C5') and two water molecules that complete the coordination spheres of the metals. This results in a total of 72 QM atoms including 5 boundary atoms. The remaining atoms were included in the MM subsystem and treated with the parm99 force field.⁴³ The charge of the triphosphate on the incoming dUTP was maintained as in our previous study.²⁷ Following the pseudobond methodology, the charges of the atoms near the boundary have been set to zero. The previously optimized structures were shown to be a good representative of Pol λ since the RMSD for all heavy atoms in the active site for both the reactant and the product is below 0.3 Å (0.2 Å for the product).²⁷

All calculations involving transition metal (TM) cations in the present work (except for Zn²⁺) were performed assuming a high spin state with all electrons on both cations unpaired, see ref. 27. Optimizations for the current work have been performed by replacing the two metal ions with the following cations: Na⁺, Ca²⁺, Co²⁺, Cr²⁺, Cu²⁺, Mg²⁺, Mn²⁺, Ni²⁺ and Zn²⁺. All structures, including the previously calculated Mg²⁺ and Mn²⁺, have been re-optimized at the B3LYP level^{44,45} with the 6-31G(d) basis on all atoms. This level of theory has been previously shown to be in good agreement with MP2 and other functionals (e.g. MO6) for ELF calculations of similar bio-systems.³⁵ All optimizations have been carried out with modified versions of Gaussian 09 and TINKER.^{46,47} All systems were optimized with an iterative method described in ref. 48 and 49 using the pseudobond method for the boundary atoms.^{50,51} The TS structures were re-optimized using the quadratric synchronous transit (QST3) method^{52,53} as in our previous study.²⁷ QM/MM simulations have been performed on a large number of enzymatic systems.^{54,55}

3.2 ELF calculations

The electron localization function, $\eta(r)$, may be interpreted as a measure of electron localization in molecular systems.^{56,57} The topological analyses of the ELF are based on the theory of gradient dynamics as applied to a scalar potential function, $\eta(r)$. In this case ELF represents a continuous and differentiable scalar field in 3D space. The topological analyses of the ELF are achieved by the partition of the molecular space *via* the gradient vector field $\nabla\eta(r)$. This partition results in a set of non-overlapping molecular volumes denoted as basins, which are localized around the maxima (attractors) of $\eta(r)$. The boundaries between basins, separatrices, are defined as surfaces which obey the zero-flux condition, that is, the normal vector of the surface at each point is 0. The quantum theory of atoms in molecules (QTAIM) employs similar methodology, except that the potential function corresponds to the molecular electronic density, $\rho(r)$.^{58,59}

The ELF basins closely match the domains for the valence shell electron pair repulsion (VSEPR) model.⁶⁰ The partition of the ELF as obtained from a computation on a 3D grid results in separation of core and valence regions for the atoms,

C() and V() respectively. The valence regions can be associated with lone pairs, V(X), or chemical bonds, V(X,Y), where X and Y represent atoms in the molecule. In addition to the partitioning of the molecular space, the ELF topological analyses provide the possibility of integration of the electronic population or distributed electrostatic moments within the basins.⁶¹

All ELF grids were set to sizes of 200³ au with a step size of 0.1 au and generated using the molecular orbitals obtained from the QM/MM calculations. That is, the molecular orbitals for the ELF calculations include the polarization from the MM environment explicitly. Since the QM subsystem is relatively large, only 15 atoms were considered for the grid generation. These include the two cations, all oxygen atoms occupying the first coordination shell of the metals, the primer terminus O3' and H and incoming nucleotide P α . The ELF calculations were performed with a modified version of the TopMod package^{61,62}.

4. Conclusions

QM/MM optimizations and ELF topological analyses have been performed on Pol λ with nine different cations, Mg²⁺, Na⁺, Ca²⁺, Zn²⁺, Co²⁺, Cr²⁺, Cu²⁺, Mn²⁺ and Ni²⁺. Distances between the metals and the atoms in the first coordination shell are seen to increase for some ligands in a cation dependent fashion. However, structural results show no significant differences that would indicate an effect on catalytic ability between the Mg²⁺ structure and any of the inhibitor or mutagenic cations. On the other hand, ELF results paint a very different picture. The “natural” cation, Mg²⁺, is the only cation observed not to split its subvalence density. Conversely, all other cations show distributed subvalence populations. This subvalence splitting may be correlated to the hardness, η_A , of the cations, with softer cations giving much stronger splitting of the subvalence domains. Population and multipolar analyses reveal that the dipolar and quadrupolar polarization of the active site for the Mg²⁺ structure, the presumed wild-type cation, is unique. All other cations present different populations and multipolar distributions compared to Mg²⁺. The most marked difference is observed on the oxygen atoms from the tri-phosphate group on the incoming nucleotide. Previously reported TS structures for Mg²⁺ and Mn²⁺ were re-optimized with a larger basis set. The resulting energy barriers are in good agreement with experiment for barrier height and cation preference. ELF analysis on the TS structures shows markedly different dipolar and quadrupolar polarization between the natural and the mutagenic Mn²⁺ cation, consistent with the reactant structures. The differences on the population, polarization and hyper-polarization on the tri-phosphate are responsible (at least in part) for the marked differences in activity for the inhibitor cations and diminished fidelity for the mutagenic cations. RVS calculations on two model active site systems show that the Mg²⁺ structure has smaller polarization interactions compared to when Zn²⁺ is present in the active site. This leads to the localization of the subvalence on the softer cations, which in turn leads to the different ELF populations and multipolar distributions. By using the ELF we have discerned the differences between a

polymerase with the natural metal, Mg^{2+} , and inhibitor or mutagenic cations in the active site of a DNA polymerase at the electronic level. These differences help explain the observed experimental results for inhibitor and mutagenic cations compared with Mg^{2+} .

Acknowledgements

This research was supported by Wayne State University. Computing time from Wayne State C&IT is gratefully acknowledged. Support from the French National Research Agency (Grant ANR-08-BLAN-0158) is acknowledged. GAC would like to thank Drs T. A. Kunkel, L. Perera and L. G. Pedersen from NIEHS for insightful discussions.

References

- D. Starcevic, S. Dalal and J. B. Sweasy, *Cell Cycle*, 2004, **3**, 98–1001.
- T. A. Kunkel, *Cancer Cell*, 2003, **3**, 105–110.
- J. Cowan, *BioMetals*, 2002, **15**, 225–235.
- T. A. Steitz, *Curr. Opin. Struct. Biol.*, 2003, **3**, 31–38.
- A. Kornberg and T. A. Baker, *DNA Replication*, W. H. Freeman & Co., San Francisco, CA, 1992.
- M. Fothergill, M. F. Goodman, J. Petruska and A. Warshel, *J. Am. Chem. Soc.*, 1995, **117**, 11619–11627.
- Y. G. Abashkin, J. W. Erickson and S. K. Burt, *J. Phys. Chem. B*, 2001, **105**, 287–292.
- K. A. Fiala, W. Abdel-Gawad and Z. Suo, *Biochemistry*, 2004, **43**, 6751–6762.
- R. Radhakrishnan and T. Schlick, *Biochem. Biophys. Res. Commun.*, 2006, **350**, 521–529.
- A. K. Showalter, B. J. Lamarche, M. Bakhtina, M.-I. Su, K.-H. Tang and M.-D. Tsai, *Chem. Rev.*, 2006, **106**, 340–360.
- T. Rungrotmongkol, A. J. Mulholland and S. Hannongbua, *J. Mol. Graphics Modell.*, 2007, **26**, 1–13.
- M. Bakhtina, M. P. Roettger, S. Kumar and M.-D. Tsai, *Biochemistry*, 2007, **46**, 5463–5472.
- Y. Xiang and A. Warshel, *J. Phys. Chem. B*, 2008, **112**, 1007–1015.
- M. D. Bojin and T. Schlick, *J. Phys. Chem. B*, 2007, **111**, 11244–11252.
- P. Lin, L. C. Pedersen, V. K. Batra, W. A. Beard, S. H. Wilson and L. G. Pedersen, *Proc. Natl. Acad. Sci. U. S. A.*, 2006, **103**, 13294–13299.
- J. Florián, M. F. Goodman and A. Warshel, *J. Am. Chem. Soc.*, 2003, **125**, 8163–8177.
- J. Florián, M. F. Goodman and A. Warshel, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 6819–6824.
- L. Wang, X. Yu, P. Hu, S. Broyde and Y. Zhang, *J. Am. Chem. Soc.*, 2007, **129**, 4731–4737.
- I. L. Alberts, Y. Wang and T. Schlick, *J. Am. Chem. Soc.*, 2007, **129**, 11100–11110.
- A. Hartwig, *Toxicol. Lett.*, 1998, **102–103**, 235–239.
- L. P. Bignold, *Carcinogenesis*, 2004, **25**, 299–307.
- M. A. Sirover and L. A. Loeb, *Science*, 1976, **194**, 1434–1436.
- H. Pelletier, M. R. Sawaya, W. Wolffe, S. H. Wilson and J. Kraut, *Biochemistry*, 1996, **35**, 12762–12777.
- J. Aqvist and A. Warshel, *J. Am. Chem. Soc.*, 1990, **112**, 2860–2868.
- P. Oelschlaeger, M. Klahn, W. A. Beard, S. H. Wilson and A. Warshel, *J. Mol. Biol.*, 2007, **366**, 687–701.
- G. Blanca, I. Shelev, K. Ramadan, G. Villani, S. Spadari, U. Hübscher and G. Maga, *Biochemistry*, 2003, **42**, 7467–7476.
- G. A. Cisneros, L. Perera, M. García-Díaz, K. Bebenek, T. A. Kunkel and L. G. Pedersen, *DNA Repair*, 2008, **7**, 1824–1834.
- R. D. Shannon, *Acta Crystallogr., Sect. A: Cryst. Phys., Diffraction Gen. Crystallogr.*, 1976, **32**, 751–767.
- M. García-Díaz, K. Bebenek, L. C. Pedersen and T. A. Kunkel, *DNA Repair*, 2007, **6**, 1333–1340.
- G. A. Cisneros, L. Perera, R. M. Schaaper, L. C. Pedersen, R. E. London, L. G. Pedersen and T. A. Darden, *J. Am. Chem. Soc.*, 2009, **131**, 1550–1556.
- B. Silvi, *J. Phys. Chem. A*, 2003, **107**, 3081–3085.
- W. Tiznado, E. Chamorro, R. Contreras and P. Fuentealba, *J. Phys. Chem. A*, 2005, **109**, 3220–3224.
- J.-P. Piquemal, J. Pilmé, O. Parisel, H. Gerard, I. Fourre, J. Berges, C. Gourlaouen, A. D. L. Lande, M.-C. V. Severen and B. Silvi, *Int. J. Quantum Chem.*, 2008, **108**, 1951–1969.
- R. Chaudret, G. A. Cisneros, O. Parisel and J.-P. Piquemal, *Chem.–Eur. J.*, 2011, **17**, 2833–2837.
- B. de Courcy, L. G. Pedersen, O. Parisel, N. Gresh, B. Silvi, J. Pilmé and J.-P. Piquemal, *J. Chem. Theory Comput.*, 2010, **6**, 1048–1063.
- R. G. Parr and R. G. Pearson, *J. Am. Chem. Soc.*, 1983, **105**, 7512–7516.
- R. G. Pearson, *Inorg. Chem.*, 1988, **27**, 734–740.
- A. Warshel, P. K. Sharma, M. Kato, Y. Xiang, H. Liu and M. H. M. Olsson, *Chem. Rev.*, 2006, **106**, 3210–3235.
- W. J. Stevens and W. H. Fink, *Chem. Phys. Lett.*, 1987, **139**, 15–22.
- M. W. Schmidt, K. K. Baldrige, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis and J. A. Montgomery Jr, *J. Comput. Chem.*, 1993, **14**, 1347–1363.
- N. Gresh, G. A. Cisneros, T. A. Darden and J.-P. Piquemal, *J. Chem. Theory Comput.*, 2007, **3**, 1960–1986.
- G. A. Cisneros, J.-P. Piquemal and T. A. Darden, *J. Phys. Chem. B*, 2006, **110**, 13682–13684.
- D. A. Case, T. E. Cheatham III, T. A. Darden, H. Gohlke, R. Luo, K. M. Merz Jr., A. Onufriev, C. Simmerling, B. Wang and R. J. Woods, *J. Comput. Chem.*, 2005, **26**, 1668–1688.
- A. D. Becke, *J. Chem. Phys.*, 1993, **98**, 5648–5652.
- C. Lee, W. Yang and R. G. Parr, *Phys. Rev. B: Condens. Matter*, 1988, **37**, 785–788.
- M. J. Frisch, G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery, Jr., J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox, *GAUSSIAN 09, Revision A.02*, Gaussian, Inc., Wallingford, CT, 2010.
- J. Ponder, *TINKER, Software Tools for Molecular Design, Version 4.2: the most updated version for the TINKER program can be obtained from J. W. Ponder's WWW site at http://dasher.wustl.edu/tinker*, Washington University, St. Louis, 2004.
- Y. Zhang, H. Liu and W. Yang, *J. Chem. Phys.*, 2000, **112**, 3483–3491.
- Y. Zhang, H. Liu and W. Yang, *Ab initio QM/MM and Free Energy Calculations of Enzyme Reactions*, in *Computational Methods for Macromolecules: Challenges and Applications*, Springer Verlag, Heidelberg, Germany, 2002, pp. 332–354.
- Y. Zhang, T. Lee and W. Yang, *J. Chem. Phys.*, 1999, **110**, 46–54.
- J. M. Parks, H. Hu, A. J. Cohen and W. Yang, *J. Chem. Phys.*, 2008, **129**, 154106.
- C. Gonzalez and H. B. Schlegel, *J. Phys. Chem.*, 1990, **94**, 5523–5527.
- H. B. Schlegel, *J. Comput. Chem.*, 2003, **24**, 1514–1527.
- H. Senn and W. Thiel, *QM/MM methods for biological systems*, in *Atomistic approaches in modern biology*, Springer Berlin/Heidelberg, Berlin, Germany, 2007, vol. 268, pp. 173–290.
- K. E. Ranaghan and A. J. Mulholland, *Int. Rev. Phys. Chem.*, 2010, **29**, 65–133.

-
- 56 A. D. Becke and K. E. Edgecombe, *J. Chem. Phys.*, 1990, **92**, 5397–5403.
- 57 B. Silvi and A. Savin, *Nature*, 1994, **371**, 683–686.
- 58 R. F. W. Bader, *J. Am. Chem. Soc.*, 1979, **101**, 1389–1395.
- 59 R. F. W. Bader, *Atoms in molecules: a quantum theory*, Clarendon Press, Oxford, UK, 1990.
- 60 R. J. Gillespie and E. A. Robinson, *Chem. Soc. Rev.*, 2005, **34**, 396–407.
- 61 J. Pilmé and J.-P. Piquemal, *J. Comput. Chem.*, 2008, **29**, 1440–1449.
- 62 S. Noury, X. Krokidis, F. Fuster and B. Silvi, *Comput. Chem.*, 1999, **23**, 597–604.