ELSEVIER

Journal of Molecular Structure: THEOCHEM 806 (2007) 247–252

THEO CHEM

www.elsevier.com/locate/theochem

$[Co(phen)_2dpq]^{3+}$-sheared DNA interactions: Investigation into the basis for major groove recognition and repair

Yan-Bo Wu [a], Zhen-Hai Xiong [b], Hui-Li Chen [a], Pin Yang [a,*]

[a] *Institute of Molecular Science, Key Laboratory of Chemical Biology and Molecular Engineering of Ministry of Education, Shanxi University, Taiyuan 030006, PR China*

[b] *School of Chemistry and Chemical Technologies, Shanghai Jiao Tong University, Shanghai 200240, PR China*

Received 21 July 2006; received in revised form 3 November 2006; accepted 30 November 2006
Available online 12 December 2006

Abstract

The interactions between the chiral metal complexes Δ- and Λ-$[Co(phen)_2dpq]^{3+}$ (phen and dpq stand for 1,10-phenanthroline and dipyrido [3,2-d:2,3-f]quinoxaline, respectively; we abbreviated Δ to "D" and Λ to "L") and sheared DNA d(CC**GAATGAG**G)$_2$ were studied with molecular modeling method. The results reveal that the interactions show obvious enantio-selectivity, groove-selectivity and site-specificity. The intercalation of D-isomer in A_4T_5/T_5A_4 region of DNA major groove was the most preferential binding mode. While when L-isomer intercalated into its preferential binding site (A_3A_4/G_6T_5 region of major groove), the conformation of sheared base pairs was converted to the parallel form, that is to say that the sheared mismatch was repair in conformational level. Further analyses show: (1) The steric interaction, especially the electrostatic part, is the determinant of the recognition events; (2) The structural incitant factor of repair events is the inevitable collision between the ancillary ligand phen and mismatched G_2 base, while the thermodynamic incitant factor is the great decrease of energy brought by the integral π-π stack among normal bases, mismatched bases and dpq ligand.

Keywords: Sheared DNA; Mismatch repair; Molecular modeling; Metal complex

1. Introduction

The study of DNA damage has been a stand-alone research goal, because the results may have important application in the detection and therapy of molecular disease, such as cancer. Recently, a large number of literatures reported the studies on the recognition and repair of base mismatch [1–9], as the base mismatch is the one of the structural modification that could cause these most serious DNA damages. However, most of the recognition and repair system is based on the bio-enzyme, while little attention has been paid to the abiological systems.

In our previous work, the interactions between D- and L-$[Co(phen)_2dpq]^{3+}$ (Fig. 1) and B-DNA were studied by the combined 2D-NMR and computational modeling method [10]. The results from the computation and from the experimental are consistent with each other, so the modeling method is accurate enough for that large system, though it is not so elaborate. In this work, we studied the recognizing characters of D- and L-$[Co(phen)_2dpq]^{3+}$ to sheared DNA, which containing double G:A mismatch, with the same modeling method. During this course, we found that the L-isomer of the complex could repair the sheared G:A mismatch in conformational level, which together with our previous studies on the interaction between D- and L-$[Co(phen)_2hpip]^{3+}$ [11] and sheared DNA discovers a new class of base mismatch repair reagents composed of abiological octahedral coordination complexes.

2. Computational

All calculations were performed in SGI workstation with Insight II software package. The main calculating

* Corresponding author. Tel./fax: +86 351 701 1022.
E-mail address: yangpin33@hotmail.com (P. Yang).

0166-1280/$ - see front matter © 2006 Elsevier B.V. All rights reserved.
doi:10.1016/j.theochem.2006.11.032

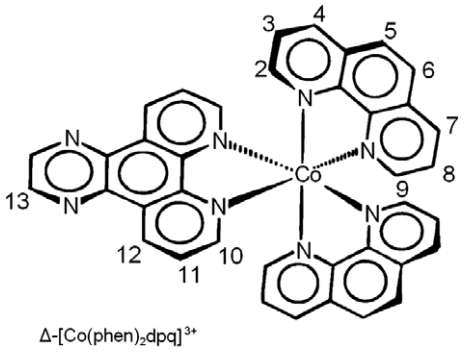

Fig. 1. The structure of D-$[Co(phen)_2dpq]^{3+}$.

program was DISCOVER 98. Default settings for that program were used unless specified otherwise. The system studied contained DNA and a Co atom with octahedral coordination structure. ESFF force field could be dealt with the system efficiently and could offer more output information for analysis, so this force field was used with its default parameters. At the beginning of optimization, the Steepest Descent method was used until the root-mean-square (RMS) derivation was less than 5 kcal/mol. Then it was switched to Conjugate Gradient method automatically by the DISCOVER 98 program. When the RMS derivation was less than 0.5 kcal/mol, optimization was stopped.

As a starting point, the metal complex was constructed in the BUILDER module of InsightII package and optimized in ESFF force field. The X-ray structure of sheared DNA 5'-$(CC\mathbf{GA}AT\mathbf{GA}GG)_2$-3' (Fig. 2) was downloaded from the National Center for Biotechnology Information [12]. First, the metal ions and all H_2O molecules in downloaded structure were eliminated and then all bonds types and atom types were reset. Second, the DNA structure was optimized in ESFF force field.

Electroneutrality of each docked structure was achieved with the addition of 15 Na^+ counterions by standard procedures to balance the 9 phosphate anions provided by each single strand of DNA and the positively charged metal complex. Each isomer was docked manually into the DNA base stack between every double base pairs except the CC/GG region at the termini, and intercalations were carried out in the major groove and minor groove, respectively. As a beginning, the dpq plane was placed nearly parallel to the base pairs plane (perpendicular to DNA helix axis) and just out of the DNA helix. This point was regarded as the first checkpoint and its intercalation depth was defined as 0. Then, Co (III) complex was docked into base stack until the dpq ligand was intercalated into the base stack entirely. We selected the checkpoint for every 0.2 nm, and the intercalation depths were thus defined as 0.2, 0.4 nm, etc. Then based on the potential energy distribution, we could acquire the optimal DNA binding model for each isomer. All interaction systems were dealt with periodic boundary conditions (PBC, usually in a parallelepiped that preserves the shape of the unit cell) and corresponding summation method was the group-based cut-off. All systems containing DNA were calculated in aqueous solution, while other systems in vacuum.

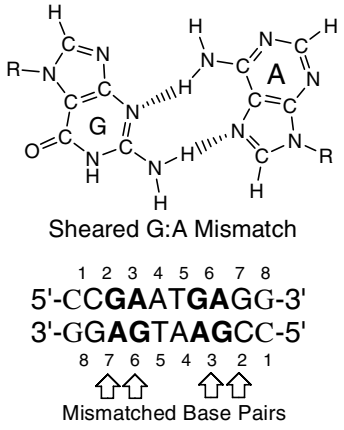

Fig. 2. Sheared G:A mismatch and sequence of sheared DNA.

3. Results

The calculating results of sheared DNA-complex interaction are listed in Tables 1 and 2. When the potential energy distribution is analyzed, the intercalation of D-isomer into A_4T_5 region (Fig. 3-1) and that of L-isomer into T_5G_6 region (Fig. 3-2) of sheared DNA major groove are turned out to be the preferential interaction modes for the two isomers, respectively. As the energy of optimal D-isomer-DNA association (3047.3 kcal/mol) is much lower than that of L-isomer-DNA association (3078.2 kcal/mol), based on our previous research results [10] (for this complex, the isomer that bind more stable with DNA will be enriched), any state of the complex (L-isomer only, D-isomer only, racemic complex or any other rate of D-isomer:L-isomer) will be converted to the state that D-isomer is the main existing form of the complex. However, when the structures of complex-DNA association are studied, the L-isomer is found that could convert the conformation of two mismatched base pairs of the DNA from the sheared form to a parallel form. This means that the sheared G:A mismatch is repaired in conformational level. As this kind of complex maybe the detection and therapy reagent for mismatch-related molecular disease, it is significant to study the mechanism of recognition characters and repair events.

4. Discussion

4.1. Selectivity and related mechanism

The $[Co(phen)_2dpq]^{3+}$-DNA interactions show obvious enantio-selectivity for the complex, the groove-selectivity and site-specificity for sheared DNA. The most preferential interaction mode is the intercalation of D-isomer of complex in A_4T_5/T_5A_4 region of sheared DNA major groove

Table 1
Intercalation of D-isomer into mismatched DNA d(CCGAATGAGG)$_2$ (Kcal/mol)

Depth	Minor groove							Major groove						
	C_1G_2	G_2A_3	A_3A_4	A_4T_5	T_5G_6	G_6A_7	A_7G_8	C_1G_2	G_2A_3	A_3A_4	A_4T_5	T_5G_6	G_6A_7	A_7G_8
00	3199.5	3202.6	3200.2	3203.1	3198.3	3202.1	3195.4	3215.1	3240.2	3176.9	3190.5	3192.5	3235.4	3217.3
02	3168.4	3185.8	3170.8	3189.2	3172.3	3176.2	3179.2	3191.7	3236.2	3152.1	3179.0	3164.0	3213.5	3186.9
04	3157.3	3178.0	3153.9	3168.3	3142.6	3168.1	3163.0	3184.5	3227.4	3145.1	3162.7	3144.5	3175.8	3175.4
06	3141.1	3153.0	**3148.7**	3148.7	3131.0	3169.2	**3166.5**	3179.9	3224.1	3128.7	3144.3	3116.1	3157.8	3180.8
08	3147.0	3147.3	3163.0	3136.3	3121.4	**3151.4**	3177.5	3179.8	**3214.9**	3119.1	3102.0	**3113.2**	**3154.0**	**3164.2**
10	3129.1	3149.8	—	3122.9	3099.9	3172.2	—	**3178.7**	3227.8	3110.0	3088.1	3117.2	3170.7	3178.8
12	3122.9	**3155.1**	—	3110.7	3085.3	—	—	3195.4	—	3109.9	3072.7	3117.4	—	—
14	**3102.5**	3162.2	—	**3108.7**	**3067.1**	—	—	—	—	**3105.8**	**3047.3**	—	—	—
15	3135.2	—	—	3143.5	3094.0	—	—	—	—	3152.0	3055.8	—	—	—

The energies in bold indicate that they are the lowest one in corresponding site.

Table 2
Intercalation of L-isomer into mismatched DNA d(CC**GA**ATG**A**GG)$_2$ (Kcal/mol)

Depth	Minor groove							Major groove						
	C_1G_2	G_2A_3	A_3A_4	A_4T_5	T_5G_6	G_6A_7	A_7G_8	C_1G_2	G_2A_3	A_3A_4	A_4T_5	T_5G_6	G_6A_7	A_7G_8
00	3200.0	3210.0	3196.7	3207.0	3205.6	3185.9	3186.5	3202.3	3221.5	3190.2	3182.1	3202.3	3225.6	3215.4
02	3168.1	3192.7	3183.1	3190.0	3190.3	3175.5	3168.8	3177.1	3215.9	3151.5	3178.7	3177.5	3222.9	3187.3
04	3160.4	3176.8	3179.6	3169.1	3170.5	3177.7	3164.2	3171.0	3190.7	3142.9	3176.1	3147.3	3196.4	3165.6
06	3144.9	3173.5	3161.0	3147.3	3154.0	**3167.8**	3152.0	3167.6	3172.4	3123.9	3171.7	3116.2	3170.8	3168.4
08	3145.4	**3172.7**	3158.9	3137.3	3139.1	3170.1	3153.4	**3156.8**	3160.6	3101.8	3152.0	3084.3	3152.6	3168.4
10	3143.5	3176.9	3149.3	3134.4	3112.9	3196.2	3146.4	3167.2	3164.9	3104.3	3130.0	**3078.2**	3152.8	**3164.2**
12	3135.6	3187.7	3130.5	3128.7	3099.5	—	3132.5	—	3156.7	3104.2	3121.8	3093.3	3153.2	3180.1
14	**3111.1**	—	**3128.9**	**3115.7**	**3093.3**	—	**3129.2**	—	**3144.8**	**3098.8**	**3109.0**	—	**3138.1**	—
15	3133.6	—	3149.9	3122.7	3101.8	—	3148.0	—	3148.0	3117.3	3112.5	—	3148.0	—

The energies in bold indicate that they are the lowest one in corresponding site.

with the dpq ligand as the intercalator. With the detailed structural analysis, the values of potential energy, which is the criterion of recognition characters, were found to have tight relation to steric interaction.

First, the enantio-selectivity and site-specificity are determined by the collisions between the ancillary ligand phen and DNA bases. In general, the purine base adenine (base A) and guanine (base G) are larger in size than the pyrimidine bases cytosine (base C) and thymine (base T), so when the phen ligand collides with purine base A and G, it will encounter more steric collision than with pyrimidine base C and T. For the situation in this work, the arrangement of mismatched purine bases will bring even more collision if the phen ligand collides with them. As shown in Fig. 4, the optimal binding model for each isomer was the intercalation of the complex to the site that phen plane could avoid colliding with the purine bases, especially with mismatched purine bases.

Second, the groove-selectivity is determined by the structural characters of DNA. In our previous studies on the complex-normal DNA interactions, the complex all intercalated DNA in minor groove because the bases in normal DNA are all stretch from minor groove to major groove as shown in Fig. 5a and b. The intercalation in

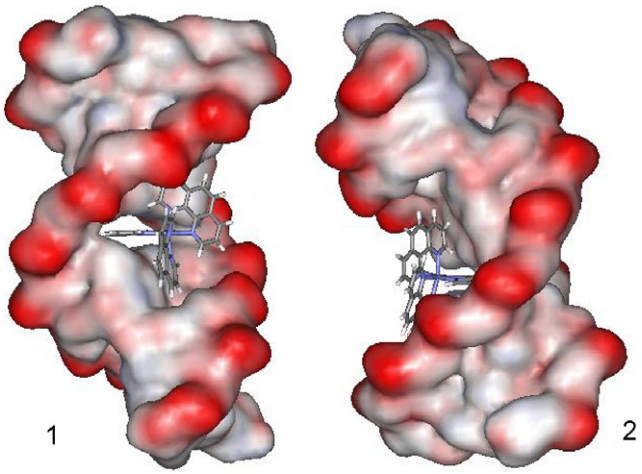

Fig. 3. The optimal structures of complex-sheared DNA interactions. (1) the D-isomer in the A_4T_5/T_5A_4 region; (2) L-isomer in the T_5G_6/A_4A_3 region.

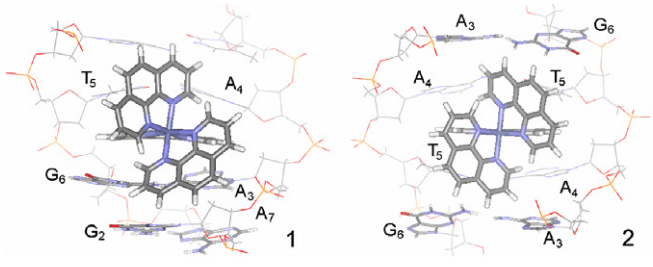

Fig. 4. The ancillary ligand phen avoided colliding with purine base, especially the mismatched bases. The complex in "1" is L-isomer, that in "2" is D-isomer (in the picture, the more ambiguous the further the atom from the eyes).

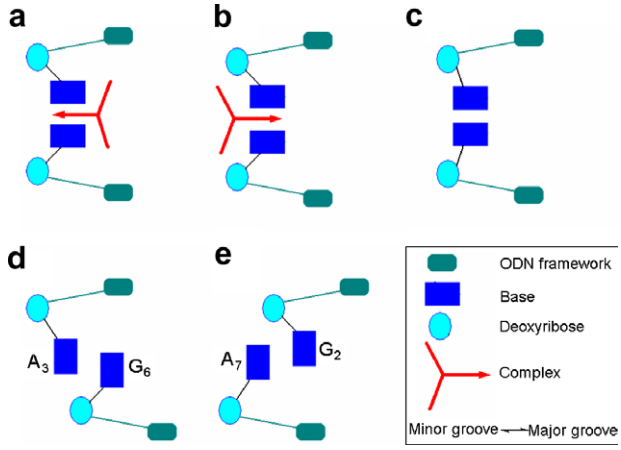

Fig. 5. The contrast of base pairs arrangement between the normal and sheared G:A mismatched DNA, a and b are the colliding condition of the complex's intercalation from normal DNA major and minor groove, respectively. c, d and e are the arrangement of normal base pairs, mismatched A_3:G_6 and G_2:A_7 base pairs, respectively.

the minor groove just likes a wedge wedging into a block. However, the structure of the mismatched DNA is much different from that of the normal DNA. As shown in Fig. 5c, d and e, the arrangement of either the normal base pairs or the mismatched base pairs in sheared DNA makes the intercalation in major and minor groove not so discriminative. In this condition, the existence of a hydrophobe cavity (as shown in Fig. 3) in sheared DNA major groove, in which its size is proper for the complex to enter, decreases the potential energy greatly when the complex enters it. Thus, both isomers interacted sheared DNA in the major groove.

Third, the detailed energy analysis is consistent with the results of structural analysis. Our previous calculation [10,11,13–15] shows that the electrostatics interactions are an important factor influencing the final results. The detailed energy items of some selected binding model are shown in Table 3. In the table, Total stands for total energy; VDW stands for Van de Waals energy; Elect stands for electrostatic energy; Non-bond stands for non-bond energy (the sum of VDW and electrostatic energies) describing the steric interactions; Internal stands for internal energy describing the bond properties. The steric (non-bond) item is more influential than the internal item, as the former is much larger than the later, which consists with the discussion described above. In addition, the elect part was found

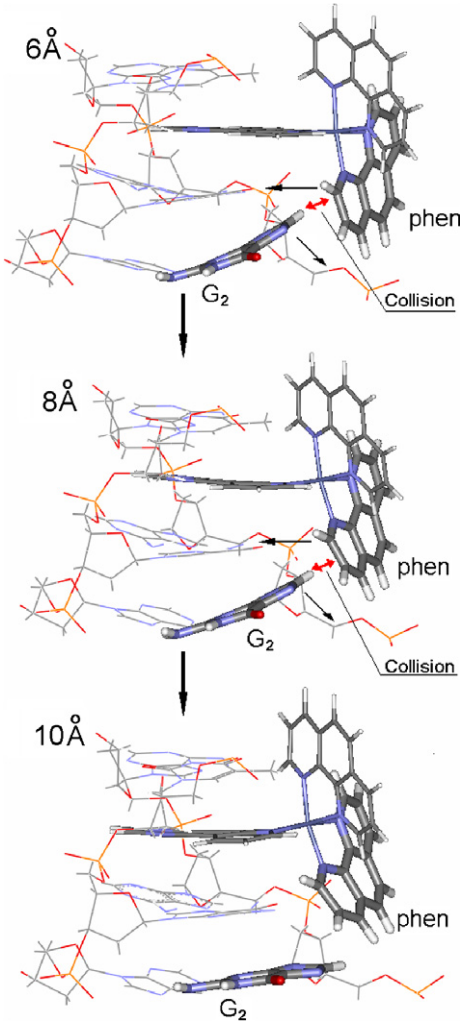

Fig. 6. The mechanism for the reason why the conformation of mismatched G_2:A_7 base pairs was converted.

to be the final determinant of the complex-DNA interactions, because the magnitude order of elect item will generally determine the magnitude order of potential energy.

4.2. Repair events and related mechanism

Generally speaking, the structure of DNA was stabilized by electrostatic interaction between positive metal ions and negative DNA backbone, the intra-base pairs hydrogen bond and the inter-base pairs π–π stacking interaction.

Table 3
Detailed energy (Kcal/mol) comparison of optimal interaction model

Subitems	D-isomer				L-isomer			
	C_1G_2	G_2A_3	A_3A_4	A_4T_5	C_1G_2	G_2A_3	A_3A_4	A_4T_5
Total	3164.19	3154.05	3105.80	**3047.32**	3156.82	3138.12	**3078.22**	3108.99
Internal	646.81	654.85	678.71	664.47	633.822	678.19	650.95	671.29
Non-bond	2517.38	2499.20	2427.09	2382.86	2522.99	2459.93	2427.27	2437.71
VDW	−1.37	5.29	12.01	2.04	4.102	15.62	−14.47	6.52
Elect	2518.75	2493.91	2415.08	2380.82	2518.89	2444.31	2441.74	2431.19

The total energies in bold indicate that they are the lowest one for L- or D-isomer.

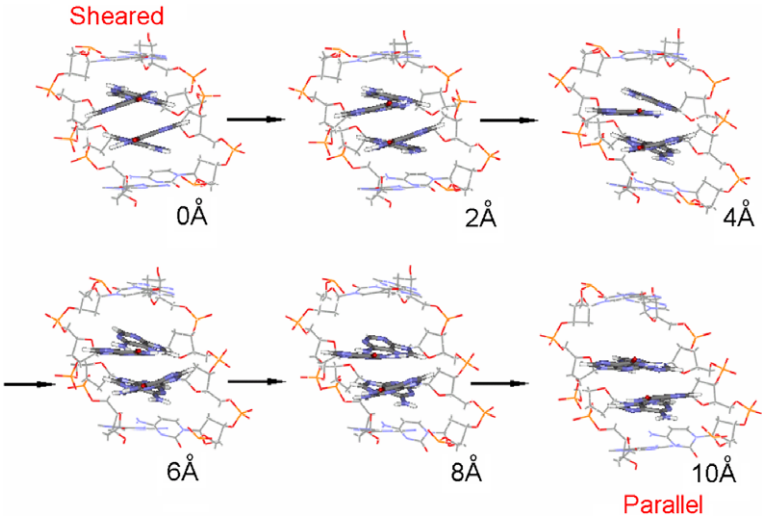

Fig. 7. The conformational conversion of mismatched bases (0, 2, 4,... correspond the intercalation depth).

The integrality of DNA π–π stack is especially important for the stabilization of DNA structure. It is no doubt that the repair of A_3:G_7 base pairs is caused by its steric interactions with dpq ligand. Remarkably, the detailed analyses of L-isomer-DNA interaction model indicate that the steric collision between ancillary ligand phen and G_2 base is the structural incitant factor of conformational repair, while the integral π–π stack concerning normal base pairs, mismatched base pairs and dpq ligand, which decreases the total energy greatly, is the thermodynamic incitant factor.

First, as shown in Table 3, when L-isomer intercalates into A_3A_4 region of sheared DNA major groove, the Elect energy containing the π–π stack is located at a relative lower level (although it is not the lowest), which indicates a well-stacked structure. The Internal energy is also relative lower and VDW is lowest in all interaction systems, revealing the less deviation of bond from the standard values and the lowest steric collision between molecules and between different parts of individual molecules. We consider that this is the thermodynamic reason of repair event.

Second, the related structural analyses reveal that when the L-isomer of the complex inserts into its recognition site and the intercalation depth reaches 6 Å, the ancillary ligand phen inevitably collides with the base G_2 at interlayer, although it could avoid the collision with base G_7 in the binding site. Consistent with the energy analysis, such collision leads to the conformational conversion of mismatched base pairs at interlayer from sheared form to parallel form (as shown in Fig. 6), which makes the chemical bonds less deviate from standard values and also decreases the steric collision between DNA bases greatly. As shown in Fig. 7, with the intercalation of dpq ligand into the base stack, the conformation of mismatched bases is changed gradually from sheared form to parallel form. The converted conformation of mismatched base pairs in complex-sheared DNA association is similar to that of base pairs in complex-normal DNA association (see the enlarged picture in Fig. 8I and II).

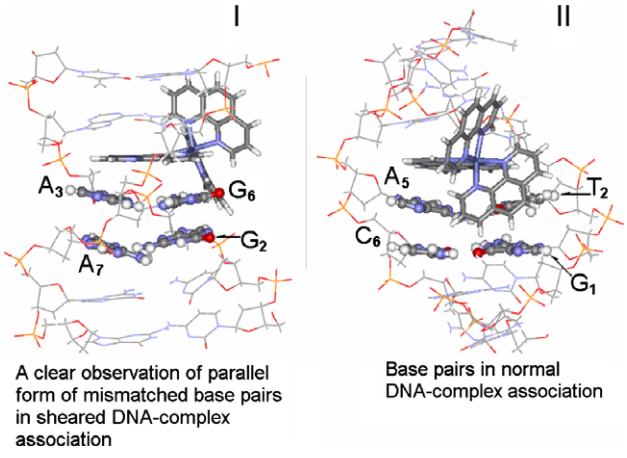

Fig. 8. The enlarged (rotate 90°) structure of parallel style of mismatched base pairs in complex-sheared DNA association (I) and base pairs in complex-normal DNA association (II).

5. Conclusion

The interactions between the $[Co(phen)_2dpq]^{3+}$ and sheared DNA show obvious enantio-selectivity for the complex, groove-selectivity and site-specificity for the DNA. The intercalation of D-isomer into the A_4T_5/T_5A_4 region of sheared DNA major groove with the dpq ligand as the intercalator is the most preferential interaction mode, thus if the complex interacts with sheared DNA, the D-isomer will be enriched. The intercalation of L-isomer into A_3A_4/G_6T_5 region of sheared DNA major groove is not the most preferential interaction mode, but L-isomer-DNA interaction promotes the conformational conversion of mismatched base pairs from the sheared form to parallel form, which together with our previous studies, discovers a new class of base mismatch repair reagents composed of abiological octahedral coordination complexes. The mechanism of recognition characters and repair events are also present. The predictions in this paper await experimental

confirmation, which would open a new branch of biological inorganic chemistry concerning base mismatch recognition and repair.

Acknowledgement

This work was supported by the National Natural Science Foundation of China (Grant No. 20601018) and Natural Science Foundation of Shanxi Province.

References

[1] B.D. Harfe, S. Jinks-Robertson, Annu. Rev. Genet. 34 (2000) 359.
[2] A.B. Buermeyer, S.M. Deschenes, S.M. Baker, R.M. Liskay, Annu. Rev. Genet. 33 (1999) 533.
[3] H.L. Chen, P. Yang, Prog. Chem. 14 (2002) 133.
[4] H.L. Chen, P. Yang, Prog. Chem. 14 (2002) 239.
[5] T.V. Chalikian, K.J. Breslauer, Curr. Opin. Struct. Biol. 8 (1998) 657.
[6] J.E. Cleaver, J.C. States, Biochem. J. 328 (1997) 1.
[7] R.J. Roberts, X. Cheng, Annu. Rev. Biochem. 67 (1998) 181.
[8] M. O'Gara, J.R. Horton, R.J. Roberts, X. Cheng, Nat. Struct. Biol. 5 (1998) 872.
[9] R. Savva, K. McAuley-Hecht, T. Brown, L.H. Pearl, Nature 373 (1995) 487.
[10] Y.B. Wu, H.L. Chen, P. Yang, Z.H. Xiong, J. Inorg. Biochem. 99 (2005) 1126.
[11] Y.B. Wu, C.P. Zhang, P. Yang, Sci. China Ser. B 49 (2006) 177.
[12] http://www.ncbi.nlm.nih.gov/Entrez.
[13] Z.H. Xiong, P. Yang, J. Mol. Struct. (Theochem) 620 (2003) 129.
[14] Z.H. Xiong, P. Yang, J. Mol. Struct. (Theochem) 582 (2002) 107.
[15] C.P. Zhang, Y.B. Wu, P. Yang, Chinese J. Chem. 24 (2006) 739.