

# Molecular modeling on recognition of sheared and normal DNA by novel metal complex $\Lambda$ - and $\Delta$ -[Co(phen)<sub>2</sub>hpip]<sup>3+</sup>

Zhenhai Xiong, Pin Yang\*

*Institute of Molecular Science, Shanxi University, Taiyuan 030006, People's Republic of China*

Received 14 June 2002; accepted 5 August 2002

## Abstract

Recognition of sheared and normal DNA by a novel metal complex [Co(phen)<sub>2</sub>hpip]<sup>3+</sup> (phen = 1,10-phenanthroline, hpip = 2-(2-hydroxyphenyl)imidazole[4,5-f][1,10]phenanthroline) is studied by molecular modeling. Calculating results indicate that, this complex can specifically recognize DNA segment of sequence –MMNMM– (M means mismatch base pairs and N means normal base pairs). Intercalating from minor groove between the middle normal duplex into the sheared DNA with the depth of 1.2 nm is of preference and enantioselectivity is observed. Comparison on the two DNA structures of optimal conformation and analysis on the interaction between DNA and the two tail ligands of the complex show that, the effect of the two neighboring mismatch duplexes on the structure of the middle normal base pairs and the steric interaction between the mismatch duplexes and the two tail ligands of the complex are the essential reason to the segment specificity. Investigation on the detailed energy terms indicate that, in effecting enantioselectivity, the electrostatic distribution of the complex is in the majority and steric interaction is at the next place. But, steric interaction is surely the only factor determining the intercalating from minor groove.

© 2002 Elsevier Science B.V. All rights reserved.

*Keywords:* Steric interaction; Ligand; Base pair

## 1. Introduction

Due to their ability to specifically recognize DNA structure, the potent of complexes composed of extended aromatic heterocyclic ligands and transition metal as photochemical structure and stereo selective probes of DNA have been extensively explored theoretically or experimentally. By far, most of the previous studies focus on the recognition of normal base pairs in normal sequence [1–3] and mismatch

base pairs in sheared sequence [4,5]. The recognition is thought site-specific.

However, as we know, due to their given environment and sequence, DNA helix presents great diversity. In previous works, it was found that, even the segments of same sequence in different environment may be of different structures [4]. In particular, because of the distinct difference between the structure between mismatch base pair and normal base pair (Fig. 1, left), the effect of them on the neighboring segments should be different. That is, even structures of the same segments in normal sequence should be distinctly different from those close to mismatch duplex. The difference of structure

\* Corresponding author. Tel./fax: +86-351-7011022.

E-mail address: yangpin@sxu.edu.cn (P. Yang).

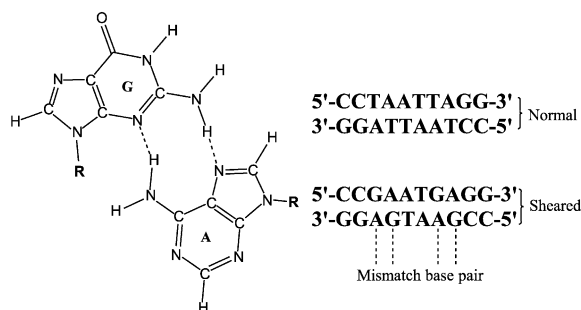


Fig. 1. The sketch of G-A mismatch base pair (left) and sequence of the DNA in this work (right).

will definitely lead to the difference of their recognition by their probe. Moreover, additional to the interaction between the intercalating site of DNA and the intercalating ligand, the two tail ligands can interact with the neighboring duplex. By these thoughts, the effect of sequence itself on the recognition should be of importance.

In this work, the recognition of a sheared DNA  $d(\text{CCGAATGAGG})_2$  and a normal B-DNA  $d(\text{CCTAATTAGG})_2$  by a novel complex  $[\text{Co}(\text{phen})_2\text{hpip}]^{3+}$  (phen = 1,10-phenanthroline, hpip = 2-(2-hydroxyphenyl)imidazole[4,5-f][1,10]phenanthroline) is studied. All difference between the two sequences lies in modification the mismatch duplex of the sheared DNA to normal as Fig. 1, right. Considering the good results of previous molecular mechanics calculated on similar recognition [3,4], we still use it here to investigate the interaction of the assembly in viewpoint of potential energy. By comparing the detailed structure of the two DNA helixes, we study the effect of the neighboring mismatch duplexes on the intercalating sites, and then discuss the interaction between the mismatch duplexes and the two tail ligands of the complex. Another part of the work is the analysis on the detailed energy terms to investigate the main factors leading to the enantioselectivity, intercalating specificity, and the other recognition characters.

Under the effect of the neighboring mismatch duplex, width of the base pairs at A5T6/T5A6 of sheared sequence should be somewhat narrower than that of normal sequence. Since the intercalating ligand of hpip is wider at the head part and narrower at the middle part than the dppz, which has been considered

a potent ligand to recognize normal DNA sequence, it is hopeful to specifically recognize the A5T6/T5A6 in mismatch sequence. Comparison of the structures of  $[\text{Co}(\text{phen})_2\text{hpip}]^{3+}$  and  $[\text{Co}(\text{phen})_2\text{dppz}]^{3+}$  is given in Fig. 2.

## 2. Method

Experiments on the recognition of DNA by similar metal complex have indicated, one of its interaction styles is intercalation as shown in Fig. 3. One of the ligands of the complex is intercalated into DNA helix from major or minor groove parallel to the base stack [1,2]. The intercalation is considered to some extent electrostatic.

As in previous work, only head-on intercalating is considered in this work. In view of the importance of static, formal charge of the central Co atom of the complex is set 3+ here. Models of the two  $[\text{Ru}(\text{phen})_2\text{dppz}]^{n+}$  isomers are constructed with the builder module of INSIGHT II and then undergone molecular dynamics and energy minimization optimization under ESFF force field. The sheared DNA structure is X-ray structure downloaded from the National Center for Biotechnology Information [6]. The sequence is  $[5'\text{-d}(\text{CCGAATGAGG})\text{-3}']$ . After downloaded, the metal ions and all  $\text{H}_2\text{O}$  molecules of the model are eliminated and all bonds types and atom types are reset. Then, the DNA structure is optimized under Amber force field. The normal sequence  $[5'\text{-d}(\text{CCTAATTAGG})_2\text{-3}']$  is constructed with biopolymer module and then undergone energy minimization optimization. In this work, molecular mechanics is used to dock the intercalator into the DNA base stacks from every double base pair except for C1C2/G1G2 and G9G10/C9C10



Fig. 2. Structure of the complexes  $[\text{Ru}(\text{phen})_2\text{dppz}]^{3+}$  (left) and  $[\text{Co}(\text{phen})_2\text{hpip}]^{3+}$  (right).

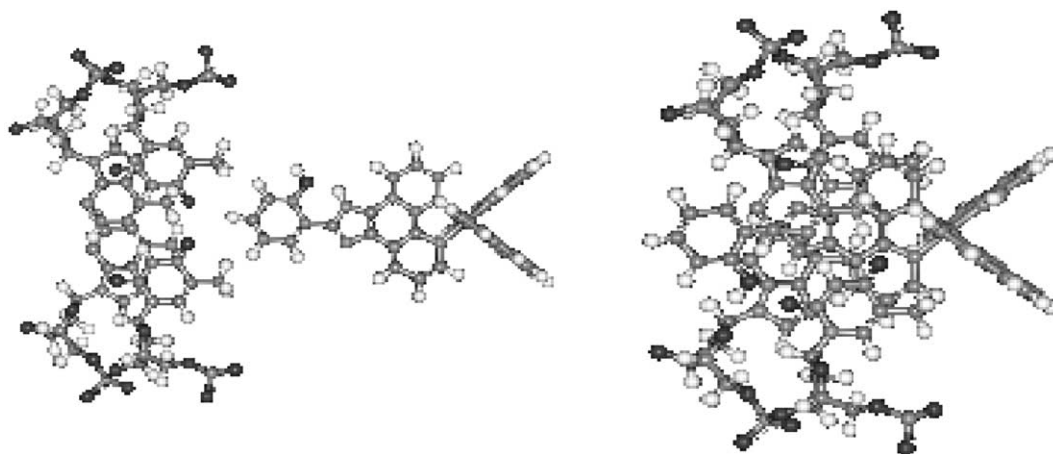


Fig. 3. The structures of the assembly at the initial site (left) and the end site (right).

(considering the terminal effects) in turn. Then, potential energy of the assembly (including DNA fragment and intercalator) at certain sites during intercalating process is calculated, respectively, after the assembly is undergone molecular dynamics. Double isomers of the intercalator, and intercalating from major groove or minor groove are considered. Then, based on the calculating results, we can get the preferential intercalating mode.

On our beginning, the hPIP plane was placed nearly parallel to the base pairs plane (perpendicular to the DNA helix axis) and just out of the DNA helix. Then, keeping the orientation of the hPIP plane changeless, along the axis of hPIP ligand, every moving the intercalator into the base stacks for 0.2 nm, we had a recording site to optimize the structure of the assembly contained DNA and intercalator. And then stopped where the ligand is mostly embedded in the base stacks. The depth of the last site is signed as 0 nm, and that of other sites is signed as 2, 4 nm as so on. The structures of the assembly at the initial site (left) and the end site (right) are listed in Fig. 3.

Boltzmann distribution is given as follows

$$\frac{N_i}{N} = \frac{g_i e^{-\varepsilon_i/kT}}{\sum (g_i e^{-\varepsilon_i/kT})}$$

where  $N_i$  and  $N$  means number of the  $i$ th energy state and their total,  $g_i$ , power of the  $i$ th energy state,  $\varepsilon_i$ , the energy of the  $i$ th energy state,  $k$ , Boltzmann constant and  $T$  is the temperature at 298 K. According to

the right formula, probabilities of all sites can be calculated. And by that, we can get the optimal site.

All calculations are taken under ESFF force field with discover3 and INSIGHTII 2000 on SGI workstation. Formal charge of all atoms but the Co is set 0 and partial charge of all atoms is calculated automatically.

### 3. Calculating results and data analysis

In all the following data table, such abbreviations are promised:

1. The title of all the table as “34”\“45”\“45”\“56”\“67”\“78” means intercalating region; “12”\“10”\“8”\“6”\“4”\“2”\“0” means intercalating depth, with the unit of nm; ‘major groove’ and ‘minor groove’ means intercalating groove.
2. The data in the table means corresponding energy. Unit is kJ/mol.
3. At the site where the energy is lowest compared to the other sites at different depth, the energy data is overstrike.
4. Total means total potential energy; internal means internal potential energy; bond means bond energy; angle means angle energy; torsion means torsion energy; OOP means out-of-plane energy; non-bond means non-bond energy; VDW means VDW energy; VDW-rep means VDW repulsive

Table 1  
Intercalating of  $\Lambda$ -[Co(phen)<sub>2</sub>hPIP]<sup>3+</sup> into mismatch sequence

		34	45	56	67	78
Major groove	12	8750.9041	<b>8325.3263</b>	8602.3193	12,884.865	8850.7144
	10	<b>8612.8211</b>	8406.3263	<b>8266.7862</b>	<b>8204.950</b>	<b>8579.6912</b>
	8	8655.2676	8453.5924	8273.9862	8488.3144	8651.1709
	6	8839.8469	8643.1224	8454.3130	8544.2320	8828.2043
	4	9208.3100	8820.1650	8653.4113	8723.8766	9156.5339
	2	9571.5092	9156.0775	8893.7656	9065.4400	9472.4926
	0	9773.6468	9446.8067	9225.1833	9348.4749	9722.6922
Minor groove	12	8901.7213	<b>8371.2109</b>	<b>7929.1811</b>	<b>8411.986</b>	8858.2281
	10	8828.9485	8462.5735	8348.8104	8564.7978	8784.4996
	8	<b>8731.6462</b>	9169.4479	8628.3052	8669.1925	<b>8701.7417</b>
	6	8839.6776	8869.3539	9037.8773	8917.4709	8851.4629
	4	9251.1601	9169.4481	9323.7941	9177.7952	9208.9349
	2	9634.8348	9850.0784	9612.8149	9496.6525	9604.2103
	0	9880.0340	9910.5225	9933.5158	9809.6275	9892.3038

energy; VDW-dis means VDW disperse energy; electrostatic means electrostatic energy.

### 3.1. Recognition of the mismatch DNA sequence by [Co(phen)<sub>2</sub>hPIP]<sup>3+</sup>

All calculating results on recognition of the mismatch sequence by both isomers of [Co(phen)<sub>2</sub>hPIP]<sup>3+</sup> are listed in Tables 1 and 2.

Based on the data in Tables 1 and 2, Boltzmann distribution of all energy states of the recognition of is

calculated. When only  $\Lambda$ -[Co(phen)<sub>2</sub>hPIP]<sup>3+</sup> is considered as intercalator, intercalating from minor groove into A5T6/T5A6 with the depth of 1.2 nm is of preference, where the potential energy is 7929.18114 kJ/mol and the possibility is 100%. Similar intercalating of  $\Delta$ -[Co(phen)<sub>2</sub>hPIP]<sup>3+</sup> is also of preference with the potential energy of 8209.96142 kJ/mol and the possibility of 100%. Moreover, when recognition of both isomers is considered to investigated the enantioselectivity, we found, intercalating from minor groove into A5T6/

Table 2  
Intercalating of  $\Delta$ -[Co(phen)<sub>2</sub>hPIP]<sup>3+</sup> into mismatch sequence

		34	45	56	67	78
Major groove	12	8973.0573	8316.8413	8270.3946	8326.8150	13,575.0724
	10	<b>8519.2675</b>	<b>8304.8389</b>	<b>8269.5432</b>	<b>8252.0206</b>	<b>8549.2299</b>
	8	8598.0565	8391.6178	8295.4198	8341.0593	8645.4239
	6	13,152.8832	8621.4988	8443.2187	8530.2626	8829.5505
	4	9189.4195	8787.2970	8696.9320	8718.3504	9181.2147
	2	9564.0721	9138.7904	8905.5646	9078.7244	9491.1393
	0	9769.6970	9436.0682	9239.0266	9328.3792	9738.8404
Minor groove	12	8808.0619	12,468.3250	<b>8209.9614</b>	<b>8433.5607</b>	8788.8421
	10	8873.0743	<b>8582.7003</b>	8383.1232	8472.5030	8761.0574
	8	<b>8764.9251</b>	8783.8387	8657.5749	8653.3564	<b>8729.3580</b>
	6	8882.3837	8965.6352	9029.2027	8965.6078	8815.3129
	4	9319.5200	9259.2851	9335.5709	9257.9185	9214.7137
	2	9722.2452	9587.8736	9613.2984	9569.3361	9671.4635
	0	9888.5064	9910.5225	9933.5426	9898.5871	9876.1442

Table 3  
Intercalating of  $\Lambda$ -[Co(phen)<sub>2</sub>hPIP]<sup>3+</sup> into normal sequence

		34	45	56	67	78
Major groove	12	5581.3016	5278.7187	4412.8283	5934.4436	4600.1172
	10	5637.8762	4813.8528	<b>4229.386</b>	5715.6167	4577.9521
	8	4553.4116	5941.3611	4265.3655	5934.4017	6142.0078
	6	4885.7926	4327.5648	4253.5581	4406.9121	4822.2165
	4	<b>4391.1778</b>	4406.7531	<b>4228.7156</b>	4384.9648	4866.6012
	2	4752.1323	<b>4279.7324</b>	4261.2090	<b>4364.6392</b>	6957.4740
	0	4420.0703	4525.2007	4329.1163	4920.7653	<b>4621.7208</b>
Minor groove	12	5854.3705	5570.5120	5537.6297	5575.7545	5879.6258
	10	5946.3619	5688.4766	5596.4237	5672.9667	5014.5249
	8	4493.8022	5692.5727	5636.6561	5702.7911	4701.4816
	6	<b>4328.4076</b>	5541.5073	5756.3393	5826.5134	4974.4308
	4	4393.6976	<b>4300.0441</b>	4316.0435	4266.4381	<b>4397.6145</b>
	2	4483.2549	4354.8249	4243.0495	4243.0789	4521.2446
	0	4602.281	4918.1451	<b>4212.4165</b>	<b>4205.5868</b>	4534.6023

T5A6 with the depth of 1.2 nm of  $\Lambda$ -[Co(phen)<sub>2</sub>hPIP]<sup>3+</sup> account for 100%. That means the whole recognition is enantioselective.

It can noted, the depth of all main intercalating styles is at 1.2 nm. That means the whole hPIP ligand is embedded in the helix. As we know, in experiment [1,2,5], it is the shield of the DNA helix from water leads the fluorescence of intercalating ligand to be observed. Thus, shielded from the water solvent by the DNA helix, fluorescence of hPIP will be observed and the recognition is got.

As some experimental results of similar recognitions, this recognition is from minor groove.

Since preferential intercalating style of each isomer accounts for possibility of 100%, it can be concluded, the recognition is rather specific.

### 3.2. Recognition of normal sequence by [Co(phen)<sub>2</sub>hPIP]<sup>3+</sup>

Data of intercalation into normal sequence by both isomers of [Co(phen)<sub>2</sub>hPIP]<sup>3+</sup> are given in Tables 3 and 4. As above, Boltzmann distribution of all energy states of the recognition of is calculated. The main Boltzmann distribution of intercalating of only  $\Lambda$ -[Co(phen)<sub>2</sub>hPIP]<sup>3+</sup> is: 99.8920% for intercalating

Table 4  
Intercalating of  $\Delta$ -[Co(phen)<sub>2</sub>hPIP]<sup>3+</sup> into normal sequence

		34	45	56	67	78
Major groove	12	5013.591667	5603.439654	4488.4537	4453.75212	5932.48692
	10	6054.306511	4687.256173	5601.72832	5834.84316	5015.22888
	8	4555.770077	4458.369663	4362.55677	6083.15094	4787.59037
	6	4499.867578	4352.884804	4369.04708	5163.30348	4641.90407
	4	5046.438066	<b>4278.529169</b>	<b>4303.0881</b>	4332.17927	5220.84894
	2	4461.894601	4345.797298	4393.52506	<b>4239.0649</b>	6941.00735
	0	<b>4432.901716</b>	4425.823143	4372.40746	4551.22409	<b>4690.85584</b>
Minor groove	12	4688.583599	5435.370161	5400.55804	5484.69743	5683.45008
	10	5431.570794	5676.376979	4537.2253	5718.01339	4506.52098
	8	5120.270169	5691.640172	4470.4367	5128.16614	4619.84372
	6	<b>4276.184068</b>	5144.268101	4339.66261	5813.69204	5860.20104
	4	4341.860152	4361.293225	4348.46999	4289.39099	<b>4253.8137</b>
	2	4371.925744	<b>4322.276582</b>	4255.54417	4297.93021	4427.25456
	0	4427.336072	4486.307222	<b>4248.8695</b>	<b>4171.89501</b>	4547.20169

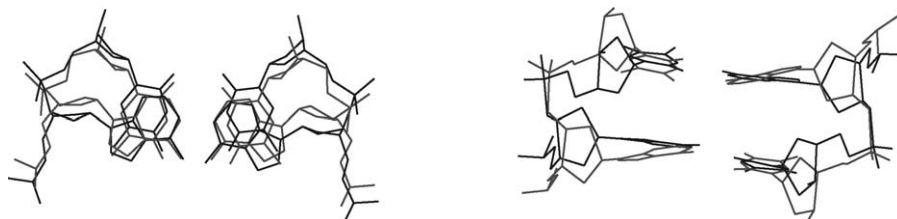


Fig. 4. Comparison of sketch of sheared and normal DNA (sheared in blue and normal in brown).

from minor groove at T6T7/A6A7 with the depth of 0 nm; 0.1080% for intercalating from minor groove at A5T6/T5A6 with the depth of 0 nm. The case of the other isomer  $\Delta$ -[Co(phen)<sub>2</sub>hpip]<sup>3+</sup> is 100% for intercalating from minor groove at T6T7/A6A7 with the depth of 0 nm.

From these data, it can be found, most of the optimal sites are with ligand hpip completely out of the DNA helix at the optimal intercalating depth, no fluorescence can be observed. That is, no recognition is shown.

By comparing recognition of both sequences by the same complex, we can find, the recognition is not simply site-specific. As mentioned above, the sequence itself is of importance to the recognition.

Different from past research, this recognition is specific to a normal duplex in a mismatch sequence. Since only the mismatch base pairs is replaced by corresponding normal base pairs, it is supposed, it is the mismatch base pairs that compose the main factors leading the difference.

### 3.3. Analysis on the effect of the neighboring mismatch base pairs on A5T6/T5A6

Fig. 4 gives the overlay of A5T6/T5A6 in two sequences from different views. It can be found, the position of corresponding atoms are rather different. Some parameters of the structures of the two A5T6/T5A6 are listed in Tables 5 and 6.

By the data in Table 5, both the average distances between P atoms of two AT base pairs in the same chain or in the different chains are somewhat wider in normal sequence than in the other. However, length of all hydrogen bond is close. That means, even the neighboring mismatch duplexes have much effect on structures of A5T6/T5A6, no hydrogen bond is broken. Thus, on one hand, stability of the whole

structure of DNA helix is maintained; on the other hand, new characters come forth.

In fact, as in Table 6, parameters of the DNA backbones and the sugar rings are also much different.  $\alpha$ ,  $\beta$ ,  $\gamma$ ,  $\delta$ ,  $\epsilon$ ,  $\zeta$  and  $\chi$  represent the corresponding dihedral angles of which the axis is P–O5', O5'–C5', C5'–C4', C4'–C3', C3'–O3' and O3'–P, respectively. For example, to base A, all difference of dihedral angle but  $\beta$  and  $\epsilon$  between the two sequences is over 20°. What causes the difference of  $\beta$  and  $\epsilon$  smaller may be, since the axis of  $\beta$  and  $\epsilon$  are bond C–O and then their own dihedral angles tend to be high, to enlarge the angles need too much energy. To be convenient, all listed parameters in Table 6 are the average of those of the two chains in same sequence. Though the two chains in same sequence are not completely same, they are not much different.

On addition, conformations of sugar rings in the two sequences are not completely same. All sugar rings in normal sequence are of 2'-endo. However, the two A base pairs in the A5T6/T5A6 of the mismatch sequence are both 3'-endo. Since the A base pairs in mismatch A/G base pairs are also 3'-endo, it is sure

Table 5  
Comparison of some distance parameters of sheared and normal DNA sketch

	Intrachain (nm)	Interchain (nm)	A/N <sub>6</sub> –T/O <sub>4</sub> (nm)	A/N <sub>1</sub> –T/N <sub>3</sub> (nm)
Sheared DNA	0.6533	1.9061	0.28995	0.2900
Normal DNA	0.68445	1.9781	0.28955	0.2874

Left: intrachain means average distance between two P atoms of A5T6/T5A6 in same chain and interchain means average distance between two corresponding P atoms of A5T6/T5A6 in different chain; right: average distance of the two hydrogen bonds between corresponding A and T bases.

Table 6  
Comparison of backbone angles of sheared and normal DNA sketch

	Sheared DNA		Normal DNA	
	A	T	A	T
$\alpha$ (deg)	-99.05	-76.1	-50.9	-49.65
$\beta$ (deg)	-177.45	174.7	-174.6	-178.4
$\gamma$ (deg)	63.6	75.45	40.5	45.1
$\delta$ (deg)	93.05	126.0	134.75	128.85
$\epsilon$ (deg)	-172.95	-175.95	177.2	179.05
$\zeta$ (deg)	-68.2	-80.05	-93.2	-95.3
$\chi$ (deg)	-144.1	-122.7	-114.6	-115.3

that the difference is caused by the existence of the neighboring mismatch base pairs.

By discussion, it can be found, structures A5T6/T5A6 in the two sequences are rather different under the effect of the neighboring mismatch base pairs. That will surely lead to different recognition results. Moreover, since all the difference between the two sequences lies in the mismatch base pairs, the recognition is not site-specific as past research, but segment-specific.

### 3.4. Analysis on the interaction of mismatch duplex and intercalator

The structures of intercalating from minor into mismatch DNA helix at A5T6/T5A6 is given in Fig. 5 where A5T6/T5A6 is in black and its neighboring mismatch base pairs are in red.

As we know, the ligand hpip can interact strongly with A5T6/T5A6 by the delocalized electrons over

the base planes and those over the hpip plane or by the hydrogen bond between the base pairs and hpip ligand. However, that is not the only interaction between DNA and the complex. It can be noted from the figure that, due to the close positions, the other two tail ligands may interact with the neighboring mismatch duplex base pairs relatively strong. However, no hydrogen bond can be formed between the two tail ligands of the complex and DNA backbone. So, the interaction between them should be non-bond interaction. That means, the steric frame and the electrostatic distribution of the mismatch duplex base pairs will necessarily be of importance to recognition.

### 3.5. Investigation on the detailed energy terms

In Ref. [4], it is found that electrostatic distribution of the assembly is the main factor affecting the recognition and steric interaction is in the next place. To seek the reason of the recognition on atom level, detailed energy results of recognition of the sheared sequence are listed in Table 7 where the intercalating depth is optimal 1.2 nm.

Since both isomers tend to intercalate from minor groove, only intercalating from minor groove is taken into account when discussing the enantioselectivity.

It can be noted that, the total energy of intercalating by  $\Lambda$ -[Co(phen)<sub>2</sub> hpip]<sup>3+</sup> is 280.7719 kJ/mol lower than that of the other isomer. But the internal energy of the former is 71.9842 kJ/mol higher. That means it is the low non-bond energy makes intercalating by  $\Lambda$ -[Co(phen)<sub>2</sub> hpip]<sup>3+</sup> preference. Non-bond energy of

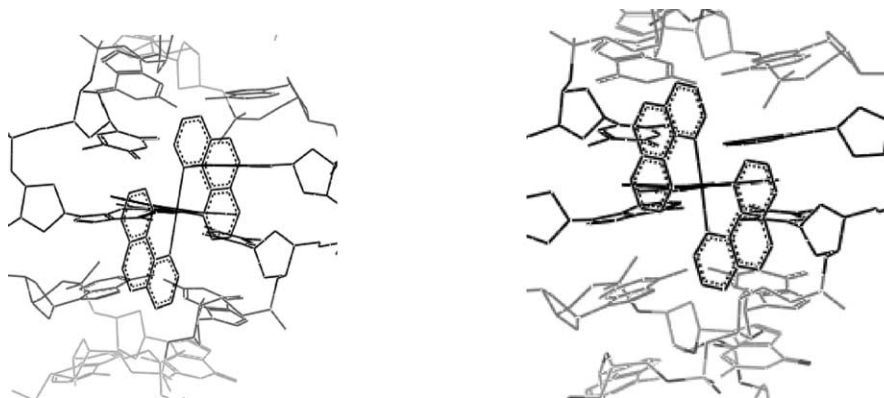


Fig. 5. Sketch of assembly of intercalating of  $\Lambda$ - and  $\Delta$ -[Co(phen)<sub>2</sub>dpphz]<sup>3+</sup> into DNA helix from minor groove at A5T6/T5A6.

Table 7  
Detailed calculating results of the both isomers with the optimal style

	$\Lambda$ -[Co(phen) <sub>2</sub> hpip] <sup>3+</sup>		$\Delta$ -[Co(phen) <sub>2</sub> hpip] <sup>3+</sup>	
	Minor groove	Major groove	Minor groove	Major groove
Total	7929.1979	8602.31721	8209.9698	8270.39379
Internal	3202.38767	3363.992618	3130.40347	3168.857195
Non-bond	4726.7809	5238.32543	5079.537	5101.53869
VDW	-265.60829	77.1256233	-130.47241	-25.3073067
VDW-rep	7019.2556	8018.28149	7541.8743	7893.43625
VDW-dis	-7284.734	-7941.1394	-7672.309	-7918.7229
Electrostatic	4992.385	5161.2001	5210.0136	5126.84629

the former is 352.7561 kJ/mol lower than the latter. Among the non-bond energy difference, as in Table 7, VDW energy accounts for about 38.31% and electrostatic energy accounts for 61.69%. That is, in the two factors affecting the enantioselectivity of recognition, electrostatic interaction is primary and VDW interaction, which means steric interaction, is at subordination. When taking the conformation of the two isomers, this is rather clearly, the symmetry leads to difference of electrostatic distribution and then affects the recognition enantioselectivity. Difference of VDW interaction may come from the fact that both DNA helix and the intercalator is symmetry.

Another problem worth of attention is from which groove to intercalate. To  $\Lambda$ -[Co(phen)<sub>2</sub>hpip]<sup>3+</sup>, the total energy of intercalating from minor groove is 673.11931 kJ/mol higher than that from major groove. Among the difference of total, that of internal energy, VDW energy and electrostatic energy account for 24.01, 50.92 and 25.07%, respectively. That means

steric interaction is in the majority determining which groove it is intercalating from. This means, existence of the two tail ligand of phen will produce huge resistance in major groove but no resistance in minor groove as Fig. 6 shows. The case of another isomer is more evident since electrostatic energy of intercalating from minor groove is higher and steric interaction becomes the only factor making intercalation from minor groove preferential.

### 3.6. Comparison on the recognition of mismatch sequences by two similar complex

In Refs. [4,7], we have studied similar recognition. Owing to different expression rule, 'G7A7/A8G8' in Ref. [4] should be changed into 'G7A8/A7G8' and 'A5T5/T6A6' in Ref. [6] should be 'A5T6/T5A6' here.

As listed in Table 8, common character among the three kinds of recognition is intercalating

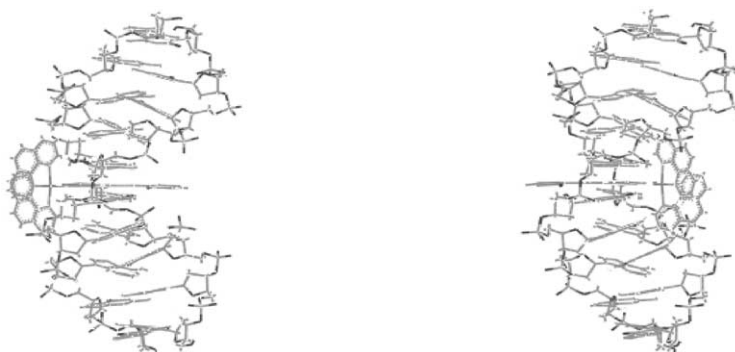


Fig. 6. Sketch of intercalating from minor groove (left) and that from major groove (right).

Table 8  
Comparison of three kinds of recognition in Refs. [4,6] and this work

	Complex	Sequence	Intercalating groove	Enantioselectivity	Intercalating duplex
Ref. [4]	[Ru(phen) <sub>2</sub> dppz] <sup>n+</sup>	d(CCGAATGAGG) <sub>2</sub>	Minor	Δ-	G7A7/A8G8
Ref. [6]	[Co(phen) <sub>2</sub> tpphz] <sup>3+</sup>	d(GTCGATCGAC) <sub>2</sub>	Minor	Δ-	A <sub>5</sub> T <sub>5</sub> /T <sub>6</sub> A <sub>6</sub>
This work	[Co(phen) <sub>2</sub> hpip] <sup>3+</sup>	d(CCGAATGAGG) <sub>2</sub>	Minor	Λ-	A5T6/A6T5
		d(CCTAATTAGG) <sub>2</sub>	None	None	None

enantioselectivity from minor groove. Difference lies in two aspects. Firstly, [Ru(phen)<sub>2</sub>dppz]<sup>n+</sup> tends to intercalate at mismatch duplex; [Co(phen)<sub>2</sub>tpphz]<sup>3+</sup> tends to A5T6/T5A6 of normal sequence; but [Co(phen)<sub>2</sub>hpip]<sup>3+</sup> tends to A5T6/T5A6 between two mismatch duplexes. Secondly, Δ-isomer of the former two is preferential; but the case is just on the contrary to the last one. The reason for these differences may come mainly from the different electrostatic distribution of the assembly and steric interaction among them.

#### 4. Conclusion

Calculating on the recognition of two DNA sequences by the novel complex indicate that, the complex can interact with the sheared DNA sequence. As other similar recognition, it is distinct enantioselective and groove-specific. Λ-[Co(phen)<sub>2</sub>hpip]<sup>3+</sup> is preferential and the intercalating tends to be from minor groove. Investigation on the detailed energy terms shows us that, the electrostatic distribution of the complex is in the majority and steric interaction is at the next place when effecting enantioselectivity. But, steric interaction is surely the only factor determining the intercalation from minor groove.

However, no recognition is found when the sequence is changed into normal B-DNA. That means, the two mismatch duplex base pairs beside the intercalating site A5T6/A6T5 is of importance to recognition of the sheared DNA sequence. After comparing the segments of A5T6/T5A6 in both sequences and investigating the interaction between the neighboring mismatch base pairs and the complex, we find, two functions of the mismatch base pairs on the whole recognition exist.

On one side, the existence of mismatch base pairs affects the structure of A5T6/T5A6; and on the other side, the mismatch base pairs can interact directly with the two tail ligands of phen of the intercalator.

Since the mismatch duplex base pairs is important to the recognition, it is reasonable to consider the recognition as segment-specific to such segment as –MMNNMM–, where M means mismatch base pairs and N means normal base pairs, but not site-specific as was thought traditionally. As Fig. 7 shows, the interaction may be divided into two parts. The first part is that between the A5T6/T5A6 duplex base pairs and the ligand hpip where hpip plane is intercalated parallel to the base pair plane into the DNA stacks. As we know, both the base pairs and the ligand hpip is a system of conjugated π bonds. The delocalized electrons on the base pairs and the hpip plane will produce strong electrostatic interaction. Moreover, some hydrogen bonds can be formed between the planes. The other part is the interaction between DNA helix and the two tail ligands of phen. In fact, here the DNA helix is

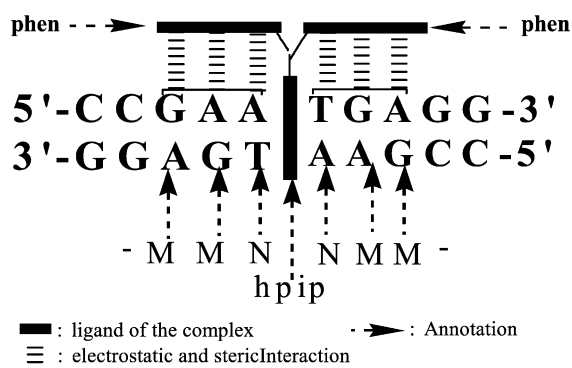


Fig. 7. Sketch map of the segment –MMNNMM– being recognized.

mainly constituted by the two mismatch duplex base pairs and the backbone of DNA. As no hydrogen bond is formed between them, the steric matching and electrostatic interaction should be in the majority.

### Acknowledgements

We sincerely thank the support of the National Natural Science Foundation of China and Provincial Natural Foundation of Shanxi.

### References

- [1] E. Tuite, P. Lincoln, B. Norden, *J. Am. Chem. Soc.* 119 (1997) 239–240.
- [2] K.E. Erkkila, D.T. Odom, J.K. Barton, *Chem. Rev.* 99 (1999) 2777–2795.
- [3] P. Yang, D. Han, Z. Xiong, *J. Mol. Struct. (THEOCHEM)* 540 (2001) 209–217.
- [4] Z. Xiong, P. Yang, *J. Mol. Struct. (THEOCHEM)* 582 (1–3) (2002) 107–117.
- [5] B.A. Jackson, J.K. Barton, *Biochemistry* 39 (2000) 6176–6182.
- [6] <http://ncbi.nlm.nih.gov/entrez>.
- [7] P. Yang, Z. Xiong, *Acta Chim. Sin.* 59 (7) (2001) 1038–1044.