American Journal of Biochemistry and Biotechnology 1 (2): 74-84, 2005
ISSN 1553-3468

Application of RNA-DNA Duplex Base Triplets to Antisense Drugs

Jie Yang, Yao Leng and Ju Bao
State Key Laboratory of Pharmaceutical Biotechnology, Nanjing University, Nanjing 210093, China

Abstract: Sixty-four sets of three-dimensional models of RNA-DNA duplex base triplets were constructed based on codons by homology modeling method using software InsightII on Indigo workstation, which should be helpful for the study of RNA-DNA annealing, the basis of nucleic acids interactions and some peculiar motifs for design antisense oligonucleotides. Our research result reveal that the energies (such as E, Ec, Eb, Et and Enr) of DNA/asRNA hybrids are lower than those of RNA/asDNA hybrids while the energies (such as En and End) of DNA/asRNA hybrids are higher than those of RNA/asDNA hybrids for most binary complex and ternary complex, especially E, Ec, Eb and En in evidence. And the total energy of GGG/CCC hybrid is the lowest of all the hybrids, the more G/C base pairs, the lower of energy of the triplet hybrid and (G/C)3<U(G/C)2<A(G/C)2 <U2(G/C)<AU(G/C)<A2(G/C) in turn; the U-including hybrid system stability, the more of number of uracil (U), the lower of energy of the triplet hybrid; the energy of GU-including hybrid is lower than that of CU-including hybrid no matter for binary complex or ternary complex. G/A/U bases often deviate from base pair planes, which can form hydrogen bonds with neighboring base pairs and affects the stabilities of triplets. Moreover, some peculiar oligodeoxynucleotide sequence motifs that could be derived from corresponding triplet hybrids are divided into two groups: four-member motif and five - member motif, where the former comprises eight motifs, namely 5'-TCTT-3',5'-TGCT-3',5'-CCTT-3',5'-CCAT-3', 5'-CATC-3', 5'-ATCT-3', 5'-GCTG-3' and 5'-GTCT-3'; and the later consists of four motifs, viz 5'-TGCTG-3', 5'-GTCTT-3', 5'-CCATC-3' and 5'-CATCT-3', which are positively correlated with antisense activities and play an important role in designing antisense drugs.

Key words: RNA-DNA hybrid triplets, peculiar sequence motif, antisense oligodeoxynucleotides

INTRODUCTION

The essential steps in rational drug design are the identification of an appropriate target responsible for a certain disease and the development of a drug with a specific affinity to that target. One of the most general approaches of drug targeting is the specific manipulation of gene expression at the DNA or RNA stage of protein synthesis. The antisense principle is based on a specific recognition of certain DNA and RNA regions by an antisense oligonucleotide, which inhibits the translation by a selective pairing of the "sense" with the complementary antisense oligonucleotide strand[1,2]. The first step of protein synthesis can be inhibited by triple helix formation and successively blockade of transcription[3,4]. Secondly, the antisense oligonucleotide may interfere with the processing proteins, which are responsible for the transformation of the primary DNA transcript into the maturated mRNA. Formation of a dual strand between the antisense oligonucleotide and the mRNA may disable the transport of the mRNA from the nucleus to the cytoplasm. Dual strand formation in the cytoplasm results in a blockade of protein synthesis in the ribosomes[2,5]. Antisense oligonucleotide technology has been a very important tool in biological research to study and control gene expression and viral functions[6]. During the past years there are substantial data in a wide range of animal models supporting the activity, specificity and therapeutic utility of antisense therapeutics. Significant progress has been reported with regard to understanding the basic acting mechanisms of antisense drugs. Generally, antisense drugs are designed to modulate the information transfer from the gene to protein. Binding of oligonucleotides to specific sequences may inhibit the interaction of the RNA or DNA with proteins, other nucleic acids or other factors required for essential steps in the intermediary metabolism of the RNA or its utilization by the cell[7]. The study of oligonucleotides, which bind to DNA-DNA duplex to form triplex region, has also been reported[8,9].

In this study, we focus on constructing three dimensional structural models for antisense oligonucleotide complex with their target sequences and provide some peculiar motifs for antisense oligonucleotides design.

Corresponding Author: Dr. Jie Yang, State Key Laboratory of Pharmaceutical Biotechnology, Nanjing University, 22 Hankou Road, Nanjing 210093, China, Tel: 086 25 83594060 Fax: 086-25-83324605

MATERIALS AND METHODS

The three-dimensional structures of the 64 RNA-DNA triplets and three antisense drugs were generated using the Biopolymer module of the commercial software packages InsightII 2000 (MSI, St Louis, MI, USA) on a Silicon Graphics Iris Indigo (SGI, Silicon, CA, USA) workstation.

The structures of A-type RNA-DNA hybrid nucleotide pairs in Biopolymer Module were used to construct the 64 triplet models. Each triplet contains two nucleic acid strands (one deoxyribonucleic acid strand and one ribonucleic acid strand). Then, one cation of K+ was added to each phosphate group to keep the whole system electrically neutral. The initial positions of the K+ ions was on the plane formed by the O-P-O atoms of phosphoric acid group and with equal distances to the two oxygen atoms. The energy of these triplets was optimized by molecular dynamics and molecular mechanics using the Discover Module. The Amber force field is adopted. The derivative was set as 0.1. We adopted steepest descents method to minimize these models for 500 steps and then adopted conjugate gradient method to minimize 1000 steps. Finally, the whole system after energy minimization was soaked in a sphere of aqueous solution (radius=5Å) that contains 312 water molecules and forms the ternary complex of water- K+-nucleotide triplet. The energy of this ternary complex was optimized by 200steps of steepest descents method following by 1000 steps of conjugate gradient method and was simulative annealed by molecular dynamics using the amber force field. A time step of 1fs was used during dynamics. The system was heated to 1000K for 1ps and then down to 300K for 50ps. The average conformation of a series of lowest energy conformations was regarded as the preponderant conformation of the complex. Following each dynamics run, the total energy was minimized via mechanics by using a steepest descent algorithm and a subsequent conjugate gradient method. Similar approaches have been used to study the energies and structural characteristics of DNA triplexes[3,9]. Two sets of data about the energy parameters and conformations of 64 triplets in different conditions were acquired to study the energy and structural characteristics of RNA-DNA duplexes.

Similarly, three antisense drugs, ISIS2922 (5'-GCGTTTGCTCTTCTTCTTGCG-3', 21nt, treating CMV- caused retinitis, ISIS Inc.)[10-12], c-myb[5,13-15] (5'-TATGCTGTGCCGGGGTCTTCGGGC-3', 24nt, inhibiting oncogene c-myb) and GEM91 (5'-CTCTCGCACCCATCTCTCTCCTTCT-3', 25nt, treating AIDS, Hybridon Inc.)[16-19] were used to construct three A-type mRNA/asDNA hybrids using InsightII/Biopolymer Modules described above[4]. Then, one cation of K+ was added to each phosphoric acid group to keep the whole system electricity equilibrium. The energy of these triplets was optimized by molecular dynamics and molecular mechanics using Discover Module. The Amber force field is adopted. At first, fixed the terminal nucleotide pairs and weight atoms. The mRNA/asDNA binary complex were optimized for 200 steps with the steepest descent minimizer and subsequently for 200 steps with the conjugate gradient minimizer. Then, the constraints were removed and computed for 1000 steps with the conjugate gradient minimizer. Secondly, the whole system after energy minimization was soaked in a sphere of aqueous solution (radius 5Å) and forms the ternary complex of water-K+-mRNA/asDNA. The energy of this ternary complex was optimized by 200steps of steepest descents method following with 2000 steps of conjugate gradient method and was simulative annealed by molecular dynamics using the amber force field. A time step of 1fs was used during dynamics integral. The system was heated to 1000K and retained 1ps and then down to 300K to keep 50ps. The average conformation of a series of lowest energy conformation was regarded as the preponderant conformation of the complex. Following each dynamics run, the total energy was minimized via mechanics by using a steepest descent algorithm and a subsequent conjugate gradient method. Similar approaches have been used to study the energies and structural characteristics of DNA triplexes[3,9].

RESULTS

The energies of 64 K+-RNA/DNA triplet binary complexes, water-K+-triplet ternary complexes and three antisense oligodeoxynucleotide complexes are listed in Table 1 and 2. Here, the coulomb energy (Ec) of the triplet complex mostly contribute to the total energy (E) of the triplet complex and non-bond dispersion energy (End), non-bond repulsion energy (Enr), phi energy (Ep), theta energy (Et), non-bond energy (En), bond energy (Eb), hydrogen bond energy (Eh) and out of plane energy (Eo) contribute to the total energy in turn. The energy unit is kcal/mol. Table 3 displays the Watson-Crick hydrogen bond types and lengths of 64 RNA-DNA duplex base triplets, table 4 shows the hydrogen bond types and lengths that are between neighboring base pairs and table 5 arranges other hydrogen bond types and lengths that are within the base pairs of 64 RNA-DNA duplex base triplets.

DISCUSSION

Analysis of DNA-RNA duplex base triplets: From Table 1, there are two types of RNA/DNA hybrids, namely RNA/asDNA and DNA/asRNA, whose difference center on nucleotide orientation

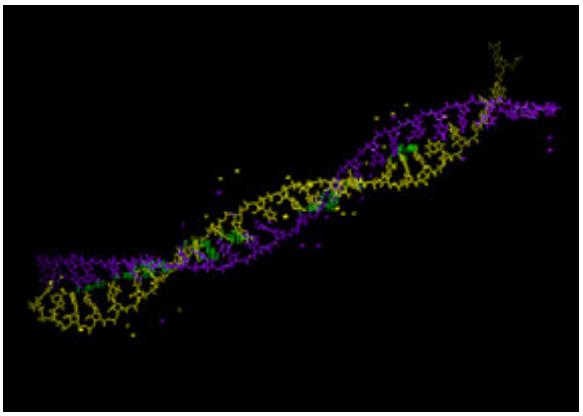


Fig. 1: The lowest energy of the water-K+-GEM91 hybrid ternary complexes

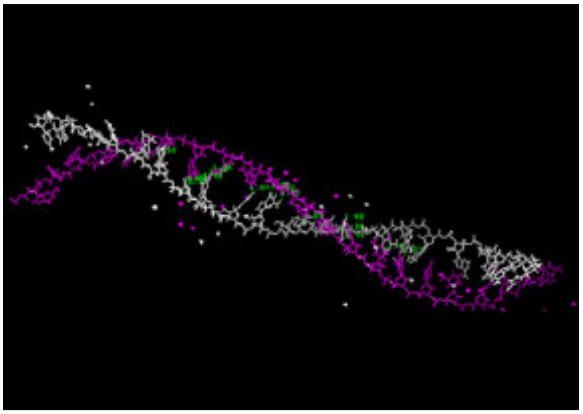


Fig. 2: The lowest energy of the water-K+-isis2922 hybrid ternary complexes

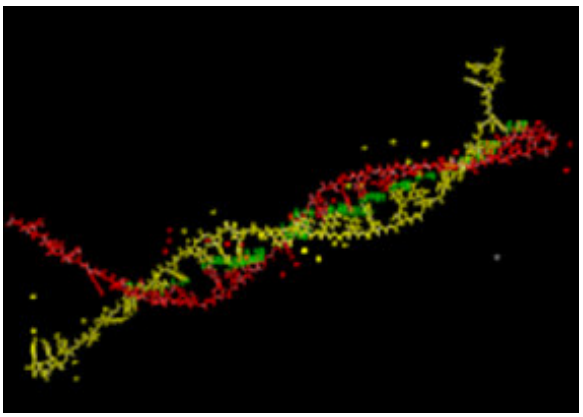
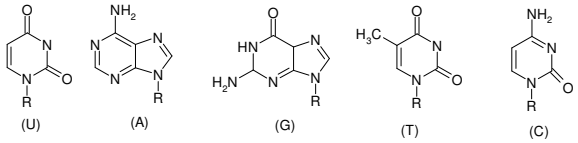


Fig. 3: The lowest energy of the water-K+-c-myb hybrid ternary complexes

(5'-3' or 3'-5'). Here, the RNA/DNA hybrids of 5'-3' RNA were defined as RNA/asDNA hybrids while those of 5'-3' DNA as DNA/asRNA hybrids. And each RNA-DNA hybrid may be regarded as RNA/asDNA or DNA/asRNA hybrids based on the orientation of RNA or DNA. Our research results reveal that the energies (such as E, Ec, Eb, Et and Enr) of DNA/asRNA hybrids are lower than those of RNA/asDNA hybrids while the energies (such as En and End) of DNA/asRNA hybrids are higher than those of RNA/asDNA hybrids for most binary complex and ternary complex, especially E, Ec, Eb and En in evidence. About Ep, Eo and Eh, there are differences in between binary complex and ternary complex, where these energies of DNA/asRNA hybrids for ternary complex are lower than those of RNA/asDNA hybrids while these energies for binary complex contrary to the ternary complex. This is because there are a great deal of hydrogen bond and hydrophobic interaction in K+-triplet binary complex surrounding by water, which lead to the water-K+-triplet ternary complex system energy down and make the system mostly stability. Moreover, the results of paired samples test show that after soaking the K+-triplet complexes in explicit water, the value of energy parameters are generally increased except the nonbond dispersion energy parameter; the energy variations are all significant, which supported by Liu's results[3].

The total energy of GGG/CCC hybrid is the lowest of all the hybrids. And the more G/C base pairs, the lower the energy of the triplet hybrid, namely (G/C)3<U(G/C)2<A(G/C)2<U2(G/C)<AU(G/C)<A2(G/C), which is because guanine and cytosine can form three hydrogen bonds leading to lower energy and structure stability. Some key triplets based on RNA are found, which are 5'-AGG-3', 5'-AGC-3', 5'-CGA-3', 5'-CAG-3', 5'-GAC-3' and 5'-GCA-3'. Moreover, the energy of AAA/UUU hybrid is lower than that of AAA/TTT hybrid. The more uracils (U), the lower of energy of the triplet hybrid, where uracil and adenine more easily form two hydrogen bonds than the formation of hydrogen bond between thymine and adenine. It is the reason that there is a space block during the formation of hydrogen bond between thymine and adenine due to 5-methyl of thymine while no at uracil. So it results the U-including hybrid system stability. And we find that the energy of GU-including hybrid based on RNA (such as 5'-UAG-3', 5'-GAU-3', 5'-AGU-3', 5'-UGA-3', 5'-AUG-3' and 5'-GUA-3') is lower than that of CU-including hybrid no matter for binary complex or ternary complex. Some useful triplet hybrids are also revealed, such as 5'-GGU-3', 5'-UGG-3', UGU, 5'-CCU-3', 5'-UCC-3', CUC, 5'-CGU-3', 5'-UGC-3', 5'-CUG-3', 5'-GUC-3', 5'-GCU-3', 5'-UCG-3', etc. These findings above are consistent with the results of cluster analysis where overall energies of 64 RNA-DNA hybrid triplets are divided into four groups according to their overall energies nearing the rescaled distance 2[20]. The most stable group is composed of triplets that only contain C and G, the group contains 2 G/C pairs in each triplet represents another stable

Table 1: The energies of RNA-DNA duplex base triplets (kcal/mol)



Complex	Triplet RNA (5'-3')	Triplet DNA (5'-3')	E	Eb	Et	Ep	Eo	Eh	En	Enr	End	Ec
Binary	AAA	TTT	-892.1274	9.0894	34.9197	61.4876	0.1866	-6.9049	-11.0848	267.4124	-278.4972	-979.8209
	UUU	AAA	-1133.3453	8.4657	32.8896	60.6247	0.2965	-6.8297	-6.7236	256.9514	-263.6751	-1222.0684
Trinary	AAA	TTT	-860.0808	10.5360	43.9527	62.7329	0.9412	-6.1982	-1.0071	292.5108	-293.5179	-971.0383
	UUU	AAA	-1099.6544	9.4066	39.9575	61.8240	0.5106	-6.1034	-1.0050	279.0758	-280.0808	-1204.2448
Binary	AAG	CTT	-1157.8763	11.9334	36.0343	60.5020	0.2250	-7.0584	-12.4284	268.5683	-280.9967	-1247.0843
	CUU	AAG	-1303.5930	11.7132	35.1836	58.0337	0.1731	-6.0410	-6.4449	256.5591	-263.0039	-1396.2108
Trinary	AAG	CTT	-1119.8012	13.0550	43.4658	61.5516	0.4433	-7.0087	-6.0903	292.3496	-298.4399	-1225.2178
	CUU	AAG	-1263.2487	13.4315	47.4848	67.5289	1.2575	-6.5394	1.5221	280.0128	-278.4907	-1387.9342
Binary	AAC	GTT	-1145.4810	12.1497	36.2204	60.4829	0.2022	-6.8514	-10.3135	268.3446	-278.6581	-1237.3713
	GUU	AAC	-1319.0221	11.5456	34.1271	61.7102	0.2013	-7.3063	-4.8644	262.1383	-267.0026	-1414.4356
Trinary	AAC	GTT	-1109.2442	13.0279	41.4041	66.3327	0.4676	-6.0384	-6.4214	293.3764	-299.7978	-1218.0167
	GUU	AAC	-1292.3654	12.0877	40.1747	62.0737	0.5410	-6.6938	1.2697	288.5729	-287.3032	-1401.8184
Binary	AAU	ATT	-972.4581	9.0340	35.5422	60.7647	0.3722	-6.3750	-8.1244	262.7800	-270.9244	-1063.6717
	AUU	AAT	-1054.0379	8.7913	33.1282	60.8283	0.3015	-6.3660	-6.9404	256.9032	-263.8436	-1143.7808
Trinary	AAU	ATT	-947.8622	9.7396	40.6070	63.0610	0.8422	-5.5878	-4.6349	280.5200	-285.1549	-1051.8894
	AUU	AAT	-1019.0130	10.0514	40.3236	58.9984	0.4387	-5.9533	1.8667	283.6095	-281.7428	-1124.7386
Binary	ACA	TGT	-1145.7127	11.9477	34.4620	61.6547	0.2495	-6.9208	-12.2844	267.2775	-279.5619	-1234.8215
	UGU	ACA	-1320.2743	11.4766	35.3600	61.6991	0.4768	-6.8474	-10.4560	265.8030	-276.2590	-1411.9834
Trinary	ACA	TGT	-1121.5019	12.0004	40.1438	63.0973	0.2828	-5.4392	-6.2058	289.5106	-295.7164	-1225.3813
	UGU	ACA	-1292.2540	12.3058	41.0985	61.9860	0.3703	-7.0490	-2.2363	283.1755	-285.4118	-1398.7293
Binary	ACC	GGT	-1396.0509	14.9238	37.8106	59.6804	0.1342	-6.9789	-12.9496	270.1318	-283.0814	-1488.6714
	GGU	ACC	-1503.7588	14.5628	37.2316	62.5853	0.1983	-7.3839	-9.7183	273.1927	-282.9110	-1601.2348
Trinary	ACC	GGT	-1367.1103	15.3913	42.3188	62.6604	0.4943	-6.1681	-7.9624	289.4795	-297.4418	-1473.8446
	GGU	ACC	-1468.9475	15.7453	37.2370	67.9968	0.4837	-7.4110	-2.3774	297.3432	-299.7206	-1580.6218
Binary	ACG	CGT	-1410.7698	14.5174	34.6074	63.1942	0.4687	-7.5621	-12.2709	269.3976	-281.6686	-1503.7227
	CGU	ACG	-1493.3193	14.5396	36.9950	60.7555	0.7594	-7.3905	-11.0348	268.9524	-279.9872	-1587.9435
Trinary	ACG	CGT	-1380.0357	15.2610	40.7140	65.2631	1.4235	-6.7155	-4.8756	286.5798	-291.4554	-1491.1062
	CGU	ACG	-1461.1223	15.9583	44.6885	59.6665	1.2314	-7.0034	-5.5327	294.7414	-300.2741	-1570.1309
Binary	ACU	AGT	-1225.9236	11.9211	34.3928	61.5605	0.3412	-6.8335	-11.1872	264.9063	-276.0936	-1316.1184
	AGU	ACT	-1240.6151	11.9053	36.0293	61.5532	0.3186	-6.8177	-11.2541	269.9493	-281.2033	-1332.3498
Trinary	ACU	AGT	-1185.4210	13.4377	38.9296	65.8886	1.2194	-7.0633	-0.6434	284.9898	-285.6332	-1297.1896
	AGU	ACT	-1212.6391	12.8144	38.2158	67.9215	0.6278	-6.1370	-2.9902	287.6146	-290.6049	-1323.0913
Binary	AGA	TCT	-1160.8003	11.9782	36.2287	60.6989	0.3159	-6.9821	-11.9218	269.4169	-281.3388	-1251.1182
	UCU	AGA	-1308.4221	11.5531	36.0364	59.6077	0.7423	-7.3085	-5.9722	261.4330	-267.4052	-1403.0809
Trinary	AGA	TCT	-1118.4633	12.7398	42.8992	60.9273	0.3866	-7.2045	-0.9438	291.8649	-292.8087	-1227.2678
	UCU	AGA	-1276.1375	12.2801	40.5856	59.0356	0.6197	-6.1368	-3.6665	282.3013	-285.9678	-1378.8551
Binary	AGG	CCT	-1422.0223	14.5080	35.8884	62.8047	0.3930	-7.6143	-11.4136	273.6138	-285.0275	-1516.5885
	CCU	AGG	-1474.2972	14.7521	37.2177	59.0694	0.4644	-7.4746	-9.3850	267.8481	-277.2331	-1568.9411
Trinary	AGG	CCT	-1391.2051	15.1589	40.3780	67.4345	0.7367	-6.9543	-2.0728	296.4839	-298.5567	-1505.8860
	CCU	AGG	-1449.5096	15.2534	47.5136	58.4900	1.4631	-6.9634	-6.0987	291.4214	-297.5201	-1559.1675
Binary	CGA	TCG	-1405.8814	14.5213	36.4390	63.7954	0.5779	-7.3514	-12.0521	269.1962	-281.2483	-1501.8114
	UCG	CGA	-1488.2767	14.4398	36.3644	62.3529	0.7565	-7.4337	-10.1267	264.4723	-274.5991	-1584.6298
Trinary	CGA	TCG	-1375.2967	15.4747	40.3063	61.8159	0.9685	-6.9191	-4.4581	295.8660	-300.3241	-1482.4850
	UCG	CGA	-1456.1094	14.9103	38.5985	64.6111	0.8073	-7.1139	0.6895	292.6628	-291.9734	-1568.6121
Binary	CGC	GCG	-1663.5803	17.4107	37.7939	60.3568	0.1179	-8.1118	-10.2878	273.7575	-284.0453	-1760.8601
	GCG	CGC	-1676.6523	17.0467	35.4025	64.3445	0.4474	-8.0839	-10.0522	274.2656	-284.3178	-1775.7573
Trinary	CGC	GCG	-1638.4409	18.5871	43.8129	61.2863	0.9603	-6.8421	-6.54260	296.7661	-303.3087	-1749.7028
	GCG	CGC	-1647.9430	17.5169	40.2483	66.1422	1.0235	-8.3819	-5.3342	299.7932	-305.1273	-1759.1575
Binary	CGG	CCG	-1670.1193	17.2313	37.4542	64.2895	0.6887	-7.9671	-11.8787	271.9730	-283.8517	-1769.9371
	CCG	CGG	-1658.1186	17.6450	39.5965	64.7853	1.8079	-7.7231	-12.8562	269.8191	-282.6753	-1761.3740
Trinary	CGG	CCG	-1644.7648	17.5733	40.7184	65.2277	1.3751	-7.8421	-4.1175	303.1527	-307.2702	-1757.6996
	CCG	CGG	-1624.6857	18.1186	45.8418	66.1265	0.9966	-6.7505	-7.0178	293.6257	-300.6435	-1742.0009
Binary	AGC	GCT	-1415.0450	14.8187	36.3189	61.1816	0.3184	-7.2520	-10.9281	272.3746	-283.3027	-1509.5024
	GCU	AGC	-1492.7558	14.5126	35.2994	61.9483	0.3063	-7.7254	-8.5125	269.2184	-277.7308	-1588.5846
Trinary	AGC	GCT	-1382.3593	15.9387	45.8472	64.6920	0.5547	-6.3991	-6.3852	289.9535	-296.3388	-1496.6075
	GCU	AGC	-1455.7270	14.9476	39.1766	65.7583	1.5904	-7.7806	-4.7159	293.9007	-298.6167	-1564.7034
Binary	UCA	TGA	-1223.3530	11.5780	35.0046	61.9437	0.5851	-7.2725	-9.4923	263.8659	-273.3582	-1315.6997
	UGA	TCA	-1236.2727	11.4294	34.4789	62.7783	0.3678	-7.0688	-9.2372	265.1698	-274.4070	-1329.0212
Trinary	UCA	TGA	-1193.6957	11.6518	38.6049	60.8299	0.3048	-5.7323	-6.8747	281.5443	-288.4191	-1292.4799
	UGA	TCA	-1194.1431	13.0838	43.8842	58.6164	0.6052	-5.0542	-3.7376	291.4676	-295.2052	-1301.5410
Binary	UCC	GGA	-1474.9908	14.5510	38.9798	62.2472	0.6055	-8.1016	-9.5836	271.6757	-281.2592	-1573.6900
	GGA	TCC	-1421.7176	14.6286	36.6951	63.5602	0.1138	-7.7026	-9.5760	274.4277	-284.0038	-1519.4368
Trinary	UCC	GGA	-1445.1504	16.0016	40.8409	62.6770	0.6588	-6.5351	-4.5378	293.4477	-297.9855	-1554.2558
	GGA	TCC	-1394.3505	15.1553	39.4757	60.8066	0.2105	-6.8124	-5.9126	296.2142	-302.1269	-1497.2736

Binary	AUA	TAT	-973.0734	8.9236	35.0759	61.1068	0.5578	-6.4511	-9.9378	261.5599	-271.4978	-1062.3487
	UAU	ATA	-1053.8305	8.5878	35.9182	61.3129	0.5384	-7.0403	-8.7586	264.2969	-273.0555	-1144.3889
Trinary	AUA	TAT	-936.8729	9.6271	38.2604	62.7930	0.3213	-5.2600	-2.1648	288.2039	-290.3687	-1040.4498
	UAU	ATA	-1023.1316	9.3245	38.3437	61.8355	0.2550	-5.7694	-4.4486	281.8163	-286.2649	-1122.6724
Binary	AUC	GAT	-1224.7037	11.8234	36.9914	60.6028	0.2361	-6.8955	-8.1171	264.5782	-272.6952	-1319.3449
	GAU	ATC	-1237.0725	11.7648	37.1660	62.1615	0.2845	-7.1714	-6.9398	268.8115	-275.7513	-1334.3380
Trinary	AUC	GAT	-1200.0112	13.0352	42.2637	58.8822	0.5789	-5.3218	-3.0866	284.8626	-287.9492	-1306.3629
	GAU	ATC	-1206.5504	13.3248	42.5296	64.7127	0.3307	-6.5142	-5.6963	292.3485	-298.0448	-1315.2377
Binary	CAA	TTG	-1137.2210	11.7938	36.7515	62.2197	0.1214	-6.5212	-11.0476	265.6309	-276.6785	-1230.5387
	UUG	CAA	-1313.6770	11.1977	34.3399	60.9879	0.2145	-5.5742	-9.1235	257.9843	-267.1078	-1405.7194
Trinary	CAA	TTG	-1104.1301	12.5543	43.4129	61.7886	1.2822	-7.0484	-2.3086	292.3968	-294.7054	-1213.8111
	UUG	CAA	-1275.3791	11.9860	40.3001	62.0144	0.3681	-6.0809	-2.7887	280.6689	-283.4576	-1381.1781
Binary	CAG	CTG	-1406.0451	14.6527	36.7495	60.6522	0.1366	-6.7157	-10.8639	265.8806	-276.7445	-1500.6565
	CUG	CAG	-1489.5260	14.3160	37.4958	58.1298	0.1154	-6.2821	-8.7982	263.2435	-272.0417	-1584.5027
Trinary	CAG	CTG	-1374.6764	15.6290	41.6675	64.0788	1.0964	-6.6761	-1.5077	292.8989	-294.4067	-1488.9644
	CUG	CAG	-1457.8070	15.0688	47.8270	62.8092	1.3891	-6.4384	1.0734	294.4763	-293.4029	-1579.5362
Binary	CAC	GTG	-1398.5179	14.9163	38.2545	60.5311	0.3076	-6.5096	-11.9173	270.2965	-282.2139	-1494.1006
	GUG	CAC	-1501.5177	14.2658	34.2684	63.2177	0.1299	-7.7642	-8.1945	269.1442	-277.3388	-1597.4407
Trinary	CAC	GTG	-1352.1016	14.9911	44.0090	60.6461	1.2062	-6.6647	-3.1681	295.9991	-299.1672	-1463.1213
	GUG	CAC	-1470.8296	15.3741	40.9590	60.7658	0.7622	-7.1714	-2.6908	294.3751	-297.0660	-1578.8285
Binary	CAU	ATG	-1223.5881	11.8050	36.9576	59.2722	0.3134	-6.4917	-8.9752	263.7956	-272.7708	-1316.4693
	AUG	CAT	-1236.2917	11.4274	35.3216	62.7886	0.4936	-7.1380	-7.7194	265.2796	-272.9991	-1331.4655
Trinary	CAU	ATG	-1200.4989	12.8648	40.7639	60.9766	0.9423	-6.2018	-4.0713	285.9463	-290.0175	-1305.7735
	AUG	CAT	-1202.3139	12.6439	41.0276	62.0796	0.8121	-6.0117	-1.8688	287.5742	-289.4431	-1310.9967
Binary	CCA	TGG	-1390.8810	14.9038	37.1698	61.7594	0.4755	-7.4437	-11.0789	268.5330	-279.6119	-1486.6668
	UGG	CCA	-1499.4501	14.0116	35.1260	63.3374	0.3958	-7.4826	-9.4568	268.2926	-277.7494	-1595.3815
Trinary	CCA	TGG	-1362.2390	14.9868	43.0640	63.1239	0.2872	-6.3047	-4.8965	289.7996	-294.6962	-1472.4995
	UGG	CCA	-1462.1552	15.1830	42.9549	63.6416	0.6811	-6.6685	0.5779	286.5116	-285.9337	-1578.5252
Binary	CCC	GGG	-1645.0131	17.7811	39.8792	59.0592	0.1949	-8.2780	-11.0986	273.5544	-284.6530	-1742.5508
	GGG	CCC	-1688.2991	17.2194	35.9042	62.6711	0.1695	-8.2730	-10.7680	275.8587	-286.6267	-1785.2223
Trinary	CCC	GGG	-1616.4314	18.4813	45.2495	63.3234	1.4127	-7.6667	-4.9316	306.7176	-311.6493	-1732.2999
	GGG	CCC	-1646.1508	18.4225	43.5392	64.8357	1.1351	-8.8193	-0.5594	311.7346	-312.2940	-1764.7046
Binary	CUA	TAG	-1225.3419	11.7493	36.9776	58.0039	0.5599	-6.7065	-7.9611	261.0013	-268.9623	-1317.9651
	UAG	CTA	-1237.7046	11.5444	36.6425	62.3331	0.8264	-7.7821	-9.7915	265.9302	-275.7217	-1331.4775
Trinary	CUA	TAG	-1199.2074	12.3475	43.5983	59.4417	1.2318	-6.6703	-1.4989	281.6425	-283.1414	-1307.6576
	UAG	CTA	-1209.8071	12.6183	41.5668	59.3542	0.3838	-6.3039	-3.9152	288.0651	-291.9803	-1313.5111
Binary	CUC	GAG	-1477.5149	14.7676	37.5578	57.2929	0.4535	-6.8166	-9.9365	264.9379	-274.8744	-1570.8336
	GAG	CTC	-1422.6949	14.6187	37.0285	61.8419	0.1737	-7.3797	-11.5478	271.5985	-283.1463	-1517.4303
Trinary	CUC	GAG	-1451.9325	15.1856	43.1510	62.4400	1.1689	-6.8902	-4.1957	292.8643	-297.0600	-1562.7921
	GAG	CTC	-1385.2455	16.0218	42.1183	61.4462	0.4542	-6.8053	-2.2388	294.8932	-297.1320	-1496.2418
Binary	UUA	TAA	-1051.6961	8.5930	34.6046	61.2998	0.7801	-6.3230	-9.3090	259.2090	-268.5180	-1141.3416
	UAA	TTA	-972.8294	8.8282	36.3497	62.0674	0.8754	-7.1382	-10.2874	263.9105	-274.1979	-1063.5245
Trinary	UUA	TAA	-1020.8118	9.7421	41.2129	63.3353	0.4709	-5.4949	-3.9571	284.6966	-288.6537	-1126.1211
	UAA	TTA	-944.0417	9.4177	43.1251	60.9558	0.8833	-6.1174	-4.4110	283.6150	-288.0260	-1047.8952
Binary	GAA	TTC	-1156.1055	11.8155	36.1643	63.4802	0.1993	-7.5743	-8.3650	272.1532	-280.5182	-1251.8255
	UUC	GAA	-1303.6707	11.6780	37.2452	58.5819	0.1321	-6.0163	-8.5372	260.5448	-269.0820	-1396.7543
Trinary	GAA	TTC	-1121.1559	12.7189	39.5492	67.4335	0.2258	-6.2932	-2.7481	290.7601	-293.5082	-1232.0421
	UUC	GAA	-1276.3310	12.0451	38.5129	63.9589	0.5940	-5.8036	-2.5272	281.8009	-284.3282	-1383.1111
Binary	GCA	TGC	-1412.0770	14.5352	35.2539	63.2274	0.3948	-7.7353	-10.2237	272.1904	-282.4140	-1507.5293
	UGC	GCA	-1494.1112	14.3888	35.6835	62.7840	0.6159	-7.2917	-11.3303	270.2159	-281.5462	-1588.9616
Trinary	GCA	TGC	-1386.4938	15.0433	39.1489	66.7229	0.5790	-7.1533	-3.4779	294.8278	-298.3057	-1497.3568
	UGC	GCA	-1464.8150	15.2892	39.4776	66.3946	1.0949	-7.6459	-5.4956	294.4848	-299.9804	-1573.9299
Binary	GCC	GGC	-1663.5803	17.4107	37.7939	60.3568	0.1179	-8.1119	-10.2878	273.7575	-284.0453	-1760.8601
	GGC	GCC	-1678.3702	17.5253	37.7322	62.1123	0.1917	-8.0183	-9.2477	276.7725	-286.0202	-1778.6658
Trinary	GCC	GGC	-1636.5380	18.6167	44.1230	57.9487	0.4089	-7.0095	-8.4192	297.0209	-305.4402	-1742.2065
	GGC	GCC	-1647.3758	18.0883	40.8863	66.4929	0.7907	-8.2442	-3.3278	294.8642	-298.1919	-1762.0621
Binary	GUA	TAC	-1238.3182	11.6792	35.6182	62.3843	0.5472	-7.2189	-8.7138	267.4294	-276.1433	-1332.6144
	UAC	GTA	-1229.8042	11.8033	36.7012	60.9056	0.7916	-7.7482	-9.1734	268.7519	-277.9253	-1323.0843
Trinary	GUA	TAC	-1211.7182	12.9848	39.2577	66.3518	0.1841	-6.2609	-3.2386	289.5713	-292.8099	-1320.9971
	UAC	GTA	-1196.2720	12.5480	41.7685	55.6180	0.1735	-6.2656	-5.0637	284.1658	-289.2295	-1295.0507
Binary	GUC	GAC	-1489.0150	14.5896	37.9956	60.0152	0.1030	-7.0297	-8.3131	269.1095	-277.4227	-1586.3756
	GAC	GTC	-1410.3326	14.7770	37.3122	62.2365	0.2280	-7.5523	-8.0140	272.3090	-280.3230	-1509.3201
Trinary	GUC	GAC	-1462.4008	15.4837	41.0829	62.7354	0.5224	-6.5095	-1.5907	293.2723	-294.8630	-1574.1250
	GAC	GTC	-1381.0779	15.5765	40.3024	65.6844	0.7219	-6.4872	-1.0659	292.0725	-293.1384	-1495.8101

Note: E (total energy), Eb (bond energy), Et (theta energy), Ep (phi energy), Eo (out of plane energy), Eh (hydrogen bond energy), En (non-bond energy), Enr (non-bond repulsion energy), End (non-bond dispersion energy), Ec (coulomb energy). The energy unit is kcal/mol

Table 2: Energy of antisense nucleotides and antigene nucleotides

Name	E	Eb	Et	Ep	Eo	Eh	En	Ec
c-myb	-31330.2857	837.8773	747.4132	637.8622	12.1066	-286.7198	4295.5745	-37574.3996
GEM91	-32206.7652	872.6766	776.2612	662.7832	12.7412	-252.5665	3714.4414	-32684.9953
ISIS2922	-27312.2499	723.2961	625.8243	554.4170	7.3332	-294.8608	4494.0106	-38730.3772

Table 3: The Watson-Crick hydrogen bond types and lengths of 64 RNA-DNA triplets

Amino acid	Triplet RNA (5'-3')	Triplet DNA (5'-3')	Base pairs (RNA-DNA)	Watson-Crick H-bond types (Å) A-T A:H61-T:O4	A-T A:N1-T: H33	A-U A:H61-U:O4	A-U A:N1-U: H33	C-G C: H 42-G:O6	C-G C:N3-G: H11	C-G C:O2-G: H22
Lys	AAA	TTT	1A-3T	2.07	1.91					
			2A-2T	1.99	1.90					
			3A-1T		1.85					
	AAG	CTT	1A-3T	2.13	1.92					
			2A-2T	2.00	1.86					
			3G-1C					1.94	1.91	1.88
Asn	AAC	GTT	1A-3T	1.85	1.96					
			2A-2T	1.92	1.93					
			3C-1G						1.84	1.97
	AAU	ATT	1A-3T	2.03	1.86					
			2A-2T	1.93	1.90					
			3U-1A			1.91	1.90			
Thr	ACA	TGT	1A-3T	1.95	1.85					
			2C-2G					1.88	1.91	1.84
			3A-1T	1.86	1.93					
	ACC	GGT	1A-3T	1.96	1.87					
			2C-2G					1.91	1.92	1.84
			3C-1G					1.9	1.91	1.89
	ACG	CGT	1A-3T	1.84						
			2C-2G					2.04	1.92	1.86
			3G-1C					1.85	1.91	1.90
	ACU	AGT	1A-3T	1.94	1.91					
			2C-2G					1.99	1.92	1.91
			3U-1A			1.90	1.91			
Arg	AGA	TCT	1A-3T	2.00	1.96					
			2G-2C					1.91	1.91	1.86
			3A-1T	2.01	1.95					
	AGG	CCT	1A-3T	2.01	1.86					
			2G-2C					1.89	1.88	1.87
			3G-1C					1.94	1.96	1.96
	CGA	TCG	1C-3G					2.45	1.87	1.89
			2G-2C					1.92	1.94	1.84
			3A-1T	1.97	1.87					
	CGC	GCG	1C-3G					2.20	1.82	1.86
			2G-2C					1.86	1.95	1.94
			3C-1G					1.84	1.97	1.90
	CGG	CCG	1C-3G					1.99	1.85	1.88
			2G-2C					1.89	1.90	1.85
			3G-1C					2.11	1.88	1.92
	CGU	ACG	1C-3G					2.46	1.85	1.86
			2G-2C						1.95	1.89
			3U-1A	1.90	1.72					
Ser	AGC	GCT	1A-3T	1.96	1.90					
			2G-2C					2.21	1.97	
			3C-1G					1.87	1.84	
	AGU	ACT	1A-3T	1.93	1.88					
			2C-2G					1.91	1.93	1.91
			3U-1A			1.86	1.89			
	UCA	TGA	1U-3A			2.05	1.85			
			2C-2G					1.86	1.90	1.92
			3A-1T	1.94	1.92					
	UCC	GGA	1U-3A			2.19	1.98			
			2C-2G					1.98	1.92	1.87
			3C-1G					1.95	1.87	1.97
	UCG	CGA	1U-3A			2.00	1.83			
			2C-2G					1.98	1.92	1.87
			3G-1C					1.93	1.92	1.88
	UCU	AGA	1U-3A				1.89			
			2C-2G					1.87	1.94	1.87
			3U-1A			1.97	1.90			
Ile	AUA	TAT	1A-3T	1.92	1.87					
			2U-2A			1.90	1.90			
			3A-1T		1.92					
	AUC	GAT	1A-3T							
			2U-2A							
			3C-1G					1.86	1.97	
	AUU	AAT	1A-3T		1.90					
			2U-2A			1.93	1.86			
			3U-1A			1.92	1.88			
Met	AUG	CAT	1A-3T		1.83					
			3G-1C					1.90	1.98	1.87
Gln	CAA	TTG	1C-3G					1.96	1.90	1.88
			2A-2T	1.95	1.85					
			3A-1T	1.98	1.86					

	CAG	CTG	1C-3G					1.98	1.87	1.90
			2A-2T	1.87	1.95					
			3G-1C					2.16	1.89	1.94
His	CAC	GTG	1C-3G					1.99	1.89	1.86
			2A-2T	1.94	1.92					
			3C-1G					1.95	1.99	
	CAU	ATG	1C-3G					1.94	1.83	1.89
			2A-2T	1.92	1.87					
			3U-1A			1.90	1.92			
Pro	CCA	TGG	1C-3G					1.96	1.84	1.91
			2C-2G					1.91	1.95	1.84
			3A-1T	1.94	1.84					
	CCC	GGG	1C-3G					1.90	1.93	1.90
			2C-2G					1.88	1.92	1.85
			3C-1G					1.94	1.96	1.91
	CCG	CGG	1C-3G					1.92	1.97	1.95
			2C-2G					1.92	1.87	1.85
			3G-1C					1.84	1.96	1.89
	CCU	AGG	1C-3G					1.88	1.93	1.91
			2C-2G					1.93	1.93	1.82
			3U-1A			1.90	1.92			
Leu	CUA	TAG	1C-3G					1.95	1.87	1.92
			2U-2A			1.89	1.88			
			3A-1T	1.88	1.89					
	CUC	GAG	1C-3G					1.88	1.90	1.84
			2U-2A			1.93	1.91			
			3C-1G					1.93	1.89	
	CUG	CAG	1C-3G					2.04	1.94	1.92
			2U-2A			1.85	2.39			
			3G-1C					1.99	1.86	1.83
	CUU	AAG	1C-3G						1.88	1.90
			2U-2A			2.01	1.84			
			3U-1A			2.00	2.03			
	UUA	TAA	1U-3A			1.96	1.89			
			2U-2A			1.90	1.89			
			3A-1T		1.96					
	UUG	CAA	1U-3A			1.97	1.86			
			2U-2A			1.96	1.90			
			3G-1C					1.88	1.94	1.88
Glu	GAA	TTC	1G-3C					1.94	1.84	1.98
			2A-2T		1.96					
			3A-1T	1.96	1.87					
	GAG	CTC	1G-3C					1.97	1.89	1.90
			2A-2T	1.94	1.90					
			3G-1C						1.94	1.91
Asp	GAC	GTC	1G-3C					2.07	1.84	1.92
			2A-2T	1.96	1.83					
			3C-1G					1.95	1.96	
	GAU	ATC	1G-3C					2.01	1.85	1.90
			2A-2T	2.02	1.86					
			3U-1A			2.17	1.90			
Ala	GCA	TGC	1G-3C					2.01	1.92	1.90
			2C-2G					1.95	1.87	1.84
			3A-1T	1.95	1.87					
	GCC	GGC	1G-3C					1.94	1.9	1.83
			2C-2G					1.9	1.92	1.89
			3C-1G					1.89	1.95	1.93
	GCG	CGC	1G-3C					1.97	1.97	1.90
			2C-2G					1.88	1.91	1.86
			3G-1C					1.93	2.01	1.86
	GCU	AGC	1G-3C					2.03	1.92	1.92
			2C-2G					1.89	1.93	1.86
			3U-1A	1.88	1.92					
Gly	GGA	TCC	1G-3C					2.00	1.88	1.92
			2G-2C					1.89	1.92	1.95
			3A-1T	1.93	1.84					
	GGC	GCC	1G-3C						1.87	1.89
			2G-2C					2.18	1.88	1.89
			3C-1G					1.91	1.89	2.18
	GGG	CCC	1G-3C						1.93	1.91
			2G-2C					1.92	1.93	1.93
			3G-1C					1.95	1.92	1.89
	GGU	ACC	1G-3C					2.04	1.85	1.95
			2G-2C					1.98	1.89	1.87
			3U-1A			1.87	1.98			
Val	GUA	TAC	1G-3C						1.89	1.85
			2U-2A			2.02	1.85			
			3A-1T	2.10	1.86					
	GUC	GAC	1G-3C					2.00	1.84	1.87
			2U-2A			2.04	1.84			

			3C-1G							1.86	1.89	1.91
	GUG	CAC	1G-3C							1.96	1.94	1.85
			2U-2A				2.04		1.87			
			3G-1C							1.88	1.88	1.90
	GUU	AAC	1G-3C							2.04	1.84	1.92
			2U-2A				2.17		1.85			
			3U-1A				1.94		1.96			
Tyr	UAC	GTA	1U-3A				2.01		1.83			
			2A-2T	1.89		1.96						
			3C-1G							1.91	2.01	1.92
	UAU	ATA	1U-3A				2.04		1.86			
			2A-2T	2.05		1.86						
			3U-1A				1.88		1.90			
Cys	UGC	GCA	1U-3A						1.87			
			2G-2C							1.97	1.92	
			3C-1G							1.92	1.93	1.91
	UGU	ACA	1U-3A				2.09		1.84			
			2G-2C									
			3U-1A				1.86		2.01	2.01	1.94	1.92
Trp	UGG	CCA	1U-3A				2.12		1.86			
			2G-2C							1.92	1.89	1.94
			3G-1C							1.93	1.93	1.89
Phe	UUC	GAA	1U-3A				2.44		1.87			
			2U-2A				2.05		1.86			
			3C-1G							1.94	1.87	
	UUU	AAA	2U-2A				2.16		1.90			
			3U-1A				1.92		2.02			
X	UAA	TTA	1U-3A				1.98		1.85			
			2A-2T	1.90		1.89						
			3A-1T	1.87		1.98						
	UAG	CTA	1U-3A				1.98		1.85			
			2A-2T	1.90		1.89						
			3G-1C							2.03	1.93	1.87
	UGA	TCA	1U-3A				1.98		1.85			
			2G-2C							1.90	1.89	

Table 4: The hydrogen bond types and lengths that are between neighboring base pairs

	Triplet			Between the base pairs (Å)													
Amino acid	RNA (5'-3')	DNA (5'-3')	Base pairs (RNA-DNA)	A: H61-G:O6	G:O6-T: H33	C: N4-A: H61	C: H42-G:O6	A: H61-A: N1	A: N1-G:O6	C: H42-A: N6	U:O4-A: H61	A: H61-T:O4	A: N6-C: H42	A: N6-G: H42	G: H22-G: N2	G:O6-A: H61	U: H33-A: N1
Thr	ACA	TGT	1A-2G	2.28													
	ACG	CGT	1A-2G	1.93													
Cys	UGC	GCA	2G-3A	1.99													
	UGU	ACA	2G-3A	2.26													
Arg	AGA	TCT	2G-3T		2.15												
	CGG	CCG	2G-1C				2.11										
Ser	UCC	GGA	2C-3A			2.31											
			3C-2G				2.36										
Gly	GGA	TCC	2G-3C				2.29										
	GGC	GCC	2G-3C				1.99										
	GGG	CCC	2G-3C				2.05										
			3G-2C				2.35										
	GGU	ACC	2G-3C				2.14										
Met	AUG	CAT	1A-2A					1.87									
			3G-2A						2.93								
Tyr	UAU	ATA	3A-2A					2.46									
Leu	CUU	AAG	1C-2A							2.41							
			2U-1A								2.09						
Phe	UUC	GAA	2U-3A								2.07						
	UUU	AAA	2U-3A								1.96						
			3U-2A								2.11						2.31
Glu	GAA	TTC	3A-2T									1.96					
Asp	GAC	GTC	2A-3C										2.03				
	GAU	ATC	2A-3C											2.42			
Ala	GCA	TGC	1G-2G												2.43		
Trp	UGG	CCA	2G-3A													2.22	
Val	GUU	AAC	3U-2A														2.42

group and the group contains 1 G/C pair in each triplet is more unstable than the two former G/C abundant groups, but more stable than the most unstable group (containing only A and U, except UUU).

So, we bring a hypothesis that the RNA/asDNA hybrid can be stably formed when G/C base pair continuously emerge more than three times, which is consistent with Zewert's research result[15] where a seven-member G/C sequence motif exist in an antisense oligodeoxynucleotides. A similar finding occurs in triplet DNA technology where the third strand oligodeoxyribonucleic acid as an antisense drug[4,9],

Table 5: Other hydrogen bond types and lengths that are within the base pairs

Amino acid	Triplet RNA(5'-3')	Triplet DNA(5'-3')	Base pairs (RNA-DNA)	Other H-bond types within the base pair (Å) C:O2-G: H11	C: N3-G: H22	C: N4-G: H11	C: N3-G:O6	C: N3-G: H11	C: N3-G: H22	C: H42-G: N1
Lys	AAG	CTT	3G-1C	2.40						
Asn	AAC	GTT	3C-1G	1.91						
Thr	ACC	GGT	3C-1G	2.49						
Arg	CGA	TCG	1C-3G	2.24						
	CGC	GCG	3C-1G	2.47						
	CGG	CCG	1C-3G	2.42						
	CGU	ACG	1C-3G	2.39						
Ser	UCC	GGA	3C-1G	2.32						
	UCG	CGA	3G-1C	2.38						
	UCU	AGA	2C-2G	2.46						
Met	AUG	CAT	3G-1C	2.26						
Gln	CAA	TTG	1C-3G	2.21						
Pro	CCA	TGG	2C-2G	2.44						
	CCC	GGG	1C-3G	2.08			2.96			
			3C-1G	2.26						
	CCU	AGG	1C-3G	2.37						
Leu	CUA	TAG	1C-3G	2.43						
	CUC	GAG	1C-3G	2.49						
	CUG	CAG	1C-3G	2.23			2.96			
	UUG	CAA	3G-1C	2.40						
Glu	GAG	CTC	1G-3C	2.42						
			3G-1C	2.40						
Asp	GAC	GTC	3C-1G							2.48
	GAU	ATC	1G-3C	2.46						
Ala	GCC	GGC	1G-3C	2.34						
			3C-1G	1.89						
	GCG	CGC	1G-3C	2.42						
			3G-1C	2.14				2.94		
	GCU	AGC	2C-2G	2.38						
Gly	GGG	CCC	1G-3C			2.38			2.16	
			3G-1C	2.48						
Tyr	UAC	GTA	3C-1G	2.38						
Trp	UGG	CCA	2G-2C		2.44					
			3G-1C	2.48						

which the antisense chain mainly recognize G/C-rich segment, such as SP1 binding site and promoter. But it is different from Matveeva's results, which GGGG, 5'-ACTG-3', 5'-CCGG-3', 5'-TAA-3' and AAA based on DNA are negatively correlated with antisense activities. Moreover, according to Table 1, some corresponding deoxynucleotide triplets could be derived from useful triplet hybrids based on RNA mentioned above, which maybe play an important role in inhibiting mRNA transcript, namely 3'-TCC-5', 3'-TCG-5', 3'-GCT-5', 3'-GTC5', 3'-CTG-5', 3'-CGT-5', 3'-ATC5', 3'-CTA-5', 3'-TCA-5', 3'-ACT-5', 3'-TAC-5', 3'-CAT-5', 3'-CCA-5', 3'-ACC-5', 3'-ACA-5', 3'-GGA-5', 3'-AGG-5', 3'-GCA-5', 3'-ACG-5', 3'-AGC-5', 3'-CGA-5', 3'-GCA-5', 3'-GAC-5', 3'-CAG-5' and 3'-GAG-5'. Here, these motifs are partly consistent with Matveeva's result[21] and are supported by Yamaguchi's result[18].

Analysis of antisense drugs: Fig. 1-3 reveal that the antisense nucleotides possessing higher G/C abundance could construct stably hybrid with mRNA while erroneous bit often appear in those of A/T pair to hybrid instability. This is because GC pair can form 3 hydrogen bonds but AT pair only form 1 or 2 hydrogen bonds. And hydrogen bond energy is another parameter that affects the dedicated structures and also plays a critical role in stabilizing the RNA-DNA helix. Although a GC pair can form 3 hydrogen bonds, the contribution of G/C in maintaining the stabilities of RNA-DNA chains is largely concentrated on the redistributions of charges (because the main contribution to overall energy is provided by coulomb energy) and the intensities and directions of chemical bonds. It also suggests that the overall energy evaluation of the complexes formed by antisense nucleotides binding with their target sequences is a reasonable estimation of their binding affinities.

There are usually Watson-Crick hydrogen bonds in 64 RNA-DNA duplex base triplets (Table 3), hydrogen bonds between neighboring base pairs (Table 4) and other hydrogen bonds within the base pairs in RNA-DNA triplet-K+-water trinary complexes (Table 5) also exist. The cluster results of hydrogen energies of triplets in aquatic solution show that 64 triplets are divided into 4 groups according to nearing rescaled distance 5, representing triplets with different hydrogen bond energies and stabilities. The stabilities of hydrogen bond energies are much more complicated due to the intercross between the neighboring base pairs and other factors. Some GC pairs may lose one of their Watson-Crick hydrogen bonds; sometimes the lengths of their hydrogen bonds are extended while decreasing the intensities and some

triplets can form some additional Hoogsteen hydrogen bonds. There are 22 triplets (based on RNA strand) that lack one or more Watson-Crick hydrogen bonds (Table 3): 7 triplets (AAA, UCU, AAC, AUG, AUU, GUA and GAA) of them are in group I and other 5 triplets (AUA, AUC, UGA, UUC and UUA) are in group II, these two groups represent the most unstable hydrogen bond types; other 10 triplets (5'-ACG-3', 5'-UGC-3', 5'-CGU-3', 5'-AGC-3', CAC, CUC, 5'-CUG-3', GAG, 5'-GAC-3' and 5'-GGC-3') belong to the stable groups, because most of them can also form other hydrogen bonds. Some triplets, including CGG, GGC, GGG and UGC, which can form other hydrogen bonds between base pairs besides Watson-Crick hydrogen bonds, acquire most stable conformations (Table 5). Here, there are two triplets, AUG and GUA, different from the results of total energy. But hydrogen bond energy weakly contribute to the total energy while coulomb energy mainly dominate the total energy by the redistributions of charges and the total energy of the triplet complexes are held by coulomb energy.

The loss of hydrogen bonds is often conducted by the base distortion which makes the base deviated from the base pair plane. U is the base that can lead to severe base distortion, especially appearing with A/U, including UUC, UUU, UGA, GUA, AUU, GUA, AUA, AUC and UUA; G abundant triplets are also inclined to deviate from the base pair plane, besides loss of hydrogen bonds, the distortion of G base can often form hydrogen bonds between base pairs, this phenomenon explains why G abundant nucleotides are not often selected as antisense drugs for they are easy to fold with themselves and the consequence is low antisense activities, with the exception of two triplets, AAG and AGA. Moreover, the two triplets also belong to stable groups in vacuum based on hydrogen bond energy, which is similar to that in water. So the relative antisense oligodeoxynucleotide triplets are 5'-CTT-3' and 5'-TCT-3' to 5'-AAG-3' and 5'-AGA-3'. Combined with some deoxynucleotide triplets mentioned above, part of these sequence motifs is found in some antisense drugs, such as GEM91[18,19] (Fig. 1), ISIS2922[10] (Fig. 2) and c-myb[5,15] (Fig. 3). Studies of identification of sequence motifs in oligonucleotides whose presence is correlated with antisense activities are proved by what Matveeva *et al* did[21]. Matveeva *et al* reported that CCAC, TCCC, ACTC, ATCC, GCCA and CTCT are motifs positively correlated with antisense activities, which is similar to our research finding.

From the finding above, it seems to suggest a same result that G/C-rich and U-including hybrid possess lower energy and some relative oligodeoxynucleotide sequence motif could be derived from useful oligonucleotide triplet hybrids above. ISIS2922 is an antisense oligonucleotide with antiviral activity against cytomegalovirus in phase III clinical test, whose nucleotide acid sequence (5'-GCGTTTGCTCTTCTTCTTGCG-3') includes nine triplets underlined, under-dotted, hatched and framed, where two peculiar motifs were found, 5'-TGCT-3' and 5'-TCTT-3'. An antisense oligodeoxynucleotide c-myb (5'-TATGCTGTGCCGGGGTCTTCGGGC-3') is corresponding to the promoters of the protooncogene c-myb, which is in phase I/II clinical tests, whose sequence includes eight triplets underlined, under-dotted, hatched and framed, where six peculiar motifs were found, 5'-TGCT-3', 5'-GCTG-3', 5'-GTCT-3', 5'-TCTT-3', 5'-TGCTG-3' and 5'-GTCTT-3'. GEM 91 (gene expression modulator) is a 25-mer oligonucleotide phosphorothioate complementary to the gag initiation site of HIV-1. It has been studied in various *in vitro* cell culture models to examine inhibitory effects on different stages of HIV-1 replication. Experiments were focused on the binding of virions to the cell surface, inhibition of virus entry, reverse transcription (HIV DNA production), inhibition of steady state viral mRNA levels, inhibition of virus production from chronically infected cells and inhibition of HIV genome packaging within virions. Experiments were also performed *in vitro* in an attempt to generate strains of HIV with reduced sensitivity to GEM 91. Using *in vitro* methods that were successful in generating HIV strains with reduced sensitivity to AZT, Yamaguchi K *et al*[18] were unable to generate strains with reduced sensitivity to GEM 91. At present GEM91, which inhibits both HIV adsorption and HIV integrase, is in phase I/II clinical trials. There are six peculiar motifs, 5'-CCAT-3', 5'-CATC-3', 5'-ATCT-3', 5'-CCTT-3', 5'-CCATC-3' and 5'-CATCT-3', in its sequence (5'-CTCTCGCACCCATCTCTCTCCTTCT-3'), which includes thirteen triplets underlined, under-dotted, hatched and framed.

CONCLUSION

Our research results reveal that the energies (such as E, Ec, Eb, Et and Enr) of DNA/asRNA hybrids are lower than those of RNA/asDNA hybrids while the energies (such as En and End) of DNA/asRNA hybrids are higher than those of RNA/asDNA hybrids for most binary complex and ternary complex, especially E, Ec, Eb and En in evidence. The total energy of GGG/CCC hybrid is the lowest of all the hybrids, the more of number of G/C base pair, the lower of energy of the triplet hybrid and (G/C)3<U(G/C)2<A(G/C)2<U2(G/C)<AU(G/C)<A2(G/C) in turn; the U-including hybrid system stability, the more of number of uracil (U), the lower of energy of the triplet hybrid; the energy of GU-including hybrid is lower than that of CU-including hybrid no matter for binary complex or ternary complex. The formation of hydrogen bonds between

neighboring base pairs often results to the distortion of base pair planes and affects the stabilities of triplets, which is often conducted by G/A/U bases. Moreover, some peculiar oligodeoxynucleotide sequence motifs that could be derived from corresponding triplet hybrids mentioned above are positively correlated with antisense activities, which are divided into two groups: four-member motif and five-member motif. The former comprises eight motifs, namely 5'-TCTT-3', 5'-TGCT-3', 5'-CCTT-3', 5'-CCAT-3', 5'-CATC-3', 5'-ATCT-3', 5'-GCTG-3' and 5'-GTCT-3'; and the later consists of four motifs, which are 5'-TGCTG-3', 5'-GTCTT-3', 5'-CCATC-3' and 5'-CATCT-3'. Matveeva *et al*[21] reported that CCAC, TCCC, ACTC, ATCC, GCCA and CTCT are motifs positively correlated with antisense activities, which is similar to our research finding. So designing antisense drugs is required not only to find a sequence with low bind energy, but also to avoid the motifs that can decrease the antisense activity.

ACKNOWLEDGMENTS

This work is supported by NSFC (Nature Science Fund of China, No. 30171094 and 30271497).

REFERENCES

1. Tamm, I., B. Dorken and G. Hartmann, 2001. Antisense therapy in oncology: new hope for an old idea? Lancet 358: 489-497.
2. Urban, E. and C.R. Noe, 2003. Structural modifications of antisense oligonucleotides. Farmaco, 58: 243-258.
3. Liu, S.Q., J. Yang, C.L. Bai and C.Q. Liu, 2002. Analysis of the energies of the triple base triplets. J. Mol. Struct.-Theochem., 578: 199-207.
4. Yang, J. and C.Q. Liu, 2000. Stability and Molecular Modeling of Triplex DNA inhibiting DNA Binding Protein Binding to the Core Promoter of Hepatitis B Virus. Prog. Biochem. Biophys, 27: 283-286.
5. Pastorino, F., D. Stuart, M. Ponzoni and T.M. Allen, 2001. Targeted delivery of antisense oligonucleotides in cancer. J. Controlled Release, 74: 69-75.
6. Stein, C.A. and A.M. Krieg, 1994. Problems in interpretation of data derived from *in vitro* and *in vivo* use of antisense oligodeoxynucleotides. Antisense Res. Dev., 4: 67-69.
7. Crooke, S.T., 1999. Molecular mechanisms of action of antisense drugs. Biochim Biophys Acta, 1489: 31-43.
8. Brion, P. and E. Westhof, 1997. Hierarchy and dynamics of RNA folding. Annu. Rev. Biophys. Biomol. Struct, 26: 113-127.
9. Liu, S.Q., J. Yang, C.L. Bai and C.Q. Liu, 2002. The comparative analysis of the energies of conjugated triple base triplets and their applications. J. Mol. Struct.-Theochem., 578: 215-219.
10. Arnedo, A., M.A. Campanero, S. Espuelas, M.J. Renedo and J.M. Irache, 2000. Determination of oligonucleotide ISIS 2922 in nanoparticulate delivery systems by capillary zone electrophoresis. J Chromatography, 871: 311-320.
11. Grooke, S.T., 1998. Antisense Nuclei Acid Drug Dev., 8: 115-122.
12. Kandimalla, E.R., A. Manning, Q. Zhao, D.R. Shaw, R.A. Byrn, V. Sasisekharan and S. Agrawal, 1997 Mixed backbone antisense oligonucleotides: design, biochemical and biological properties of oligonucleotides containing 2'-5'-ribo- and 3'-5'-deoxyribonucleotide segments. Nucleic Acids Res., 25: 370-8.
13. Gewirtz, A.M., D.L. Sokol and M.Z. Ratajczar, 1998. Nucleic acid therapeutics: state of the art and future prospects. Blood, 92: 712-736.
14. Roush, W., 1997. Antisense aims for a renaissance. Sci., 276: 1192-1193.
15. Zewert, T.E., U.F. Pliquett, R. Langer and J.C. Weaver, 1995. Transdermal transport of DNA antisense oligonucleotides by electroporation. Biochem. Biophy. Res. Communications, 212: 286-292.
16. Akhtar, S. and S. Agrawal, 1997. *In vivo* studies with antisense oligonucleotides. Trends Pharmacol. Sci., 18: 12-18.
17. Wagner, R.W., 1994. Gene inhibition using antisense oligodeoxynucleotides. Nature, 372: 333-335.
18. Yamaguchi, K., B. Papp, D. Zhang, A.N. Ali, S. Agrawal and R.A. Byrn, 1997. The multiple inhibitory mechanisms of GEM 91, a gag antisense phosphorothioate oligonucleotide, for human immunodeficiency virus type 1. AIDS Res Hum Retroviruses, 13: 545-554.
19. Zhang, R., J. Yan, H. Shahinian, G. Amin, Z. Lu, T. Liu, M.S. Saag, Z. Jiang, J. Temsamani and R.R. Martin *et al*., 1995. Pharmacokinetics of an anti-human immunodeficiency virus antisense oligodeoxynucleotide phosphorothioate (GEM 91) in HIV-infected subjects. Clin Pharmacol. Ther., 58: 44-53.
20. Wang, Y.L. and J. Yang, Energy and structural analysis of double nucleotide acid triplets. J. Theoretical. Biol., in press.
21. Matveeva, O.V., A.D. Tsodikov, M. Giddings, S.M. Freier, J.R. Wyatt, A.M. Spiridinov, S.A. Shabalina, R.F. Gesteland and J.F. Atkins, 2000. Identification of sequence motifs in oligonucleotides whose presence is correlated with antisense activity. Nucleic Acids Res., 28: 2862-2865.