

Supporting Information

Conformational Stabilities of Iminoallantoin and its Base Pairs in DNA: Implications for Mutagenicity

N.R. Jena^{1*}, Manju Bansal², P.C. Mishra³

¹Discipline of Natural Sciences, Indian Institute of Information Technology, Design and
Manufacturing, Jabalpur-482005

²Molecular Biophysics Unit, Indian Institute of Science, Bangalore-560012

³Department of Physics, Banaras Hindu University, Varanasi-221005

*Corresponding Author's Email Address: nrjena@iiitdmj.ac.in

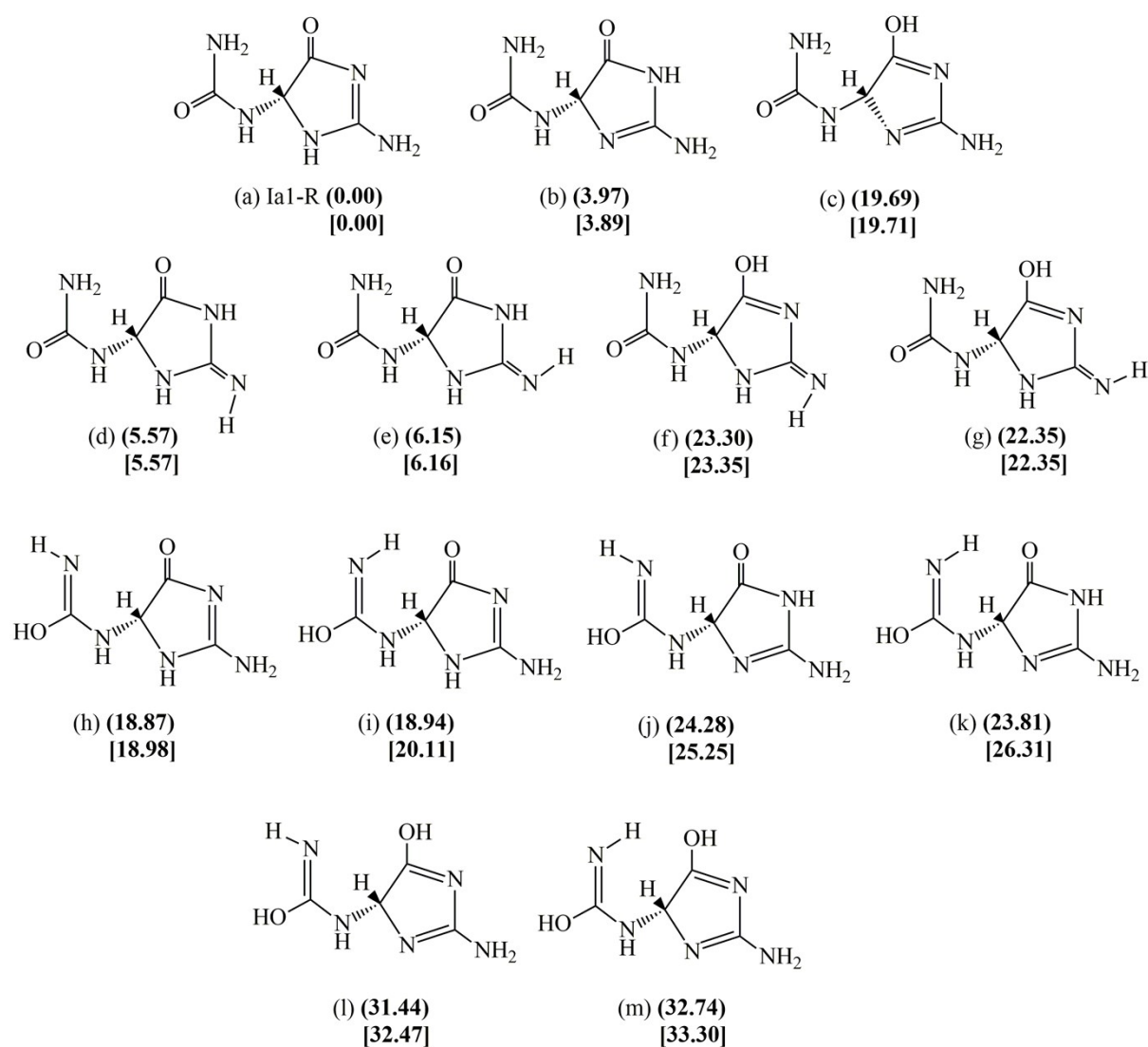


Fig. S1: Structures of different tautomers of Ia1 in the R-stereoisomeric conformation (Ia1-R). The relative ZPE-corrected total energies (kcal/mol) calculated with respect to (a) by employing B3LYP/6-31+G* level of theory are shown in parentheses. The energies shown in brackets correspond to the equivalent tautomers of Ia1 in the S-stereoisomeric conformation (Ia1-S).

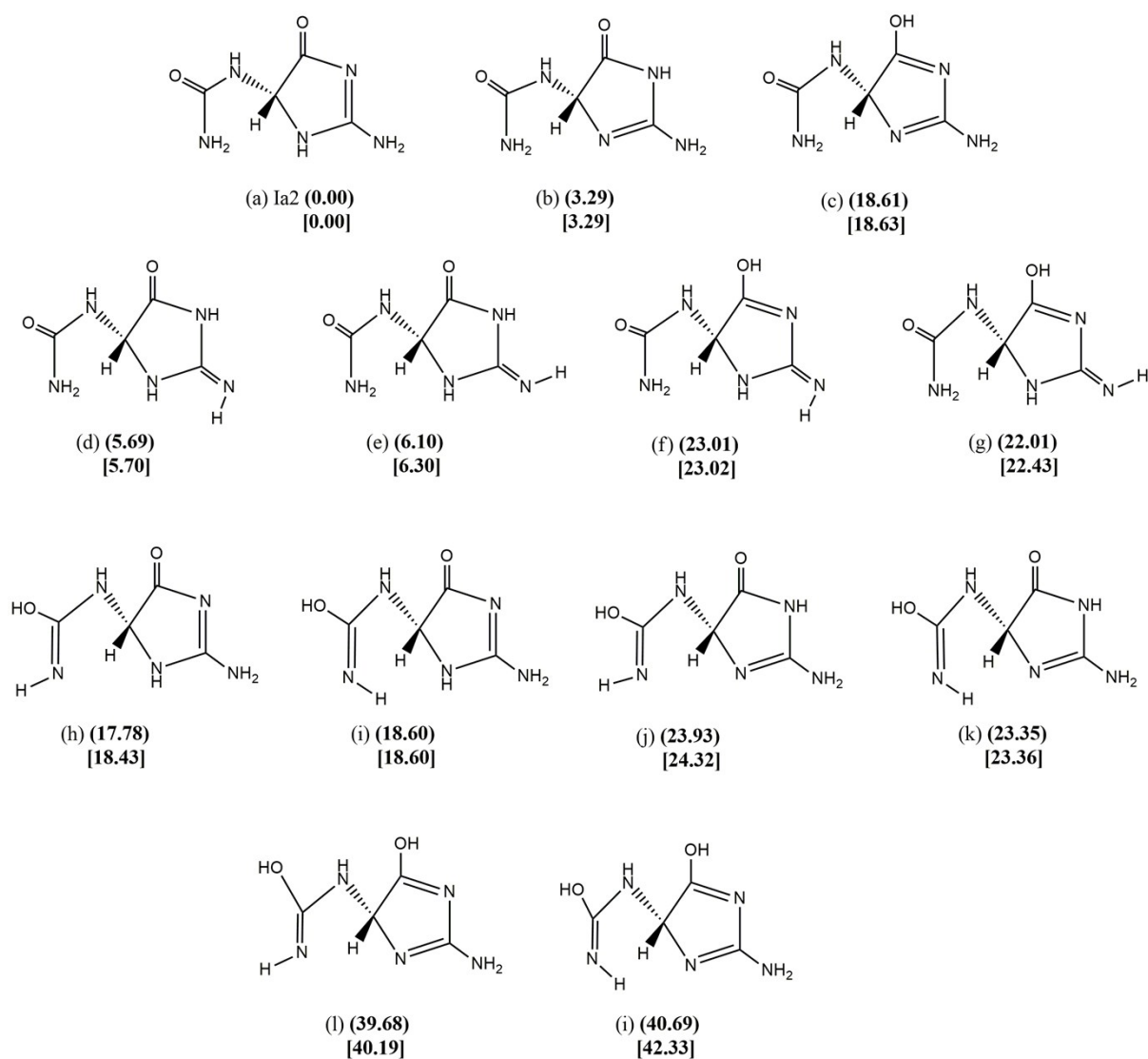


Fig. S2: Structures of different tautomers of Ia2 in the R-stereoisomeric conformation (Ia2-R). The relative ZPE-corrected total energies (kcal/mol) calculated with respect to (a) by employing B3LYP/6-31+G* level of theory are shown in parentheses. The energies shown in brackets correspond to the equivalent tautomers of Ia2 in the S-stereoisomeric conformation (Ia2-S).

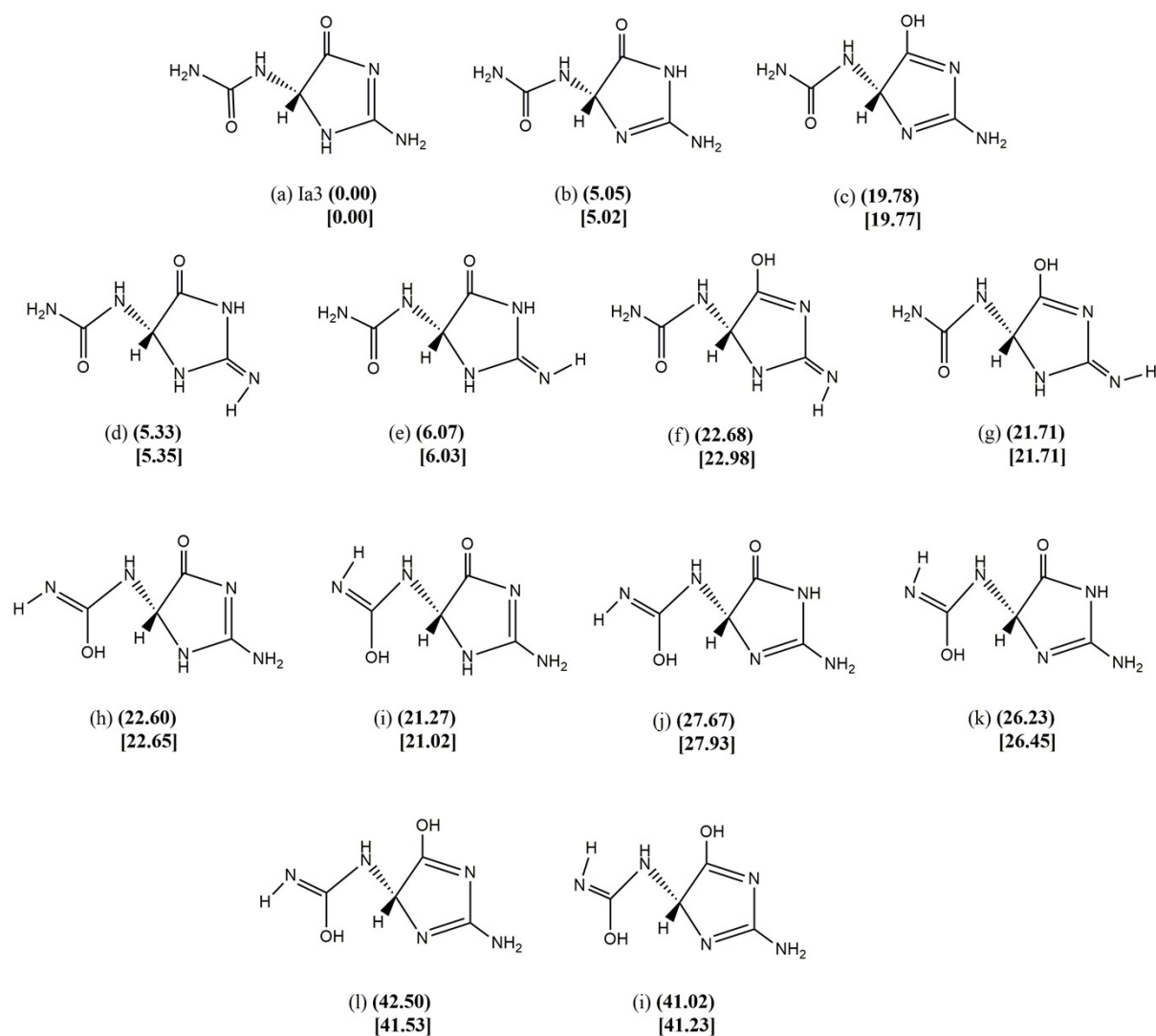


Fig. S3: Structures of different tautomers of Ia3 in the R-stereoisomeric conformation (Ia3-R). The relative ZPE-corrected total energies (kcal/mol) calculated with respect to (a) by employing B3LYP/6-31+G* level of theory are shown in parentheses. The energies shown in brackets correspond to the equivalent tautomers of Ia3 in the S-stereoisomeric conformation (Ia3-S).

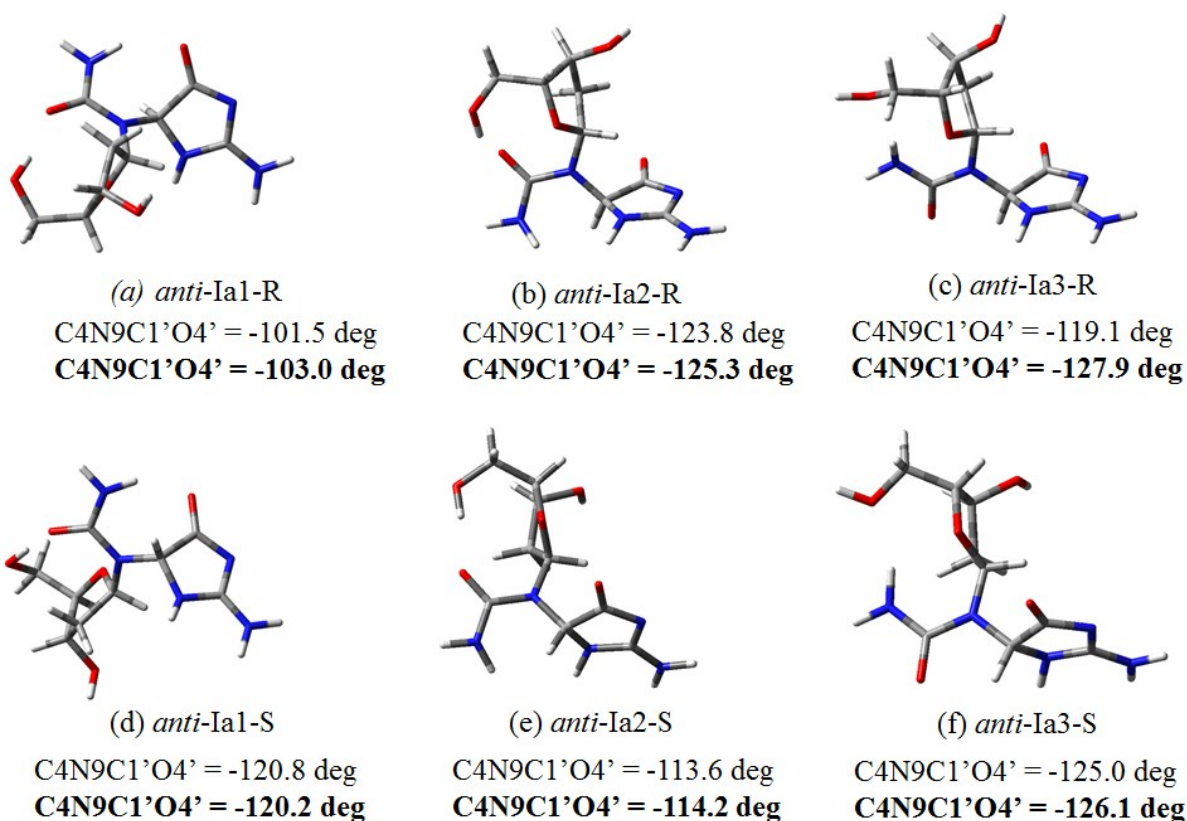


Fig. S4: Optimized structures of different rotamers of the most stable tautomer of 2'-deoxyIa in the R- and S-diastereoisomeric conformations in aqueous medium. The χ values of these rotamers obtained at the B3LYP and B3LYP-D3 (in bold) methods are shown for comparison. As the χ values are just higher than the permissible value of a *syn*-conformation in DNA ($\chi = 0$ to 90 deg), *N*-glycosidic bond rotation from *anti*-Ia to *syn*-Ia would be easy in DNA.

Table S1: The ZPE-corrected binding energies of different complexes involving the R-stereoisomer of Ia (Ia-R) in the *anti*- and *syn*-conformations as obtained in aqueous medium by employing different level of theories. The binding energies of T:G, G:C and T:A complexes are shown in parentheses for comparison. The absence of entry shows that the corresponding complex does not exist.

Complex	Method	<i>anti</i> -Ia-R							<i>syn</i> -Ia-R							
		Ia11	Ia12	Ia13	Ia21	Ia22	Ia31	Ia32	Ia11	Ia12	Ia13	Ia21	Ia22	Ia31	Ia32	Ia33
Ia:G (T:G)	B3LYP	-6.84 (-7.41)	-5.07	-10.12	-6.80	-10.29	-7.40	-10.59	-7.61	-2.93	-5.10	-5.46	-3.59	-7.68	-5.85	-7.60
	ωB97XD ^a	-9.88 (-10.82)	-8.42	-14.64	-9.80	-14.84	-10.56	-15.20	-10.81	-6.00	-7.46	-7.45	-6.51	-10.90	-8.53	-10.20
	ωB97XD	-10.19 (-9.99)					-10.19	-14.11						-10.26		
	B3LYP-D3	-10.81 (-11.12)					-10.81	-15.26						-11.37		
Ia:C (G:C)	B3LYP	-7.21 (-11.63)			-7.01		-6.80		-3.50			-6.09		-5.96	-10.78	-2.42
	ωB97XD ^a	-10.69 (-16.26)			-10.49		-10.61		-6.53			-9.08		-8.92	-15.90	-4.74
	ωB97XD	-10.23 (-15.16)					-10.23								-14.75	
	B3LYP-D3	-10.87 (-16.44)					-10.87								-15.95	
Ia:A (T:A)	B3LYP	-6.80 (-6.57)	-2.64	-3.74	-6.73		-6.76		-2.31			-4.84	-3.43	-5.75	-5.58	-1.47
	ωB97XD ^a	-9.97 (-10.30)	-4.30	-6.09	-9.82		-9.96		-4.29			-8.47	-6.44	-8.36	-9.48	-3.98
	ωB97XD	-9.21 (-9.30)					-9.21									
	B3LYP-D3	-10.03 (-10.28)					-10.03									
Ia:T	B3LYP	-6.91	-6.80		-4.88		-6.89		-6.29			-6.59	-5.40	-6.53	-6.20	-6.90
	ωB97XD ^a	-10.03	-9.96		-10.22		-11.59		-8.82			-9.13	-8.46	-9.19	-8.92	-10.12
	ωB97XD						-6.91									-9.96
	B3LYP-D3						-7.90									-11.30

^aObtained by single-point energy calculations by employing B3LYP/6-31+G* geometry.

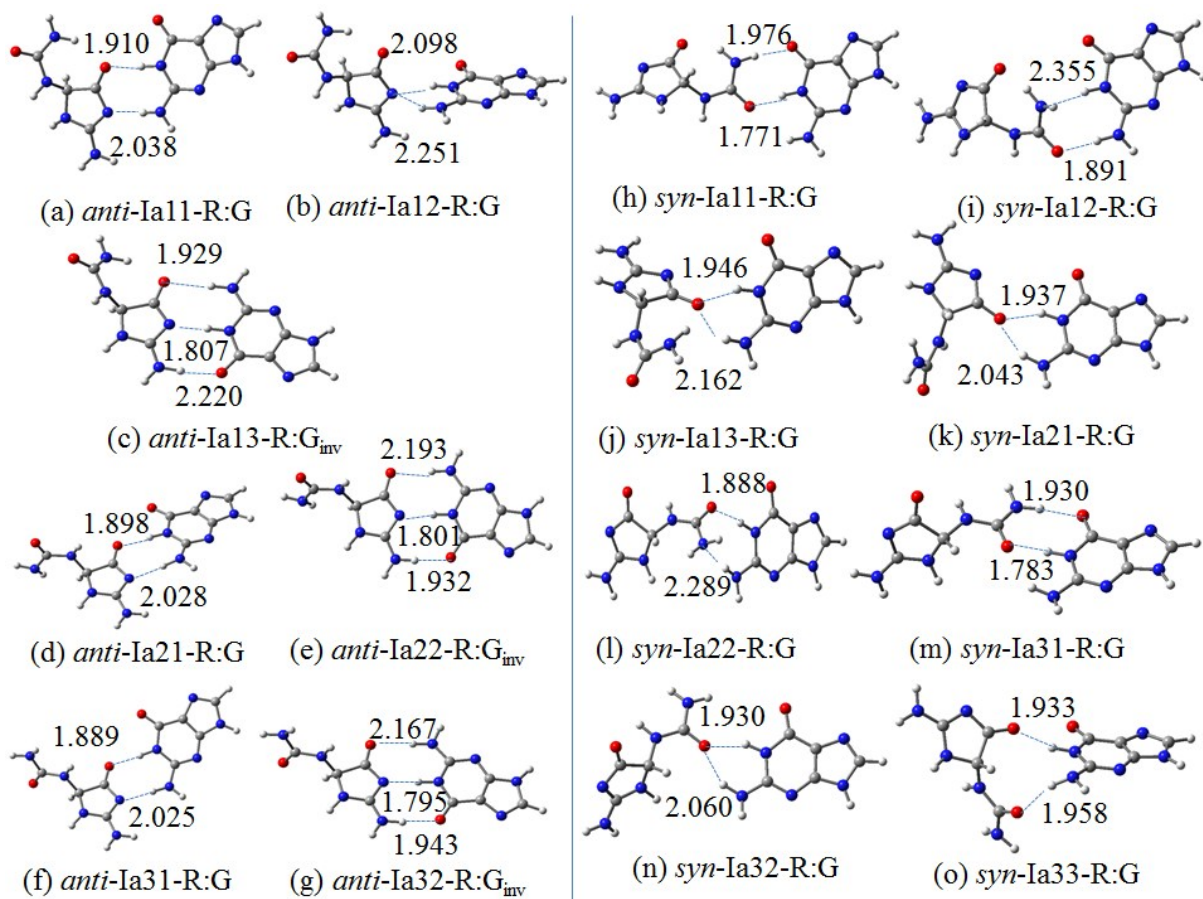


Fig. S5: Different optimized structures of Ia-R:G complexes as obtained in aqueous medium by employing B3LYP/6-31+G* level of theory. In (c), (e), and (g) G binds with *anti*-Ia-R in the inverted orientation (G_{inv}). We noted that these complexes can also be formed by binding of G with the inverted *anti*-Ia-R (*anti*-Ia-R_{inv}).

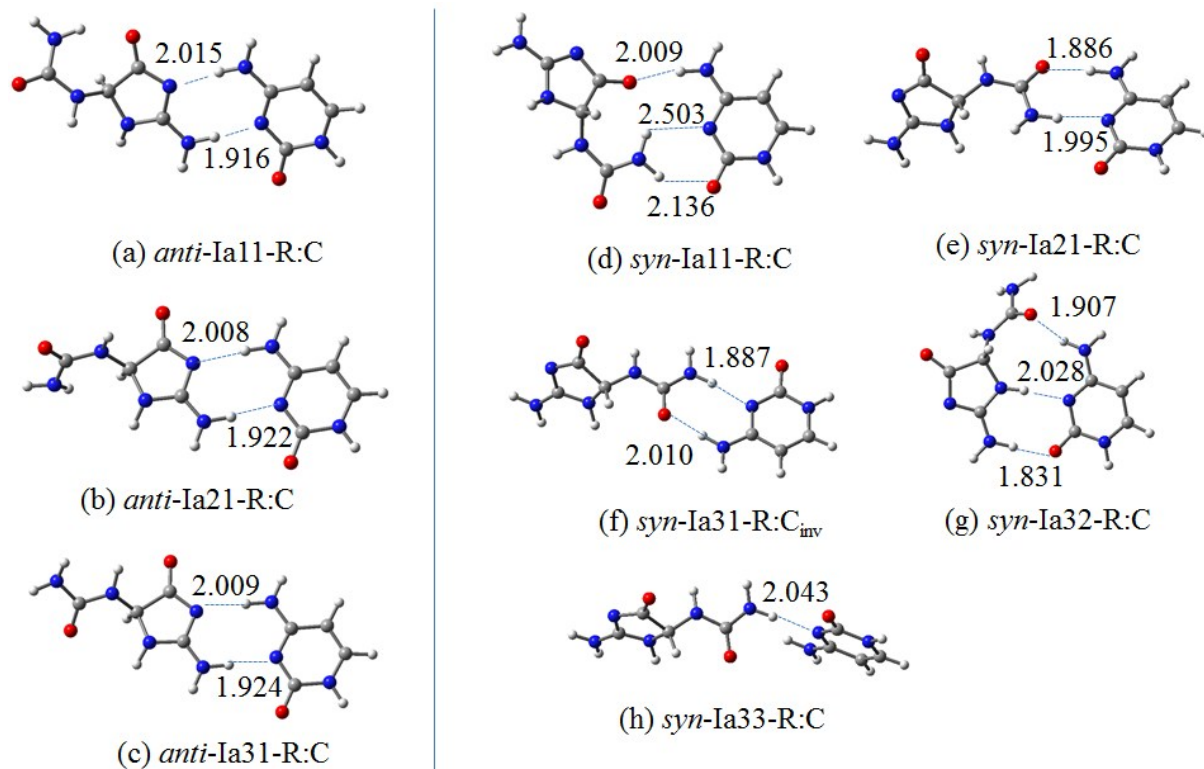


Fig. S6: Different optimized structures of Ia-R:C complexes as obtained in aqueous medium by employing B3LYP/6-31+G* level of theory. In (f) C binds with *syn*-Ia-R in the inverted orientation (C_{inv}). We noted that this complex can also be formed by binding of C with the inverted *syn*-Ia-R (*syn*-Ia-R_{inv}).

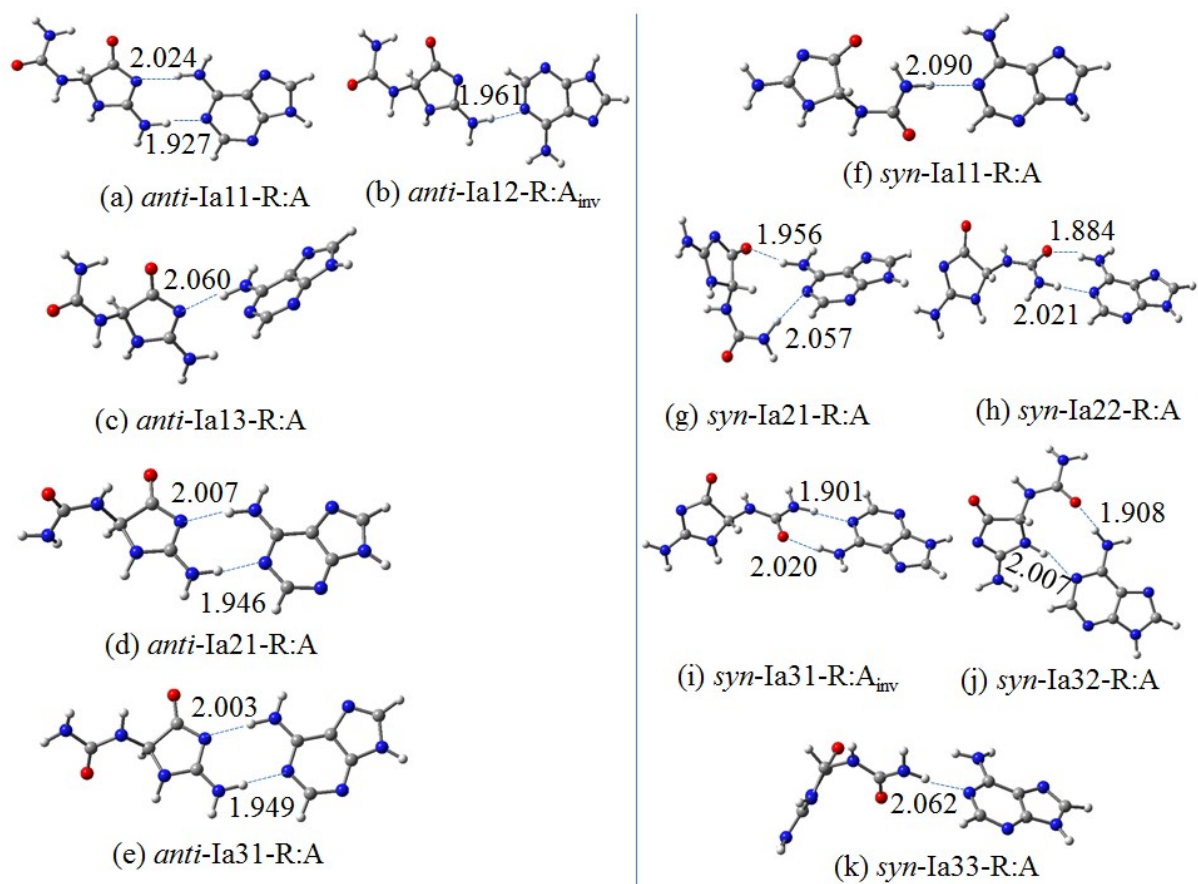


Fig. S7: Different optimized structures of Ia-R:A complexes as obtained in aqueous medium by employing B3LYP/6-31+G* level of theory. In (b) and (i) A binds with *anti*-Ia-R and *syn*-Ia-R in the inverted orientation (A_{inv}) respectively. We noted that these complexes can also be formed by binding of A with the inverted Ia-R in the *anti*- (*anti*-Ia-R_{inv}) and *syn*- (*syn*-Ia-R_{inv}) conformations respectively.

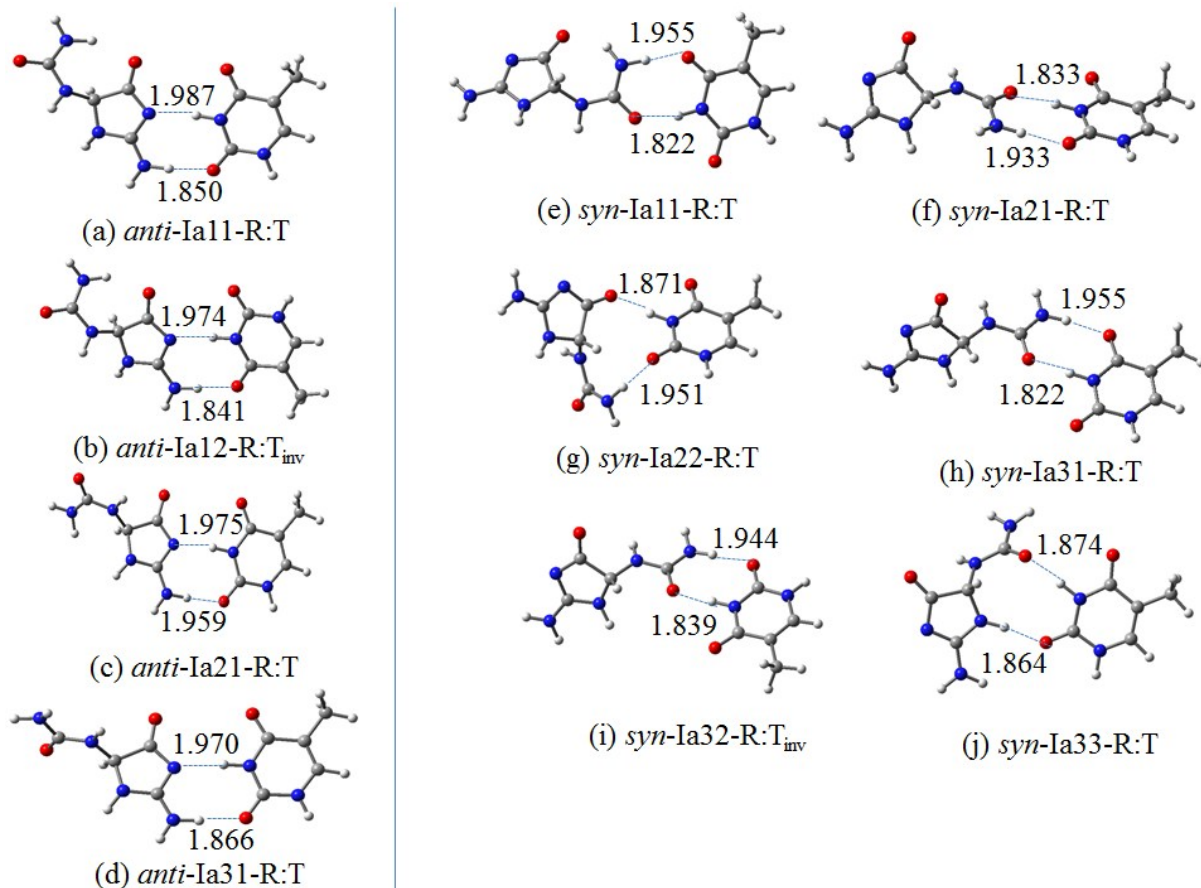


Fig. S8: Different optimized structures of Ia-R:T complexes as obtained in aqueous medium by employing B3LYP/6-31+G* level of theory. In (b) and (i) T binds with *anti*-Ia-R and *syn*-Ia-R in the inverted orientation (T_{inv}) respectively. We noted that these complexes can also be formed by binding of T with the inverted Ia-R in the *anti*- (*anti*-Ia-R_{inv}) and *syn*- (*syn*-Ia-R_{inv}) conformations respectively.

Table S2: The XYZ-coordinates of the *anti*-Ia3_{inv}-R:G complex

C	5.65010300	-0.38529400	0.15799200
N	4.33245600	-0.23493300	0.53782500
C	3.33159200	0.22429400	-0.37417500
C	1.97863200	-0.53739300	-0.26207500
O	1.86261100	-1.76554300	-0.28933100
N	0.96647800	0.35635500	-0.13253500
C	1.50855200	1.59180200	-0.11743800
N	2.86501000	1.59793700	-0.14402800
N	6.49182700	-0.83791800	1.13998900
H	3.73912300	0.11943600	-1.38383100
H	7.48042600	-0.76494700	0.93815200
H	6.24017400	-0.75682300	2.11727600
H	3.40801600	2.40707500	-0.42137500
O	6.03334300	-0.17611300	-1.00065600
N	0.74586800	2.67688400	-0.05993600
H	1.14243300	3.60589000	-0.01374800
H	-0.27424400	2.55805300	-0.02652900
H	4.06652000	-0.39544700	1.50186000
N	-5.70013800	-0.87829700	0.15672100
C	-6.15188800	0.42248300	0.24987500
C	-4.33394700	-0.82602100	0.05805000
H	-7.20577200	0.64906700	0.33743300
N	-5.17805100	1.30237200	0.21843600
C	-4.02498100	0.54013900	0.09816400
N	-3.49302000	-1.87760900	-0.05266400
C	-2.65086900	0.91999300	0.02255700
C	-2.22014500	-1.50789400	-0.12978500
N	-1.81432100	-0.19707200	-0.09334600
H	-6.26904500	-1.71615300	0.16053000
N	-1.24868200	-2.44138400	-0.28127000
H	-0.25632200	-2.21330100	-0.22109000
H	-1.51507200	-3.41149200	-0.18424100
O	-2.15303500	2.06601100	0.04942300
H	-0.79387200	0.00679300	-0.12326200

Table S3: The XYZ-coordinates of the *anti*-Ia3-R:G complex

C	4.81871700	-1.74555100	0.07470200
N	3.71986800	-0.98423900	0.41754900
C	3.33468500	0.16156200	-0.34340300
C	1.79670200	0.35789000	-0.42067100
O	1.01281100	-0.53727800	-0.75071200
N	1.47000200	1.62831900	-0.06737300
C	2.62189100	2.24893800	0.25837700
N	3.71871100	1.45297900	0.23504400
N	5.08082300	-2.80153500	0.90718000
H	3.75102100	0.05364200	-1.34904000
H	5.98187400	-3.24450400	0.78442300
H	4.70305300	-2.83306600	1.84605200
H	4.66258300	1.81405400	0.16749700
O	5.48518100	-1.52122400	-0.94437900
N	2.66405900	3.53597300	0.59227100
H	3.52102500	3.99330600	0.87404400
H	1.80743700	4.07356100	0.60419500
H	3.23866500	-1.15858700	1.29149100
N	-5.71873200	0.27322300	0.30166100
C	-6.06767400	-1.06052300	0.37196200
C	-4.36430500	0.32533800	0.09387200
H	-7.09206500	-1.36823800	0.53199800
N	-5.03793300	-1.86259100	0.22434700
C	-3.95660100	-1.01097200	0.04824600
N	-3.60895400	1.43863400	-0.03015900
C	-2.56929000	-1.29379700	-0.16036400
C	-2.32335900	1.17352200	-0.22809600
N	-1.81743800	-0.10638500	-0.29260100
H	-6.34341000	1.06556900	0.38760300
O	-2.00643200	-2.39683500	-0.23221500
N	-1.44886900	2.19376500	-0.42078200
H	-0.44191000	2.05581400	-0.27885200
H	-1.80676000	3.11696600	-0.21128700
H	-0.80934400	-0.24125700	-0.46739400

Table S4: The XYZ-coordinates of the *syn*-Ia3-R:G complex

C	1.29533400	-1.10601700	-0.14663600
N	2.64284100	-1.24842800	0.07043400
C	3.45099000	-0.11245500	0.40732700
C	4.85415000	-0.53000700	0.93789300
O	5.01340000	-1.37982800	1.81290900
N	5.82130100	0.17410300	0.28136000
C	5.18892400	0.93396000	-0.62651400
N	3.84243900	0.73885200	-0.71654600
N	0.61371200	-2.20562300	-0.54787600
H	2.92550100	0.48137500	1.16077400
H	-0.40839400	-2.15419400	-0.59669100
H	1.03328100	-3.12540400	-0.50792500
H	3.21601100	1.45250800	-1.07034600
O	0.74412500	0.00826700	0.01419400
N	5.83584700	1.80555000	-1.40111100
H	5.36863300	2.33886200	-2.12173700
H	6.83909600	1.89569400	-1.31274300
H	3.11192300	-2.09351700	-0.23199700
N	-5.92842800	0.81082400	0.22747900
C	-6.35863800	-0.44683000	-0.14408700
C	-4.55905800	0.79393400	0.19859500
H	-7.41108600	-0.68952300	-0.19893900
N	-5.36644100	-1.26613400	-0.40762400
C	-4.22488800	-0.50676400	-0.19967700
N	-3.73173000	1.82071700	0.49906300
C	-2.84163000	-0.84257600	-0.31473100
C	-2.45456500	1.49496800	0.37862900
N	-2.02115900	0.25130300	-0.00495300
H	-6.51286100	1.60001900	0.47480200
N	-1.48476400	2.42940800	0.59163900
H	-1.76523100	3.26319900	1.09169800
H	-0.53362100	2.11798700	0.75177900
O	-2.32424600	-1.93076800	-0.63652400
H	-0.99852500	0.08834600	-0.05116500

Table S5: The ZPE-corrected binding energies of different complexes involving the S-stereoisomer of Ia (Ia-S) in the *anti*- and *syn*-conformations as obtained in aqueous medium. The absence of entry shows that the corresponding complex does not exist.

Complex	Method	<i>anti</i> -Ia-S							<i>syn</i> -Ia-S							
		Ia11	Ia12	Ia13	Ia21	Ia22	Ia31	Ia32	Ia11	Ia12	Ia13	Ia21	Ia22	Ia31	Ia32	Ia33
Ia:G	B3LYP	-6.96	-4.94	-10.12	-7.02	-10.39	-7.40	-10.51	-7.63	-3.32	-5.11	-6.76	-3.11	-7.68	-5.80	-7.26
	ωB97XD ^a	-10.03	-8.26	-14.64	-11.23	-14.94	-10.56	-15.09	-11.30	-6.52	-7.45	-7.44	-6.05	-10.90	-8.39	-9.72
	ωB97XD						-10.19	-14.09						-10.27		
	B3LYP-D3						-10.81	-15.25						-11.37		
Ia:C	B3LYP	-7.21			-7.01		-7.00		-2.26			-6.09	-5.90	-5.97	-10.71	-2.48
	ωB97XD ^a	-10.69			-10.49		-10.51		-4.64			-9.08	-8.21	-8.93	-15.13	-4.81
	ωB97XD						-10.35								-14.68	
	B3LYP-D3						-10.87								-15.89	
Ia:A	B3LYP	-6.80	-3.74		-6.73		-6.77		-2.31			-5.95	-4.49	-5.75	-6.03	-2.75
	ωB97XD ^a	-9.97	-6.09		-9.82		-9.96		-4.28			-8.48	-7.97	-8.36	-9.74	-4.48
	ωB97XD						-9.21									
	B3LYP-D3						-10.02									
Ia:T	B3LYP	-6.92	-6.79		-4.88		-6.90		-6.30			-6.29	-5.71	-9.46	-6.19	-6.65
	ωB97XD ^a	-10.03	-9.95		-10.22		-10.11		-8.83			-8.71	-8.84	-11.10	-8.79	-9.67
	ωB97XD						-9.56							-8.71		
	B3LYP-D3						-10.35							-9.65		

^aObtained by single-point energy calculations by employing B3LYP/6-31+G* geometry.

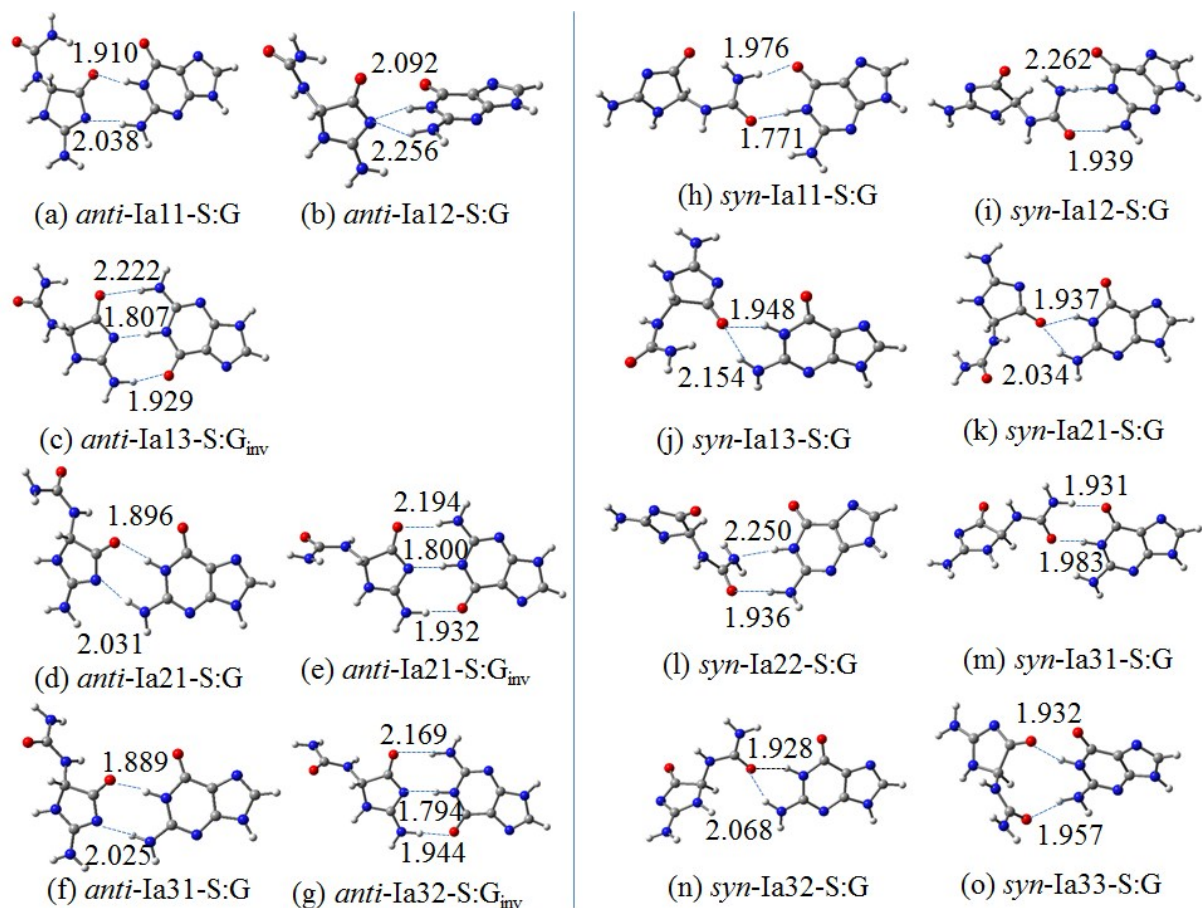


Fig. S9: Different optimized structures of Ia-S:G complexes as obtained in aqueous medium by employing B3LYP/6-31+G* level of theory. In (c), (e), and (g) G binds with *anti*-Ia-S in the inverted orientation (G_{inv}). We noted that these complexes can also be formed by binding of G with the inverted *anti*-Ia-S (*anti*-Ia-S_{inv}).

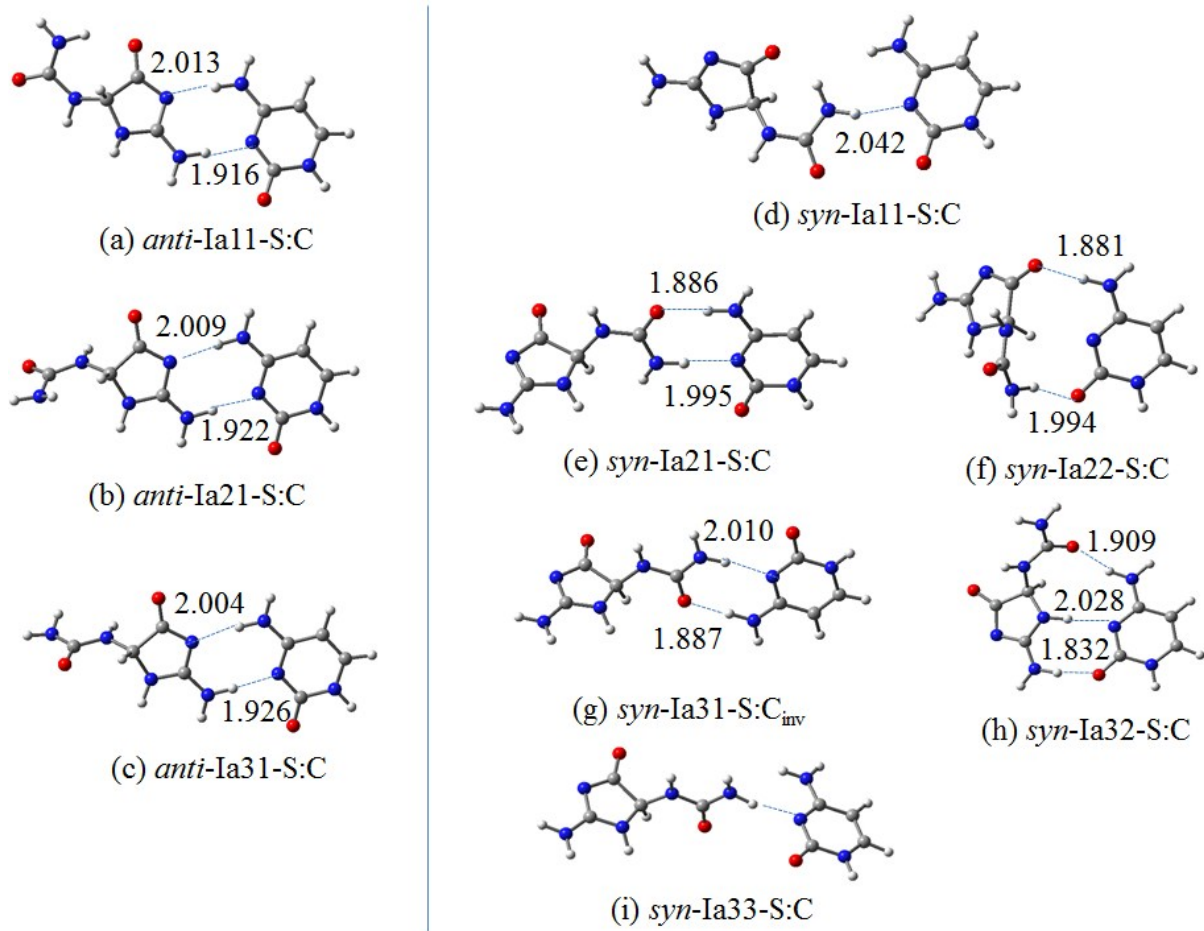


Fig. S10: Different optimized structures of Ia-S:C complexes as obtained in aqueous medium by employing B3LYP/6-31+G* level of theory. In (g), C binds with *syn*-Ia in the inverted orientation (C_{inv}). We noted that this complex can also be formed by binding of C with the inverted *syn*-Ia-S (*syn*-Ia-S_{inv}).

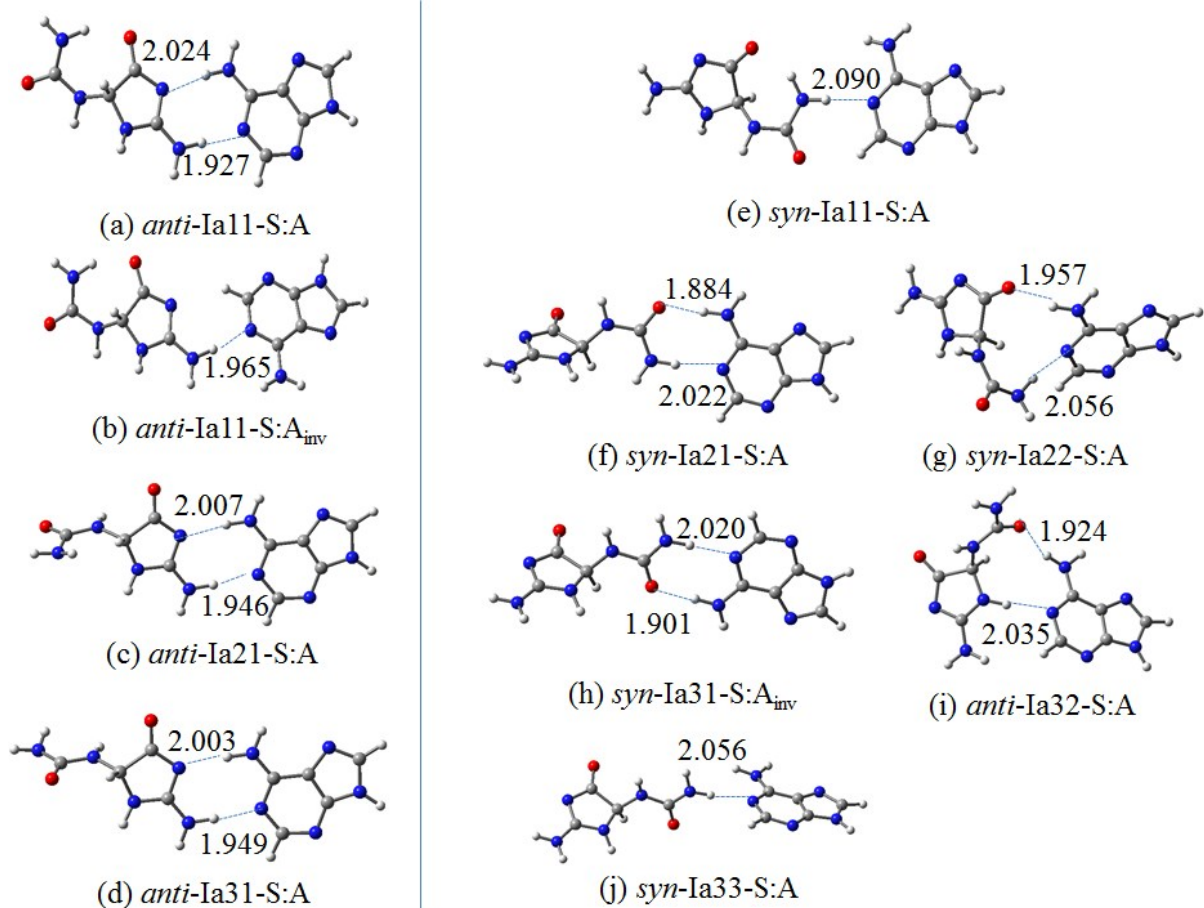


Fig. S11: Different optimized structures of Ia-S:A complexes as obtained in aqueous medium by employing B3LYP/6-31+G* level of theory. In (b) and (h), A binds with *anti*-Ia-S and *syn*-Ia-S in the inverted orientation (A_{inv}) respectively. We noted that these complexes can also be formed by binding of A with the inverted Ia-S in the *anti*- (*anti*-Ia-S_{inv}) and *syn*- (*syn*-Ia-S_{inv}) conformations respectively.

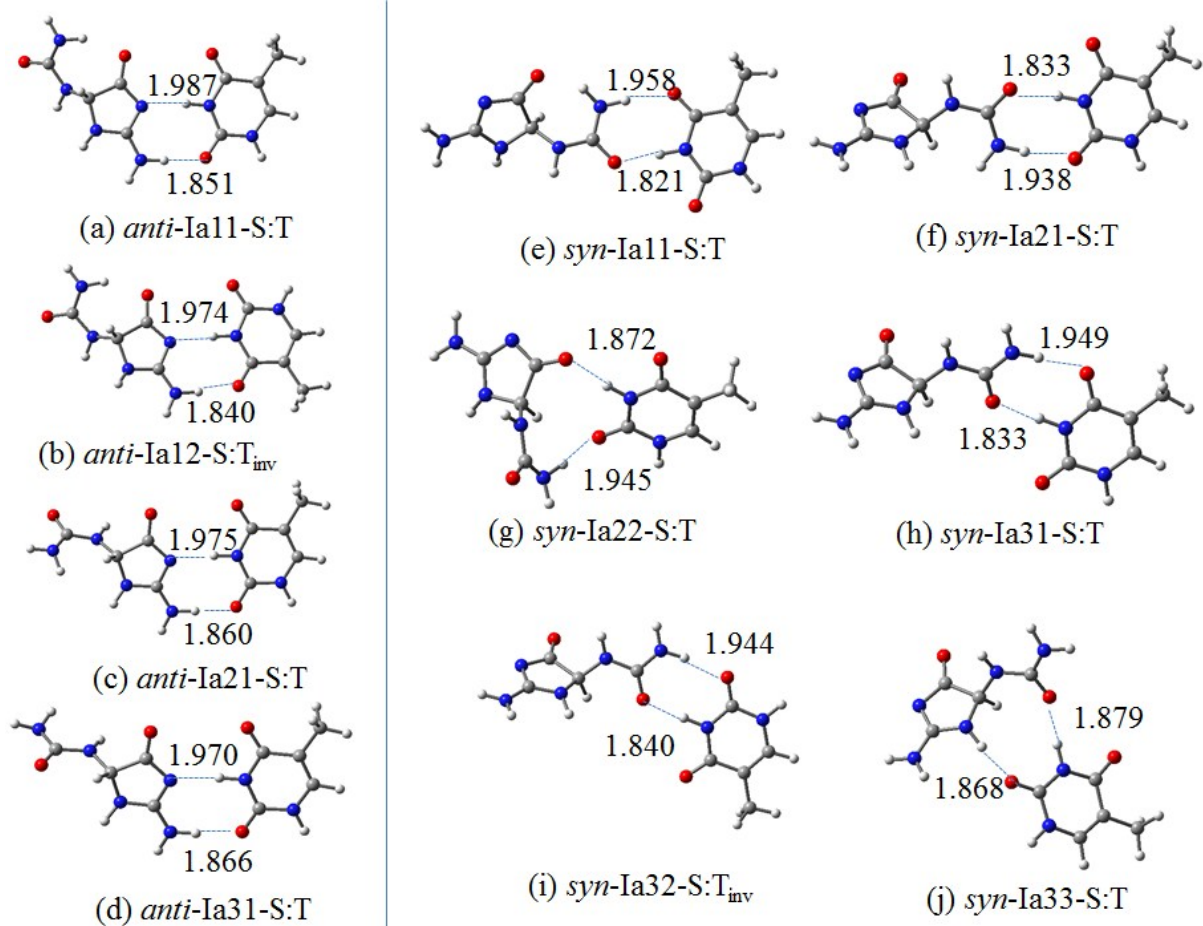


Fig. S12: Different optimized structures of Ia-S:T complexes as obtained in the aqueous medium by employing B3LYP/6-31+G* level of theory. In (b) and (i) T binds with *anti*-Ia-S and *syn*-Ia-S in the inverted orientation (T_{inv}) respectively. We noted that these complexes can also be formed by binding of T with the inverted Ia-S in the *anti*- (*anti*-Ia-S_{inv}) and *syn*- (*syn*-Ia-S_{inv}) conformations respectively.