

Università degli Studi della Calabria

Dipartimento di Chimica

Tesi di Dottorato di Ricerca in
Metodologie Chimiche Inorganiche (XX ciclo)

AREA 03 – Scienze Chimiche, SSD-CHIM03/Chimica generale ed inorganica

**Investigazione teorica delle proprietà delle unità
strutturali presenti negli acidi nucleici e delle
interazioni di queste con ioni metallici e molecole
di acqua**

Dott.ssa Donatella Mazzuca

Supervisore

Ch.ma Prof.ssa Marirosa Toscano

Coordinatore

Ch.mo Prof. Nino Russo

A.A. 2006-2007

Indice	1
Introduzione	3
CAPITOLO 1 – GLI ACIDI NUCLEICI	5
1.1 Introduzione	5
1.2 Stabilità dei nucleotidi e formazione del legame fosfodiesterico	8
1.3 Struttura degli acidi nucleici	8
1.3.1 La struttura primaria	8
1.3.2 La struttura secondaria	9
1.3.3 La struttura terziaria	12
1.4 La trascrizione e la traduzione	13
1.5 Il tautomerismo e l'accoppiamento delle basi negli acidi nucleici	14
1.6 I metalli e il DNA	16
1.6.1 I metalli alcalini	17
1.6.2 I metalli alcalino-terrosi	18
1.6.3 I metalli trivalenti	18
1.6.4 I metalli di transizione	18
1.7 L'acqua e gli acidi nucleici	19
Bibliografia	22
CAPITOLO 2 – METODI TEORICI	24
2.1 La teoria del funzionale della densità	24
2.1.1 Ricerca di minimi, massimi e punti di sella	28
2.2 La teoria perturbativa di Møller-Plesset	30
2.2.1 L'approssimazione RI	31
Bibliografia	33
CAPITOLO 3 – PRESENTAZIONE DEI RISULTATI	34
3.1 Interazione di ioni metallici con le subunità strutturali degli acidi nucleici	34
3.1.1 Interazione dello ione Al^{3+} con le basi degli acidi nucleici (U, T, C, G, A) e i nucleotidi monofosfato (UMP, dTMP, dCMP, dGMP, dAMP)	34
3.1.2 Interazione con il catione nudo: i complessi delle nucleobasi e dei nucleotidi monofosfato	35
3.1.3 Interazione con il catione idratato: complessi con le nucleobasi e i mononucleotidi	36
3.1.4 Affinità dei leganti per lo ione Al^{3+}	38

3.2 Interazione in fase gassosa dello ione Zn^{2+} con le basi puriniche e pirimidiniche del DNA e dell'RNA	39
3.2.1 Affinità delle nucleobasi per lo ione Zinco	41
3.3 Interazione in fase gassosa degli ioni Ba^{2+}, Sr^{2+}, Cs^+ e Rb^+ con le basi puriniche e pirimidiniche del DNA e dell'RNA	42
3.4 Processi di tautomerizzazione della citosina catalizzata da molecole di acqua	46
3.5 Microidratazione della coppia di base guanina-guanina	49
3.5.1 Sistemi hydrogen bonded, stacked e T-forms isolati	50
3.5.2 Sistemi hydrogen bonded, stacked e T-forms monoidratati	54
3.5.3 Sistemi hydrogen bonded, stacked e T-forms biidratati	61
Bibliografia	67
ARTICOLO 1	70
ARTICOLO 2	71
ARTICOLO 3	72

INTRODUZIONE

Lo studio sistematico delle subunità strutturali degli acidi nucleici e delle loro interazioni con ioni metallici o con molecole di acqua costituisce argomento di notevole attrattiva ormai da molto tempo sia dal punto di vista sperimentale che teorico. E' infatti noto che tali specie possono influenzare la struttura, la conformazione e quindi le funzioni del DNA e dell'RNA.

La prima evidenza del ruolo giocato dagli ioni metallici risale a circa 50 anni fa, quando in presenza di ioni mercurio furono evidenziati cambiamenti nelle proprietà chimico-fisiche del DNA. Da allora in poi, una serie di studi ha dimostrato che gli ioni metallici si legano ai gruppi fosfato dei nucleotidi, agli atomi di ossigeno dello zucchero e alle nucleobasi in relazione alla loro concentrazione. Questo evento ha come conseguenza la stabilizzazione o la destabilizzazione del DNA. L'effetto più immediato dei metalli è quello di neutralizzare la carica negativa sul gruppo fosfato attraverso interazioni elettrostatiche non specifiche che stabilizzano la doppia elica. Al contrario, la principale fonte di destabilizzazione è dovuta alla interazione con le nucleobasi, in seguito alla quale si verifica la separazione dei due filamenti che comporta la rottura o la modifica dei legami a idrogeno tra le coppie di basi.

La potenzialità di alcuni metalli a disturbare i processi di replicazione può collegarsi alla loro abilità di stabilizzare i tautomeri delle basi del DNA incompatibilmente con la formazione delle coppie di base di Watson-Crick. Con l'aggiunta del metallo, l'acidità dei gruppi funzionali contenenti idrogeno delle basi è alterata e questo si traduce in un cambiamento dell'energetica della tautomerizzazione.

L'interazione degli acidi nucleici o dei loro costituenti con molecole di acqua è stata oggetto di studi più recenti ma sta diventando un tema sempre più attraente. Grazie alle sue particolari proprietà solventi, l'acqua costituisce un mezzo intracellulare ed extracellulare universale. Questa proprietà è dovuta alla sua tendenza a formare legami idrogeno e al suo carattere dipolare. La descrizione dell'effetto solvente in questi sistemi altamente polari riveste un ruolo dominante sulla struttura e sulle funzioni che essi svolgono e può cambiare la stabilità dei diversi tautomeri.

La micro-idratazione diventa particolarmente importante quando si considerano le barriere di interconversione tra i vari tautomeri come elemento di prova o diniego della loro coesistenza, in alcuni casi verificata in via sperimentale.

L'idea di studiare il comportamento delle unità costituenti gli acidi nucleici e le loro interazioni con ioni metallici e/o con molecole di acqua a livello teorico ha lo scopo di fornire alcuni descrittori per l'interazione che grazie al loro contenuto di informazioni possono essere d'aiuto alla razionalizzazione delle risposte biologiche osservate. Infatti, l'ammontare dei dati accumulati finora sulla chimica degli acidi nucleici, in seguito a studi di tipo sperimentale, ha mostrato che tutti i metodi di analisi impiegati sono soggetti a limitazioni. Così, i risultati acquisiti per via teorica possono costituire un complemento significativo per le misure offrendo una comprensione meccanicistica ad un livello di dettaglio molecolare.

L'aumentata disponibilità di computer veloci, unitamente al miglioramento delle metodologie computerizzate, permettono oggi l'indagine di caratteristiche strutturali e funzionali delle biomolecole, consentendo ai ricercatori di fare un passo avanti nella comprensione dei meccanismi alla base delle loro interazioni con le più svariate specie.

Nel presente lavoro di tesi, come verrà ampiamente illustrato di seguito, sono state considerati a livello teorico alcuni casi di interazione fra le più importanti subunità strutturali degli acidi nucleici con diversi ioni metallici e con l'acqua. La nostra opinione è che la razionalizzazione del ruolo effettivo della metallazione o dell'idratazione nella biofisica degli acidi nucleici, debba partire dall'investigazione degli effetti che questi processi hanno sulle unità minimali e isolate degli stessi.

In particolare, la ricerca ha riguardato:

- L'interazione delle nucleobasi e dei nucleotidi monofosfato con lo ione Al^{3+} nudo e idratato;
- L'interazione delle nucleobasi con alcuni cationi dei metalli alcalini e alcalino terrosi e con lo zinco;
- Il processo di tautomerizzazione della citosina in presenza di molecole di acqua esplicite e non;
- L'interazione della coppia di base rara guanina-guanina in presenza di una e due molecole d'acqua.

CAPITOLO 1

GLI ACIDI NUCLEICI

1.1 Introduzione

Gli acidi nucleici sono i costituenti più importanti e fondamentali delle cellule viventi. È probabile che la vita stessa abbia iniziato la sua evoluzione dagli acidi nucleici poiché soltanto questi, tra i composti biologici, possiedono potenzialità per l'autoduplicazione [1].

Il ruolo degli acidi nucleici è quello di custodire e trasmettere l'informazione genetica nelle cellule, nei tessuti e negli organismi. Il progetto completo di un organismo è codificato nel suo acido nucleico.

Gli acidi nucleici furono scoperti nel 1868 quando Miescher riuscì ad estrarre dai nuclei delle cellule di pus una sostanza che chiamò "nucleina". In seguito, a tale sostanza fu dato il nome di acido deossiribonucleico. L'importanza biologica degli acidi nucleici fu messa in luce nel 1944 da Avery [2] che stabilì che l'acido deossiribonucleico è la sostanza che trasmette le caratteristiche ereditarie.

In tutti i viventi esistono due tipi di acidi nucleici: il deossiribonucleico o DNA e il ribonucleico o RNA [3,4]. Il primo è tipicamente nucleare e costituisce il materiale genetico dei cromosomi; è legato a proteine basiche dette istoni formando le nucleoproteine dei cromosomi. L'RNA, presente nel nucleolo ed in certe parti dei cromosomi, è caratteristico del citoplasma dove interviene nella sintesi delle proteine. Ciascuno di essi è una catena polimerica con unità monomeriche simili unite da legami covalenti. In entrambi i casi, l'unità monomerica contiene uno zucchero a cinque atomi di carbonio, il deossiribosio per il DNA e il ribosio per l'RNA. La differenza tra i due zuccheri risiede unicamente nel gruppo 2' ossidrilico del ribosio (fig. 1.1).

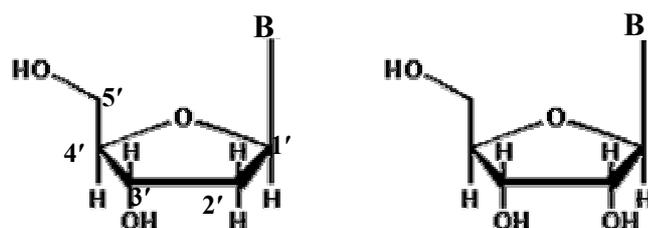


Figura 1.1

Negli acidi nucleici, il collegamento tra unità monomeriche successive è realizzato attraverso un residuo fosfato attaccato all'ossidrile sul carbonio 5' di una unità e un gruppo 3' ossidrilico dell'unità successiva. Si forma un legame fosfodiesterico tra i due residui (fig. 1.2).

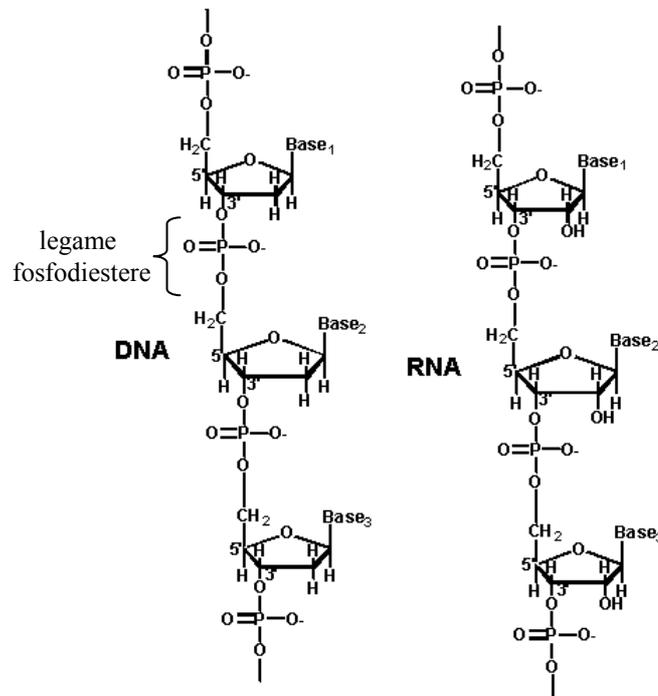


Figura 1.2

Il gruppo fosforico è un acido forte ($pK_a \sim 1$); questo è il motivo per cui DNA e RNA sono denominati acidi nucleici. A pH fisiologico, ogni residuo di una catena di DNA o RNA porta una carica netta negativa.

I residui di zucchero, uniti dai legami fosfodiesterici, formano lo scheletro della molecola dell'acido nucleico. Di per se stesso, lo scheletro è una struttura ripetitiva, incapace di codificare l'informazione genetica. L'importanza degli acidi nucleici nell'immagazzinamento e nella trasmissione dell'informazione genetica deriva dal fatto che essi sono eteropolimeri. Ciascun monomero della catena porta un gruppo basico, sempre attaccato al carbonio 1' dello zucchero (vedi fig. 1.1).

Nel DNA si trovano le basi puriniche adenina e guanina e le basi pirimidiniche citosina e timina; l'RNA [5] contiene le stesse basi, ad eccezione della timina sostituita dall'uracile. Le strutture delle basi che si trovano negli acidi nucleici sono mostrate in figura 1.3.

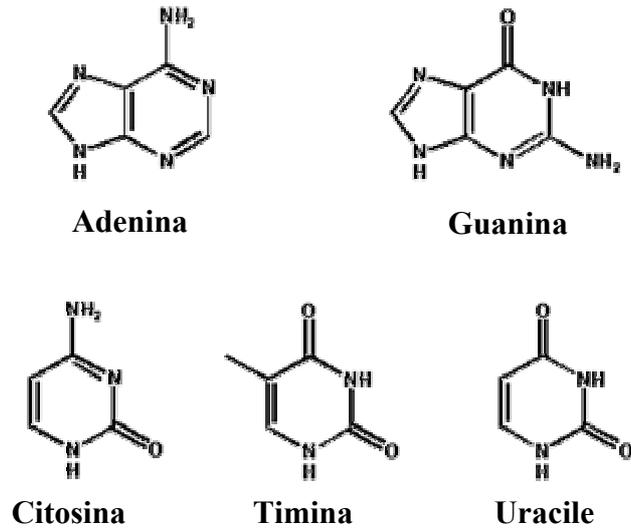


Figura 1.3

I monomeri sopra citati sono detti nucleotidi. Ciascun nucleotide è un derivato 5' monofosforilato di un addotto zucchero-base denominato nucleoside (fig.1.4).

I nucleotidi sono acidi molto forti; la ionizzazione primaria del gruppo fosfato avviene con un pK_a di circa 1. Sia la ionizzazione secondaria del fosfato che la protonazione di alcuni gruppi delle basi nei nucleotidi si osservano a valori di pH molto vicini alla neutralità.

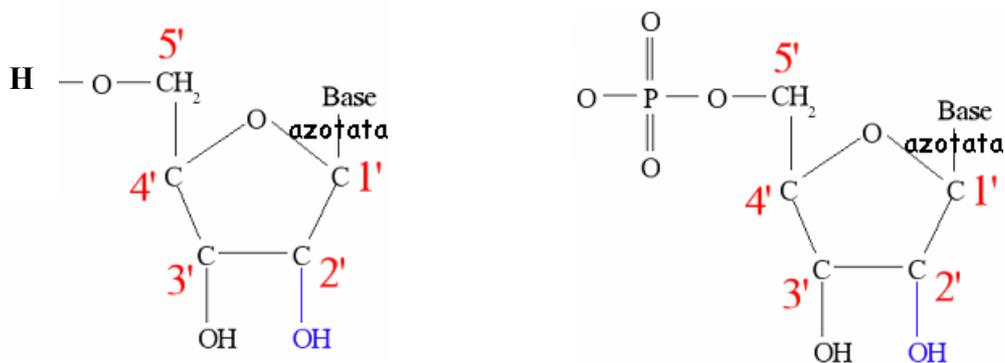


Figura 1.4

1.2 Stabilità dei nucleotidi e formazione del legame fosfodiesterico

Poiché gli acidi nucleici possono essere considerati polimeri di nucleotidi, essi vengono spesso chiamati con il nome generico di polinucleotidi.

In linea di principio, ogni polinucleotide viene prodotto a partire dai suoi monomeri nucleotidici per eliminazione di una molecola di acqua per ciascuna coppia di monomeri. La reazione di disidratazione appena descritta è, però, piuttosto positiva, circa +25 kJ/mol; l'equilibrio in un ambiente acquoso è molto spostato verso l'idrolisi [1].

Sebbene i polinucleotidi siano in vivo termodinamicamente instabili, la loro idrolisi è estremamente lenta quando non è catalizzata. Questa caratteristica è della massima importanza in quanto assicura che il DNA nelle cellule sia sufficientemente stabile per fungere da depositario dell'informazione genetica.

Infatti, il DNA è così stabile che ne sono stati recuperati frammenti da alcuni fossili.

Viene spontaneo chiedersi: come vengono allora sintetizzati i polinucleotidi? La loro sintesi coinvolge i nucleosidi trifosfati ad alto contenuto energetico. Nelle cellule viventi, il nucleoside monofosfato da aggiungere alla catena nascente viene inserito come nucleoside trifosfato (ATP o deossiATP) mentre nella reazione viene rilasciato il pirofosfato. La variazione di energia libera per questa reazione, considerata come la somma di due reazioni, l'idrolisi di un nucleoside trifosfato e la formazione di un legame fosfodiesterico per eliminazione di una molecola di acqua, è negativa ($\Delta G = -6$ kJ/mol).

Inoltre, la reazione è ulteriormente favorita dall'idrolisi del pirofosfato a fosfato con $\Delta G = -33$ kJ/mol.

Per riassumere, in questa reazione ogni monomero viene presentato come nucleoside trifosfato per essere aggiunto alla catena. La rottura di tale nucleoside fornisce l'energia libera che rende la reazione termodinamicamente favorita.

1.3 Struttura degli acidi nucleici

1.3.1 La struttura primaria

Come evidente dalla figura 1.2, una catena polinucleotidica possiede un verso; in base ad esso, il legame fosfodiesterico è localizzato tra il carbonio 3' di un monomero e il carbonio 5' di quello successivo. È così possibile distinguere le due

estremità della catena polinucleotidica; un'estremità porterà, infatti, un gruppo 5' fosfato libero mentre l'altra porterà un gruppo 3' ossidrilico libero.

Inoltre, una catena polinucleotidica ha una propria individualità, determinata dalla sequenza nucleotidica. Tale sequenza rappresenta la struttura primaria di quel particolare acido nucleico. Per convenzione, le sequenze sono scritte in modo che l'estremità 5' sia sempre a sinistra e quella 3' sempre a destra.

La struttura primaria del DNA è di fondamentale importanza perché in essa viene immagazzinata l'informazione genetica.

Il DNA, nelle sue lunghe catene di nucleotidi, conserva l'informazione genetica e, così, i geni sono porzioni di DNA cellulare che nel suo insieme costituisce il genoma di una cellula.

Solo negli ultimi anni del XIX secolo, alcuni scienziati cominciarono a sospettare che il DNA potesse essere il materiale genetico. Gli studi mostravano che il DNA conteneva solo quattro tipi di monomeri, strutture troppo semplici che portavano ad escludere il fatto che esso potesse avere un ruolo tanto complesso.

Nella seconda metà del XX secolo, tra gli anni '40 e '60, si comprese prima che l'informazione genetica è contenuta nel DNA poi si decifrò il modo in cui l'informazione scritta nella molecola di DNA viene ad essere espressa.

L'espressione genica è il primo momento di quel processo che porta la potenzialità insita nell'informazione genetica depositata nel genoma a manifestarsi nelle forme e funzioni caratteristiche di ogni organismo vivente.

La comprensione della struttura del DNA permise di capire non solo come l'informazione genetica sia trasmessa fedelmente da una generazione alla successiva ma come essa possa variare. Gli studi di Watson-Crick e altri [2-5] sulla struttura tridimensionale del DNA consentirono di conoscere come il materiale genetico fosse trasmesso alla cellula e come potesse replicarsi durante la divisione cellulare.

1.3.2 La struttura secondaria

La struttura secondaria del DNA è la ben nota doppia elica. Fu dedotta dall'analisi degli spettri di diffrazione di fibre di DNA. Gli spettri di diffrazione delle fibre non contenevano le informazioni molecolari trovate negli spettri di diffrazione del singolo cristallo, ciononostante sia gli spettri di diffrazione che le

proprietà delle fibre fornirono informazioni essenziali per dedurre la struttura della doppia elica:

- gli spettri mostravano la presenza di un'elica
- la spaziatura tra le macchie mostrava che c'erano 10 basi per giro d'elica
- la densità delle fibre indicava la presenza di due molecole per elica.

Watson e Crick proposero che la doppia elica fosse stabilizzata dalla formazione di legami a idrogeno tra le basi presenti sui filamenti opposti con lo schema A-T e G-C. Secondo il loro modello, lo scheletro idrofilo costituito dalla sequenza zucchero-fosfato si trovava all'esterno a contatto con l'acqua, mentre le basi di carattere idrofobico erano impacchettate nell'interno perpendicolarmente all'asse dell'elica.

I due filamenti di DNA sono complementari e si sviluppano in direzione opposta, risultando così antiparalleli.

Sapere che i due filamenti di DNA sono complementari e che A si appaia sempre con T e G sempre con C è di fondamentale importanza.

Nell'articolo in cui presentarono il loro modello, Watson e Crick espressero la loro idea con queste parole: "non è sfuggito alla nostra attenzione che la specificità dell'appaiamento che abbiamo postulato suggerisce immediatamente un possibile meccanismo di replicazione del materiale genetico". La loro idea può essere considerata la previsione più ricca di conseguenze che sia stata fatta sino ad oggi.

Infatti, se i filamenti potessero essere separati e si potesse sintetizzare del nuovo DNA lungo ciascuno di essi, si otterrebbero due molecole di DNA a doppio filamento ciascuna delle quali sarebbe una copia esatta dell'originale. Questa è la proprietà di autoreplicazione che il materiale genetico deve possedere: quando una cellula si divide, devono essere prodotte due copie complete dell'informazione genetica portata dalla cellula originale. La replicazione [6,7] completa di una molecola di DNA produce due molecole a doppio filamento ognuna delle quali presenta per metà DNA parentale e per metà nuovo materiale. Questo tipo di replicazione è detto semiconservativo perché metà del materiale originale è conservato in ognuna delle due copie. Esiste, inoltre, una replicazione conservativa in cui una delle due eliche è quella parentale conservata, mentre l'altra è sintetizzata ex novo ed una replicazione dispersiva in cui il materiale parentale è trasmesso alle molecole figlie senza un ordine preciso.

Benché le basi si trovino all'interno, esse possono interagire con altre molecole attraverso le due spirali profonde, dette solco maggiore e solco minore [1].

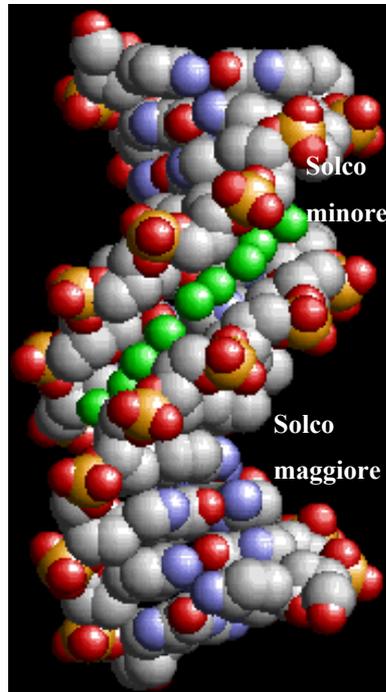


Figura 1.5

La formazione dell'elica è favorita da due fattori dominanti: le interazioni idrofobiche derivanti dall'impaccamento delle basi e dalla formazione di legami idrogeno tra le coppie di basi. Tre sono, invece, i fattori destabilizzanti: la repulsione elettrostatica dei fosfati che si trovano sullo scheletro dei filamenti, l'entropia conformazionale e la formazione di legami idrogeno con l'acqua.

Il modello di doppia elica proposto da Watson e Crick rappresenta la struttura del B-DNA, la forma più frequente nel mezzo acquoso. Esiste, in condizioni di bassa umidità una struttura diversa detta A-DNA [8].

Da un punto di vista biologico, l'elica A è molto importante, basti pensare che le molecole di RNA a doppio filamento si trovano sempre nella forma A così come le molecole ibride DNA-RNA che si formano mediante appaiamento di un filamento di DNA con uno di RNA. Questo è dovuto al fatto che nell'RNA il ribosio porta un gruppo ossidrilico in più rispetto al deossiribosio. Questo ossidrile interferisce stericamente con la forma B per la sua vicinanza al gruppo fosfato e al C8 della base adiacente. Da tutto ciò consegue che l'RNA non può assumere la forma B nemmeno in condizioni in cui l'idratazione la favorirebbe.

Nel DNA, il gruppo ossidrilico lascia il posto ad un solo atomo di idrogeno per cui l'impedimento sterico non esiste.

Le forme A e B sono molto diverse sebbene entrambe siano eliche destrorse.

Nella prima, le basi giacciono vicine all'asse dell'elica e questo asse passa tra i legami idrogeno; nell'elica A le basi si trovano più verso l'esterno e sono molto inclinate rispetto all'asse dell'elica. Inoltre, nell'elica B il solco maggiore e il solco minore sono ben distinti, mentre nell'elica A essi sono molto simili in termini di profondità.

Gli studi cristallografici danno una spiegazione del perché in acqua l'elica B [9] sia favorita. Nella forma B del DNA esiste uno spazio per una fila di molecole di acqua (cerchietti verdi in fig. 1.5) che si dispongono nel solco minore. I legami idrogeno tra queste molecole d'acqua e il DNA possono conferire stabilità alla forma B.

Recentemente, è stato trovato che alcune sequenze di acidi nucleici possono assumere una conformazione elicoidale sinistra a doppia elica che per l'andamento a zig-zag della catena fosfodiesterica è stata chiamata forma Z. La forma Z porta i gruppi fosfato negativi a distanze corte e, quindi, l'elica sinistra è stabile in condizioni di altissima forza ionica certamente non fisiologiche.

L'interesse per la forma Z consiste nel fatto che, avendo un avvolgimento contrario a quello del B-DNA, la coesistenza di forme B e Z lungo la stessa polinucleotidica rende topologicamente più semplice lo svolgimento della doppia elica in processi biochimicamente importanti come la replicazione cellulare o l'espressione genica.

1.3.3 La struttura terziaria

Una struttura tridimensionale che comporta un ripiegamento di ordine superiore di elementi della struttura secondaria è detta struttura terziaria di un polimero. La maggior parte delle molecole di DNA che si trovano in natura sono circolari. Un aspetto particolare di tali molecole è di essere superavvolte [10-12].

Oltre all'avvolgimento dei filamenti di DNA uno sull'altro, una molecola superavvolta presenta ulteriori avvolgimenti rispetto all'asse dell'elica che incrocia una o più volte. Il superavvolgimento non è raro tant'è che lo è la maggior parte delle molecole di DNA circolari chiuse. In natura, le strutture di DNA che si trovano più frequentemente sono supereliche sinistrorse ma si trovano

anche molecole di DNA con avvolgimenti destrorsi. Quando si introducono dei superavvolgimenti in una molecola di DNA allo stato rilassato, si introduce una distorsione nella molecola. Per far ciò è necessario spendere dell'energia.

Un caso più complesso di struttura terziaria si trova nell'RNA transfer (tRNA), un tipo di RNA impegnato nella sintesi proteica. Per esso, si trova non solo una struttura secondaria di tipo A data dal ripiegamento della catena su se stessa ma ancora un ripiegamento più complesso delle eliche su se stesse. La molecola di tRNA ha così una struttura terziaria ben determinata che le dà una forma definita e un'organizzazione interna necessarie per svolgere le sue funzioni.

1.4 La trascrizione e la traduzione

L'espressione dell'informazione genetica presente nel DNA richiede la trascrizione della sequenza di DNA in una sequenza di RNA messaggero, mRNA.

Nella formazione di mRNA è coinvolto l'accoppiamento delle basi complementari tra i ribonucleotidi e la sequenza del DNA.

La trascrizione [13,14] avviene tramite particolari enzimi, chiamati RNA polimerasi [15-17]. Questi enzimi utilizzano nucleosidi trifosfati per formare l'RNA.

L'RNA polimerasi si lega al DNA tramite particolari sequenze, dette promotori, che non sono trascritte. Dal promotore iniziano ad inserirsi i nucleosidi trifosfato per formare una sequenza di nucleotidi che sarà complementare al filamento di DNA in via di trascrizione. Dopo l'individuazione del promotore, la RNA polimerasi rende il DNA adatto alla trascrizione. Il filamento di RNA inizia così ad allungarsi attraverso l'aggiunta di un nucleotide per volta. La trascrizione termina quando nel DNA si incontrano particolari sequenze di basi, poste alla fine di ogni gene.

La traduzione [18] (fig. 1.6), o sintesi proteica, rappresenta la seconda fase del processo di espressione genica, processo in cui l'informazione contenuta nel DNA viene convertita in proteine che nella cellula svolgono diverse funzioni. Nella sintesi proteica un filamento di mRNA [19] è usato come stampo per la produzione di una specifica proteina.

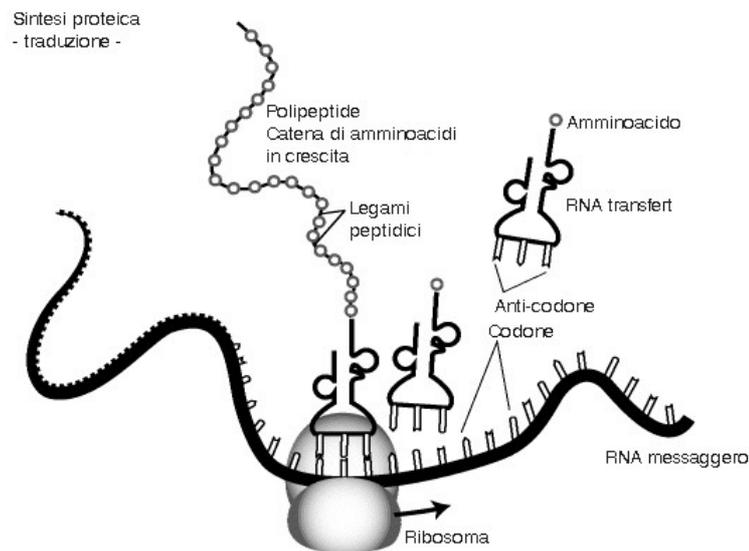


Figura 1.6

L'informazione codificante le proteine viene letta dalla cellula in blocchi di tre residui nucleotidici, codoni, ognuno dei quali specifica un diverso amminoacido. Per convertire l'informazione dalla sequenza genica del DNA alla sequenza amminoacidica della proteina sono utilizzati come intermediari particolari molecole di RNA. Copie complementari di geni che devono essere espressi sono trascritte dal DNA sotto forma di molecole di RNA messaggero, che ha il compito di portare l'informazione dal DNA all'apparato cellulare per la sintesi delle proteine. Nell'apparato si trovano molecole di tRNA, particolari enzimi, e i ribosomi complessi RNA-proteine sui quali avviene la sintesi di nuove proteine il principio base della traduzione è il seguente: una molecola di RNA messaggero si lega a un ribosoma e le molecole di tRNA portano gli amminoacidi al ribosoma, uno alla volta. Ciascun tRNA identifica il codone appropriato sull'RNA messaggero e aggiunge il proprio amminoacido alla catena proteica nascente. Il ribosoma scorre lungo la molecola di mRNA in modo che tutto il messaggio genetico possa essere letto e tradotto in proteina.

1.5 Il tautomerismo e l'accoppiamento delle basi negli acidi nucleici.

A seconda della diversa posizione degli atomi di idrogeno su una base, le basi degli acidi nucleici danno origine al tautomerismo. I tautomeri rari possono essere coinvolti in diversi processi biochimici e la loro presenza può portare ad un accoppiamento anomalo delle basi da cui può scaturire una mutazione genetica.

Tuttavia, evidenze dirette della loro presenza nelle biomolecole sono piuttosto rare e le basi degli acidi nucleici sono presenti in maniera dominante nella forma canonica più stabile.

Nei precedenti paragrafi, abbiamo analizzato in che modo il DNA memorizza e trasferisce l'informazione genetica. La grande fedeltà della replicazione del DNA si deve all'alta specificità dell'accoppiamento delle basi: la guanina lega la citosina e l'adenina si lega alla timina o all'uracile nell'RNA.

L'alta specificità viene assicurata dai legami idrogeno tra le basi. La coppia GC possiede tre forti legami idrogeno e la coppia AT/U ne possiede due. Queste sono le strutture planari di Watson-Crick.

Oltre all'accoppiamento delle basi proposto da Watson e Crick, esistono altre coppie di basi che contribuiscono alla variabilità conformazionale del DNA chiamate di Hoogsteen, di Hoogsteen invertite e di Watson-Crick invertite.

Le interazioni fra le coppie di base giocano un ruolo fondamentale nella stabilizzazione della struttura secondaria degli acidi nucleici.

Nelle coppie di base esistono due tipi di interazioni, planari e verticali, che portano alla formazione di strutture legate tramite legami idrogeno e stacked. La natura delle due interazioni è diversa. Le prime sono dominate da contributi energetici coulombiani che si originano dal carattere altamente dipolare delle basi degli acidi nucleici. Il legame idrogeno è altamente specifico e direzionale e questo tipo di interazione è responsabile della stabilizzazione di biomacromolecole ma anche di fenomeni importanti come il trasferimento del codice genetico attraverso la formazione di coppie di basi complementari.

Le interazioni di stacking sono non specifiche e sono originate principalmente da forze di dispersione di London dovute alla grande polarizzabilità delle basi. Per molto tempo tali interazioni sono state considerate piuttosto deboli e solo recentemente è stato dimostrato quanto, in particolari condizioni, siano molto simili in energia alle interazioni date dai legami idrogeno.

1.6 I metalli e il Dna

Una delle più importanti interazioni nei sistemi biologici è quella tra i cationi metallici e gli acidi nucleici. Essa gioca un ruolo chiave nella determinazione della struttura e delle funzioni di queste macromolecole [20-23]. Per esempio, il processo di fusione del DNA manifesta una considerevole sensibilità alla presenza di cationi metallici sia divalenti sia polivalenti. L'abilità legante dei diversi cationi metallici verso gli acidi nucleici può essere valutata attraverso la temperatura di fusione del DNA [24,25].

La diversa selettività dei cationi metallici verso gli acidi nucleici è una funzione della carica, della stechiometria, della struttura elettronica oltre che della loro naturale abbondanza. Essi hanno un comportamento simile a quello degli ioni idrogeno ma mentre i primi possono esibire valori alti di concentrazioni anche a pH neutro, la concentrazione degli ioni idrogeno in soluzioni neutre è bassa per definizione [26].

Spesso possiedono cariche positive maggiori di uno e hanno volumi ionici così grandi da poter ospitare molti leganti intorno a loro.

Il numero di coordinazione di uno ione metallico, cioè, il numero di atomi leganti intorno ad esso, va visto in termini di sfere concentriche: la sfera interna contiene quegli atomi a diretto contatto con lo ione metallico, la sfera esterna contiene quelli a contatto con gli atomi leganti la prima sfera. Il numero di atomi in queste sfere dipende dalle dimensioni dello ione metallico e da quelle degli atomi leganti [27,28].

Un esempio della diversa reattività dei cationi metallici è data dalla loro abilità di legare o perdere molecole di acqua. La velocità con cui avviene lo scambio delle molecole di acqua coordinate direttamente al catione con le molecole di acqua del solvente di "bulk" permette la classificazione dei cationi in quattro gruppi [29]. In linea del tutto generale tale velocità è legata al rapporto carica-raggio: se questo rapporto è alto allora la velocità di scambio è lenta.

Gli ioni metallici possono legarsi al DNA direttamente o indirettamente tramite legami a ponte idrogeno che instaurano grazie alle molecole di acqua di coordinazione. Studi spettroscopici hanno mostrato che l'attacco degli ioni metallici carichi positivamente avviene sui siti caratterizzati da alta densità elettronica o su residui carichi negativamente. Tali siti possono essere i gruppi fosfato dello scheletro delle due catene e gli atomi elettrone-donatori sulle basi

come l' N7 e l' O6 della guanina, l' N7 e l' N6 per l'adenina e l'atomo N3 per le pirimidine [30].

È noto che i metalli che si legano sulla base distruggono i legami a ponte idrogeno delle coppie di base e destabilizzano la doppia elica. Al contrario, gli ioni metallici neutralizzano l'eccesso di carica negativa dei gruppi fosfato e questo si traduce in una stabilizzazione dell'elica.

Il legame degli ioni metallici al DNA e all'RNA può influenzare indirettamente la conformazione dello zucchero con conseguenti anomalie nei processi di sintesi e replicazione.

Gli ioni metallici rivestono un ruolo fondamentale all'interno del corpo umano. Quando presenti in tracce, risultano essenziali per diverse reazioni cellulari e per svariate funzioni metaboliche e fisiologiche comportandosi da catalizzatori. Le deviazioni dai normali livelli di concentrazione possono invece portare a disfunzioni o a malattie.

1.6.1 I metalli alcalini

I metalli alcalini sono estremamente reattivi e in natura esistono solo come cationi. Per quanto riguarda il loro legame con gli acidi nucleici, essi interagiscono per lo più con il gruppo fosfato, neutralizzando le cariche negative e stabilizzando la doppia elica, ma le loro interazioni con le basi comportano, comunque, il medesimo effetto [30-34].

L'interazione di uno ione alcalino specifico con gli acidi nucleici è controllato dalla forza del legame tra lo ione metallico e i possibili centri donatori sulle basi. I siti conosciuti per la coordinazione sono principalmente gli atomi N7, N9 e O6 per la guanina e l'adenina, l'atomo O4 per la timina e l'uracile e gli atomi N3 e O2 per la citosina, quando essi non siano coinvolti in legami idrogeno con molecole di acqua.

I metalli monovalenti preferiscono interagire nella regione ricca di AT del solco minore del DNA ma per far ciò, di solito rilasciano le molecole di acqua coordinate e interagiscono direttamente con le basi. Dati cristallografici trovano che lo ione Rb^+ preferisce legarsi nella sfera interna di coordinazione sugli atomi di ossigeno O2 delle timine dei due filamenti opposti mentre per lo ione Na^+ , simulazioni di dinamica molecolare suggeriscono che la geometria di coordinazione nel solco minore dipende dalla posizione del catione [35-38].

1.6.2 I metalli alcalino-terrosi

I metalli alcalino-terrosi sono più reattivi verso il DNA rispetto a quelli alcalini e possono esibire mono- e bi-coordinazione. Tra gli ioni metallici di questo gruppo, il magnesio è quello più presente nei vari processi di attivazione del DNA e dell'RNA [30].

I cationi bivalenti, solitamente, interagiscono con il gruppo fosfato ma sono favoriti anche i siti che si trovano sulle basi. La doppia carica sul metallo può indurre trasferimento protonico tra i siti N1 e N3 nella coppia di base di Watson-Crick guanina-citosina, può favorire la formazione di tautomeri rari e gli accoppiamenti anomali, può inoltre interagire con il sistema π impedendo l'accoppiamento delle basi [39-41].

I metalli bivalenti stabilizzano le strutture di DNA e RNA e regolano le transizioni tra le forme A e B della struttura secondaria del DNA.

1.6.3 I metalli trivalenti

La letteratura relativa alla chimica bioinorganica è ricca di articoli riguardanti l'interazione tra molecole di interesse biologico e metalli mono- e bivalenti ma gli ioni trivalenti sono stati fino ad oggi poco studiati.

La presenza di ferro trivalente negli addotti ione metallico-nucleotide suggerisce che altri ioni metallici altamente carichi possono avere possibilità di coordinare al DNA sebbene non vi siano evidenze a riguardo [42].

I lantanidi e gli attinidi sono gli ioni metallici trivalenti più studiati dal punto di vista biologico. I dati di letteratura suggeriscono che la loro luminescenza viene accentuata quando si legano agli acidi nucleici. Essi possono funzionare da co-fattori ed esibire azione enzimatica in alcune reazioni [43].

1.6.4 I metalli di transizione

La formazione dei legami non covalenti con gli ioni dei metalli di transizione è controllata da forze di attrazione ione-dipolo e dalle forze di repulsione di Pauli tra il metallo e il legante. Il contributo elettrostatico al legame aumenta da sinistra verso destra della tavola periodica perché il raggio ionico dei metalli diminuisce. A parità di altri fattori, i metalli sulla sinistra del periodo formano, generalmente, legami più forti rispetto a quelli collocati a destra.

I metalli di transizione hanno orbitali d parzialmente pieni e questo li classifica come radicali liberi. Essi interagiscono con il DNA in più siti dando luogo a interazioni molto complesse [30, 44].

I metalli di transizione perdono le loro molecole di acqua molto facilmente e danno complessi coordinati attraverso la sfera interna. Essi solitamente legano direttamente sulle basi e indirettamente sul gruppo fosfato. Tra tutti i metalli di transizione lo zinco, elemento in tracce indispensabile per la salute umana, ha un ruolo importante nella stabilizzazione del DNA [45-49].

Lo zinco lega sugli atomi di ossigeno del gruppo fosfato e sulle nucleobasi. I dati spettroscopici a disposizione, tuttavia, sottolineano la preferenza dello zinco per il sito N7 della guanina.

Questo è un dato che conferma la veridicità dei dati nello studio da noi condotto, nel quale viene evidenziato come lo zinco, che forma complessi stabili con tutte le forme tautomeriche considerate, manifesti una spiccata preferenza per il sito N7 della guanina e il complesso risultante rappresenti il minimo globale.

1.7 L'acqua e gli acidi nucleici.

La vita nasce dall'acqua e l'acqua è vita!

L'acqua è un bene primario per la vita e una risorsa rinnovabile del nostro pianeta. Ogni forma di vita è legata all'acqua. Ogni attività umana è vincolata alla possibilità di accedere all'acqua.

Tutte le culture nella storia del nostro pianeta hanno attribuito all'acqua il più elevato valore simbolico e le hanno riconosciuto il ruolo centrale nella vita umana.

La vita nelle forme in cui la conosciamo non potrebbe esistere senza acqua. Tutte le cellule, sia di organismi acquatici che terrestri, contengono più del 70% di acqua. La sua molecola ha proprietà inusuali che le permettono di essere indispensabile per la vita sulla terra e per la stessa esistenza dell'uomo.

I processi chimici e fisici della vita esigono che le molecole siano capaci di muoversi, di incontrarsi, di cambiare partner frequentemente nei complicati processi del metabolismo e delle biosintesi; richiedono, cioè, che un'ampia quantità di ioni e di molecole si muovano gli uni in prossimità degli altri e che siano solubili in un mezzo comune. Grazie alle sue particolari proprietà solventi, l'acqua costituisce un mezzo intracellulare ed extracellulare universale [1].

Queste sue proprietà vengono spiegate dalla sua tendenza a formare legami idrogeno e dal suo carattere dipolare.

Un ambiente fluido consente una certa mobilità molecolare e l'acqua, non solo è il fluido più abbondante sulla terra, ma è anche stupefacentemente adatto allo scopo.

L'acqua (figura 1.7a) è una molecola **polare** e questa importantissima caratteristica fa sì che essa interagisca fortemente con qualunque soluto abbia carattere ionico come, ad esempio, un sale che si dissocia in ione positivo e negativo, oppure un filamento di DNA che è carico negativamente. Le molecole di acqua possono formare fra loro dei legami ad idrogeno che formano un *network* (figura 1.7b) in continuo movimento e che fa sì che, anche attorno a una molecola di soluto, l'acqua si disponga con determinate orientazioni che comportano un contributo energetico favorevole ma entropicamente sfavorevole (l'entropia è, in un certo senso, legata al disordine del sistema, e per il secondo principio della termodinamica, spontaneamente non può che aumentare).



FIGURA 1.7

Secondo il modello di Watson Crick, lo scheletro idrofilo del DNA costituito dalla sequenza zucchero-fosfato si trova all'esterno a contatto con l'acqua, mentre le basi di carattere idrofobico sono impacchettate perpendicolarmente all'asse dell'elica nell'interno. La formazione della doppia elica è usualmente sfavorita dalla formazione di legami ad idrogeno con l'acqua ma nel caso della forma B, quella predominante in natura, la struttura molecolare suggerisce la presenza di una spina di molecole d'acqua (in verde nella figura 1.5), giacente nel solco minore, atte ad instaurare legami ad idrogeno che aiutano a stabilizzare l'elica in questa forma.

Ciò probabilmente si deve al fatto che tale conformazione non permette l'accesso dell'acqua nell'interno idrofobico della molecola.

Come già descritto precedentemente, le basi degli acidi nucleici esistono in diverse conformazioni. Per alcune di esse, la differenza energetica tra i vari tautomeri è molto piccola e, così, l'energia termica, in soluzione, può trasformare facilmente l'uno nell'altro a temperatura ambiente.

I clusters formati tra molecole di solvente e le biomolecole sono molto utili per studiare gli effetti della solvatazione sulla struttura e sulla reattività delle biomolecole stesse. I clusters acqua-basi degli acidi nucleici sono un eccellente esempio.

Le molecole di solvente, oltre a influenzare la struttura e le proprietà delle biomolecole, possono influenzare profondamente la reattività alterando l'intorno del soluto o possono partecipare direttamente in processi dinamici in reazioni chimiche mediate dal solvente [50].

I legami idrogeno con molecole di acqua influenzano notevolmente le altezze delle barriere per alcune reazioni di tautomerizzazione [51-53]

Le molecole di solvente possono agire da catalizzatori e gli effetti dell'azione catalitica si evidenzia nell'alterazione della stabilità cinetica dei tautomeri.

Bibliografia

- [1] C. K. Mathews, K. E. van Holde *Biochimica*, seconda edizione, Casa Editrice Ambrosiana, 1998.
- [2] Avery, O. T.; MacLeod, C. M.; McCarty, M. J. *Exp. Med.*, **1944**, *79*, 137.
- [3] Watson, J. D.; Crick, F. H. C. *Nature*, **1953**, *171*, 737.
- [4] Wilkins, M. H. F.; Stokes, A. R.; Wilson, H. R. *Nature*, **1953**, *171*, 738.
- [5] Crick, F. *Nature*, **1974**, *248*, 766.
- [6] Kornberg, A.; Baker, T. A. *DNA Replication*, 2th ed, W. H. Freeman & Co., San Francisco, **1992**.
- [7] Mariani, K. J. *Annu. Rev. BioChem*, **1992**, *61*, 719.
- [8] Hackl, E. V.; Kornilova, S. V.; Blagoi, Y. P. *Int. J. of Biol. Macromol.*, **2005**, *35*, 175.
- [9] Wing, R. M.; Drew, H. M.; Takano, T.; Brodka, C.; Tanaka, S; Itakura, K; Dickerson, R. E. *Nature*, **1980**, *287*, 755.
- [10] Bauer, W. R.; Crick, F. H. C.; White, J. H. *Sci. Am.*, **1980**, *243*, 118.
- [11] Lilley, P. M. J. *Biochem. Soc. Trans.*, **1984**, *12*, 127.
- [12] Paleček, E. *Crit. Rev. Biochem. Mol. Biol.*, **1991**, *26*, 161.
- [13] Heintz, N. *Trends Biochem. Sci.*, **1991**, *15*, no.11.
- [14] Chamberlin, M. J. *Harvey Lect.*, **1994**, *88*, 1.
- [15] Kassvetis, G. A.; Geiduschelk, E. P. *Science*, **1993**, *259*, 944.
- [16] Sousa, R. ; Chung, Y. J., Rose, J. P.; Wang, B. *Nature*, **1992**, *364*, 593.
- [17] von Hippel, P. H.; Yager T. D. *Science*, **1992**, *255*, 809.
- [18] Crick, F. H. C. *Symp. Soc. Exp. Biol.*, **1958**, *12*, 138.
- [19] Brenner, S.; Jacob, F.; Meselson, M. *Nature*, **1961**, *190*, 576.
- [20] Zhang, Y.; Huang K. *J. of Molec. Struct.: THEOCHEM*, **2007**, *812*, 51.
- [21] Sigel, H.; Sigel A.; Eds, *Metal ion in Biological Systems: Interactions of Metal Ions with Nucleotides, Nucleic Acids and their Constituent*, vol. 32, Marcel Dekker, New York, **1996**.
- [22] Adams, P. L.; Stahley, M. R.; Kosek, A. B.; Wang, J.; Strobel, S. A. *Nature*, **2004**, *430*, 45.
- [23] Lippert, B. *Coord. Chem. Rev.*, **2000**, *487*, 200.
- [24] Eichhorn, G. L.; Shin, Y. A. *J. Am. Chem. Soc.*, **1968**, *90*, 7323.
- [25] Narasimhan, V.; Bryan A. M. *Inorg. Chem.*, **1984**, *91*, 39.
- [26] Glusker, J. P.; Katz, A. K.; Charles, W. B. *Rig, J.*, **1999**, *16*, 8.

- [27] Shannon, R. D.; Prewitt, C. T. *Acta Cryst.*, **1969**, *B25*, 925.
- [28] Brown, I. D. *Acta Cryst.*, **1988**, *B44*, 545.
- [29] Silva, J. J. R. F. D.; Williams, J. P. *The Biological Chemistry of Elements*, Clarendon Press, Oxford, **1991**.
- [30] Anastassopoulou, J. J. *J. Mol. Struct.*, **2003**, *651-653*, 19.
- [31] Hanzilik, R. P. *Inorganic Aspects of Biological and Organic Chemistry*, Academy Press, New York, **1976**.
- [32] Frausto da Silva, J. J. R.; Williams, R. J. P. *The Biological Chemistry of Elements. The Inorganic Chemistry of Life*, Clarendon Press, Oxford, **1994**.
- [33] Rodgers, M. T.; Armentrout, P. B. *Mass Spectr. Rev.*, **2000**, *19*, 215.
- [34] Dickerson, D. H. *J. Mol. Biol.*, **1981**, *151*, 535.
- [35] Russo, N.; Toscano, M.; Grand, A. *JACS*, **2001**, *123*, 10272.
- [36] Russo, N.; Toscano, M.; Grand, A. *J. Phys. Chem. B*, **2001**, *105*, 4735.
- [37] Armentrout, P. B.; Rodgers, M. T. *J. Phys. Chem. A*, **2000**, *104*, 2238.
- [38] Cerda, B. A.; Wesdemiotis, C. *JACS*, **1995**, *118*, 11884.
- [39] Russo, N.; Toscano, M.; Grand, A.; *J. Phys. Chem. A*, **2003**, *107*, 11533.
- [40] Šponer, J.; Sabat, M.; Burda, J. V.; Leszczynsky, J.; Hobza, P. *J. Phys. Chem. B*, **1999**, *103*, 2528.
- [41] Burda, J. V.; Šponer, J.; Hobza, P. *J. Phys. Chem.*, **1996**, *100*, 7250.
- [42] Hettich, R. H. *Int. J. Mass Spectrom.*, **2001**, *204*, 55.
- [43] Geyer, C. D.; Dipankar, S. *J. Mol. Biol.*, **1998**, *275*, 483.
- [44] Holland, P. M.; Castleman, A. W. *J. Chem. Phys.*, **1982**, *76*, 4195.
- [45] Sabat, M.; Lippert, B. *Metal ions in Biological Systems*, vol. 33, Sigel, A.; Sigel H.; Eds.; Marcel Dekker: New York, **1996**; 143.
- [46] Sigel, H. *Chem. Soc. Rev.*, **1993**, *18*, 32.
- [47] Bray, T. M.; Bettger, W. *J. Free Radic. Biol. Med.*, **1990**, *8*, 281.
- [48] Powell, S. R. *J. Nutr.*, **2000**, *130*, 1447S.
- [49] Ho, E. *J. Nutr. Biochem.*, **2004**, *15*, 572.
- [50] Anh, d.-s.; Lee, S.; Kim, B. *Chem. Phys. Lett.*, **2004**, *390*, 384.
- [51] Gorb, L.; Leszczynsky, J. *JACS*, **1998**, *120*, 5024.
- [52] Wang, X. C.; Nichols, J.; Feyereisen, M.; Gutowsky, M.; Boatz, J.; Haymet, D. J.; Simons, J. *J. Phys. Chem.*, **1991**, *95*, 10419.
- [53] Alhambra, C.; Gao, J.; Corchado, J.; Villa, J.; Truhlar, D. G.; *JACS*, **1999**, *121*, 2253.

CAPITOLO 2

METODI TEORICI

2.1 La teoria del funzionale della densità

I metodi tradizionali sulla struttura elettronica, in particolare la teoria di Hartree-Fock [1], sono basate su complicate funzioni d'onda multi-elettroniche. Il principale obiettivo della teoria del funzionale della densità (DFT) [2] è quello di sostituire la funzione d'onda multi-elettronica con la densità elettronica come quantità di base.

Mentre la funzione d'onda è dipendente da $3N$ variabili, tre variabili spaziali per ognuno degli N elettroni, la densità è funzione unicamente di tre variabili ed è più semplice da trattare sia a livello concettuale sia a livello pratico.

La densità elettronica $\rho(\mathbf{r})$ è definita come il numero di elettroni per unità di volume e tramite essa è possibile risalire alla funzione d'onda:

$$\rho(r_1) = N \int \dots \int |\Psi(x_1, x_2, \dots, x_N)|^2 d\sigma_1 dx_2 \dots dx_N \quad (2.1.1)$$

La densità restituisce il numero totale di elettroni, infatti:

$$\int \rho(r) dr = N \quad (2.1.2)$$

ed è dipendente dalle sole coordinate spaziali.

Il predecessore della teoria del funzionale della densità è stato il modello di Thomas-Fermi [3,4] sviluppato nel 1927, che valuta l'energia cinetica di un gas uniforme di elettroni non interagenti. La teoria fallisce laddove vengano considerati sistemi più complessi a causa della mancata stima, da parte dei due scienziati, dell'energia di scambio come conclusione del principio di esclusione di Pauli.

Sebbene la teoria del funzionale della densità abbia le sue radici concettuali nel modello di Thomas e Fermi, la sua nascita si fa risalire al momento in cui Hohenberg e Kohn [5] stipularono i teoremi che portano il loro nome.

Per un sistema elettronico, l'energia dello stato fondamentale e la funzione d'onda dello stato fondamentale vengono determinati minimizzando il funzionale dell'energia $E[\Psi]$. L'hamiltoniano che descrive tale sistema è:

$$\hat{H} = \hat{T} + \hat{V}_{ee} + \hat{V}_{ext} \quad (2.1.3)$$

dove \hat{V}_{ee} è potenziale di repulsione elettronica, \hat{T} è l'energia cinetica del sistema e \hat{V}_{ext} è un generico potenziale esterno espresso come:

$$\hat{V}_{ext} = \sum_{i=1}^N v(r_i) \quad (2.1.4)$$

Quindi, per un sistema N-elettronico, il numero totale di elettroni e il potenziale esterno determinano tutte le proprietà dello stato fondamentale.

Il primo teorema di Hohenberg e Kohn, usando la densità elettronica come variabile di base, afferma che il potenziale esterno è determinato, a meno di una costante, dalla densità elettronica. Dato che $\rho(\mathbf{r})$ determina il numero totale di elettroni, essa determina anche la funzione d'onda dello stato fondamentale e tutte le altre proprietà elettroniche del sistema.

Di conseguenza, il valore di aspettazione di ciascun operatore che compare nella (2.1.3) può essere espresso come un unico funzionale della densità.

Tale funzionale la cui forma è di seguito evidenziata,

$$F[\rho] = T[\rho] + V_{ee}[\rho] \quad (2.1.5)$$

è indipendente dal potenziale esterno ed è definito potenziale universale.

A questo punto è possibile scrivere l'energia totale del sistema come la somma di più contributi esprimibili in funzione di $\rho(\mathbf{r})$:

$$E[\rho] = T[\rho] + J[\rho] + \tilde{E}_{xc}[\rho] + V_{ext}[\rho] \quad (2.1.6)$$

dove $J[\rho]$ è

$$J[\rho] = \frac{1}{2} \iint r_{12}^{-1} \rho(r_1) \rho(r_2) dr_1 dr_2 \quad (2.1.7)$$

V_{ext} , rappresenta il valore medio dell'operatore \hat{V}_{ext} ed è

$$V_{ext} = \int \rho(r)v(r)dr \quad (2.1.8)$$

Riscrivendo la (2.1.6), si ottiene:

$$E[\rho] = F[\rho] + \int \rho(r)v(r)dr \quad (2.1.9)$$

Se si conoscesse l'esatta forma del funzionale universale $F[\rho]$, si potrebbe determinare correttamente la forma del funzionale dell'energia dello stato fondamentale.

Il secondo teorema di Hohenberg e Kohn rappresenta il principio variazionale e afferma che, data una densità tentativo $\tilde{\rho}(r)$, minimizzando il funzionale della (2.1.9), si trova che

$$E_0 \leq E_v[\hat{\rho}] \quad (2.1.10)$$

dove E_0 rappresenta l'energia corretta ed $E_v[\hat{\rho}]$ il funzionale dell'energia corrispondente ad un dato potenziale esterno $v(r)$.

Assumendo che $E_v[\rho]$ sia differenziabile, il principio variazionale (2.1.10) richiede che la densità dello stato fondamentale soddisfi la condizione di stazionarietà:

$$\delta \left\{ E_v[\rho] - \mu \left[\int \rho(r)dr - N \right] \right\} = 0 \quad (2.1.11)$$

da cui si ricava l'equazione di Eulero-Lagrange:

$$\mu = \frac{\delta E_v}{\delta \rho(r)} = v(r) + \frac{\delta F[\rho]}{\delta \rho(r)} \quad (2.1.12)$$

dove μ rappresenta il potenziale chimico.

Se conoscessimo esattamente $F[\rho]$, allora la (2.1.11) sarebbe l'esatta equazione per la densità elettronica dello stato fondamentale.

La più comune implementazione della teoria del funzionale della densità avviene tramite il metodo di Kohn-Sham [6]. In tale metodo, il problema degli

elettroni interagenti in un potenziale esterno statico è ridotto a un problema di elettroni non interagenti che si muovono in un potenziale effettivo. Questo tipo di potenziale include un potenziale esterno e gli effetti delle interazioni coulombiane tra gli elettroni, cioè le interazioni di scambio e correlazione.

$$\left[-\frac{1}{2}\nabla^2 + v_{eff}(r) \right] \varphi_j = \varepsilon_j \varphi_j \quad (2.1.13)$$

Queste rappresentano le N equazioni monoelettroniche di Kohn-Sham che devono essere risolte per trovare la densità esatta del sistema.

Modellare le due ultime interazioni è la parte più difficile del metodo.

La più semplice approssimazione fu proposta dagli stessi Kohn e Sham ed è nota come approssimazione della densità locale, LDA [7]. Essi associarono la densità ad un gas uniforme di elettroni, la cui energia di scambio $E_x(\rho)$ è:

$$E_x(\rho) = -C_x \rho(r)^{1/3} \quad (2.1.14)$$

$$\text{con } C_x = \frac{3}{4} \left(\frac{3}{\pi} \right)^{1/3}$$

Il potenziale di scambio è:

$$v_x^{LDA}[\rho] = -\left(\frac{3}{\pi} \right)^{1/3} \rho(r)^{1/3} \quad (2.1.15)$$

Per ciò che concerne la forma di $E_{corr}[\rho]$, un prezioso contributo è stato fornito dagli accurati calcoli Monte Carlo di Ceperley e Alder [8,9] che hanno fornito dati dai quali Vosko, Wilk e Nusair (VWN) [10] hanno estrapolato una forma analitica per $\varepsilon_{corr}[\rho]$.

Il metodo LDA è applicabile a sistemi in cui la densità si può considerare omogenea e fallisce laddove vengano considerati sistemi non omogenei, non trattabili, cioè, alla stregua di un gas uniforme di elettroni.

In chimica, un potenziale noto e largamente usato è il BLYP, dal nome di Beck [11] per la parte di scambio e Lee, Yang e Parr [12] per la parte di correlazione.

Il B3LYP è un funzionale ibrido in cui l'energia di scambio è combinata con l'esatta energia proveniente dalla teoria Hartree-Fock.

L'equazione che descrive tale funzionale è la seguente

$$E_{xc}^{B3LYP} = A \cdot E_X^{HF} + (1 - A) \cdot E_X^{Slater} + B \cdot E_X^{B88} + C \cdot E_C^{LYP} + (1 - C) \cdot E_C^{VWN} \quad (2.1.16)$$

2.1.1 Ricerca di minimi, massimi e punti di sella

Lo studio dell'energetica di una reazione chimica avviene tramite la costruzione di una superficie di energia potenziale in cui vengano caratterizzati i punti stazionari, ossia i minimi, i massimi e i punti di sella. Considerando una funzione a più variabili, le derivate prime rispetto ad ogni variabile generano un vettore gradiente, mentre le derivate seconde rappresentano gli elementi da cui si costruisce la matrice hessiana.

Un punto stazionario, per una funzione ad una sola variabile, è il punto in cui la derivata prima si annulla, mentre per una funzione a più variabili rappresenta il punto in cui si annulla la derivata prima rispetto a tutte le variabili.

La natura dei punti stazionari è stabilita per sistemi monodimensionali, dal segno della derivata seconda, positiva per un minimo e negativa per un massimo, mentre per sistemi multidimensionali è stabilita dagli autovalori della matrice hessiana. Se gli autovalori sono tutti positivi, il punto stazionario trovato è un minimo, se sono tutti negativi si tratta di un massimo, se c'è un solo valore negativo, il punto trovato è un punto di sella di ordine uno, cioè un massimo in una direzione ed un minimo in tutte le altre direzioni perpendicolari ad essa e corrisponde ad uno stato di transizione.

A proposito degli stati di transizione, poiché la possibilità di individuare sperimentalmente le loro strutture è limitata a pochi casi isolati e fortunati, i metodi teorici rappresentano l'unico mezzo per costruire superfici di energie complete e un indispensabile strumento per caratterizzarli e definirne le proprietà.

Possiamo immaginare una semplice reazione come una trasformazione che passi da un minimo energetico, il reagente, ad un altro minimo, il prodotto attraverso una struttura di massimo energetico, cioè uno stato di transizione.

Dato che possono esistere diversi cammini che connettono due minimi strutturali, lo stato di transizione è definito come il massimo lungo la curva a minima energia che connette reagenti e prodotti.

La variazione di energia in una tale trasformazione è funzione di numerose variabili e, come tale, dovrebbe essere rappresentata in uno spazio multidimensionale. Ciò nonostante, la variazione di energia può essere riportata in funzione di un'unica variabile, la coordinata di reazione, che proprio per la sua definizione è un oggetto difficile da individuare e definire analiticamente.

Per minimizzare una funzione non lineare di più variabili, sono stati messi a punto diversi metodi¹⁹ la cui bontà non può essere stabilita a priori senza considerare la particolare situazione in esame. Tra le caratteristiche che un algoritmo di minimizzazione deve possedere possiamo comprendere la velocità di convergenza verso il minimo, la stabilità e il costo computazionale, cioè il tempo di calcolo. I diversi metodi possono essere raggruppati in base al tipo di informazione sulle derivate; si hanno così:

- 1) metodi che non fanno uso delle derivate;
- 2) metodi che fanno uso delle derivate prime o del gradiente
- 3) metodi che fanno uso delle derivate prime e seconde.

Gli ultimi due metodi possono, inoltre, essere ulteriormente classificati in base alla possibilità di dell'utilizzo di gradienti e derivate seconde numerici o analitici.

Per quanto riguarda i metodi per la ricerca degli stati di transizione, in linea di principio, è valido lo stesso tipo di classificazione. Il fatto, però, che essi siano punti di sella di ordine uno, quindi, minimi in tutte le direzioni tranne che in una, rende la loro individuazione più complessa. Infatti, la direzione in cui il punto di sella è un massimo non è nota in anticipo e deve essere determinata durante la ricerca stessa e, mentre, la ricerca di un minimo locale è quasi sempre coronata dal successo, altrettanto non è detto che avvenga per gli stati di transizione. Per di più, la corrispondenza tra lo stato di transizione localizzato e quello effettivo deve sempre essere verificata mediante il calcolo dell'hessiano e dei suoi autovalori.

Per questa serie di motivi, numerosissimi sono gli algoritmi che sono stati proposti nel corso degli anni, basati sugli approcci più vari, tra i quali i più efficaci risultano quelli che fanno uso dell'hessiano, che può essere stimato, valutato numericamente o, nelle situazioni più fortunate, calcolato analiticamente.

2.2 La teoria perturbativa di Møller-Plesset

In questo capitolo, considereremo la teoria perturbativa di Møller e Plesset [13] (MP) in cui, volendo ottenere l'espansione della perturbazione per l'energia di correlazione, scegliamo l'Hamiltoniano di Hartree-Fock come Hamiltoniano di ordine zero [1].

La funzione d'onda Ψ^{HF} calcolata con il metodo HF definisce l'occupazione degli orbitali molecolari e lo stato di spin del sistema. Tale funzione d'onda determina una configurazione elettronica del sistema. Il metodo MP si basa sull'idea di generare una funzione d'onda Ψ che descriva meglio il sistema combinando tra loro diverse configurazioni, ovvero diversi stati elettronici della molecola: esso è detto, perciò, multiconfigurazionale e permette di ottenere energie complessive del sistema più basse rispetto al metodo HF e di comprendere nel calcolo il contributo dovuto alla correlazione elettronica.

Nella teoria MP, l'operatore hamiltoniano elettronico \hat{H}_{el} viene approssimato dalla somma degli operatori monoelettronici di Fock, con l'aggiunta di un operatore perturbativo \hat{P} pesato dal parametro λ :

$$\hat{H}_{el} = \sum_i \hat{f}_i + \lambda \hat{P} \quad (2.2.1)$$

Il primo termine dell'equazione, ovvero la somma degli operatori di Fock, è l'operatore approssimato della teoria HF che nella teoria MP viene corretto mediante l'applicazione dell'operatore perturbativo \hat{P} [1].

Sulla base di tale approccio, la funzione d'onda Ψ viene espressa nella teoria MP come sviluppo in serie di Taylor della funzione d'onda proveniente da un precedente calcolo HF,

$$\Psi = \Psi^{HF} + \lambda \left. \frac{\partial \Psi^{HF}}{\partial \lambda} \right|_{\lambda=0} + \frac{1}{2!} \lambda^2 \left. \frac{\partial^2 \Psi^{HF}}{\partial \lambda^2} \right|_{\lambda=0} + \frac{1}{3!} \lambda^3 \left. \frac{\partial^3 \Psi^{HF}}{\partial \lambda^3} \right|_{\lambda=0} + \frac{1}{4!} \lambda^4 \left. \frac{\partial^4 \Psi^{HF}}{\partial \lambda^4} \right|_{\lambda=0} + \dots \quad (2.2.2)$$

Si dimostra, in questo modo, che anche la teoria MP genera una nuova funzione d'onda Ψ combinando quella proveniente dal metodo HF con altre possibili configurazioni del sistema derivanti da singoli, doppi, tripli, ecc, trasferimenti elettronici da orbitali molecolari occupati ad orbitali molecolari vuoti.

I vari metodi MP si distinguono in funzione del troncamento dell'espansione in serie di Ψ . Considerando solo il primo termine dell'espansione (MP1) si dimostra che l'energia risultante per il sistema è identica all'energia ottenuta dal metodo Hartree-Fock. I primi miglioramenti nel calcolo dell'energia totale del sistema si hanno troncando l'espansione al secondo termine (MP2) [14-16]. Il metodo Mp2 è l'applicazione più diffusa della teoria di Møller-Plesset e, mediamente, corregge l'energia di correlazione di circa il 50%. Un altro metodo diffuso implica il troncamento al quarto termine dell'espansione (MP4): tale metodo è computazionalmente più impegnativo ma riesce a correggere solitamente più del 95% dell'energia di correlazione. Il troncamento al terzo termine (MP3) non è molto diffuso, poiché è stato osservato che non comporta grossi vantaggi rispetto al meno impegnativo livello MP2.

Per rendere i metodi MP computazionalmente meno dispendiosi, il procedimento multiconfigurazionale viene applicato solo agli elettroni presenti nel guscio valenza, tralasciando gli elettroni del core. Tale approssimazione è nota come "frozen core" [17].

È, inoltre, importante sottolineare che le caratteristiche del set di basi impiegato sono fondamentali per la costruzione accurata degli orbitali molecolari e, di conseguenza, della funzione d'onda che descrive il sistema. Intuitivamente, ci si può rendere conto che un metodo multiconfigurazionale, dovendo gestire contemporaneamente più funzioni d'onda è ancora più sensibile del metodo HF alle caratteristiche del set di base.

Piuttosto spesso ai metodi multiconfigurazionali (in particolare quelli basati sulla teoria "coupled-cluster") sono associati set di basi [17-19] indicati con le sigle cc-pVnZ ("correlation-consistent polarized valence n-zeta") dove la n indica il numero di funzioni di basi impiegate per ogni orbitale atomico: ad esempio, il cc-pVDZ è un set di base double- ζ , il cc-pVTZ è triple- ζ , e così via. Tali set di basi, pur avendo come principale campo di applicazione i metodi multiconfigurazionali, vengono occasionalmente impiegati anche con altri metodi, in particolare quelli basati sulla teoria DFT.

2.2.1 L'approssimazione RI

Il metodo MP2 occupa un importante ruolo nella teoria della struttura elettronica molecolare. Esso rappresenta la procedura computazionale meno

dispendiosa che tiene conto di forze di dispersione che sono trascurate nella più popolare Teoria del Funzionale della Densità (DFT). L'ostacolo per un uso più diffuso del metodo MP2 è dovuto ad esigenze computazionali piuttosto alte.

Recentemente sono state avanzate due proposte con lo scopo di ridurre i costi computazionali. La prima [20] usa metodi di correlazione locale, riducendo le dimensioni dell'espansione di configurazione tramite due approssimazioni: ogni coppia elettronica localizzata ha un proprio spazio virtuale e la debole correlazione degli elettroni distanti può essere trascurata. La seconda proposta [21] si basa su una rappresentazione approssimata di prodotti di funzioni orbitaliche di basi ν, μ o di prodotti di orbitali molecolari come combinazione lineare di funzioni di basi ausiliarie P rispetto ad atomo centrato:

$$\rho_{\nu\mu}(r) \equiv \nu(r)\mu(r) = \tilde{\rho}_{\nu\mu}(r) = \sum c_{\nu\mu,P} P(r) \quad (2.2.3)$$

Il principale vantaggio dell'equazione (2.2.3) è una considerevole riduzione dei dati [19]. Minimizzando l'energia di auto interazione di Coulomb, gli integrali a quattro centri $(\mu\nu | k\lambda)$ sono approssimati nel seguente modo:

$$(\mu\nu | k\lambda) \approx \sum_{P,Q} (\nu\mu | P)(P | Q)^{-1}(Q | k\lambda) \quad (2.2.4)$$

Il lato destro dell'equazione coinvolge solo integrali a tre centri $(\mu\nu | P)$ che sostituiscono quello a quattro centri. Tale equazione rappresenta l'approssimazione RI (risoluzione dell'identità) [22,23].

L'approssimazione degli integrali a quattro centri nel metodo della risoluzione dell'identità da buoni risultati con un costo computazionale piuttosto basso e quasi nessuna perdita in accuratezza dei calcoli rispetto all'esatto metodo MP2. È stato dimostrato che l'accuratezza delle energie molecolari è migliore di 1 mE_h per atomo e l'errore nella distanza interatomica è minore di 1 pm.

Bibliografia

- [1] Levine, I. Quantum Chemistry, Prentice Hall, Inc. New Jersey **2000**, (5th ed)
- [2] Parr, R. G.; Yang, W Density Functional Theory of Atoms and Molecules, Oxford Univ. Press., New York **1989**.
- [3] Thomas, L. H. Proc. Cambridge Phil. Soc. **1926**, 23, 542.
- [4] Fermi, E. Z. Phys. **1926**, 48, 542.
- [5] Hohenberg, P.; Kohn, W. Phys. Rev. **1964**, 136, B864.
- [6] Kohn, W.; Sham, L. J. Phys. Rev. **1966**, 145, 561.
- [7] Slater, J. C. Phys. Rev. **1951**, 81, 385.
- [8] Ceperley, D. M. Phys. Rev. B **1978**, 18, 3126.
- [9] Ceperley, D. M; Alder B. J. Phys. Rev. Lett. **1980**, 45, 566.
- [10] Vosko, S. H.; Wilk, L.; Nusair, M. Can. J. Phys. **1980**, 58, 1200.
- [11] Becke, A. D. J. Chem. Phys. **1993**, 98, 5648.
- [12] Lee, C.; Yang, W.; Parr, R.G. Phys. Rev. B **1988**, 37, 785.
- [13] Møller, C.; Plesset M. S. Phys. Rev. **1934**, 46, 618.
- [14] Bartlett, R. J.; Silver, D. M. J. Chem. Phys., **1975**, 62, 3258.
- [15] Pople, J. A.; Binkley, J. S.; Seeger, R., Int. Quant. Chem. Symp., **1976**, 10, 1.
- [16] Hehre, W. J.; Schleyer, P. V. R.; Pople, J. A. Ab initio molecular orbital theory, **1986**, Wiley, New York.
- [17] Schäfer, A.; Horn, H.; Ahlrichs, R. J. Chem. Phys., **1992**, 97, 2571.
- [18] Schäfer, A.; Huber, C.; Ahlrichs, R. J. Chem. Phys., **1994**, 100, 5829.
- [19] Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. Theor. Chem. Acc., **1997**, 97, 119.
- [20] El Azhary, A.; Rauhut, G.; Pulay, P.; Werner, H. J. J. Chem. Phys., **1998**, 108, 5185.
- [21] Vahtras, O.; Almlöf, J.; Feyereisen, M. W. Chem. Phys. Lett., **1993**, 213, 514.
- [22] Weigend, F. ; Häser, M. Theoret. Chem. Acc. **1997**, 97, 331.
- [23] Weigend, F. ; Häser, M.; Patzelt, H.; Ahlrichs, R. Chem. Phys. Lett. **1998**, 294, 143.

CAPITOLO 3

PRESENTAZIONE DEI RISULTATI

3.1 Interazione di ioni metallici con le subunità strutturali degli acidi nucleici

Capire come avvengono le interazioni fra gli ioni metalli e le biomolecole complesse quali il DNA o l'RNA può fornire un mezzo che collega numerose aree di applicazione. Per far ciò è necessario scendere ad un livello minimale, e indagare cioè, su come le diverse subunità (amino acidi, peptidi, proteine, basi, nucleotidi, etc..) interagiscono con i metalli. I caratteri strutturali, spettroscopici, elettronici e, inoltre, i trasferimenti di carica e la natura del legame tra metallo e legante sono tutti aspetti fondamentali per una dettagliata comprensione dei meccanismi alla base delle suddette interazioni.

Nei paragrafi successivi verranno brevemente illustrati i risultati ottenuti nei casi esaminati nel presente lavoro di tesi.

3.1.1 Interazione dello ione alluminio con le basi degli acidi nucleici (U, T, C, A, G) e con i nucleotidi monofosfato (UMP, dTMP, dCMP, dAMP, dGMP). ARTICOLO 1¹

In questo lavoro, sono stati determinati i caratteri strutturali ed elettronici di tutti i possibili complessi risultanti dall'attacco dello ione alluminio nudo e idratato sui vari siti dei diversi tautomeri delle nucleobasi e dei nucleotidi monofosfato selezionati. Lo studio su tali sistemi è stato condotto a livello teorico tramite la Teoria del Funzionale della Densità (DFT), usando il potenziale di scambio e correlazione B3LYP e il set di basi 6-311+G(2df,2p).

La scelta dell'alluminio si deve alla sua tossicità che, come suggerito da numerosi studi sperimentali [1,2], può comportare malfunzionamento renale causa poi di anemia e di encefalopatia, al suo probabile ruolo nell'insorgenza di malattie come il morbo di Alzheimer, la sclerosi amiotrofica laterale, la sindrome di Down e al fatto che i cationi trivalenti, come accennato nel capitolo 1, non hanno finora ricevuto l'attenzione che meritano.

¹ Le tabelle e le figure per questo lavoro sono disponibili nell'articolo 1 allegato alla fine del capitolo.

3.1.2 Interazione con il catione nudo: i complessi delle nucleobasi e dei nucleotidi monofosfato.

I siti di coordinazione dello ione metallico sui leganti sono stati scelti in base ai dati in letteratura riguardanti le interazioni di Al^{3+} o altri cationi con biomolecole. Fra le varie possibilità di legame è stata, inoltre, considerata l'interazione tra il catione e il sistema π dell'anello aromatico sia per le purine che per la pirimidine. Questo tipo di coordinazione, sebbene rara, è stata riscontrata per altri metalli e comunque mai per le basi puriniche. In accordo a ciò, i nostri risultati hanno confermato l'esistenza di addotti π solo per uracile, timina e citosina. Tali addotti sono però forme energetiche altamente sfavorite in tutti i casi e presentano distorsioni tali da distruggere i cicli aromatici con cui interagiscono.

In base a considerazioni riguardanti la stabilità delle forme tautomeriche delle nucleobasi, indice della loro concreta presenza, è stato possibile appurare che i complessi realmente esistenti si formano sempre a partire dalla forma tautomerica canonica pur essendo, a volte, energeticamente meno stabili degli addotti provenienti da tautomeri enolici.

Analizzando i valori di affinità per lo ione alluminio (Al³⁺), si può dedurre che la guanina è la base più suscettibile d'attacco da parte di questo catione. Seguono la citosina, la timina, l'adenina ed infine l'uracile.

Lo ione appare essere monocoordinato all'atomo O4 della timina e dell'uracile, bicoordinato agli atomi N3 e O2 della citosina, e ancora bicoordinato agli atomi N7 e N6 di entrambi guanina ed adenina.

L'analisi NBO e i valori della carica netta sullo ione Al^{3+} in tutti i sistemi studiati danno un'indicazione della natura del legame con i leganti considerati.

Nel caso della mono-coordinazione sull'atomo di ossigeno, il legame sembra essere covalente con un consistente contributo ionico. Nella bicoordinazione N,O, il legame N- Al^{3+} è essenzialmente covalente con un piccolo contributo ionico, mentre, il legame O- Al^{3+} è praticamente ionico.

Per l'interazione dello ione alluminio con i nucleotidi monofosfato sono stati presi in esame tutti gli addotti che sottintendono il legame del catione sia con i siti di attacco presenti sulla base in forma canonica che con il gruppo fosfato. Per quanto ci è dato sapere, la letteratura non riporta mai evidenza di un legame che coinvolga direttamente lo zucchero.

La presenza del gruppo zucchero-fosfato in posizione N1 per le basi puriniche e in posizione N9 per quelle pirimidiniche, fa sì che le forme tautomeriche dei nucleotidi siano minori rispetto a quelle delle nucleobasi. Inoltre, l'ingombro sterico di questo gruppo rende alcuni siti inaccessibili allo ione.

L'ottimizzazione di tutte le forme complesse esaminate per ciascun nucleotide sottolinea che quando si parte da strutture recanti lo ione sui siti della base, compreso il caso dell'interazione con il sistema π , si ottiene sempre un addotto in cui il catione sembra essere coordinato simultaneamente al gruppo fosfato e agli atomi della nucleobase esaminata volta per volta. In pratica avviene sempre una sorta di ripiegamento che coinvolge entrambe le unità strutturali.

Questi ripiegamenti coinvolgenti il fosfato fino alla cattura dello ione non sono plausibili nel DNA e nell'RNA, tuttavia, essi contribuiscono a dare qualche indicazione sull'affinità dello ione alluminio verso le diverse unità strutturali dei nucleotidi. In accordo ai suggerimenti sperimentali, la maggiore affinità si trova per il gruppo fosfato. Per questo motivo, l'analisi NBO è stata effettuata solo sui complessi in cui l'alluminio è posizionato su di esso.

La carica netta naturale sull'alluminio complessato indica che c'è un trasferimento di carica dal legante al metallo. Il piccolo contributo covalente al legame si origina dall'orbitale molecolare formato dalla sovrapposizione tra gli orbitali ibridi sp e spd degli atomi di ossigeno e dello ione alluminio, rispettivamente.

3.1.3 Interazione con il catione idratato: i complessi con le nucleobasi e i mononucleotidi

La chimica dello ione alluminio in acqua è complicata dalla presenza di specie mono- e poli-nucleari che dipendono dai valori di pH e dalla concentrazione di Al^{3+} [3]. A pH più acidi di 5 lo ione non è idrolizzato e la principale specie nucleare è $[\text{Al}(\text{H}_2\text{O})_6]^{3+}$. Per uno ione alluminio idratato coordinato al gruppo fosfato di un nucleotide o su un atomo donatore sulle nucleobasi, l'alta carica positiva del piccolo ione responsabile del carattere più acido delle molecole di acqua ad esso coordinate, dovrebbe essere parzialmente neutralizzata e i complessi idrossido, presenti in quantità rilevanti a valori di pH fisiologico, dovrebbero essere meno probabili. Così, per studiare l'effetto della

microidratazione del metallo sul tipo di coordinazione e sulla natura del legame abbiamo scelto il modello $[Al(H_2O)_5-L]^{3+}$.

I tentativi per trovare complessi stabili nel caso dell'interazione dello ione alluminio idratato con il sistema π dell'anello delle basi sono falliti.

L'uracile e la timina danno origine a due addotti mono-coordinati le cui strutture non sono diverse da quelle trovate per lo ione nudo. Il catione assume una geometria vicina a quella ottaedrica in entrambi i casi.

Per la citosina, il prodotto finale assume caratteristiche interessanti. Infatti, contrariamente a quanto accade per il catione nudo, l'alluminio idratato è mono-coordinato all'atomo O2 della base. Inoltre, a causa di un effetto polarizzante residuo dello ione alluminio sulle molecole di acqua della prima sfera, l'idrogeno di una molecola di acqua è trasferita su N3 sebbene rimanga ancora ancorata al gruppo -OH con un legame idrogeno.

La sfera di solvatazione non modifica il tipo di coordinazione del catione alluminio sulle basi puriniche.

I risultati per guanina e adenina mostrano che la specie $[Al(H_2O)_5]^{3+}$ espelle una molecola di acqua in modo da stabilire due legami simultanei con gli atomi N7 e O6 per la guanina e N7 e N6 per l'adenina. In entrambi i casi la molecola di solvente rimane in prossimità dei complessi grazie a qualche legame idrogeno con la sfera interna.

L'analisi NBO indica che l'alluminio idratato stabilisce legami ionici con le nucleobasi. Questo può essere dedotto comparando la carica totale sul catione prima e dopo l'interazione (2.1 contro 1.9 |e|).

Poiché nell'interazione con i nucleotidi il catione nudo è risultato sempre principalmente attratto dal gruppo fosfato, nel caso del catione idratato abbiamo considerato solo questo sito di attacco. In realtà, dal momento che i calcoli sono eseguiti in fase gassosa dove si è in presenza di grande libertà conformazionale, l'eventuale coinvolgimento delle altre unità strutturali è reso comunque possibile durante il processo di ottimizzazione.

Contrariamente a quanto trovato per i complessi dello ione nudo in cui il centro metallico era bi-coordinato, ora le cinque strutture mostrano un alluminio penta-idratato legato ad un solo atomo di ossigeno del fosfato. Questo significa che il catione preferisce tenere una molecola di acqua piuttosto che legarsi ad un altro atomo di ossigeno. Tutte le strutture esibiscono un numero elevato di legami

a ponte idrogeno tra gli atomi di ossigeno delle molecole di acqua e quelli del gruppo fosfato, che conferiscono loro grande stabilità.

Il legame tra lo ione idratato e l'atomo di ossigeno del fosfato è ionico come rivelato comparando i valori della carica netta sull'alluminio prima e dopo l'interazione (2.1 contro 1.9 |e|).

3.1.4 Affinità dei leganti per lo ione Alluminio.

I valori di affinità delle nucleobasi per lo ione alluminio (Al³⁺) hanno indicato il seguente ordine di reattività preferenziale:



e si ottengono in corrispondenza dei complessi bi-coordinati di ciascuna base ad eccezione dell'adenina, a causa della debole interazione dell'alluminio con l'atomo N6.

L'aggiunta di una sfera di coordinazione al centro metallico riduce drasticamente i valori di Al³⁺. Questo non sorprende se consideriamo che l'alluminio idratato ha una disponibilità orbitalica molto diversa da quella del catione nudo e che la forza con cui esso lega leganti addizionali decresce gradualmente.

Il nuovo andamento dei valori di Al³⁺ è: $C > G > A \geq T \geq U$.

A prima vista, questa sequenza sembra completamente diversa da quella ottenuta nel caso precedente ma i valori di Al³⁺ di A, T e U non sono sufficienti a stabilire un andamento rigoroso a causa della loro somiglianza.

L'inversione dei valori di guanina e citosina può essere attribuita al cambiamento del tipo di coordinazione passando dal catione nudo a quello idratato e alla protonazione dell'atomo N3 della citosina da parte di un atomo di idrogeno proveniente da una molecola di acqua.

I valori di Al³⁺ per i nucleotidi sono stati calcolati prendendo come riferimento i complessi in cui il catione metallico è legato al gruppo fosfato.

Ciò può essere giustificato dal fatto che, sebbene queste non siano le specie più stabili a causa dell'alta flessibilità dei nucleotidi in fase gassosa, esse diventano più significative quando i nucleotidi sono mantenuti distesi (senza possibilità di ripiegamenti) dal resto della molecola di acido nucleico.

Nel caso dell'alluminio nudo, i valori di AIIA sono molto alti grazie al legame che lo ione metallico forma con i due atomi di ossigeno del fosfato su cui è delocalizzata la carica negativa.

L'affinità di Al^{3+} segue l'ordine: $dGMP \geq dTMP > dAMP > dCMP > UMP$.

Nessun paragone è possibile per i valori ottenuti per le nucleobasi poiché nell'interazione sono coinvolte altre unità strutturali.

I nucleotidi interagiscono con lo ione idratato meno fortemente che con quello nudo. I valori di AIIA hanno grandezze completamente diverse da quelle ottenute in assenza di molecole di acqua. Questo si può spiegare tenendo presente che, non solo il catione solvatato ha una minore capacità di coordinazione, ma anche esso stabilisce un legame di natura diversa col legante rispetto al catione nudo.

Il nuovo ordine per le AIIA è: $dGMP > dAMP \geq dCMP > dTMP > UMP$.

I valori di affinità ottenuti indicano che il catione alluminio ha una grossa tendenza a legarsi agli acidi nucleici e suggeriscono una probabile spiegazione per la sua tossicità.

3.2 Interazione in fase gassosa dello ione Zinco con le basi puriniche e pirimidiniche del DNA e dell'RNA. ARTICOLO 2²

In questo lavoro, è stata studiata l'interazione delle nucleobasi del DNA e dell'RNA con lo ione zinco.

Come è noto, questo catione metallico è coinvolto in diversi processi cellulari che includono la proliferazione, la riproduzione, la difesa immunologia in generale e contro i radicali liberi. Non più di 10 anni fa, è stato provato che lo ione zinco controlla i cambiamenti conformazionali che avvengono nel DNA e che inducono distorsioni della coppia di base Watson-Crick guanina-citosina [4,5].

Inoltre, esso influenza i processi di replicazione, di trascrizione e riparazione di questo acido nucleico e di molti enzimi che lo contengono. La letteratura suggerisce che, come altri cationi bivalenti lo zinco si lega alle nucleobasi più favorevolmente che con il gruppo fosfato esibendo grande affinità per le prime.

² Le tabelle e le figure contenenti i dati numerici sono disponibili nell'articolo 2 alla fine del capitolo.

Lo studio, che fa parte di ricerca sistematica sull'interazione ione metallico-legante biologico, è stato condotto a livello DFT utilizzando il funzionale B3LYP di scambio e correlazione e il set di basi 6-311+G(2df,2p).

Il protocollo metodologico seguito ricalca quello usato nel precedente studio con l'alluminio, dove le ottimizzazioni dei vari complessi metallo-legante sono state precedute da una selezione dei tautomeri a più basse energie delle nucleobasi e dei siti più adatti per l'attacco dello ione. A questo proposito, è stato evidenziato che la formazione delle forme tautomeriche non canoniche per uracile, timina e adenina sono improbabili, mentre nel caso della citosina e della guanina, le strutture enoliche sono molto vicine in energia alle strutture cheto-amminiche (canoniche), come suggerito anche sperimentalmente.

Lo ione zinco forma complessi con tutte le nucleobasi.

È interessante sottolineare che i tautomeri liberi di uracile, timina e adenina differiscono notevolmente in energia e che generalmente il processo di metallazione avviene preferenzialmente su un tautomero ad alta energia della base isolata. Questo è quello che si riscontra anche nel caso dello ione zinco. Tenendo però presente che questi dati sono solo una misura della stabilità intrinseca dei sistemi e che la probabilità di esistenza di tautomeri ad alta energia è praticamente nulla, i soli complessi reali sono quelli provenienti dalle forme canoniche di queste basi. Questa conclusione è supportata anche dal fatto che spesso la separazione energetica fra i complessi di queste nucleobasi è piccola al contrario di quella esistente fra i tautomeri delle basi libere. Da questo punto di vista, lo ione zinco fa eccezione, cioè gli addotti Zn^{2+} -legante sono energeticamente abbastanza differenti in tutti i casi, ma come dato di fatto rimane che i tautomeri nella forma enolica sono altamente improbabili.

La citosina e la guanina mostrano un comportamento diverso rispetto alle altre nucleobasi. Infatti, sebbene i quattro tautomeri della base libera siano molto vicini in energia, i corrispondenti complessi con lo zinco hanno una stabilità che è molto differente.

Dall'analisi strutturale dei complessi realmente esistenti, risulta che il catione predilige la monocoordinazione all'atomo O4 nel caso della timina e dell'uracile, la bicoordinazione agli atomi N3 e O2 nella citosina, ed ancora la bicoordinazione alle coppie di atomi N7, O6 e N7, N6, rispettivamente nella guanina e nell'adenina.

Per tutti i complessi di minimo globale, l'analisi NBO mostra nell'interazione ione-legante non esistono legami covalenti. Il trasferimento di carica di circa $0.4 |e|$ dal legante al catione sottolinea un prevalente contributo ionico al legame.

3.2.1 Affinità delle nucleobasi per lo ione Zinco

In linea di principio, i valori di MIA dovrebbero essere riferiti ai complessi più stabili che il catione forma con una particolare nucleobase. Però, come abbiamo accennato precedentemente, il complesso più stabile può originare dall'interazione dello ione con il tautomero che non sempre è il più stabile.

Il problema può essere poco significativo quando i tautomeri cadono in un intervallo energetico piccolo, come nel caso di citosina e guanina, perchè come regola generale, la loro contemporanea presenza dovrebbe garantire l'esistenza di tutti i loro complessi, così che la scelta dei valori più affidabili di affinità dovrebbe dipendere solo dall'ordine di stabilità di questi complessi. In altre parole, la scelta del valore di AIIA più probabile dipende dall'ordine di stabilità dei complessi che segue quello dei tautomeri liberi.

Per l'uracile, la timina e l'adenina bisogna invece considerare i complessi originatesi in seguito alla metallazione del tautomero canonico.

L'ordine di affinità delle nucleobasi per lo ione zinco trovato indicano I valori che abbiamo proposto come i più affidabili indicano nella guanina la base maggiormente soggetta all'attacco dello ione seguita da citosina, adenina, timina e uracile.

La presenza in letteratura di studi riguardanti i metalli di transizione Mn^{2+} e Cu^{2+} ha permesso il confronto con i dati ottenuti nel presente lavoro. Poiché, l'affinità delle nucleobasi per gli ioni è il parametro energetico più rilevante, in quanto dipende dalla geometria di coordinazione, dalle caratteristiche elettroniche del sito di attacco e dalla stabilità del complesso che si forma, la comparazione dei risultati è stata fatta solo in relazione a esso. In particolare, seguendo i trend di affinità nei tre casi, la regola della diminuzione di affinità con il numero atomico del catione sembra non essere confermata. Il catione che mostra più affinità è il Cu^{2+} seguito dallo Zn^{2+} e dal Mn^{2+} . La forza del legame sembra piuttosto dipendere dalla particolare configurazione elettronica dei cationi e sfavorisce i sistemi open shell. La nucleobase privilegiata dall'attacco di questi ioni metallici è

la guanina. L'adenina, di solito sfavorita, possiede per tali elementi affinità pari alla citosina e maggiore dell'uracile e della timina.

3.3 Interazione in fase gassosa degli ioni Ba^{2+} , Sr^{2+} , Cs^+ e Rb^+ con le basi puriniche e pirimidiniche del DNA e dell'RNA.

Lo studio in questione è da inquadrare sempre nell'ambito della interazione metallo-legante [6-10].

In questo caso l'attenzione è stata rivolta ai cationi bivalenti Ba^{2+} e Sr^{2+} e a quelli monovalenti Cs^+ e Rb^+ di cui sono stati esaminati i complessi con le forme tautomeriche più stabili delle nucleobasi. I dati ottenuti in questo studio, condotto a livello B3LYP facendo uso di basi estese 6-311+G** per gli atomi dei non metalli e descrivendo i metalli pesanti con lo pseudopotenziale SDD, hanno permesso in combinazione con informazioni derivanti da precedenti lavori su altri cationi, di tracciare degli andamenti sistematici di valori di affinità per le nucleobasi all'interno dei gruppi di appartenenza dei rispettivi elementi metallici.

Nella figura 1 sono riportati gli andamenti delle affinità per le nucleobasi ottenute per i metalli alcalini.

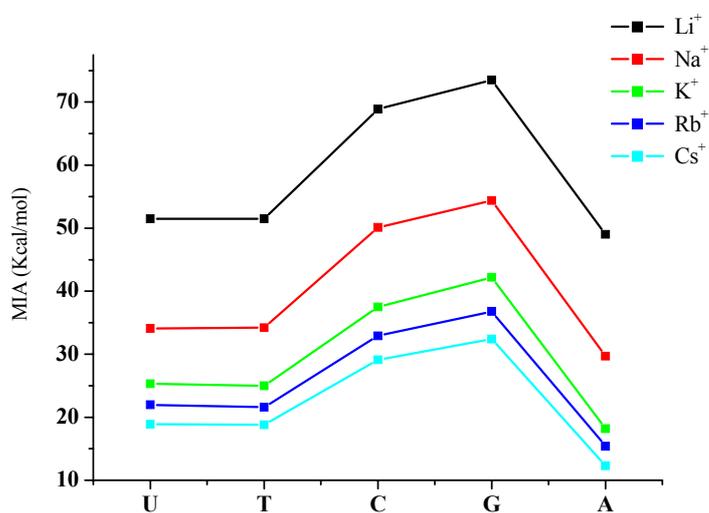


Figura 1. Andamento dei valori di affinità (MIA) per gli ioni alcalini.

In particolare, per i metalli alcalini Li^+ , Na^+ , K^+ , Cs^+ e Rb^+ [11,12] le conclusioni generali sono che i valori di affinità (MIA) per le basi diventano sempre più piccoli al crescere del numero atomico sebbene per gli elementi più pesanti la diminuzione sia più graduale. Tutti questi ioni mostrano, come base preferita per l'attacco, la guanina seguita dalla citosina, dalla timina e dell'uracile che spesso presentano valori molto simili di affinità e infine dell'adenina.

I cationi del gruppo danno luogo a complessi monocoordinati con la timina e l'uracile e bicoordinati negli altri casi.

Le strutture di minimo dei vari complessi metallati degli ioni rubidio e cesio sono illustrate nelle figure 2 e 3 dove sono state anche incluse le distanze (in Å) di coordinazione ai siti favoriti.

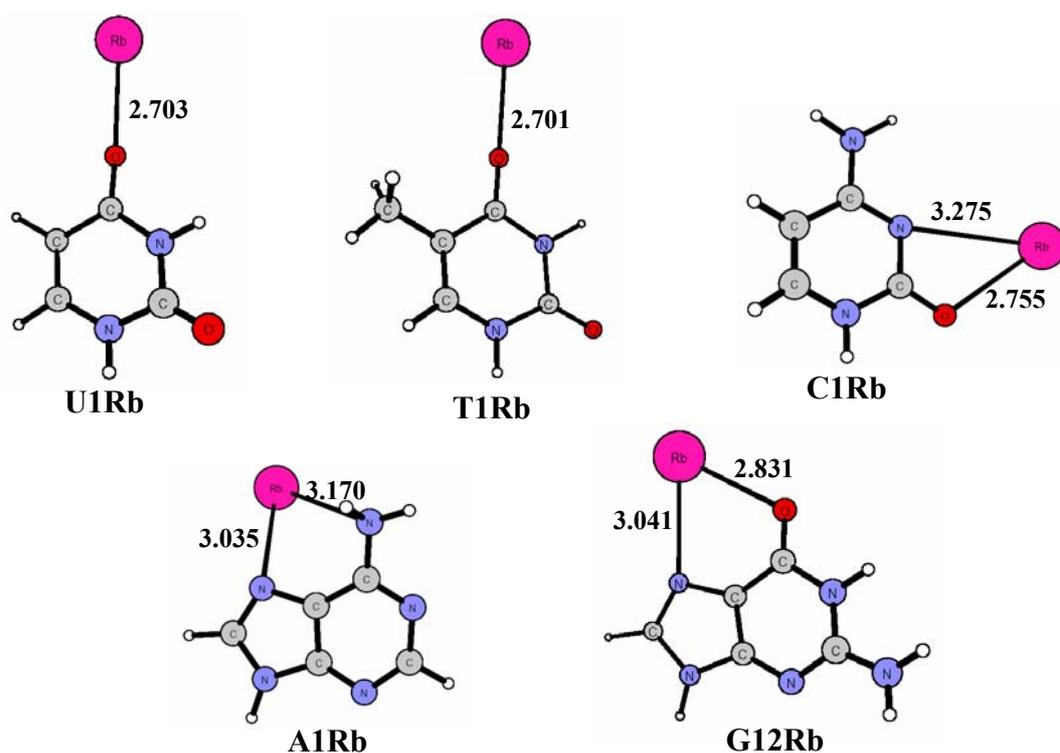


Figura 2. Strutture di minimo dei complessi dello ione rubidio con le nucleobasi.

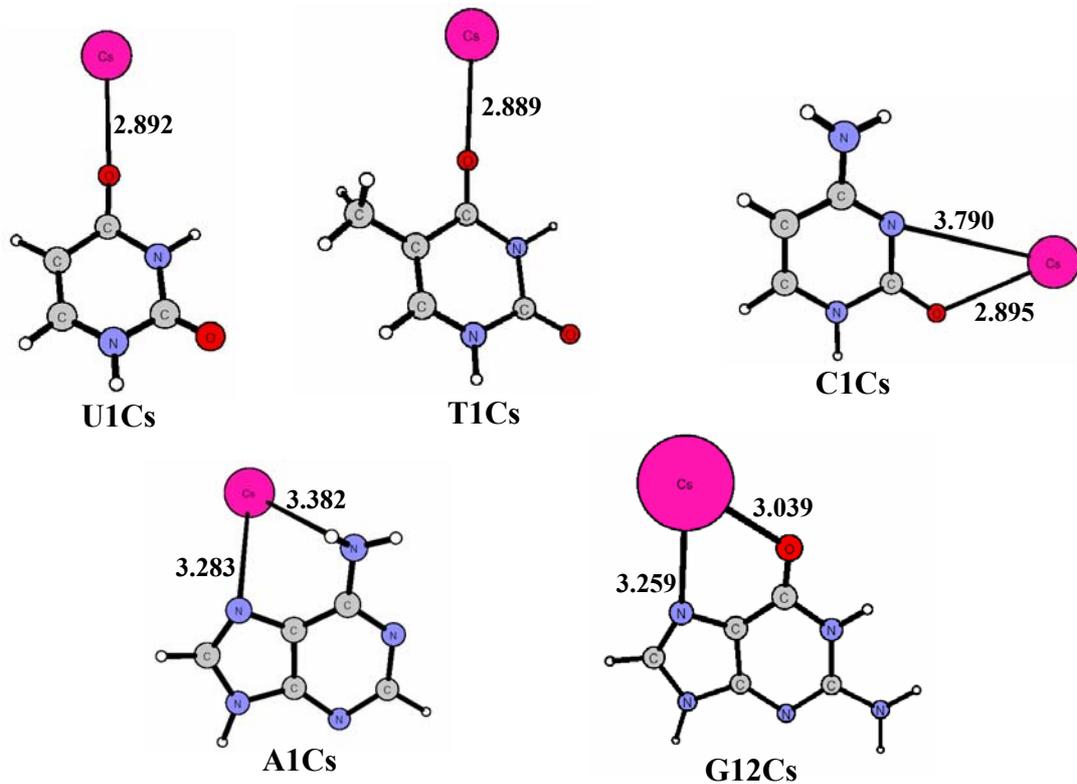


Figura 3. Strutture di minimo dei complessi dello ione cesio con le nucleobasi.

Per i metalli alcalino terrosi Mg^{2+} , Ca^{2+} , Ba^{2+} e Sr^{2+} [13] sono stati determinati andamenti molto simili a quelli degli alcalini sia per quanto riguarda le MIA viste in funzione della dimensione dei cationi, sia per il tipo di coordinazione alle basi (vedi figura 4).

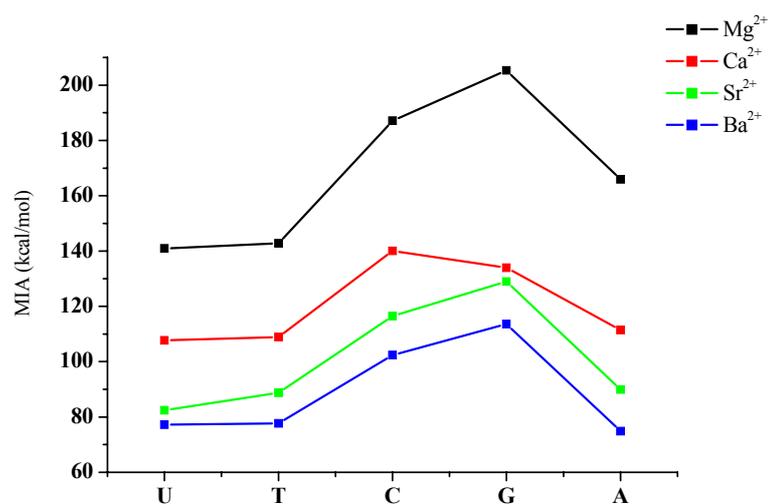


Figura 4. Andamento dei valori di affinità (MIA) per gli ioni alcalini terrosi.

Come si può osservare i valori di MIA assoluti sono più alti che nel caso dei metalli alcalini. La guanina, risulta anche in questo caso più suscettibile di attacco.

Le geometrie di coordinazione nelle strutture di minimo dei complessi dello ione bario e dello stronzio possono essere dedotte osservando le figure 5 e 6.

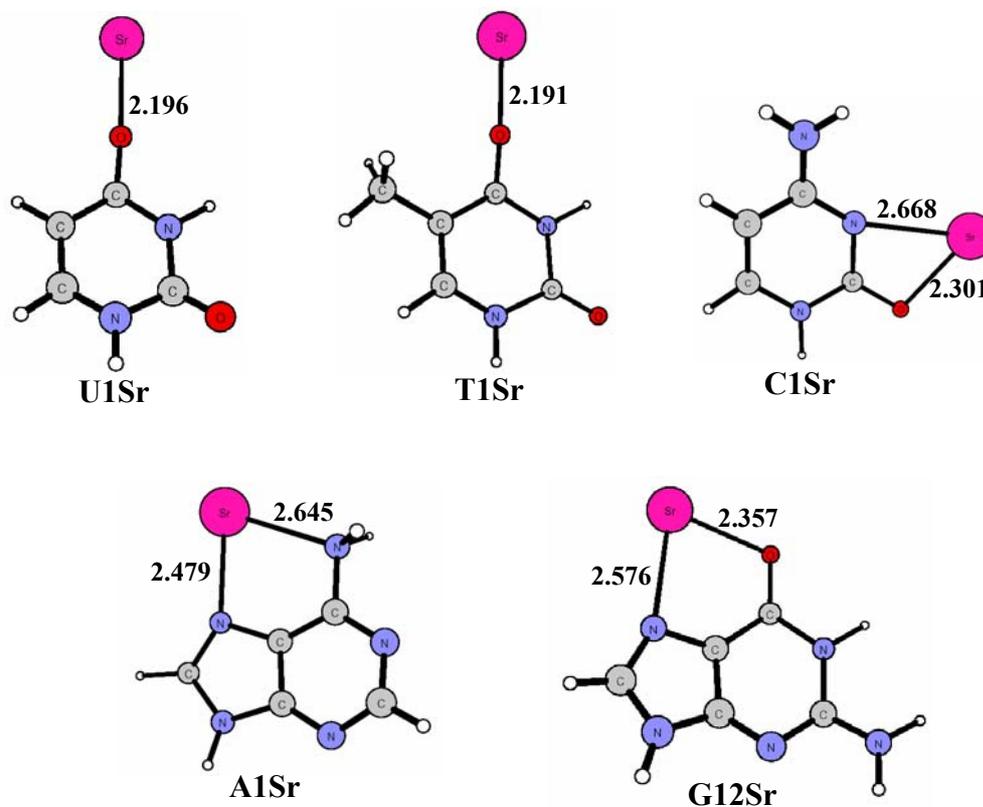


Figura 5. Strutture di minimo dei complessi dello ione bario con le nucleobasi.

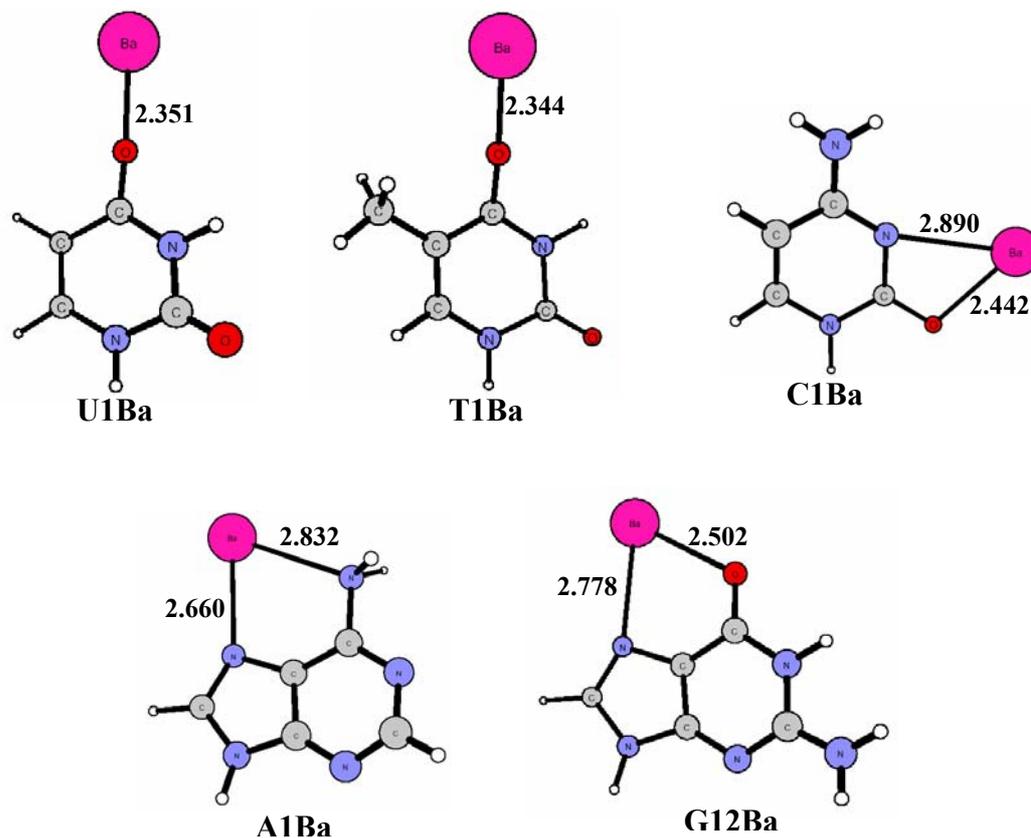


Figura 6. Strutture di minimo dei complessi dello ione stronzio con le nucleobasi.

3.4 Processi di tautomerizzazione della citosina catalizzati da molecole di acqua. ARTICOLO 3³

Il tautomerismo è un fenomeno importante per tutte le basi del DNA e dell'RNA ma acquista particolare significato per la citosina e la guanina per cui i tautomeri più stabili sono racchiusi in un intervallo energetico molto ristretto.

Studi sperimentali, condotti in condizioni di vaporizzazione termica, hanno accertato che alcuni tautomeri di queste due nucleobasi sono coesistenti [14-17]. Tuttavia, nonostante la comparabile stabilità, la loro interconversione in fase gassosa richiede energie molto alte. Questo fatto dà luogo ad una serie di interrogativi fra cui il più immediato pone l'accento su quali possano essere i meccanismi rendono possibile la contemporanea presenza di due o più forme tautomeriche di una nucleobase.

Il lavoro presente ha cercato di far luce proprio su questa problematica. D'altra parte la comprensione degli equilibri tautomerici delle basi degli acidi nucleici, è fondamentale di per sé, essendo i ponti a idrogeno che danno origine

³ Le tabelle e le figure contenenti i dati numerici sono disponibili nell'articolo 3 alla fine del capitolo

all'accoppiamento di base e molti modelli di mutazione spontanea basati su di essi.

Nello studio, sono stati presi in considerazione sette tautomeri diversi della citosina scelti fra quelli più stabili ed analizzati i meccanismi di interconversione fra loro valutando volta per volta geometrie, caratteristiche elettroniche e barriere di attivazione. Sono stati esplorati diversi canali di tautomerizzazione per la base libera e monoidrata in quanto, studi precedenti, hanno attribuito alla presenza di molecole di acqua, un' azione catalitica nei riguardi dell'interconversione nel caso della guanina. Il processo di tautomerizzazione cheto-amminico→enol-amminico è stato indagato più dettagliatamente allo scopo di valutare l'influenza di una seconda molecola di acqua di coordinazione sull'altezza della barriera di questo processo. Infine sono stati calcolati gli effetti del solvente acqua d'intorno simulato con una costante dielettrica $\epsilon = 78$.

Le ottimizzazioni delle geometrie sono state ottenute utilizzando il protocollo B3LYP/6-311+G(2df,2p). Per esplorare l'effetto dell'idratazione ci siamo avvalsi del metodo polarizable continuum model (PCM).

In accordo ai dati riportati in precedenza da altri studi teorici, nel caso della citosina isolata i processi di interconversione fra i vari tautomeri, che coinvolgono prevalentemente shift protonici, presuppongono spese energetiche molto elevate comprese nell'intervallo 34-50 kcal/mol che di fatto impediscono qualunque tipo di trasformazione da un tautomero all'altro.

Una volta selezionati i centri idrofilici sull'anello della citosina, sono state determinate le strutture idratate più stabili.

Le situazioni energeticamente favorite sono risultate quelle in cui le molecole di acqua agiscono simultaneamente da accettori e donatori di idrogeno.

Da un punto di vista della stabilità, i sistemi idratati seguono lo stesso andamento osservato in assenza di acqua ed l'energia di separazione fra le forme mono-solvate rimane abbastanza piccola da permettere, in linea di principio, una loro possibile interconversione .

I dati ottenuti, sottolineano che, eccetto per due processi che coinvolgono gli addotti idratati più instabili, ogni singola tautomerizzazione richiede un'energia di circa 11 kcal/mol. Questo, insieme al fatto che i trasferimenti protonici coinvolgono spesso tunnelling quanto-meccanici che rendono i processi più

veloci, suggerisce che la cinetica della tautomerizzazione sia fortemente influenzata dalla presenza della molecola di acqua.

Per verificare se l'aggiunta di una seconda molecola di acqua coordinata ai tautomeri della citosina possa avere qualche ulteriore effetto positivo sulle altezze delle barriere, abbiamo studiato il processo di interconversione tra le forme bi-idratate dei due tautomeri a più bassa energia.

La barriera si abbassa di sole 0.1 kcal/mol rispetto al caso precedente.

Al contrario di quanto accade per la citosina, nel caso della guanina l'aggiunta di una seconda molecola d'acqua di coordinazione abbassa ulteriormente ed in maniera diversa le varie barriere di interconversione. Una terza molecola d'acqua sortisce invece un effetto contrario. Ciò porta a ipotizzare che il numero di molecole di acqua discrete da coordinare alla base per avere qualche beneficio sulla velocità del processo di tautomerizzazione dipende dalle caratteristiche strutturali della base e dal tipo di trasformazione che avviene.

L'effetto del solvente di bulk applicato alla tautomerizzazione fra i due sistemi bi-idratati si traduce in un niente di fatto dal punto di vista cinetico ma il processo diventa termodinamicamente sfavorito.

Da questo ultimo risultato si deduce che l'azione delle molecole di acqua esplicite è sminuito dal campo elettrico del mezzo solvente. Questo concetto, trasferito in una soluzione reale, potrebbe significare che le molecole di acqua legate direttamente alla citosina perdono parzialmente la loro abilità di donatori e accettori di atomi di idrogeno perché attratte dalla presenza di altre molecole di acqua nella sfera di coordinazione esterna.

Così, il ruolo delle molecole di acqua è quello di mediare, grazie alle loro proprietà di donatori e accettori di atomi di idrogeno, i trasferimenti protonici che solitamente richiedono grandi quantità di energia.

Dalle indicazioni provenienti da questo studio e da lavori precedenti, noi pensiamo che, l'abbassamento della barriere osservate nella tautomerizzazione che coinvolge specie monoidratate, insieme a considerazioni sulle loro reciproche concentrazioni e sulla possibilità di un tunnelling quanto-meccanico, possano spiegare la coesistenza, provata sperimentalmente, di alcune forme tautomeriche della citosina, almeno ad alte temperature.

3.5 Microidratazione della coppia di base rara guanina-guanina.

Le interazioni fra le nucleobasi giocano un ruolo cruciale nella stabilizzazione degli acidi nucleici [18,19]

L'accoppiamento di queste nel DNA è altamente specifico, infatti, nella maggior parte dei casi la guanina si lega alla citosina e l'adenina alla timina dando luogo alle ben note strutture di Watson-Crick.

Nell'RNA o negli ibridi DNA-RNA possono essere presenti differenti coppie di base che sono conosciute come Watson-Crick inversa, Hoogsteen e Hoogsteen inversa [20].

I vari arrangiamenti spaziali delle basi prevedono disposizioni parallele (hydrogen bonded), strutture incolonnate (stacked) e a forma di T (T-shaped).

Ognuna di queste è differente sia per la forma che per la natura delle interazioni. Le strutture parallele, tenute insieme da ponti a idrogeno, sono principalmente stabilizzate da forze di natura elettrostatica fra cui dominano le interazioni di tipo dipolo-dipolo. Le coppie stacked sono invece governate da interazioni di dispersione di London.

Fra le differenti conformazioni, le strutture planari parallele sono le più stabili in fase gassosa [21] e in solventi di natura non polare [22-24], mentre nei solventi polari è più facile trovare strutture stacked [25-27]. Questo sottintende che l'intorno può essere determinante nel favorire una data struttura sulle altre.

Sebbene il ruolo dell'acqua sui caratteri strutturali e sulle funzioni degli acidi nucleici sia un argomento di notevole interesse, gli studi sperimentali sono ancora abbastanza scarsi. Ciò è principalmente legato alla difficoltà di portare le nucleobasi in fase gassosa e realizzare la formazione di aggregati con l'acqua [28-32]. Le investigazioni teoriche condotte a vari livelli di sofisticazione sono invece, più numerose [33-38].

Da questi studi precedenti è stato stabilito che il DNA è idratato [39,40] e che l'idratazione influenza sia la sua struttura che le sue funzioni. Inoltre, è stato ipotizzato che la popolazione delle coppie di base planari parallele e di quelle incolonnate possono essere influenzate dalla coordinazione di un piccolo numero di molecole di acqua.

In particolare, Kabeláč e al. [33] in uno studio riguardante la coppia AT, hanno mostrato che il grado di stabilizzazione delle forme stacked dipende dal numero di molecole di solvente molto di più di quanto succede per le forme legate

con i ponti ad idrogeno [26-27, 41]. Inoltre, in un lavoro riguardante la coppia di base rara guanina-guanina, è stato stabilito che l'aumento da 2 a 4 del numero di molecole d'acqua coordinate alle basi fa crescere la popolazione delle strutture stacked dal 20 al 90% [42].

Tuttavia, come frequentemente puntualizzato, è ancora non completamente chiaro se le transizioni dalle forme planari a quelle incolonnate siano controllate da interazioni idrofobiche dovute all'intorno del solvente acqua o da interazioni idrofile dovute ad un esiguo numero di molecole di questo solvente

Nel presente lavoro, abbiamo scelto di investigare la coppia di base rara guanina-guanina (GG) indicata da precedenti studi teorici come un sistema di stabilità comparabile e addirittura più stabile delle più comuni coppie di Watson-Crick CG e AT, rispettivamente [41, 43-44].

Allo scopo di verificare il ruolo della microidratazione, sono stati considerati tutti i possibili dimeri planari e non della GG e valutate le stabilità relative prima in fase gassosa ed poi in presenza di una e due molecole di acqua esplicite coordinate alla coppia di base.

Per ognuna delle forme di minimo globale appartenenti ad una specifica categoria, sono stati simulati inoltre gli spettri IR che abbiamo utilizzato per completare le informazioni derivanti da uno studio UV-IR di Abo-Riziq e al. [42] condotto sul sistema GG per risolvere alcune strutture mono e bi-idratate della coppia di base.

Le geometrie sono state ottimizzate utilizzando il metodo resolution of identity second-order Møller-Plesset (RI-MP2) [45-48] insieme con il set di base di cc-pVDZ implementato nel codice di calcolo TURBOMOLE [49-51]. Le frequenze armoniche vibrazionali sono state valutate usando lo stesso set di base e poi scalate di un fattore universale pari a 0.956 per fittare i valori sperimentali dei modi di stretching N-H della guanina isolata.

3.5.1 Sistemi hydrogen bonded, stacked e T-forms isolati

I dimeri planari più stabili guanina-guanina sono stati ottenuti ottimizzando tutti i possibili arrangiamenti della coppia con le unità monometriche nella forma cheto-amminica prevalente in fase gassosa. Il nome GG1 è stato attribuito ai sistemi caratterizzati da tre legami a ponte idrogeno, HNH...N, NH...N e O=C...HNH fra le due unità di guanina, mentre la sigla GG2 si riferisce ai sistemi

che presentano due ponti idrogeno simmetrici coinvolgenti i gruppi C=O...NH e NH...O=C. Poiché la guanina è presente in natura sotto due forme tautomeriche praticamente isoenergetiche, differenti solo per la posizione di un atomo di idrogeno sugli atomi di azoto N7 e N9 dell'anello a cinque termini, le combinazioni diverse dei due monomeri nella coppia sono portano l'indicazione dei tautomeri che le hanno generate.

I due arrangiamenti isomerici planari più stabili sono illustrati in figura 1, mentre in tabella 1 sono riportate le energie relative di tutti i dimeri planari stabili.

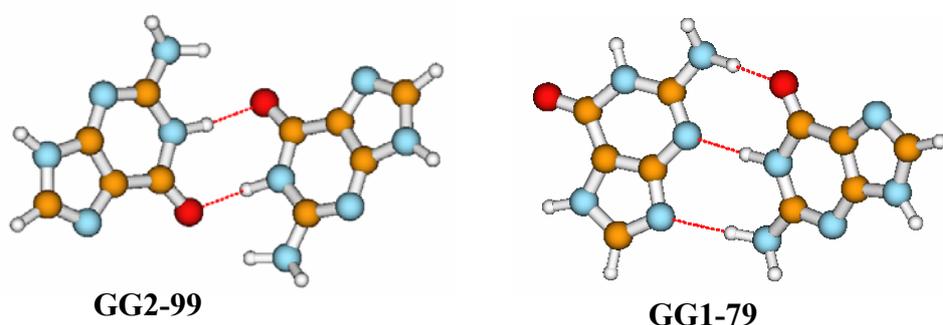


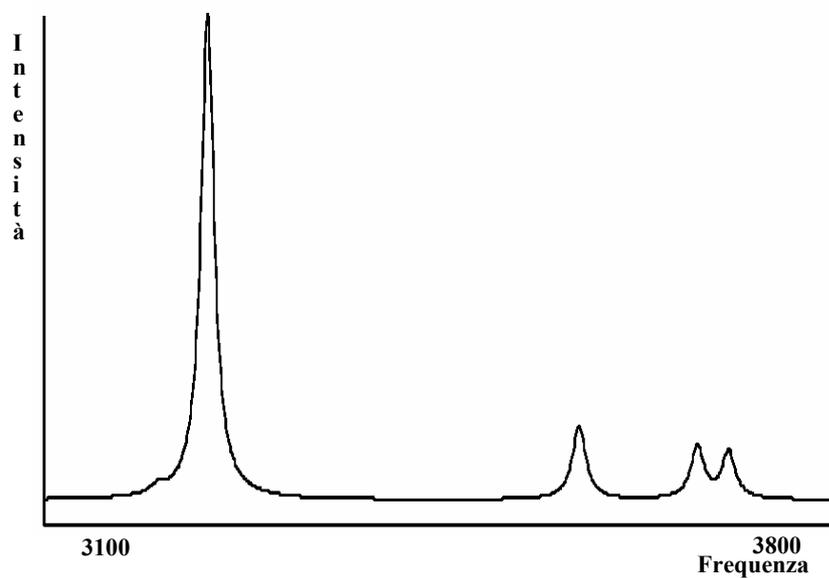
Figura 1. Strutture planari più stabili

Tabella 1: Energie assolute (E_{SCF} in a.u.) e relative (ΔE in kcal/mol) dei vari dimeri della coppia di base GG.

Dimeri planari	$E_{SCF+MP2}$	ΔE
GG2-99	-1082.176360	0.0
GG1-79	-1082.175523	0.5
GG1-77	-1082.172620	2.3
GG2-79	-1082.172022	2.7
GG2-77	-1082.168470	5.0
Dimeri stacked	$E_{SCF+MP2}$	ΔE
N	-1082.1168875	4.7
I	-1082.153627	14.3
Dimeri T-form	$E_{SCF+MP2}$	ΔE
T	-1082.159111	10.8
T1	-1082.157152	12.0
T2	-1082.156389	12.5

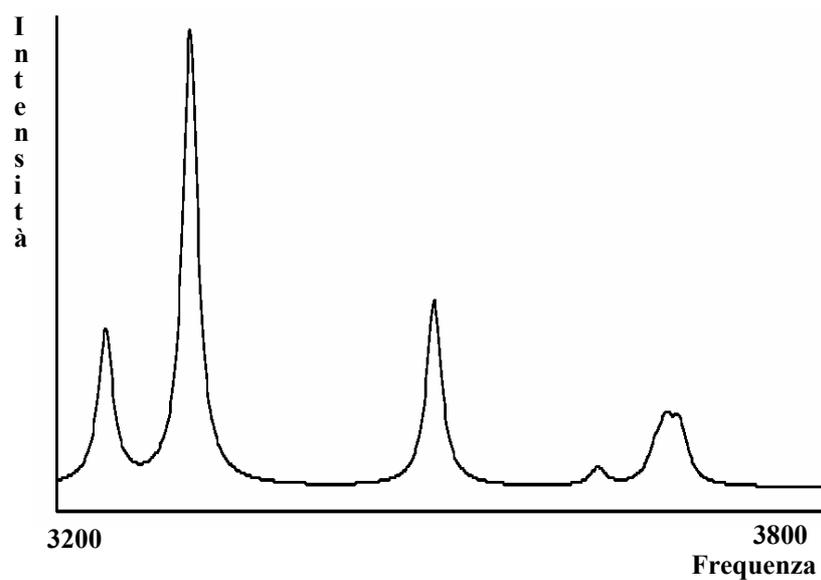
I minimi GG2-99 e GG1-79 sono separati da un gap energetico di sole 0.5 kcal/mol, il che implica che probabilmente esse sono coesistenti. Tuttavia anche gli altri sistemi non sono così distanti in energia.

Gli spettri IR simulati per i due dimeri GG2-99 e GG1-79, sono riportati in figura 2 e 3 e confrontati con la controparte sperimentale [42,52]. La regione spettrale di maggiore interesse è quella compresa fra 3100 e 3800 cm^{-1} in cui solitamente cadono le frequenze di stretching dei gruppi OH, NH and NH_2 .



	NH_2^a		$\text{N}_9\text{-H}$		$\text{N}_1\text{-H}$		NH_2^s	
	a	b	a	b	a	b	a	b
RI-MP2	3708	3708	3681	3681	3201	3245	3575	3573
*0.956	3545	3545	3519	3519	3060	3102	3418	3416
Exp ^a	3564	3564					3444	3444

Figura 2: Simulazione dello spettro IR per il dimero GG2-99 e relative frequenze⁴.



	NH ₂ ^a		N ₉ -H	N ₇ -H	N ₁ -H		NH ₂ ^s	
	a	b			a	b	a	b
RI-MP2	3667	3688	3685	3676	3623	3237	3304	3494
*0.956	3506	3526	3523	3514	3464	3091	3159	3340
Exp^a				3504			3253	3374
Exp^b	3530	3536	3511	3504	3441			3354

Figura 3: Simulazione dello spettro IR per il dimero GG1-79 e relative frequenze⁴

Per quanto riguarda il sistema GG1-79, i risultati ottenuti mostrano un ottimo accordo con le frequenze misurate e forniscono indicazioni su quelle non assegnate sperimentalmente.

Per GG2-99 non esistono dati di letteratura per la comparazione, tuttavia, vista la bontà dei dati ottenuti per il precedente sistema, possiamo essere confidenti nella validità dei risultati anche in questo caso.

Per una coppia di basi interagenti, ci sono due tipi di incolonnamento energeticamente favoriti che vengono classificati come strutture stacked normale e invertita.

La differenza consiste nel fatto che nella struttura normale le due unità hanno gli ossigeni carbonilici opposti l'uno all'altro mentre la disposizione di questi gruppi nel dimero invertito li vede affacciati.

Nella struttura stacked normale della coppia GG, più stabile rispetto all'invertita di circa 10 kcal/mol, sono presenti due legami a ponte idrogeno di pari lunghezza (2.450 Å) che coinvolgono i gruppi -NH₂ e l'atomo N7 sulle due unità monometriche. Tali strutture sono riportate nella figura 4 e indicate con N e I, rispettivamente.

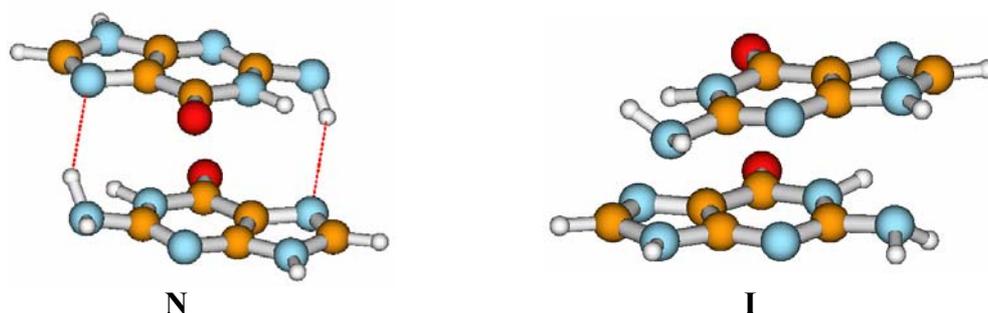


Figura 4. Strutture stacked più stabili per la coppia di basi GG

⁴ Gli esponenti a e b nelle tabelle delle frequenze si riferiscono agli articoli 42 e 52 alla fine del capitolo.

I parametri di stacking delle due basi calcolati concordano con quelli esistenti per il DNA B. In particolare la distanza (R) fra i centri delle due guanine risulta essere pari a 3.872 Å, la distanza verticale (d) fra i piani incolonnati è circa 3.5 Å e l'angolo di proiezione (θ) fra i due legami H-N9 è 179.6°.

L'osservazione dei dati in tabella 1, evidenzia che le forme stacked della coppia guanina-guanina sono quasi tutte ad energia più elevata rispetto ai dimeri planari in accordo con gli studi precedenti [18]. Fa eccezione solo la struttura stacked normale che tuttavia si trova solo 0.3 kcal/mol più stabile del dimero planare meno favorito GG2-77.

In base a queste considerazioni energetiche, le forme stacked non dovrebbero essere presenti in condizioni normali di temperatura, così sarebbe piuttosto improbabile rilevarle attraverso determinazioni IR. Questa è la ragione per cui abbiamo ritenuto superfluo simulare i loro spettri vibrazionali.

L'ottimizzazione delle strutture T-form esaminate, propone tre isomeri di energia molto simile ma decisamente meno stabili dei dimeri planari e di quelli stacked, fatta eccezione per la struttura stacked invertita (vedi tabella 1).

Le geometrie di equilibrio per le coppie di base suddette sono riportate in figura 5.

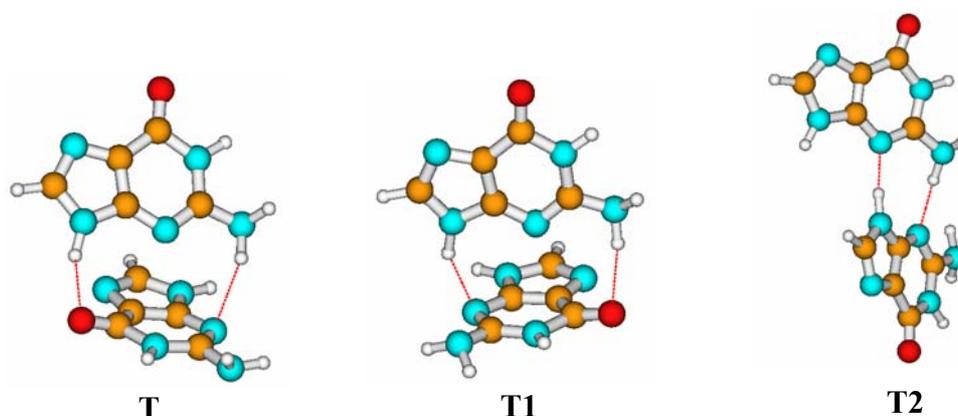


Figura 5. Strutture T-forms più stabili per la coppia di basi GG

3.5.2 Sistemi hydrogen bonded, stacked e T-forms monoidratati.

Le diverse disposizioni planari e non della coppia di base guanina-guanina, offrono una grande varietà di siti di coordinazione per la molecola di acqua essendo la nucleobase ricca di centri idrofili. L'acqua può, in linea di principio, monocoordinarsi su ogni atomo di azoto e di ossigeno di una delle guanine o a ponte fra le due unità nelle strutture stacked e T-forms. Per tale motivo il numero delle strutture monoidratate risulta veramente alto. Tuttavia, in questo studio,

sono state esaminate tutte le possibilità, poiché uno degli scopi di questo lavoro è di accertare il ruolo dell'acqua nella stabilizzazione dei vari sistemi anche se questi non risultano molto favoriti rispetto ad altri quando sono isolati.

Le stabilità dei sistemi monoidratati sono riportate in tabella 2.

Tabella 2: Energie assolute (in au) e relative (in kcal/mol) dei vari dimeri monoidratati della coppia di base GG.

	$E_{\text{SCF+MP2}}$	ΔE
GG1-79(H₂O)	-1158.432579	0.0
GG2-99(H₂O)	-1158.432480	0.06
GG1-77(H₂O)	-1158.429550	1.9
GG1-77(H₂O)-1	-1158.428810	2.4
GG2-99(H₂O)-1	-1158.427800	3.0
GG1-79(H₂O)-1	-1158.427177	3.4
GG1-79(H₂O)-2	-1158.427170	3.4
GG2-79(H₂O)	-1158.426973	3.6
GG2-79(H₂O)-1	-1158.426778	3.7
GG1-79(H₂O)-3	-1158.424415	5.1
GG1-77(H₂O)-2	-1158.424460	5.1
GG2-99(H₂O)-2	-1158.424470	5.1
GG2-77(H₂O)	-1158.424200	5.3
GG2-77(H₂O)-1	-1158.424200	5.3
GG1-79(H₂O)-4	-1158.423414	5.7
GG2-79(H₂O)-2	-1158.423155	5.9
GG2-79(H₂O)-3	-1158.420381	7.6
GG2-77(H₂O)-2	-1158.419030	8.5
N(H₂O)	-1158.426314	3.9
N(H₂O)-1	-1158.423173	5.9
I(H₂O)	-1158.421117	7.2
I(H₂O)-1	-1158.407986	15.4
TFORM(H₂O)up	-1158.412759	12.4
TFORM(H₂O)down	-1158.412700	12.5
TFORM1(H₂O)down	-1158.410480	13.9
TFORM1(H₂O)up	-1158.410309	14.0
TFORM2(H₂O)down	-1158.410260	14.0
TFORM2(H₂O)up	-1158.410170	14.1

Come si può osservare dai dati, le strutture planari idratate sono sempre le più favorite sebbene si possa notare un lieve recupero a carico della forma stacked normale. Le energie relative delle T-forms rimangono sempre molto elevate. Questo lascia intendere che la monoidratazione non ha un effetto molto significativo nel favorire una specie rispetto ad un'altra.

In figura 6 sono illustrate le strutture dei sette dimeri monoidratati a più bassa energia.

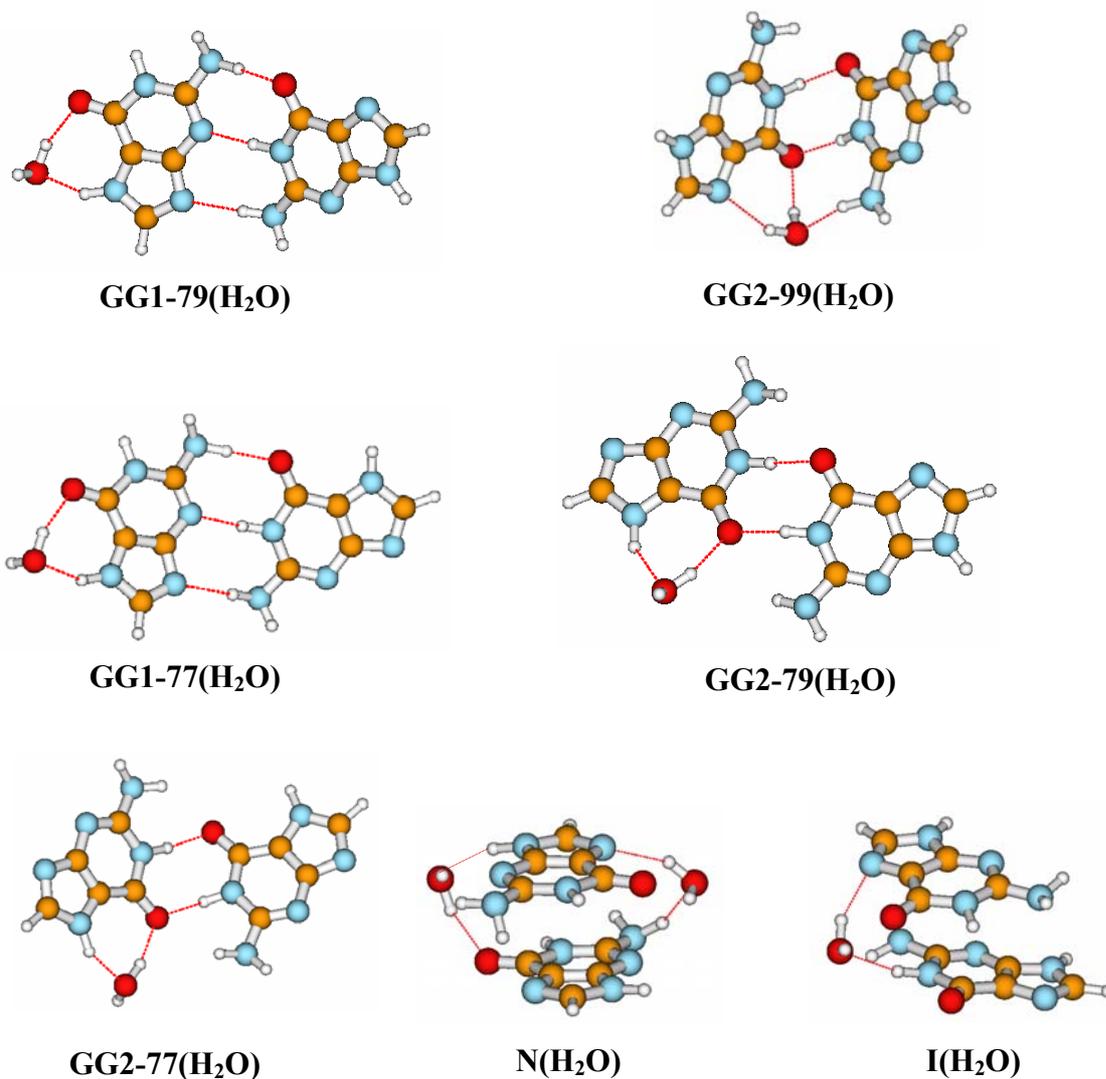


Figura 6. Strutture monoidratate planari e stacked più stabili per la coppia di basi GG

In ognuna delle forme planari riportate, la molecola d'acqua risulta sempre coordinata tramite due legami a ponte idrogeno con il sito N7 e l'ossigeno carbonilico di una delle guanine. Nelle forme stacked i due atomi di idrogeno dell'acqua si legano il primo all'N7 di una guanina e il secondo all'ossigeno carbonilico dell'altra guanina.

Comunque, sebbene siano state discusse solo le forme in figura 6, le altre elencate nella tabella, caratterizzate da una diversa posizione di coordinazione

dell'acqua, presentano dei valori di stabilità relativa molto prossimi a quelli evidenziati che lascia loro una possibilità di recupero energetico quando considereremo l'effetto di una seconda molecola di acqua.

Per ciascuna delle strutture monoidratate planari e stacked più stabili della coppia di basi GG, sono stati simulati gli spettri IR.

Questi, riportati di seguito nelle figure 7-13 forniscono dati nuovi che possono facilitare l'interpretazione degli spettri registrati sperimentalmente.

Il confronto dello spettro simulato con l'unico ottenuto sperimentalmente per il sistema GG1-79(H₂O), mostra, come può essere osservato dai valori delle frequenze risolte, un ottimo accordo.

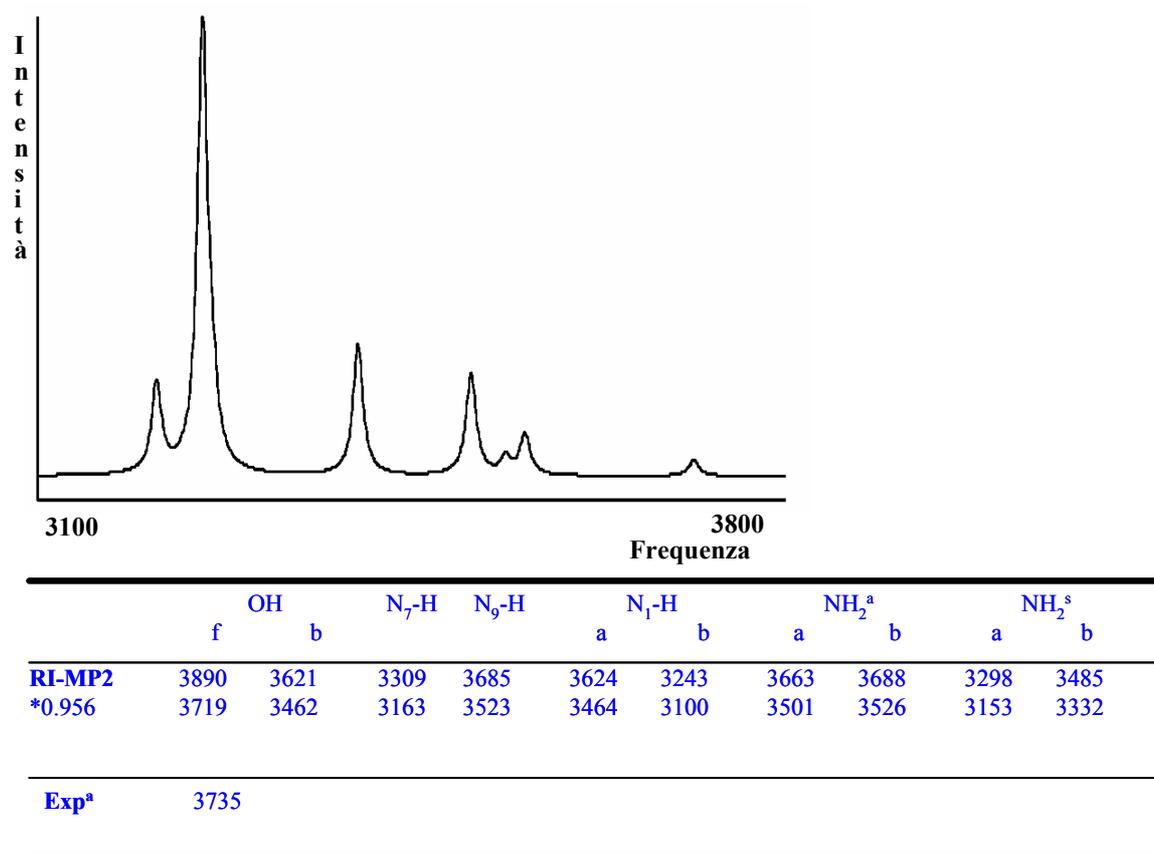


Figure 7. Spettro IR simulato per il sistema GG1-79(H₂O) e relative frequenze⁴

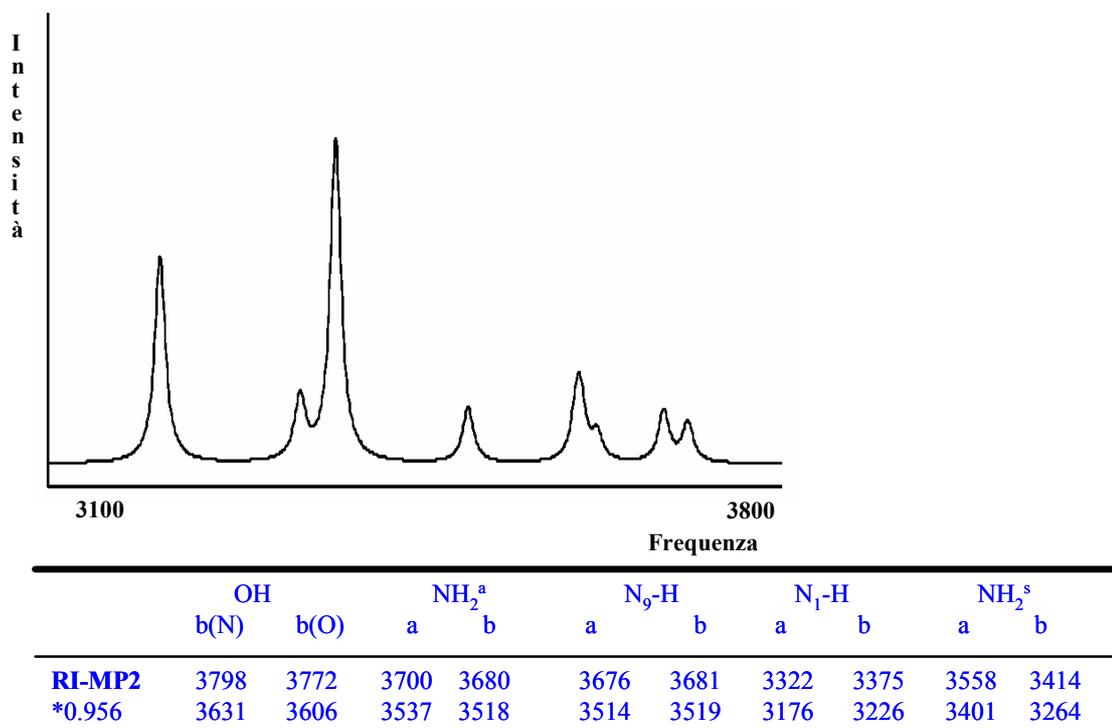


Figure 8. Spettro IR simulato per il sistema GG2-99(H₂O) e relative frequenze

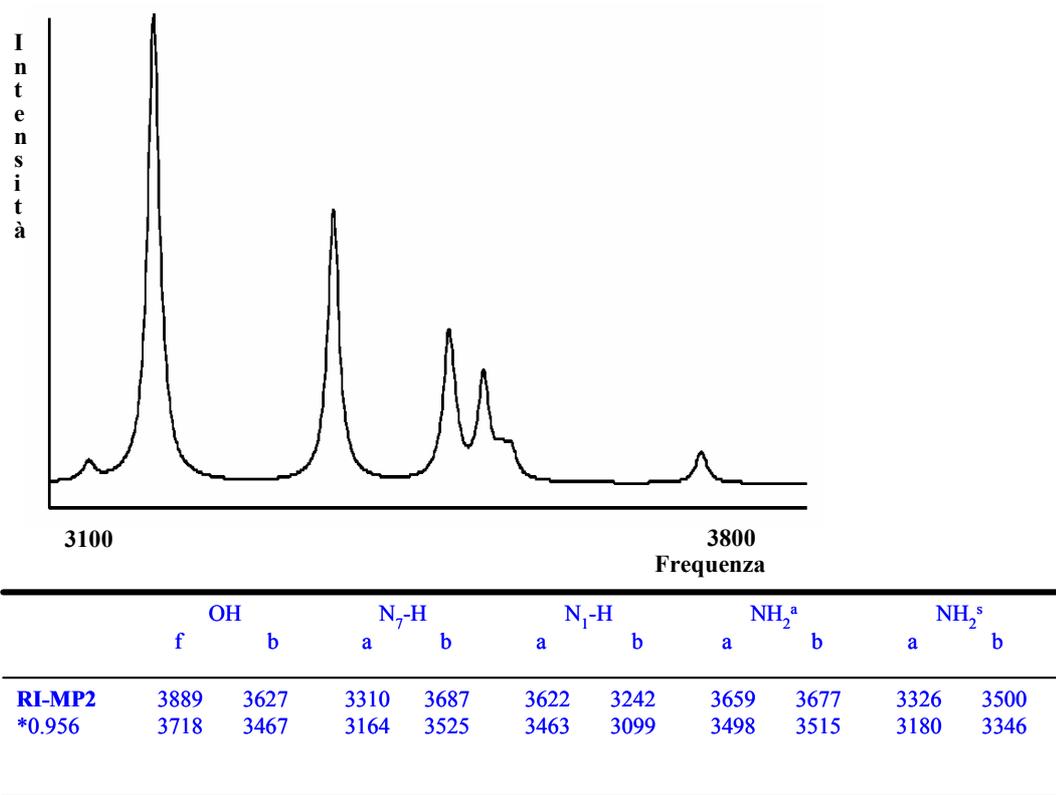
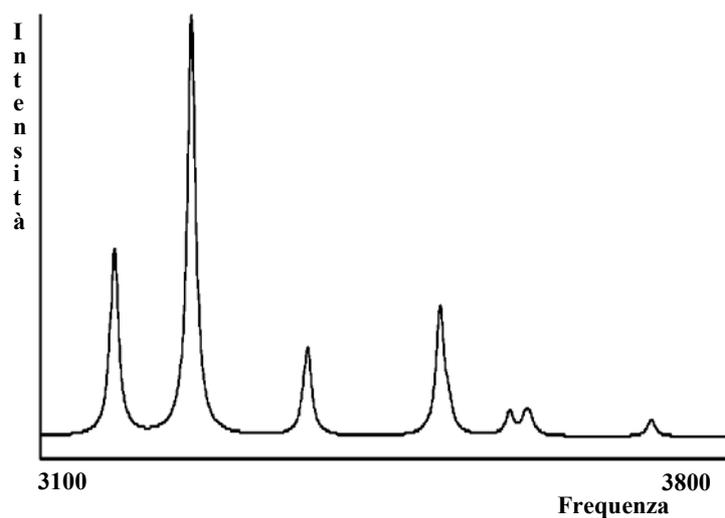
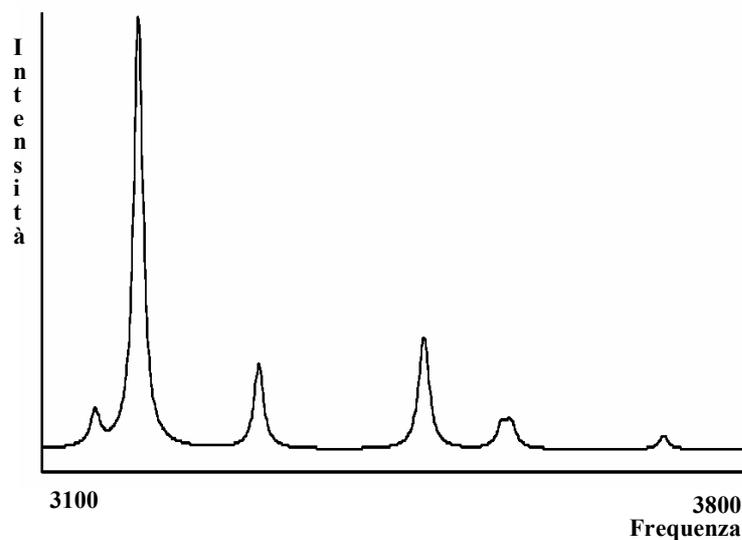


Figure 9. Spettro IR simulato per il sistema GG1-77(H₂O) e relative frequenze



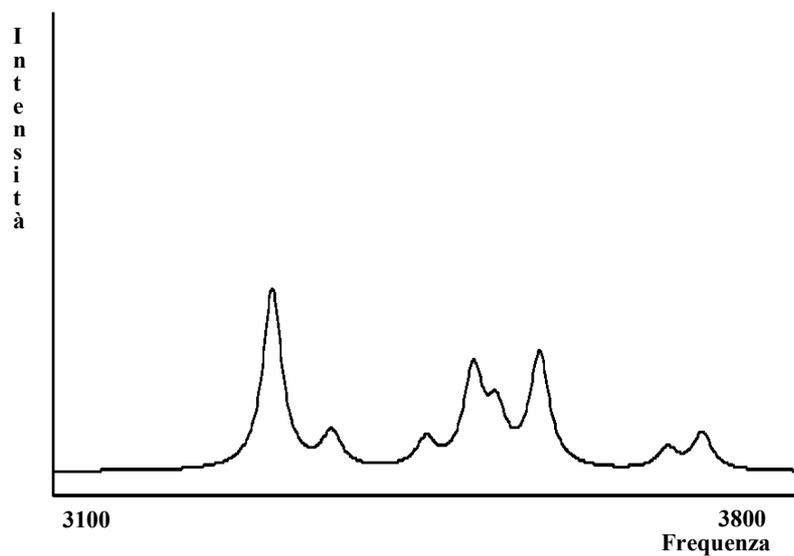
	OH		NH ₂ ^a		N ₉ -H	N ₇ -H	N ₁ -H		NH ₂ ^s	
	f	b	a	b			a	b	a	b
RI-MP2	3886	3579	3710	3702	3680	3387	3107	3219	3592	3591
*0.956	3715	3421	3547	3539	3518	3238	2970	3077	3434	3433

Figure 10. Spettro IR simulato per il sistema GG2-79(H₂O) e relative frequenze



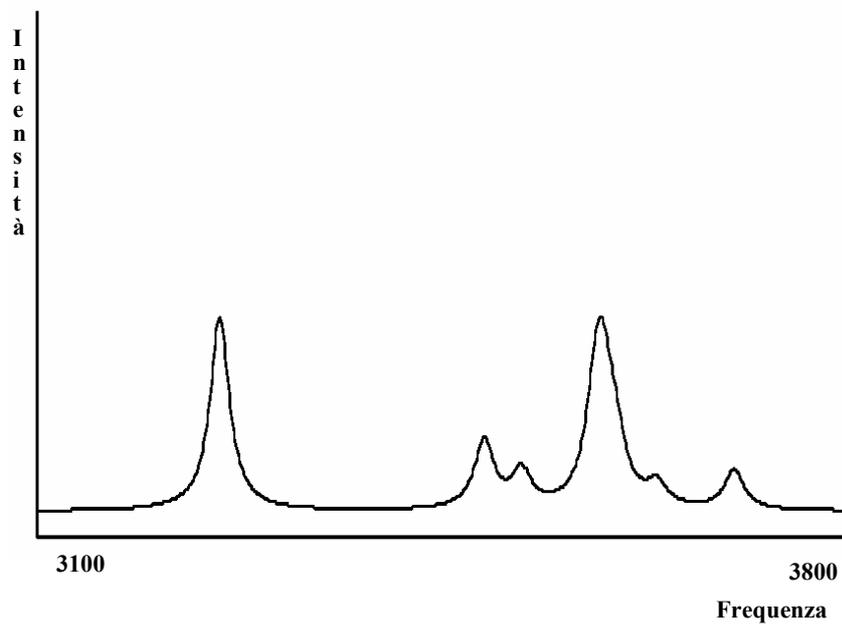
	OH		NH ₂ ^a		N ₇ -H		N ₁ -H		NH ₂ ^s	
	f	b	a	b	a	b	a	b	a	b
RI-MP2	3888	3587	3696	3691	3375	3683	3167	3222	3581	3583
*0.956	3717	3429	3533	3539	3528	3521	3027	3080	3423	3425

Figure 11. Spettro IR simulato per il sistema GG2-77(H₂O).



	OH		N ₉ -H		N ₁ -H		NH ₂ ^a		NH ₂ ^s		
	a	s	up	down	up	down	up	down	up	down	
RI-MP2	3814	3784	3673	3672	3615	3615	3635	3574	3492	3440	
*0.956	3646	3617	3511	3510	3456	3455	3475	3379	3305	3289	
Exp ^a	3735										

Figure 12. Spettro IR simulato per il sistema N(H₂O) e relative frequenze⁴



	OH		N ₉ -H		N ₁ -H		NH ₂ ^a		NH ₂ ^s		
	f	b	up	down	up	down	up	down	up	down	
RI-MP2	3799	3688	3683	3679	3584	3357	3732	3698	3616	3584	
*0.956	3632	3526	3521	3517	3426	3209	3568	3535	3457	3426	
Exp ^a	3735										

Figure 13. Spettro IR simulato per il sistema I(H₂O) e relative frequenze⁴

3.5.3 Sistemi hydrogen bonded, stacked e T-forms biidratati.

Le diverse disposizioni reciproche delle due molecole di acqua sui vari dimeri della GG conducono ad un numero molto elevato di forme biidratate.

Le energie relative di tutti i possibili minimi sono riportate in tabella 3

Il ruolo della seconda molecola di acqua è chiaramente evidenziato dalla nuova sequenza di stabilità. Ad eccezione dei dimeri nella forma T, che risultano sempre molto meno stabili di tutti gli altri, si può ora notare come alcuni sistemi stacked risultino decisamente stabilizzati.

Tabella 3: Energie assolute (in au) e relative (in kcal/mol) dei vari dimeri biidratati della coppia di base GG.

	$E_{\text{SCF+MP2}}$	ΔE
I(H₂O)₂	-1234.688801	0.0
GG1-79(H₂O)₂	-1234.684405	2.8
GG1-79(H₂O)₂-1	-1234.681140	4.8
GG1-79(H₂O)₂-2	-1234.680561	5.2
GG1-79(H₂O)₂-3	-1234.679000	6.1
GG1-79(H₂O)₂-4	-1234.676090	8.0
GG1-77(H₂O)₂	-1234.685710	1.9
GG1-77(H₂O)₂-1	-1234.680740	5.1
GG1-77(H₂O)₂-2	-1234.680410	5.3
GG1-77(H₂O)₂-3	-1234.675360	8.4
GG2-79(H₂O)₂	-1234.685343	2.2
GG2-79(H₂O)₂-1	-1234.677876	6.8
GG2-79(H₂O)₂-2	-1234.675221	8.5
GG2-99(H₂O)₂	-1234.684220	2.9
GG2-99(H₂O)₂-1	-1234.679710	5.7
GG2-77(H₂O)₂	-1234.681880	4.3
GG2-77(H₂O)₂-2	-1234.674550	8.9
N(H₂O)₂	-1234.685449	2.1
N(H₂O)₂-1	-1234.682360	4.0
N(H₂O)₂-2	-1234.679208	6.1
N(H₂O)₂-3	-1234.677397	7.2
I(H₂O)₂-1	-1234.683230	3.5
I(H₂O)₂-2	-1234.670280	11.6
TFORM(H₂O)	-1234.666300	14.1
TFORM2(H₂O)	1234.664040	15.6
TFORM1(H₂O)	-1234.663850	15.7

La struttura stacked invertita biidratata diventa quella favorita in assoluto ma ad energia leggermente superiore, in un range di meno di 3 kcal/mol, esistono un buon numero di sistemi planari e la forma stacked normale.

Le sei strutture selezionate giacenti in questo intervallo energetico sono illustrate in figura 14.

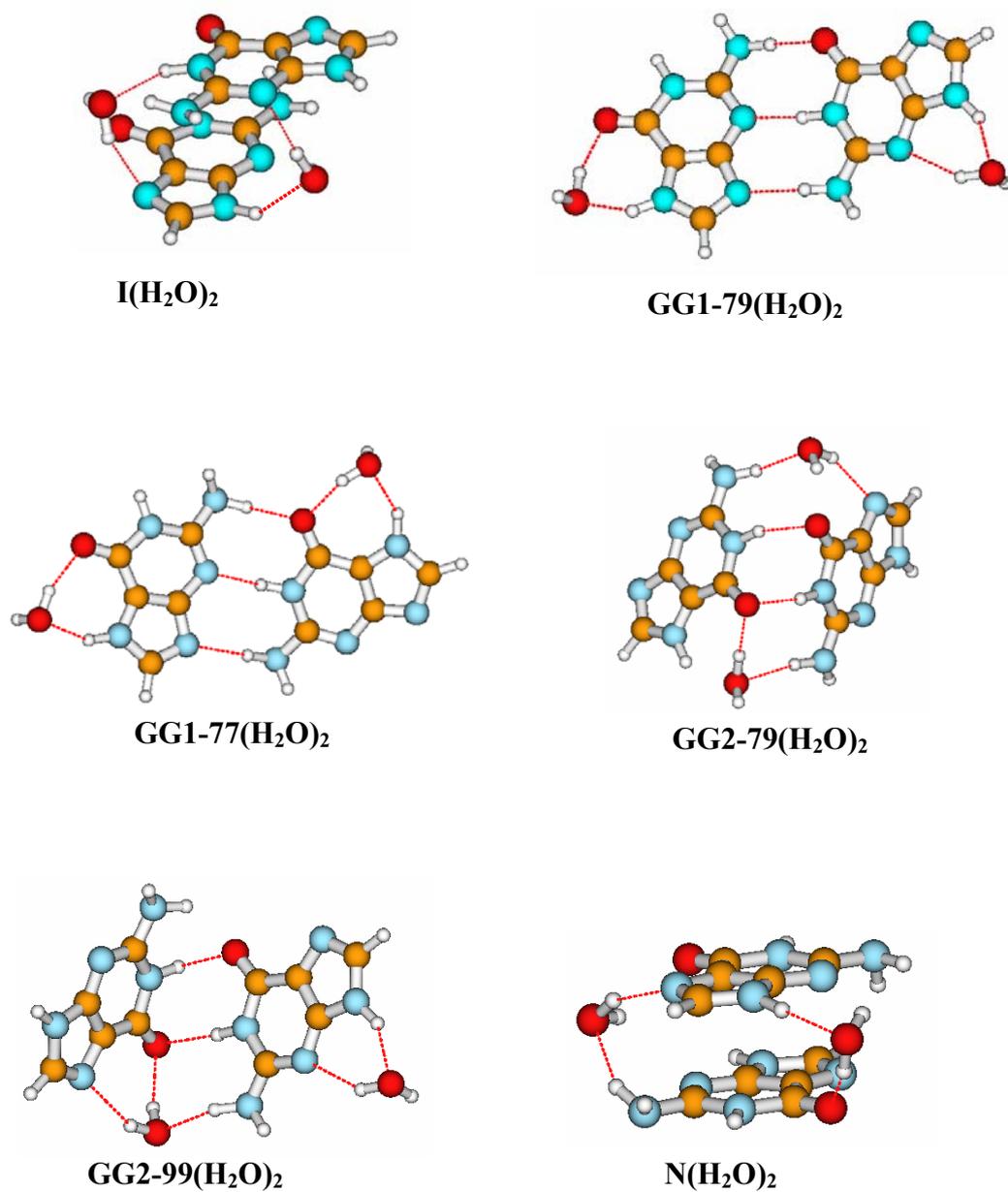
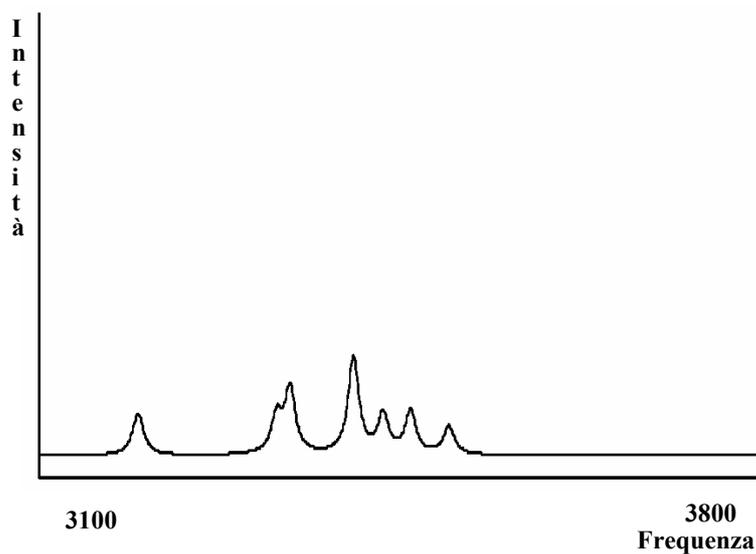


Figura 14. Strutture biidratate planari e stacked più stabili per la coppia di basi GG

I caratteri geometrici di questi dimeri biidratati a più bassa energia mostrano che non si può estrarre una informazione comune sui siti privilegiati dall'attacco delle due molecole di acqua. Ad esempio, la struttura stacked invertita $I(H_2O)_2$ propone dei siti di coordinazione insoliti rispetto a quelli solitamente coinvolti nelle forme monoidratate.

Le simulazioni degli spettri IR con la relativa assegnazione delle frequenze possono essere trovate nelle figure 15-20.



	OH(O)		OH		N ₉ -H		N ₁ -H		NH ₂ ^a		NH ₂ ^s	
	f	b	f	b	up	down	up	down	up	down	up	down
RI-MP2	3825	3688	3759	3707	3668	3479	3465	3613	3720	3656	3583	3550
*0.956	3657	3526	3594	3544	3507	3326	3312	3454	3556	3495	3425	3394

Figure 15. Spettro IR simulato per il sistema $I(H_2O)_2$ e relative frequenze

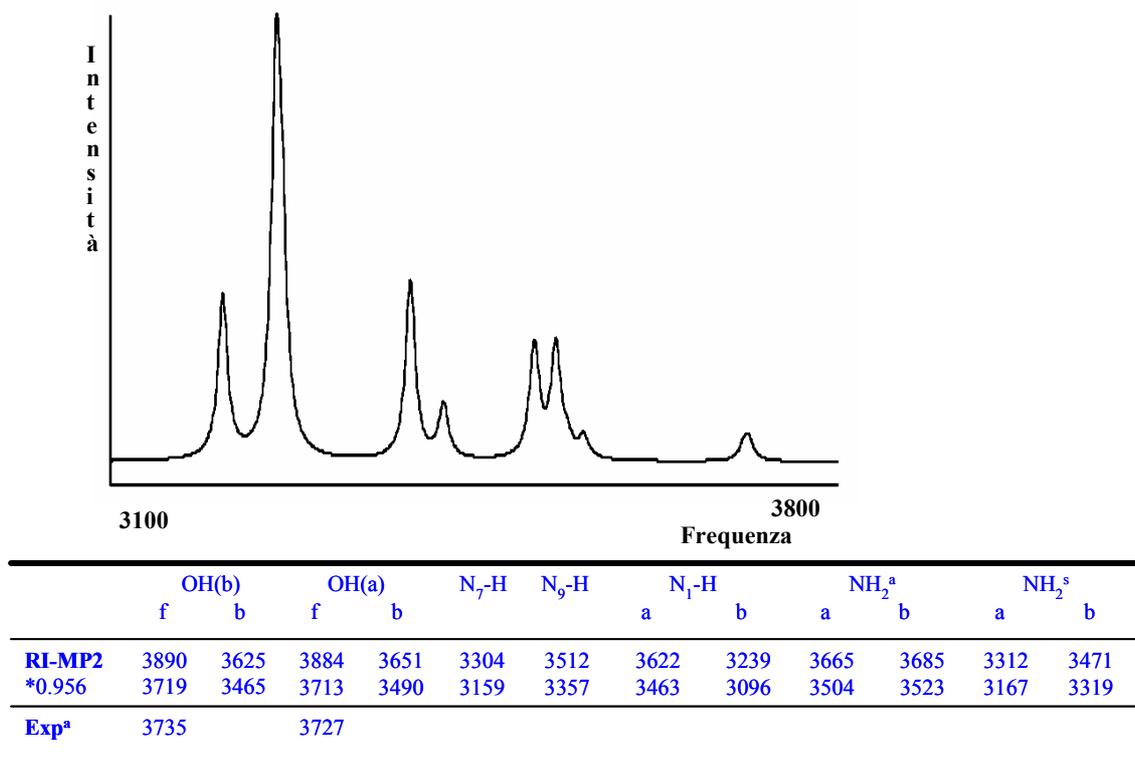


Figure 16. Spettro IR simulato per il sistema GG1-79(H₂O)₂ e relative frequenze⁴

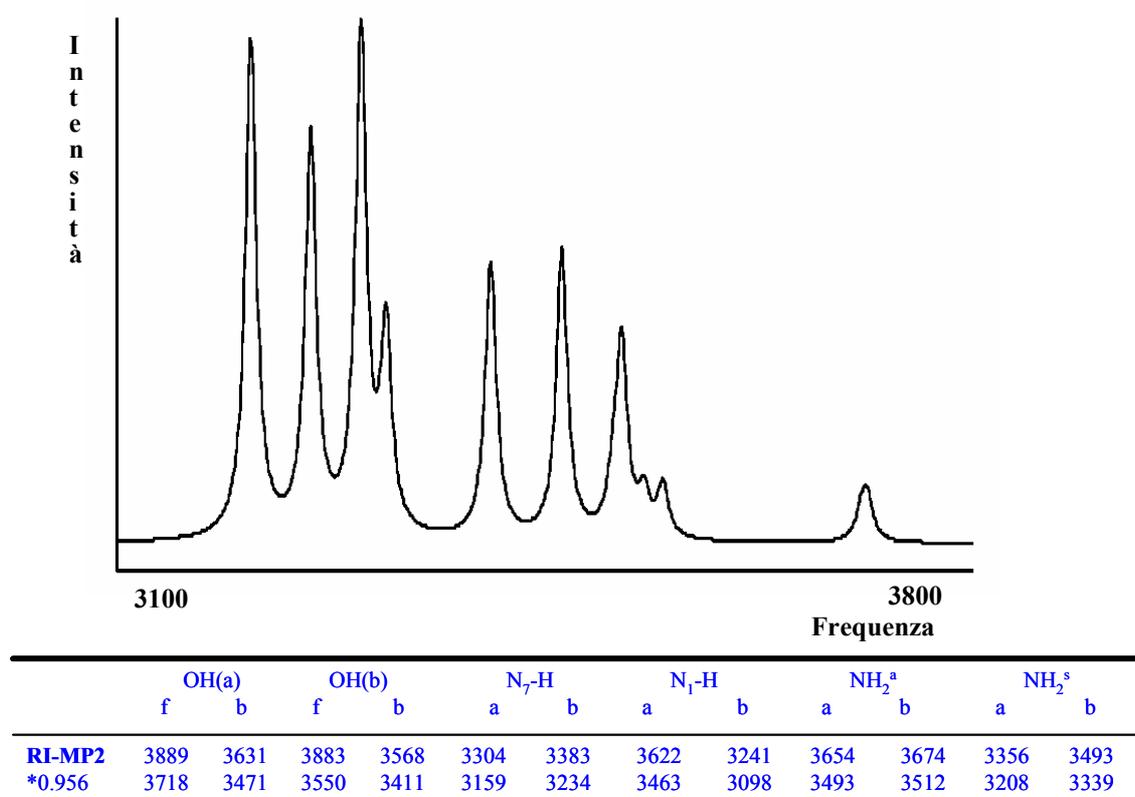


Figure 17. Spettro IR simulato per il sistema GG1-77(H₂O)₂ e relative frequenze

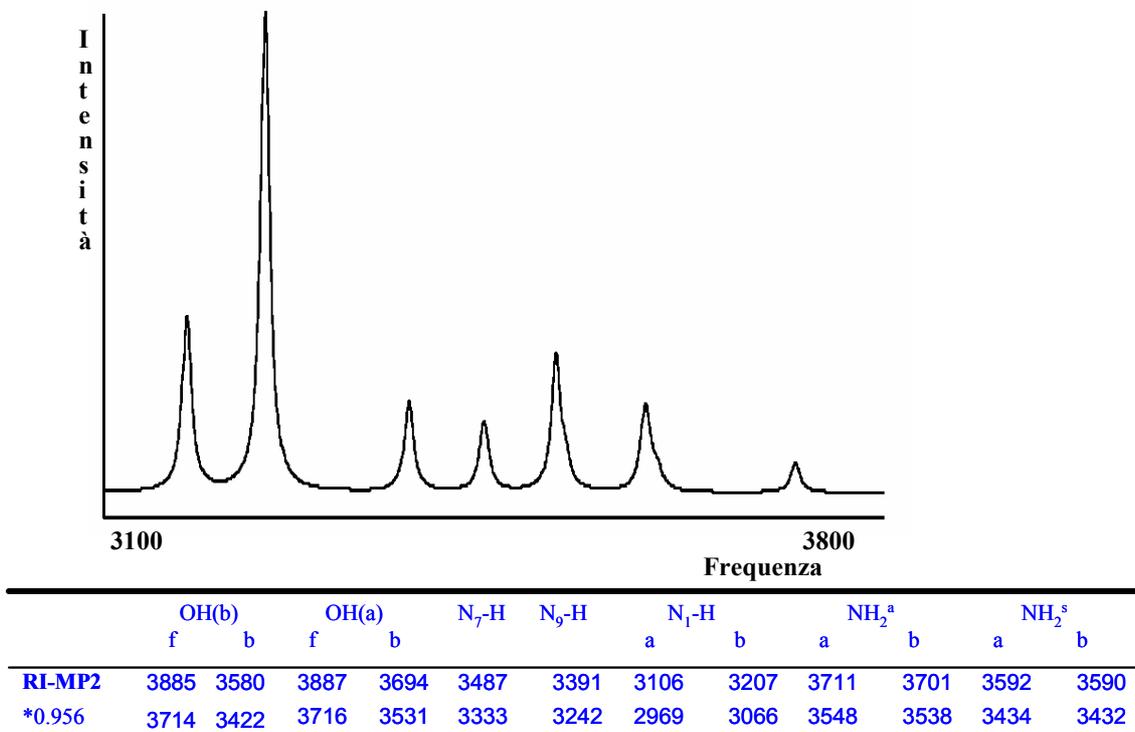


Figure 18. Spettro IR simulato per il sistema GG2-79(H₂O)₂ e relative frequenze

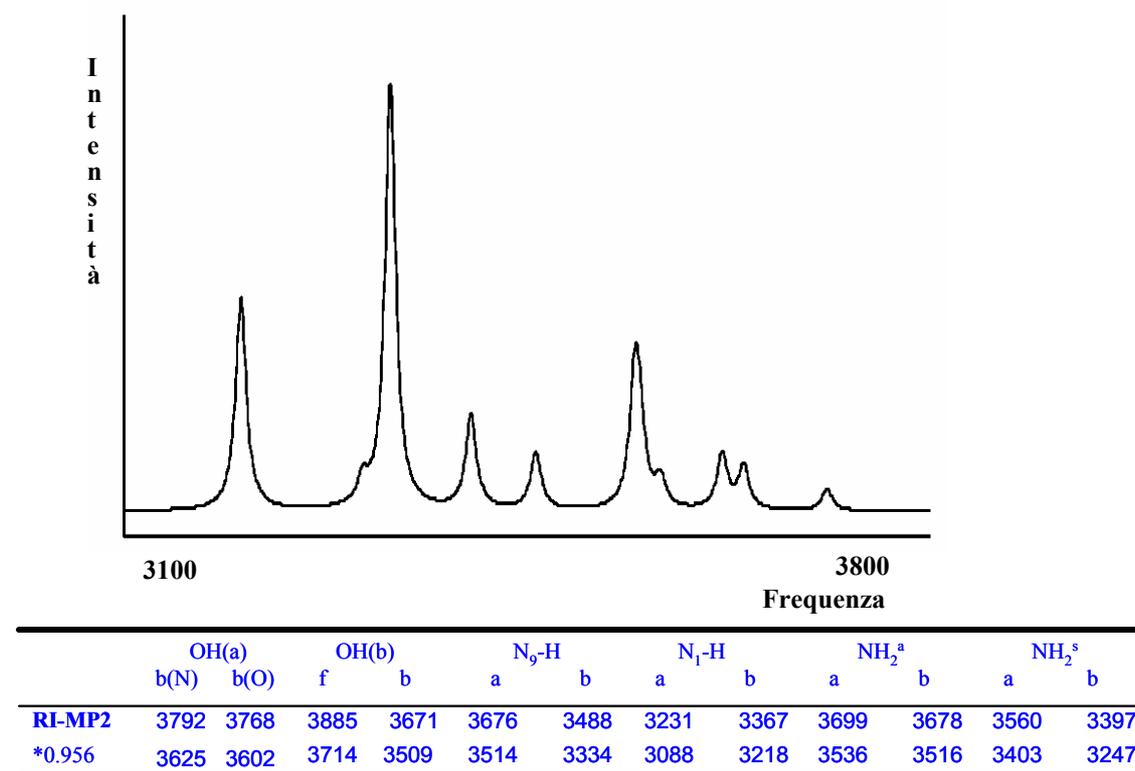
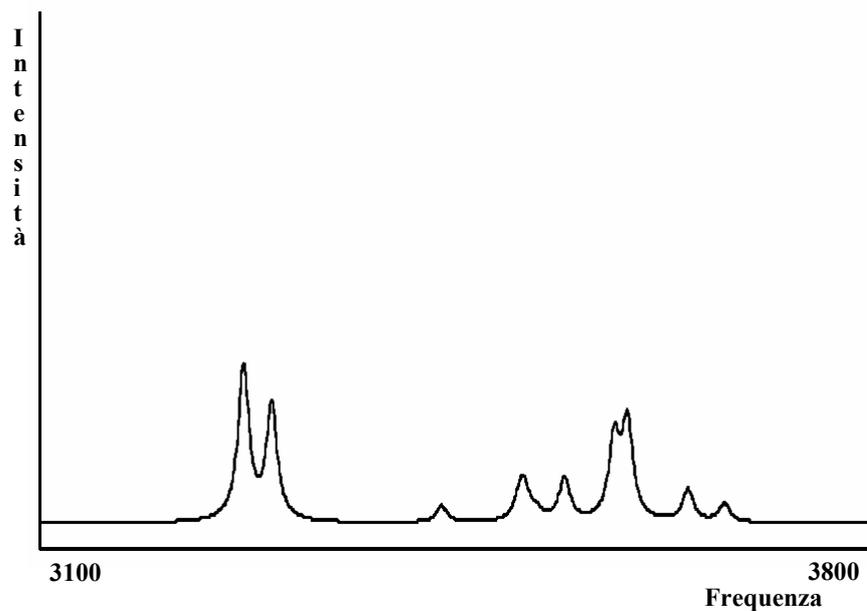


Figure 19. Spettro IR simulato per il sistema GG2-99(H₂O)₂ e relative frequenze



	OH(O)		OH		N ₉ -H		N ₁ -H		NH ₂ ^a		NH ₂ ^s	
	f	b	f	b	up	down	up	down	up	down	up	down
RI-MP2	3802	3723	3842	3737	3320	3669	3625	3620	3622	3638	3535	3351
*0.956	3635	3559	3673	3573	3174	3507	3465	3461	3462	3478	3380	3204

Figure 20. Spettro IR simulato per il sistema $N(H_2O)_2$ e relative frequenze

Se le nostre previsioni teoriche sono attendibili, c'è una grande probabilità che i campioni della coppia di base clusterizzati con due molecole di acqua, usati nelle determinazioni sperimentali, possano essere costituiti da una miscela di tutte le forme di energia comparabile.

Questo sottolinea l'importanza di disporre delle varie simulazioni spettrali IR per l'assegnazione della struttura prevalente.

Bibliografia

- [1] Zatta, P.; Canadese, C.; Costantini, S.; Gallieni, M.; Andriani, M.; Berthon, G.; Boggio-Bertinet, D.; Domingo, J.; Flaten, T.; Golub, V.; Goto, N.; Kawahara, M.; Kiss, T.; Lukiw, W.; Markesbery, W.; Milacic, R.; Ronco, C.; Sandstead, H.; Taylor, A. *Interdisciplinary Approach to the study of Aluminium Toxicity*, **2000**.
- [2] Giordo, P. *Vaccini al veleno*, Incanta: webmagazine di carattere informativo, **2003**.
- [3] Rubini, P., Lakatos, A.; Champmartin, D.; Kiss, T. *Coord. Chem. Rev.*, **2002**, 228, 137.
- [4] Vallee, B.; Falchuk, K. H. *Physiol. Rev.*, **1993**, 73, 79.
- [5] Kang, S.; Wells, R. D. *J. Biol. Chem.*, **1994**, 269, 9528.
- [6] Rodgers, M. T.; Armentrout, P. B. *JACS*, **2000**, 122, 8548.
- [7] Cerda, B. A.; Wesdemiotis, C. *JACS*, **1996**, 118, 11884.
- [8] Russo, N.; Sicilia, E.; Toscano, M.; Grand, A. *J. Int. J. Quant. Chem*, **2002**, 90, 903.
- [9] Marino, T.; Russo, N.; Toscano, M.; Grand, A. *Int. J. Quant. Chem*, **2004**, 98, 347.
- [10] Russo, N.; Toscano, M.; Grand, A. *J. Mass Spectrom*, **2003**, 38, 265.
- [11] Russo, N.; Toscano, M.; Grand, A. *J. Phys. Chem. B*, **2001**, 105, 4735.
- [12] Russo, N.; Toscano, M.; Grand, A. *JACS*, **2001**, 123, 10272.
- [13] Russo, N.; Toscano, M.; Grand, A. *J. Phys. Chem. A*, **2003**, 107, 11533.
- [14] Ahn, D.-S.; Lee, S.; Kim, B. *Chem. Phys. Lett.*, **2004**, 390, 384.
- [15] Nir, E.; Müller, M.; Grace, L. I.; de Vries, M. S. *Chem. Phys. Lett*, **2000**, 317, 437.
- [16] Gorb, L.; Leszczynsky, J. *Int. J. Quant. Chem.*, **1998**, 70, 855.
- [17] Alemán, C. *Chem. Phys.*, **2000**, 253, 13.
- [18] Kratochvíl, M.; Šponer, J.; Hobza, P. *JACS*, **2000**, 122, 3495.
- [19] Šponer, J.; Leszczynsky, J.; Hobza, P. *J. Phys. Chem.*, **1996**, 100, 1965.
- [20] Saenger, W. *Principle of Nucleic Acid Structure*, Springer-Verlag: New York, **1984**.
- [21] Yanson, I. K.; Teplitsky, A. B.; Sukhodub, L. F. *Biopolym.*, **1979**, 18, 1149.
- [22] Miller, J. H.; Sobell, H. M. *J. Mol. Biol.*, **1967**, 24, 345.
- [23] Binford, J. S.; Holloway, D. M. *J. Mol. Biol.*, **1968**, 31, 91ù
- [24] Katz, L.; Penman, S. *J. Mol. Biol.*, **1966**, 15, 220.

- [25] Kyogoku, Y.; Lord, R. C.; Rich, A. *Science*, **1966**, *154*, 518.
- [26] Solie, T. N.; Schellman, J. A. *J. Mol. Biol.*, **1968**, *33*, 61.
- [27] Imoto, T. *Biochim. Biophys. Acta*, **1977**, *475*, 409.
- [28] Kim, S. K.; lee, W.; Herschback, D. H. *J. Phys. Chem.*, **1996**, *100*, 7933.
- [29] Kang, H.; Lee, K. T.; Kim, S. K. *Chem. Phys. Lett.*, **2002**, *359*, 213.
- [30] He, Y. G.; Wu, C. Y.; Kong, W. J. *J. Phys. Chem. A*, **2004**, *108*, 943.
- [31] Piuzzi, F.; Mons, M.; Dimicoli, L.; Tardivel, B.; Zhao, Q. *Chem. Phys.*, **2001**, *270*, 205.
- [32] Chin, W.; Mons, M.; Piuzzi, F.; Tardivel, B.; Dimicoli, L.; Gorb, L.; Leszczynsky, J. *J. Phys. Chem. A*, **2004**, *108*, 8237.
- [33] Kabeláč, M.; Ryjáček, F.; Hobza, P. *Phys. Chem. Chem. Phys.*, **2000**, *2*, 4906.
- [34] Kabeláč, M.; Hobza, P. *Chem. Eur. J.*, **2001**, *7*, 2067.
- [35] Hanus, M.; Ryjáček, F.; Kabeláč, M.; Kubař, T.; Bogdan, T. V.; Trygubenko, S. A.; Hobza, P. *JACS*, **2003**, *125*, 7678.
- [36] Shishkin, O. V.; Sukhanov, O. S.; Gorb, L.; Leszczynsky, J. *Phys. Chem. Chem. Phys.*, **2002**, *4*, 5359.
- [37] Zendlová, L.; Hobza, P.; Kabeláč, M. *Chem. Phys. Chem.*, **2006**, *7*, 439.
- [38] Kabeláč, M.; Hobza, P. *Chem. Phys. Phys. Chem.*, **2007**, *9*, 903.
- [39] Wolf, B.; Hanlon, S. *Biochemistry*, **1975**, *14*, 1661.
- [40] Branders, R.; Rupprecht, A.; Kearns, D. R. *J. Biophys.*, **1989**, *56*, 683.
- [41] Hobza, P.; Šponer, J. *Chem. Rev.*, **1999**, *99*, 3247 e riferimenti.
- [42] Abo-Riziq, A.; Crews, B.; Grace, L.; de Vries, M. S. *JACS*, **2005**, *127*, 2374.
- [43] Tsuzuki, S.; Honda, K.; Uchimaru, T.; Mikami, M.; Tanabe, K. *JACS*, **2000**, *122*, 3746.
- [44] Pittner, J.; Hobza, P. *Chem Phys. Lett.*, **2004**, *390*, 496.
- [45] Møller, C.; Plesset M. S. *Phys. Rev.* **1934**, *46*, 618.
- [46] Bartlett, R. J.; Silver, D. M. *J. Chem. Phys.*, **1975**, *62*, 3258.
- [47] Pople, J. A.; Binkley, J. S.; Seeger, R., *Int. Quant. Chem. Symp.*, **1976**, *10*, 1.
- [48] Hehre, W. J.; Schleyer, P. V. R.; Pople, J. A. *Ab initio molecular orbital theory*, **1986**, Wiley, New York.
- [49] Schäfer, A.; Horn, H.; Ahlrichs, R. *J. Chem. Phys.*, **1992**, *97*, 2571.
- [50] Schäfer, A.; Huber, C.; Ahlrichs, R. *J. Chem. Phys.*, **1994**, *100*, 5829.
- [51] Eichkorn, K.; Weigend, F.; Treutler, O.; Ahlrichs, R. *Theor. Chem. Acc.*, **1997**, *97*, 119.

[52] Nir, E.; Janzen, C.; Imhof, P.; Kleinermanns, K.; de Vries, M. S. *Phys. Chem. Chem. Phys.*, **2002**, *4*, 740.

ARTICOLO 1

On the interaction of bare and hydrated aluminium ion with nucleic acid bases (U, T, C, A, G) and monophosphate nucleotides (UMP, dTP, dCMP, dAMP, dGMP)

On the Interaction of Bare and Hydrated Aluminum Ion with Nucleic Acid Bases (U, T, C, A, G) and Monophosphate Nucleotides (UMP, dTMP, dCMP, dAMP, dGMP)

Donatella Mazzuca,[†] Nino Russo,^{*,†} Marirosa Toscano,[†] and André Grand[‡]

Dipartimento di Chimica and Centro di Calcolo ad Alte Prestazioni per Elaborazioni Parallele e Distribuite-Centro d'Eccellenza MURST, Università della Calabria, I-87030 Arcavacata di Rende (CS), Italy, and Département de Recherche Fondamentale sur la Matière Condensée, Service de Chimie Inorganique et Biologique, CEA-Grenoble, 17 Rue des Martyrs, 38054 Grenoble Cedex 9, France

Received: September 15, 2005; In Final Form: March 16, 2006

The B3LYP/6-311+G(2df,2p) density functional approach was used to study the interaction that aluminum trication, in the bare and hydrated forms, establishes with the nucleic acid bases and the corresponding monophosphate nucleotides. In this investigation, we determine equilibrium geometry of all possible complexes resulting from the attachment of the ion on the different binding sites selected on each ligand. The relative energies of complexes and metal ion affinities are also given. The most meaningful aspect was found to lie in the energetics of this interaction that underlines a very high affinity of aluminum ion for the examined biological systems.

1. Introduction

The interaction of metal cations with nucleic acids is topical in the bioinorganic field because of its possible effects on synthesis, replication, and structural integrity of DNA and RNA.^{1–4}

The interaction with phosphate group was more frequently observed, although evidence exists also for the direct attack of some ions on purine and pyrimidine bases. Generally, there are not simple rules to use to determine if an ion prefers to bind to these bases rather than to phosphate, but it is known that the binding to the last group, contrarily to what occurs when the metal ions bind to the bases, tends to stabilize the DNA double helix.

In the recent literature, numerous studies in which alkali,^{5–16} alkaline earth,^{17–22} and some transition mono- and divalent metal ions^{20,23–37} interact with DNA and RNA nucleobases or base pairs are reported. Much less work was devoted to trivalent cations, such as Al(III), and their triply charged biological complexes.^{38–44}

The toxicity of aluminum (III) was presumed 30 years ago.⁴⁵ Since this date, many authors^{46,47} suggested that an high concentration of Al³⁺ in the human body, can be responsible for a renal malfunction that causes further pathologies such as anaemia, and encephalopathy. Furthermore, Al³⁺ was hypothesized to have a role in the onset of Alzheimer's disease,⁴⁸ amyotrophic lateral sclerosis, and Down's syndrome.⁴⁹ Uptake and retention of aluminum in tissues are favored by zinc deficiency⁵⁰ that is a possible condition of pregnant women, children, and the elderly.⁵¹ In the studies on the aqueous solution of some metallic ions, Birchall⁵² suggested that silicon deficiency also can enhance the uptake and toxicity of aluminum. Al³⁺ is able to compete with other metal ions, such as magnesium and iron, for preferential coordination to biological systems because of its high Lewis acidity and slow exchange of the ligands.^{53–55} For instance, previous investigations demonstrated that the replacement of Mg²⁺ with Al³⁺ has an

inhibitory effect on the enzyme-induced phosphate transfer reactions.⁵⁴ The localization of aluminum ion in cells occurs preferentially in the neuron nuclei, and affects the conformational variations of DNA necessary for genetic expression.⁵⁶ Some work, concerning the interaction of aluminum ion with biological systems, suggested its great affinity for phosphate in DNA.⁵⁷ Because of its hard acid nature, aluminum ion reacts favorably with hard bases that are generally oxygen-containing ligands.⁴¹ Martin,⁵⁸ in contrast to the previous opinions, proposed that, in the cell nucleus, Al³⁺ binds to the ATP rather than to DNA. In the investigation of Zhang et al.,⁵⁹ the interaction between Al³⁺ and DNA was studied by UV spectroscopy and cyclic voltammetry (CV) under neutral pH conditions using [Co(phen)₃]^{3+/2+} as indicator. The results of this study showed that the binding site for the aluminum cation is the phosphate group of the DNA backbone and that it binds to the nucleic acid much more strongly than the indicator. The large affinity of Al³⁺ for DNA was confirmed by the UV spectrum. An electrostatic nature of the binding of Al³⁺–DNA was also supposed.

No theoretical analysis was attempted until now to describe the interaction between Al³⁺ and DNA or RNA nucleobases and monophosphate nucleotides.

In this paper, we report the first computational study in this respect. The aim of this work is the determination of structural, conformational, and energetic properties of the complexes that Al³⁺ ion forms with the ligands reported in the Chart 1. The aluminum ion affinity values for these systems will be useful to deduce which nucleobase is more susceptible to the ion attack and which is the preferred binding site on the ligands. Furthermore, this work will provide a contribution to clarify why aluminum may be toxic for cells.

2. Computational Details

We have performed a density functional (DF) study of all the species involved in the reactions

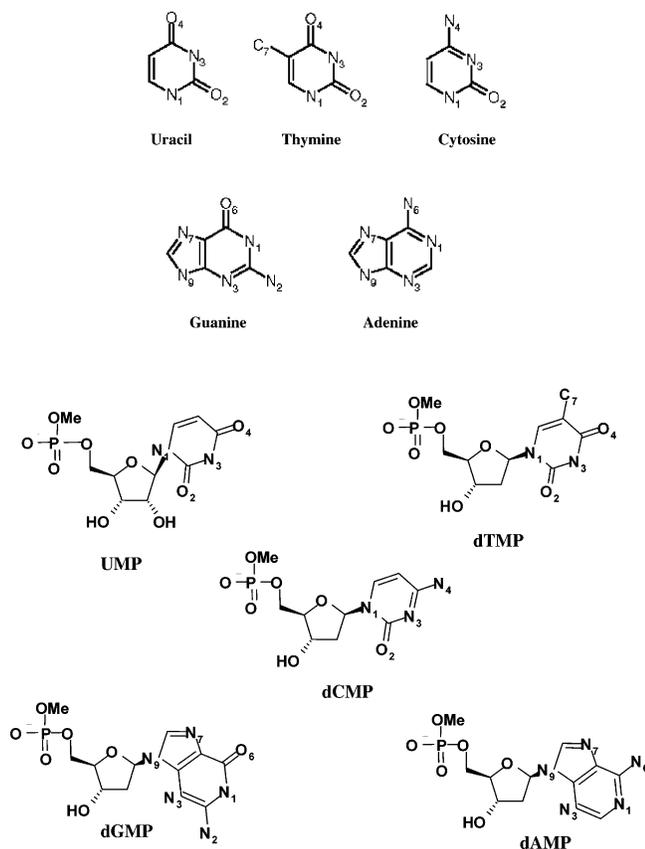


* To whom correspondence should be sent. E-mail: nrusso@unical.it.

[†] Università della Calabria.

[‡] CEA-Grenoble.

CHART 1



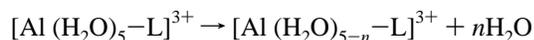
where L = uracil (U), thymine (T), cytosine (C), adenine (A), guanine (G), uridine monophosphate (UMP), deoxythymidine monophosphate (dTMP), deoxycytidine monophosphate (dCMP), deoxyadenosine monophosphate (dAMP), and deoxyguanosine monophosphate (dGMP).

All computations were done using Gaussian 03 code⁶⁰ with the hybrid Becke exchange (B3)⁶¹–Lee, Yang, and Parr correlation (LYP)⁶² functionals. Optimizations were performed with the orbital 6-311G**⁶³ triple- ζ type set. Harmonic vibrational frequencies were computed, at this same level of theory, to verify the nature of minima of obtained complexes. A valence triple- ζ 6-311+G(2df,2p) basis set that contains one set of diffuse and two sets of polarization functions on every atom⁶³ was used for a better estimate of energies. Zero point energy (ZPE) corrections were enclosed in all energetic data.

Aluminum ion affinity (AIIA) was evaluated as negative variation of ΔH for reaction 1 and reported at 298 K.

The strategy adopted in the calculations with solvated aluminum was suggested by the results obtained in the initial calculations with the bare cation in which we have ascertained the existence of several mono- and bicoordinated Al^{3+} -L adducts for each nucleobase and nucleotide.

In the supposition that hydrated aluminum is also at least monocoordinated to the particular ligand, five water molecules were considered to complete its inner coordination sphere. This approach allowed us to determine the possible variation of the coordination mode on going from bare to hydrated metal ion. In fact, water molecules and binding sites on the ligand have the possibility to compete in the coordination process of aluminum ion, on the basis of the following reaction scheme:

TABLE 1: B3LYP/6-311+G(2df,2p) Absolute (E in a.u.) and Relative (ΔE in kcal/mol) Energies for Free Tautomers of Nucleobases and Their Complexes with Al^{3+} Cation at 0 K

	L		$\text{Al}^{3+}\text{-L}$			
	E	ΔE	E	ΔE	AIIA ^a	
U1	-414.884518	0.0	U1Al	11.7	424.5	
			U1Al-ap	35.0		
			U2Al	6.0		
U2		11.7	U2Al	6.0		
U3		18.9	U3Al	0.0		
T1	-454.188940	0.0	T1Al	-695.335482	0.0	466.9
			T1Al-ap	54.5		
			T2Al	15.3		
T2		12.6	T2Al	15.3		
T3		18.3	T3Al	12.3		
C1	-394.980359	0.0	C1Al	-636.132818	0.0	470.6
			C1Al-ap	71.2		
			C2Al	38.2		
C2		1.3	C2Al	38.2		
C3		2.1	C3Al	22.0		
C4		2.3	C4Al	32.6		
G1	-542.631348	0.0	G1Al	40.2		
			G2Al	-783.864968	0.0	522.2
			G3	38.0		
G2		0.7	G2Al	-783.864968	0.0	522.2
G3		1.8	G3Al	38.0		
G4		2.3	G4Al	76.0		
G5		4.4	G5Al	28.8		
A1	-467.369809	0.0	A1Al	50.3	457.2	
			A2Al	23.7		
			A3Al	-708.581175	0.0	
A2		8.1	A2Al	23.7		
A3		18.5	A3Al	-708.581175	0.0	

^a Aluminum ion affinity values (AIIA in kcal/mol) are given at 298 K.

where n takes the value of 0 in the presence of monocoordination, and the value of 1 if the hydrated ion is bicoordinated.

NBO analysis⁶⁴ was performed for all obtained complexes.

3. Results and Discussion

3.1. Interaction with Bare Cation. *3.1.1. Nucleobases Complexes.* An interesting aspect of nucleobases is their tendency to tautomerize.

In agreement with experimental findings,^{65–69} previous theoretical works^{70–79} demonstrated, especially for cytosine and guanine nucleobases, that many tautomeric forms have an energy that falls in a narrow range of values. The computations concerning this issue are reported in Table 1 in the first two columns.

The four cytosine tautomers, as well as the five for guanine, and the three for uracil, thymine, and adenine, which we have selected (see Figure 1), are the most stable forms. The literature data concerning the evaluation of energetic barriers for the tautomerization of nucleobases⁶ in which proton shifts are usually involved, underline the difficulty of the interconversion between all tautomeric forms. However, it has been proven experimentally^{65–69} that C1 and C2 tautomers of cytosine and G1 and G2 of guanine coexist. For uracil, thymine, and adenine, the rare tautomers (enol structures) L2 and L3 lie at energy values much higher with respect to the canonical form (keto structure) L1; consequently, U1, T1, and A1 are substantially prevalent irrespective of conditions. This is not the case of cytosine and guanine, for which the previous experience^{5,6,17,23–25} suggests that the tautomeric forms L1 and L2 (for cytosine also L3) could have a crucial role in determining the correct metal ion affinity value.

The coordination sites for the metal ion were chosen on the basis of literature data concerning the interactions of Al^{3+} or other cations with biomolecules.^{5,6,17,23–25,80} They involve the oxygen (both on the nucleobases and phosphate group) and nitrogen atoms for the monocoordination, and the $-\text{N}$, O^- , $-\text{O}$, O^- , and $-\text{N}$, N^- pairs of atoms for the bicoordination. The out-of-plane interaction between the cation and the aromatic

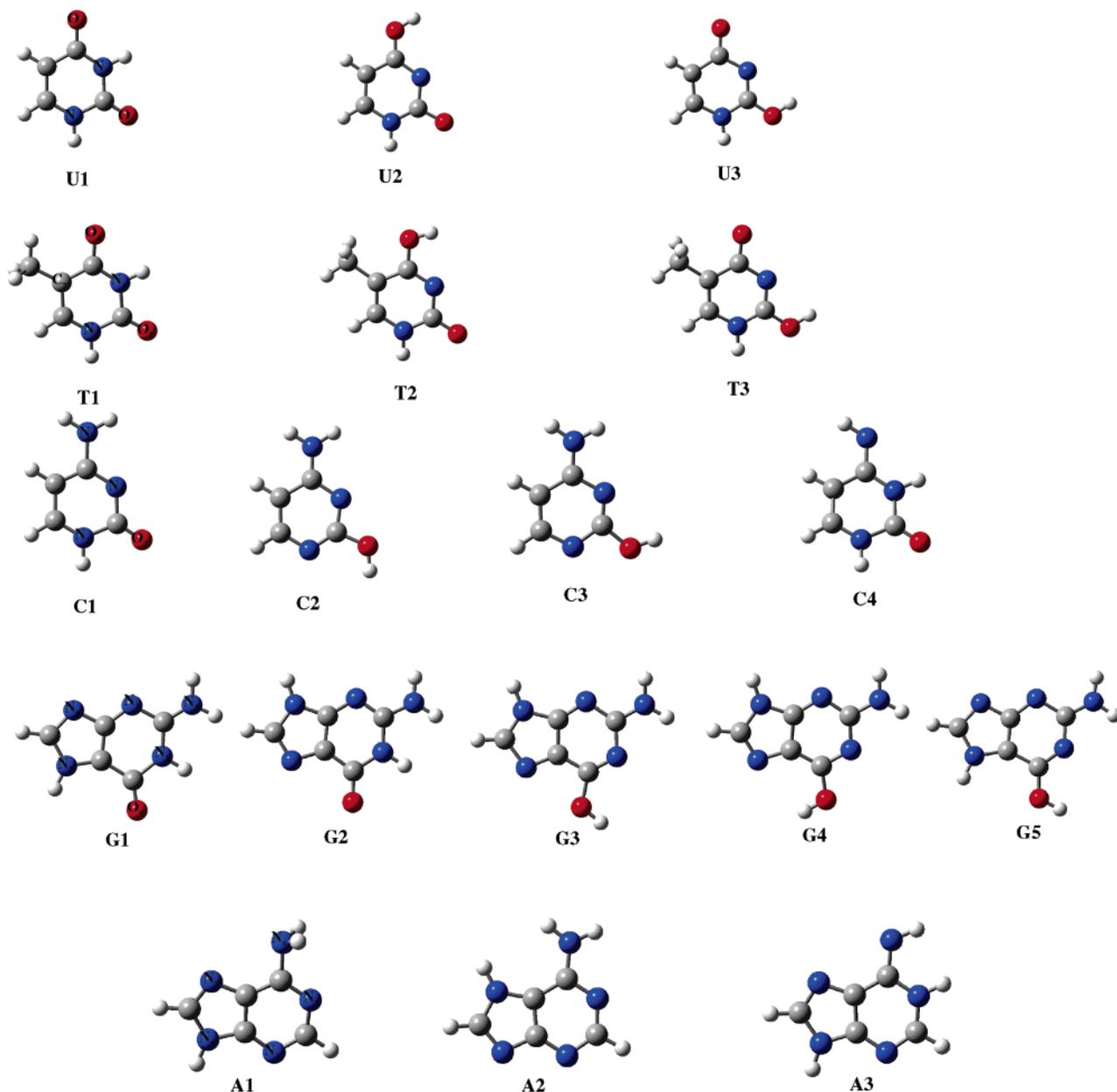


Figure 1. Low-lying tautomeric forms for each nucleobase.

rings was studied for both purines and pyrimidines, but π -complexes were obtained only in the cases of uracil, thymine, and cytosine. On the other hand, this type of interaction was never found for adenine and guanine.

A very large number of systems was examined in this phase of the study. In Figure 2, we have shown only the picture of the most likely complex Al^{3+} -L for each base. The other complexes and their main dimensions can be found in the Supporting Information (see Figure 1S). However, in Table 1, we have reported the energetic data concerning all complexes obtained with each tautomer of Figure 1.

Al^{3+} forms four complexes with uracil tautomers. The rare forms U2 and U3 give rise to the most stable compounds (U2Al, U3Al) in which the cation appears to be bicoordinated to N3 and to the ketonic oxygen on duty.

The attack of Al^{3+} on the π -system of U1 generates a complex (U1Alap) whose rearrangements after the optimization rule out a simple electrostatic interaction. In fact, aluminum cation inserts in the C-N3 bond changing the uracil cycle nature. However,

this system lies at very high energy (35 kcal/mol) with respect to the global minimum; thus, its formation is quite out of the question.

Taking into account that U1 is the energetically favored tautomeric form of uracil, we can indicate the U1Al system as the most probable one although it lies at 11.7 kcal/mol above the most stable complex U3Al. In U1Al, the cation is mono-coordinated to the O4 atom with a bond length of 1.780 Å. The Al-O4-C valence angle is 157°.

For thymine, the number of complexes was the same as that of uracil. The stability order of these systems is, however, quite different. The global minimum T1Al is a monocoordinated complex in which the Al-O4 distance is 1.741 Å and the Al-O4-C angle of 176° is closer to linearity than the analogous angle in U1Al.

The less favored adduct is again the apical complex T1Alap. It lies at 54.7 kcal/mol above the most stable one, and as in the case examined before, the aluminum cation inserts in the cycle. The formation of this system and that of the bicoordinated

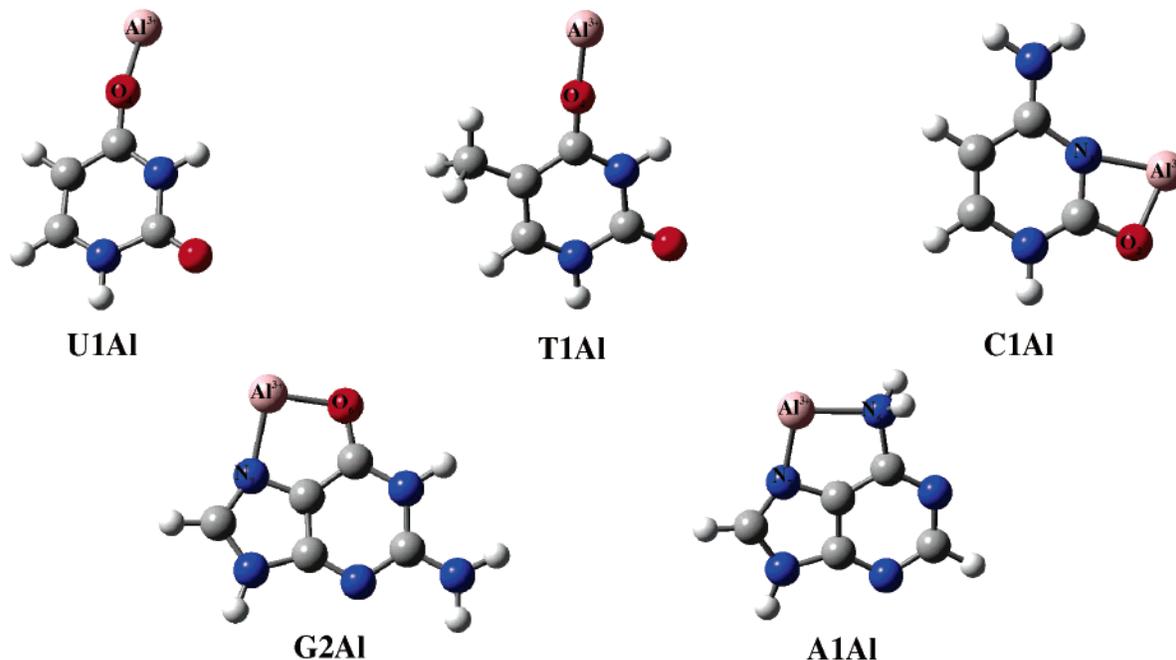


Figure 2. Equilibrium structures of the most likely complex $\text{Al}^{3+}\text{-L}$ for each nucleobase.

complexes T2AI and T3AI, lying at 15.5 and 12.4 kcal/mol, respectively, is again improbable.

As far as the Al^{3+} -cytosine complexes are concerned, we found the following stability order: $\text{C1AI} > \text{C3AI} > \text{C4AI} > \text{C2AI} \gg \text{C1Alap}$. As is evident from Table 1, the energetic separation between complexes is larger than that between free tautomers. This means that, despite the coexistence of some free tautomers, the aluminum affinity value for this base will depend only on the stability of metalated species.

In the most stable system C1AI, Al^{3+} cation is coordinated to both N3 and O2 atoms with bond lengths of 1.865 and 1.743 Å, respectively. This same result was previously obtained in a combined experimental-theoretical study on the interaction between aluminum atom and cytosine.⁸¹

The Al^{3+} in apical position on the C1 canonical tautomer destroys the cytosine ring planarity forming two σ -bonds with the N3 and C5 atoms (see Figure 1S). This rearrangement requires high energy (71.2 kcal/mol); thus, we can discard with certainty the real existence of this species.

The interactions of Al^{3+} with the π -system of pyrimidine bases are therefore purely fictitious, and the data obtained in this work do not confirm some literature findings that suggested this type of interaction for other metal ions.⁸²⁻⁸⁴

The aluminum-guanine complexes appear to be separated by large energy gaps.

The G2 tautomer forms the most stable complex G2AI. This species shows the cation coordinated to N7 and O6 with bond lengths of 1.939 and 1.819 Å, respectively.

The other canonical tautomer (G1) gives rise to an adduct that is located at 40 kcal/mol above the global minimum.

A previous experimental/theoretical work devoted to the determination of the neutral aluminum effect on the photochemistry of guanine⁸⁵ cannot be considered useful for a comparison because the neutral atom and ion behave differently.

The three tautomers of adenine interact with Al^{3+} cation, giving as many complexes.

As for other metal ions,^{5,17,23-25} the most stable system is formed with A3 owing to the presence of imino nitrogen that is highly reactive toward positively charged species. However, A3, and likewise the A2 tautomer, is decidedly less stable than

A1; thus, considering also the noticeable height of the inter-conversion barriers (45.86 and 38.78 kcal/mol are required for the $\text{A1} \rightarrow \text{A3}$ and $\text{A1AI} \rightarrow \text{A3AI}$ transformation processes, respectively), the most likely complex should be A1AI. In this last system, the aluminum is practically bicoordinated to N7 and N6 atoms with distances of 1.833 and 2.013 Å, respectively. Although the distance between Al^{3+} and N6 gives an indication of a weak interaction, the absence of this bond leads to a structure that, like with other metal ions,¹⁷ lies at very high energy.

Some theoretical work^{18,20} has suggested that the interaction of adenine with cations should generate the N7-monocoordinated species. We think that the disagreement with our findings is because, in all these investigations, the authors never consider that the $-\text{NH}_2$ group on the base can be pyramidal. This is of fundamental importance for determining the most stable structure of the cation-adenine complex since an amino group lying in the molecular plane introduces a strong repulsion toward the incoming ion preventing the bicoordination.

The NBO analysis and the net charge values on Al^{3+} in all the studied systems provide some indication about the nature of the bond with the considered ligands. In the cases of monocoordination on oxygen atom, the bond appears to be covalent with a consistent ionic contribution. The oxygen participates in the covalent bond with a 2p orbital (100%) and aluminum ion with the 3s (96%) and a 3d (4%). In the N,O bicoordination, the $\text{N}-\text{Al}^{3+}$ bond is characterized by a slight ionic contribution but is essentially covalent, while the $\text{O}-\text{Al}^{3+}$ bond is basically ionic. In the $\text{N}-\text{Al}^{3+}$ interaction, a nitrogen p orbital (100%) and the 3s (88.5%) and 3d (11.5%) orbitals of the aluminum ion are involved. The charge transfer occurs always from ligand to metal ion. The net charge value on Al^{3+} after the complex formation is on the order of $1.2|e|$.

3.1.2. Nucleotide Complexes. The subject of this part of the study is the interaction of aluminum ion with monophosphate nucleotides. To our knowledge, no literature information about the sugar involvement in the interaction exists.

The monophosphate nucleotides considered here are reported in Chart 1.

TABLE 2: B3LYP/6-311+G(2df,2p) Absolute (E in a.u.) and Relative (ΔE in kcal/mol) Energies for Free Nucleotides and Their Complexes with Al^{3+} Cation at 0 K

	L		$\text{Al}^{3+}\text{-L}$		
	E		E	ΔE	AIIA ^a
UMP	-1517.774964	UMP-al		88.4	813.0
		UMP-phos		68.4	
dTMP	-1481.832937	UMP-ap	-1759.582069	0.0	852.4
		dTMP-al		55.7	
dCMP	-1422.621170	dTMP-phos		27.4	830.0
		dTMP-ap	-1723.637498	0.0	
dGMP	-1570.271800	dCMP-al		124.2	856.3
		dCMP-phos		88.1	
dAMP	-1495.009712	dCMP-ap	-1664.486384	0.0	844.1
		dGMP-al	-1812.175941	0.0	
		dGMP-phos		85.9	
		dAMP-al	-1736.874311	0.0	
		dAMP-phos		73.3	

^a Aluminum ion affinity values (AIIA in kcal/mol) are given at 298 K.

The presence of the sugar-phosphate group in N1/N9 position (in the case of uracil, thymine, and cytosine/guanine, and adenine) makes the number of the tautomeric forms of nucleotides lower than that of nucleobases. Furthermore, the steric hindrance due to the dimension of this group makes some attachment sites inaccessible by metal ion. These arguments, together with previous results indicating that the most stable/likely $\text{Al}^{3+}\text{-L}$ complexes derive from L1 tautomers, suggest that only calculations concerning the most usual canonical forms of nucleotides are meaningful. For each of them, we have explored all suitable attachment sites. Data were reported in Table 2, whereas structures of the most stable complexes are in Figure 3. For the other complexes, we refer the reader to both Supporting Information (Figure 2S) and Figure 4.

In the case of uridine monophosphate, three different positions were considered for the attack of cation. These involve the O4 atom and the π -system on uracil, and the oxygen atoms of the

phosphate group. The O2 atom results in being sterically inaccessible because of the sugar proximity.

From Table 2, it can be observed that the optimization produces three minima. Among them, the most stable structure seems to derive from the interaction of Al^{3+} in apical position on the π -system of uracil ring. Really, the final complex UMP-ap does not conserve any original characteristics after the geometry optimization. The cation appears to be coordinated simultaneously to phosphate group and O4 atom because of a folding that involves both the first and second structural units of nucleotide (see Figure 4). Since such a type of folding should not actually occur in RNA, this finding must be considered valid only for an isolated UMP nucleotide.

The interaction with O4 atom on uracil was found to be the less favored one. The complex obtained in this case lies at 88.4 kcal/mol above the mentioned fictitious minimum. Thus, the most probable adduct in RNA should be that in which Al^{3+} is coordinated to the oxygen atoms of phosphate with an average distance of 1.849 Å.

This same situation arises again in the cases of thymidine and cytidine monophosphates.

The different magnitude of the energetic gaps (ΔE) between the various dTMP and dCMP complexes, obtained by selecting the same attachment sites as in UMP, can be ascribed to the degree of stabilization of the apical system that in turn depends on the number and strength of bonds that the cation forms with the different structural units in each nucleotide (see Figure 4).

For dGMP and dAMP, the apical complexes do not exist, so we have selected only two suitable attachment sites for the cation: the phosphate group for both nucleotides, and the N7-O6 and N7-N6 positions for guanosine and adenosine, respectively.

As far as these two nucleotides are concerned, we again found results that must be explained critically.

In fact, as can be observed in Figure 4, the complexes in which we should expect the simultaneous interaction of cation

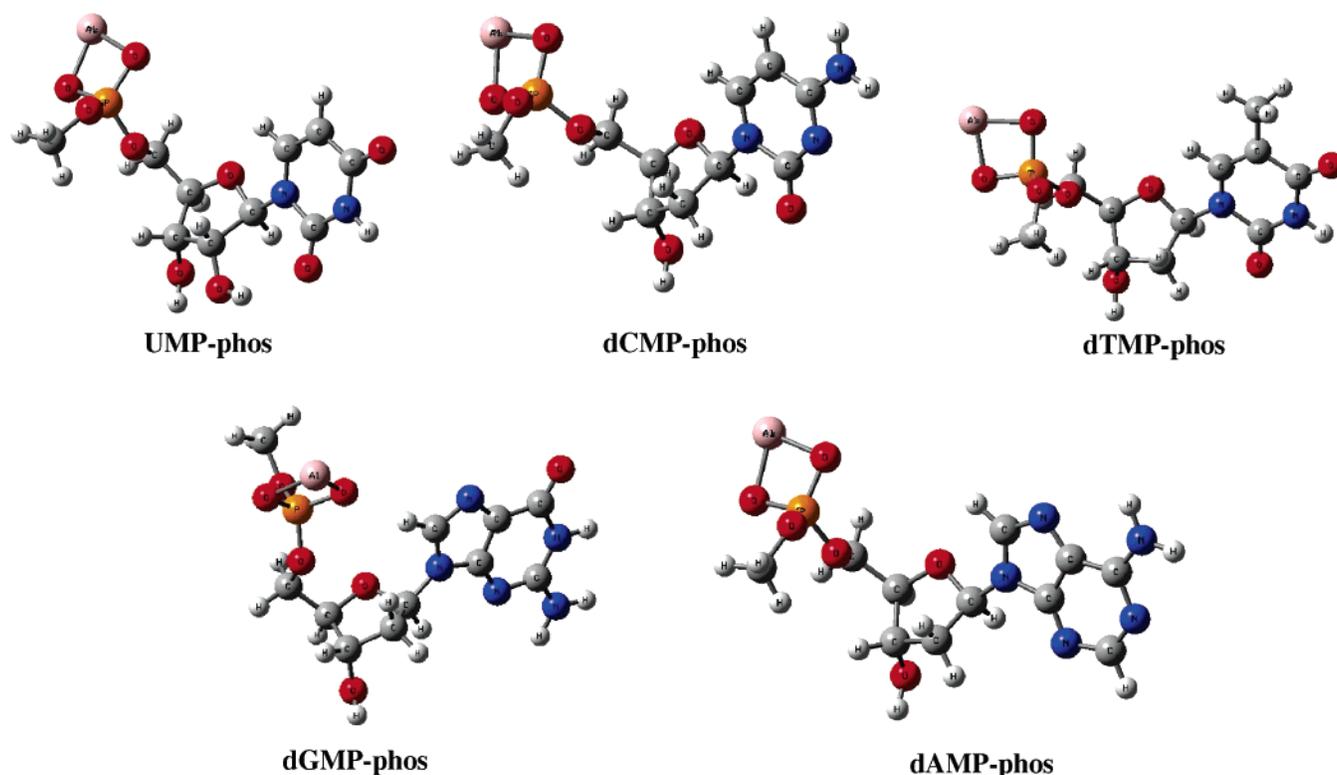


Figure 3. Equilibrium structures showing the interaction of Al^{3+} on phosphate group of nucleotides.

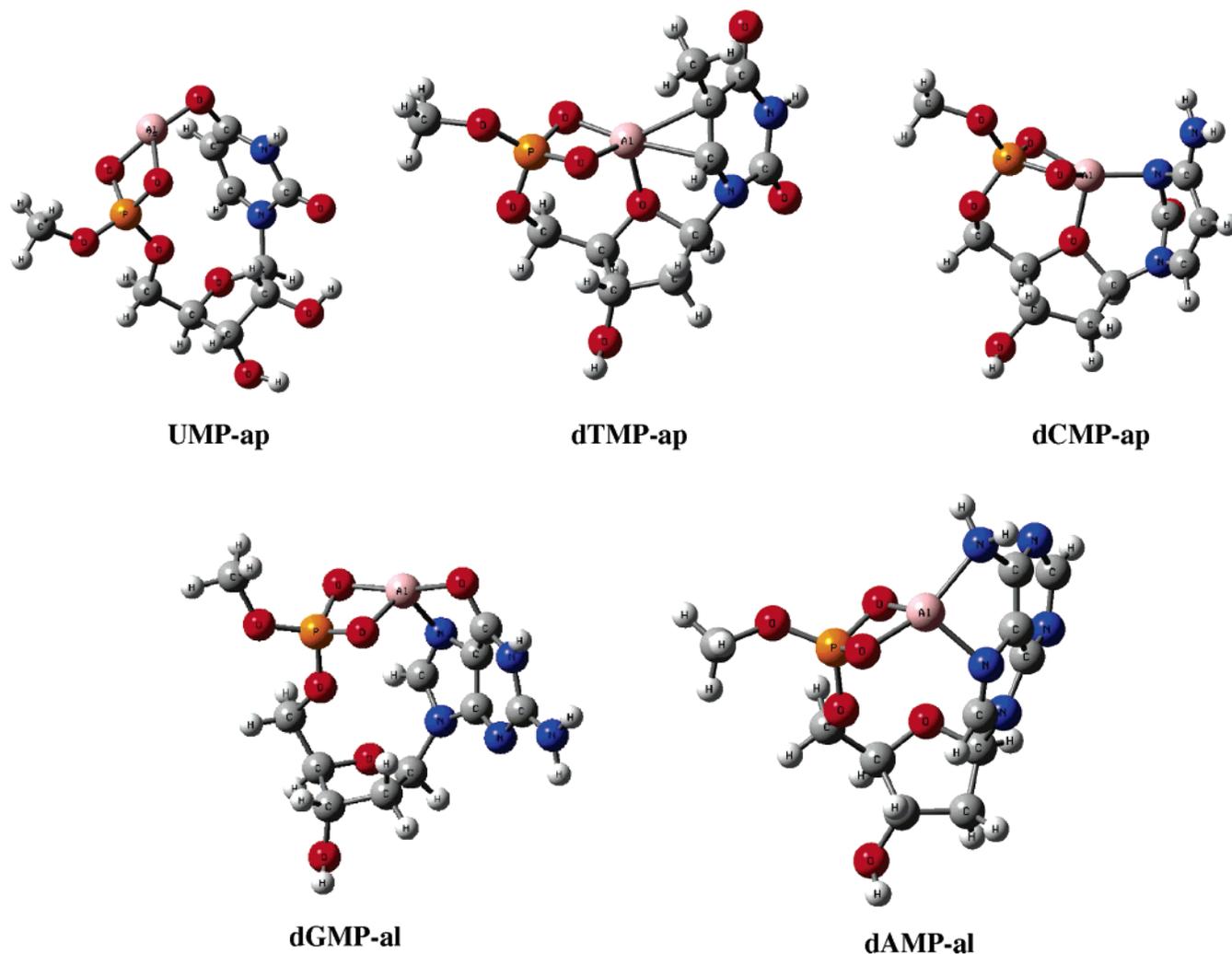


Figure 4. Foldings in some Al^{3+} -monophosphate nucleotide complexes.

with guanosine N7–O6 and adenosine N7–N6 atoms, show also an involvement of the phosphate group. This additional interaction stabilizes the systems by 85.2 and 73.3 kcal/mol, respectively, with respect to those in which Al^{3+} is coordinated just to the phosphate.

Although these foldings involving the phosphate until the capture of the ion are not plausible in RNA and DNA, they nevertheless contribute to give some indication about the affinity of aluminum ion toward the various structural units of the nucleotides. In agreement with experimental suggestions, the largest affinity is found for the phosphate group.

For this reason, NBO analysis was performed only for the complexes of Figure 3.

Natural net charge on the complexed aluminum indicates that, in all cases, an average of $0.6|e|$ is transferred from the ligand to the cation. The slight covalent contribution to the bond arises from the molecular orbital formed by the overlap between the sp (p character 86.11%) and spd (s character 49.3, p character 48.3) hybrid orbitals of oxygen and aluminum ion, respectively.

As suggested by Puškar et al.,⁸⁶ the interactions of aluminum ion with σ -bonding ligands in the gas phase are not always well described using the electrostatic model alone. They have demonstrated that the ligands that act as good σ -electron donors having the capacity to accept the back-donation of electron density from Al^{3+} into antibonding orbitals give generally very stable complexes.

In our case, owing to the complexity of the ligands, it is not easy to do a more detailed analysis of the orbital situation to determine the possible back-donation degree.

However, the ligand–cation charge transfer and the molecular orbitals that we found can be considered as a further confirmation of the mixed covalent–ionic nature of the bond that aluminum forms with some ligands.

3.2. Interaction with Hydrated Cation. The aqueous chemistry of aluminum ions is complicated by the presence of various mono- ($[\text{Al}(\text{OH})]^{2+}$, $[\text{Al}(\text{OH})_2]^+$, soluble $[\text{Al}(\text{OH})_3]$, etc.) and polynuclear species depending on pH values and concentrations of Al^{3+} .⁸⁰ At a pH more acidic than ~ 5 , the ion remains unhydrolyzed, and the main mononuclear species is $[\text{Al}(\text{H}_2\text{O})_6]^{3+}$. At a physiological pH, the soluble hydroxo complexes of Al^{3+} should be taken into account to describe the solution state of this ion. However, for a hydrated Al^{3+} coordinated to the negative phosphate group of nucleotides or to the donor atoms of nucleobases, the high positive charge of this small ion, responsible for the more acidic character of water molecules, is partially neutralized and the hydroxo complexes are less likely. Thus, we have chosen the model $[\text{Al}(\text{H}_2\text{O})_5\text{-L}]^{3+}$ to study the effect of the metal microhydration on the coordination type and on the bond nature. On the other hand, this model was already used in some previous works^{87,88} concerning the interaction of hydrated aluminum ion with biological molecules.

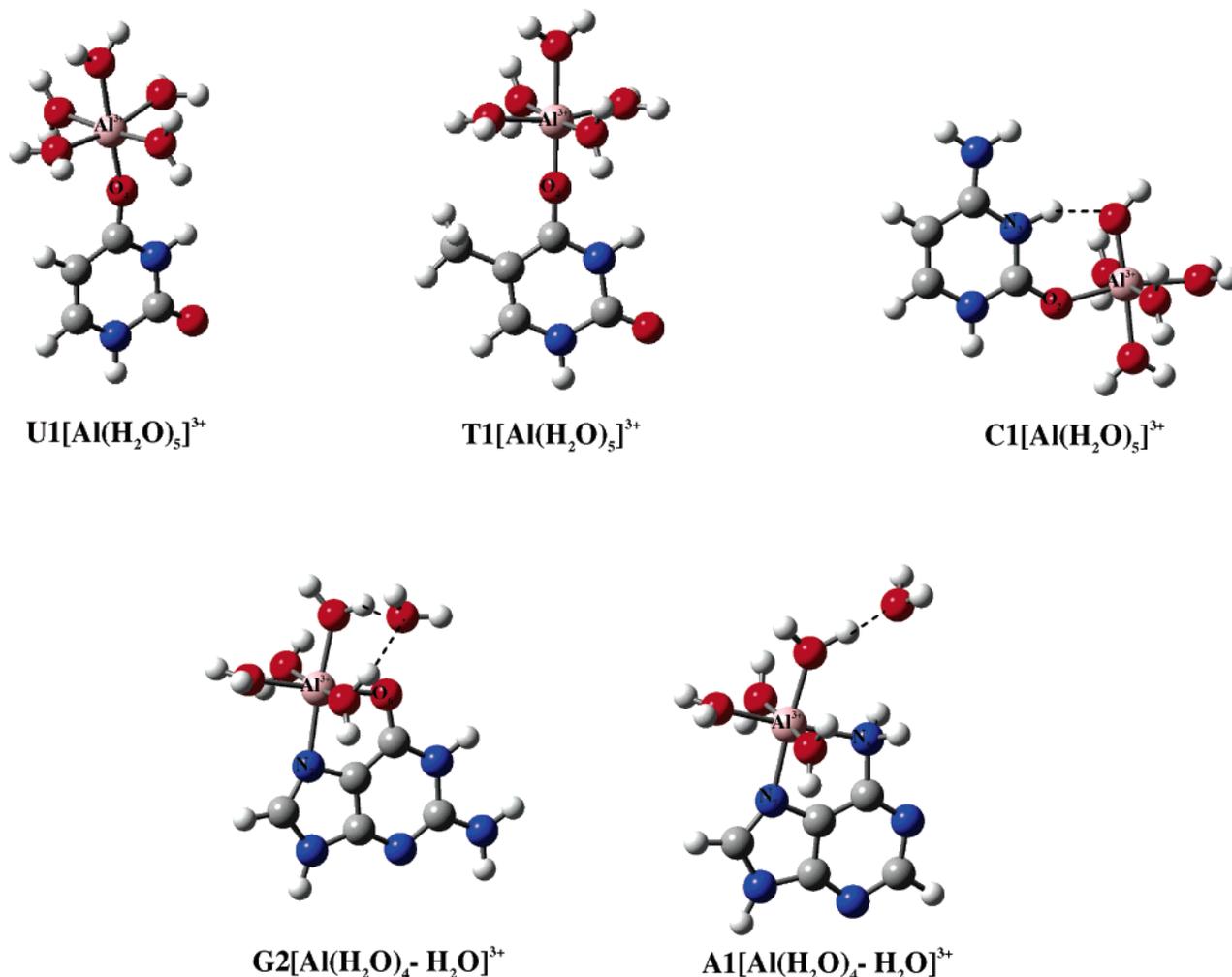


Figure 5. Equilibrium structures of the five $[\text{Al}(\text{H}_2\text{O})_5]^{3+}$ -nucleobase complexes.

TABLE 3: B3LYP/6-311+G(2df,2p) Absolute (E in a.u.) Energies for Nucleobases Complexes with the $[\text{Al}(\text{H}_2\text{O})_5]^{3+}$ Cation at 0 K

$[\text{Al}(\text{H}_2\text{O})_5]^{3+}\text{-L}$	E	AlAl^a
U1 $[\text{Al}(\text{H}_2\text{O})_5]^{3+}$	-1038.674211	103.6
T1 $[\text{Al}(\text{H}_2\text{O})_5]^{3+}$	-1077.979872	104.4
C1 $[\text{Al}(\text{H}_2\text{O})_5]^{3+}$	-1018.852272	155.2
G2 $[\text{Al}(\text{H}_2\text{O})_5]^{3+}$	-1166.496908	151.9
A1 $[\text{Al}(\text{H}_2\text{O})_5]^{3+}$	-1091.161075	104.6

^a Aluminum ion affinity values (AlIA in kcal/mol) are given at 298 K.

Al^{3+} hexa- and penta-aqua complexes were found to have a Th and an irregular triangular bipyramidal symmetry, respectively. The water affinity for the $[\text{Al}(\text{H}_2\text{O})_5]^{3+}$ species was computed to be 48.6 kcal/mol. To take into account the loss of a further water molecule, which is the case for a cation bicoordinated to a nucleobase, we have evaluated also the affinity of the water molecule for the tetrahydrated $[\text{Al}(\text{H}_2\text{O})_4]^{3+}$ system. The obtained value was 50.7 kcal/mol.

3.2.1. *Nucleobase Complexes.* Complexes resulting from hydrated cation-nucleobase interactions were reported in Figure 5. Energetic parameters were collected in Table 3.

All attempts to find a stable complex in the case of the interaction of aluminum ion with the π -system of the nucleobase rings failed. Only for uracil did we obtain a minimum that is, however, identical to that arising from the interaction of the cation with the O4 atom of nucleobase (see Figure 5).

Uracil and thymine gave rise to two monocoordinated adducts whose structures are not so different from those found in the previous computations for the bare ion. Hydrated Al^{3+} establishes with the O4 atom of the considered nucleic acid bases a bond of 1.756 and 1.757 Å, respectively. The cation assumes a nearly octahedral geometry.

For cytosine, the final product presents interesting features. Contrary to what happens for bare cation, hydrated aluminum is monocoordinated to the O2 atom of the base with a distance of 1.913 Å. The hydrogen of one water molecule is transferred to N1 although it remains anchored to the -OH group with a hydrogen bond of 1.730 Å. This can be due to the residual polarizing effect of the aluminum ion on the water molecules of the first sphere.⁸⁹ The analogous complex in which this solvent molecule conserves its hydrogen atom lies at 8.1 kcal/mol.

The solvation sphere does not modify the coordination type of aluminum cation to purine bases.

The results for guanine and adenine show that the $[\text{Al}(\text{H}_2\text{O})_5]^{3+}$ species expels one water molecule in order to establish two simultaneous bonds with N7 and O6 and N7 and N6 atoms, respectively. In both cases, the ejected solvent molecule remains in the proximity of the complexes thanks to some hydrogen bonding with the inner sphere (see Figure 5).

In the guanine- $\text{Al}[(\text{H}_2\text{O})_4\text{H}_2\text{O}]^{3+}$ system, the bond that aluminum forms with oxygen is shorter than that with nitrogen (1.914 versus 2.040 Å). This is normal for aluminum because of its large affinity for oxygen containing species. Instead, the

TABLE 4: B3LYP/6-311+G(2df,2p) Absolute (E in a.u.) Energies for Nucleotides Complexes with $Al[(H_2O)_5]^{3+}$ Cation on Phosphate Group, at 0 K.

	E	AlAI
UMPphos $[Al(H_2O)_5]^{3+}$	-2141.988293	369.4
dTMPphos $[Al(H_2O)_5]^{3+}$	-2106.084851	393.6
dCMPphos $[Al(H_2O)_5]^{3+}$	-2046.909101	416.2
dGMPphos $[Al(H_2O)_5]^{3+}$	-2194.575419	426.1
dAMPphos $[Al(H_2O)_5]^{3+}$	-2119.300890	418.3

^a Aluminum ion affinity values (AlAI in kcal/mol) are given at 298 K.

larger basicity of N7 with respect to N6 explains the different strength and length of N7– Al^{3+} (1.987 Å) and N6– Al^{3+} (2.124 Å) bonds in the adenine complex.

NBO analysis indicates that hydrated aluminum establishes an ionic bond with nucleobases. This can be deduced comparing the total charge on the cation before and after the interaction (2.1 vs 1.9|e|, respectively).

3.2.2. Nucleotide Complexes. Since in the interaction with nucleotides the naked aluminum ion is always mainly attracted by the phosphate group, in the case of hydrated cation we considered only this attachment site. Really, since the computations are performed in the gas phase, where nucleotides enjoy a wide conformational freedom, the potential involvement of the other structural units is however possible during the optimization process.

The energetic data and the equilibrium geometries of the stable adducts $Al[(H_2O)_n]^{3+}$ –nucleotides were collected in Table 4 and Figure 6, respectively.

Contrary to what was found in the complexes of the bare ion in which we had the bicoordination of the metal center, now all five structures show the pentahydrated aluminum linked to a single oxygen atom of the phosphate. This means that the cation prefers to keep a water molecule rather than to bind another oxygen atom. All the structures are characterized by a number of hydrogen shifts or hydrogen bond interactions, between the oxygen atoms of water molecules and those of phosphate group, that introduce stabilizing effects. In the dTMPphos– $Al[(H_2O)_5]^{3+}$ complex also the sugar moiety appears to be involved in the hydrogen bond network.

The bond between the hydrated ion and the oxygen atom of the phosphate is again completely ionic as revealed by the comparison between total net charge on free and coordinated hydrated cation (2.1 vs 1.9|e|, respectively).

3.3. Aluminum Ion Affinity. The aluminum ion affinity (AlAI) values reported in Table 1 indicate that the nucleobases interact very favorably with the naked cation in the order $G > C > T > A > U$.

The highest values of AlAI are obtained in correspondence with bicoordinated complexes. The reason that adenine does not follow this trend strictly can lie, as mentioned before, in the weak interaction between Al^{3+} and N6 atom.

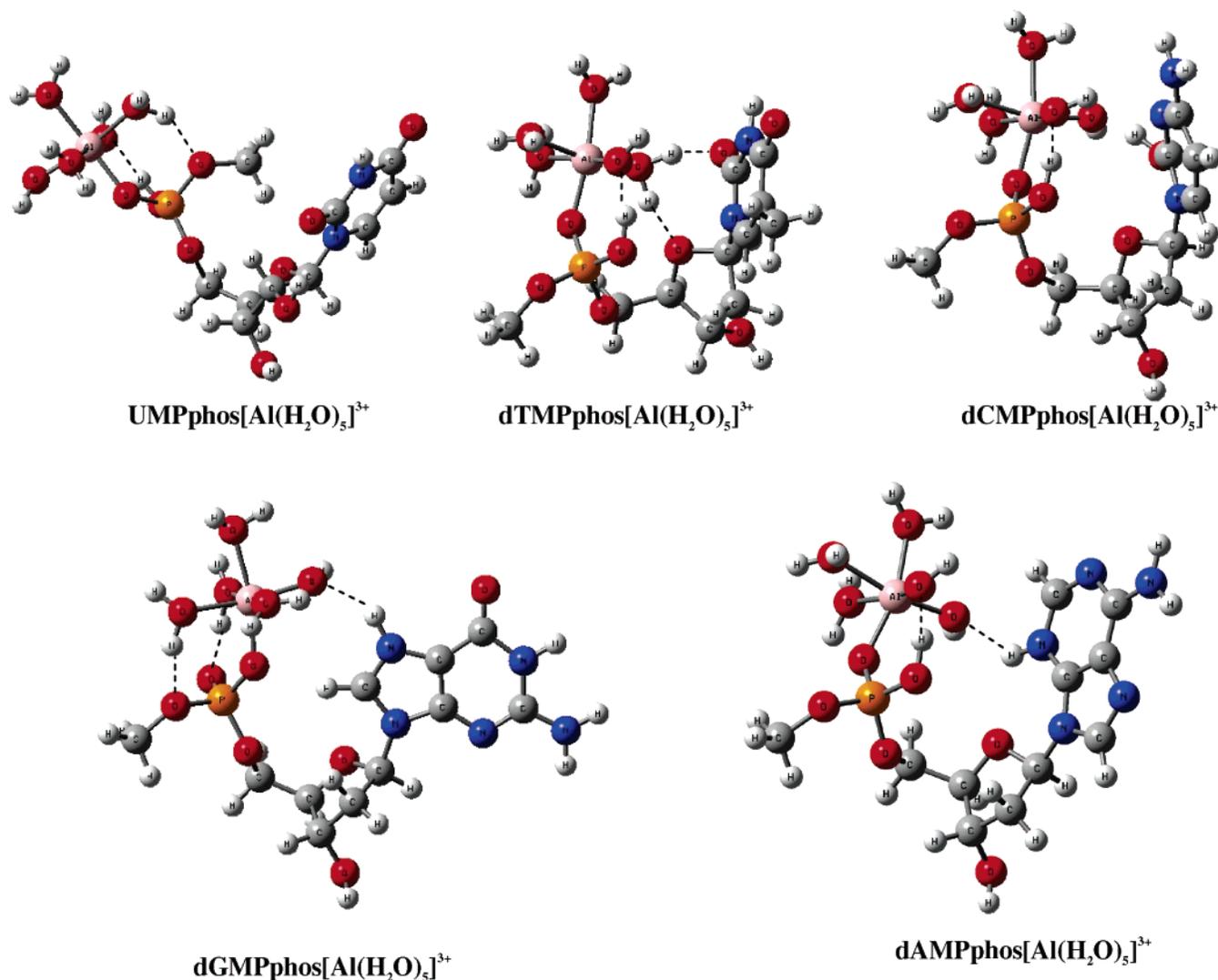


Figure 6. Equilibrium structures of the five $[Al(H_2O)_5]^{3+}$ –nucleotide complexes.

The addition of a solvation shell to the metal center reduces drastically (see Table 3) the AIIA values. This is not surprising if we take into account that hydrated aluminum has an orbital availability very different from that of bare cation and that the strength with which it can bind additional ligands decreases gradually.

Actually, the loss of the second water molecule from the Al-[(H₂O)₆]³⁺ species, requires an energetic cost that is larger than that required for the loss of the first solvent molecule (101.4 vs 48.6 kcal/mol).

The new trend of AIIA values is: C > G > A ≥ T ≥ U.

At a first glance, this sequence seems to be totally different from that obtained in the previous case, but the AIIA values of A, T, and U are not sufficient to establish a rigorous trend because of their similarity. The inversion of the AIIA values of guanine and cytosine can be attributed to the change of the coordination type and to the protonation of N3 on cytosine. AIIA values for nucleotides were computed taking as references the complexes in which the metal cation is linked to the phosphate group. In fact, we think that although these are not the most stable species because of the high flexibility of nucleotides in gas phase, they probably become the most significant ones when the nucleotides are kept "strained" in DNA and RNA helices.

In the case of naked aluminum ion, AIIA values are very high thanks to the bond that the metal ion forms with the two oxygen atoms on which the negative charge of phosphate group is delocalized.

The Al³⁺ affinity follows the order dGMP ≥ dTMP > dAMP > dCMP > UMP.

No comparison is possible with values obtained for nucleobases since another structural unit is involved in the interaction.

Nucleotides interact with hydrated cation less strongly than with the naked one. The AIIA values, as expected, have a magnitude completely different from that obtained in the absence of water molecules. This can be explained not only by the fact that the solvated cation has a minor capacity of coordination but also by the different nature of the bond and by the different metal coordination type. Thus, a new order of AIIA values is obtained: dGMP > dAMP ≥ dCMP > dTMP > UMP.

4. Conclusions

This paper collects the results of the first computational study concerning the interaction of the bare and hydrated aluminum cation with nucleobases and monophosphate nucleotides.

The most significant aspects of the investigation are reported here.

Aluminum ion forms stable complexes with all the considered ligands.

The coordination type on nucleobases remains essentially the same when the hydration shell is added to the naked cation, but binding energy values decrease significantly in accordance with the different orbital availability of the metal cation and with the change of bond nature.

Except for the sugar moiety, all other selected sites can be suitable for the cation attack on nucleotides. The phosphate group is practically always involved in the interaction.

Naked ion appears to be coordinated to both oxygen atoms of phosphate, while hydrated aluminum ion binds only one of these two atoms since the other is often involved in some hydrogen bonding with water molecules of the metal solvation shell.

Also in this case and for the same previous reasons, binding energies decrease on going from bare to hydrated cation.

Very high binding energies involved in this interaction demonstrate the elevated affinity of aluminum trication for both nucleobases and nucleotides.

Acknowledgment. We thank the Computer Center of CENG-CEA of Grenoble and Università degli Studi della Calabria and Regione Calabria (POR Calabria 2000/2006, misura 3.16, progetto PROSICA) for grants of computer time.

Supporting Information Available: Figure containing equilibrium geometries for all relative minima for each nucleobase. Figure containing the high lying Al³⁺—monophosphate nucleotide complexes. This material is available free of charge via the Internet at <http://pubs.acs.org>.

References and Notes

- (1) (a), *Interactions of Metal Ions with Nucleotides, Nucleic Acids and their Constituents*; Sigel, A., Sigel, H., Eds.; Metal Ions in Biological Systems 32; Marcel Dekker: New York, 1996. (b) *Probing of Nucleic Acids by Metal Ion Complexes of Small Molecules*; Sigel, A., Sigel, H., Eds.; Metal Ions in Biological Systems 33; Marcel Dekker: New York, 1996.
- (2) Nakano, S. I.; Fujimoto, M.; Hara, H.; Sugimoto, N. *Nucleic Acids Res.* **1999**, *27*, 2957.
- (3) Potaman, V. N.; Soyfer, V. N. *J. Biomol. Struct. Dyn.* **1994**, *11*, 1035.
- (4) Guschlbauer, W.; Chantot, J. F.; Thiele, D. *J. Biomol. Struct. Dyn.* **1990**, *8*, 491.
- (5) Russo, N.; Toscano, M.; Grand, A. *J. Phys. Chem. B* **2001**, *105*, 4735.
- (6) Russo, N.; Toscano, M.; Grand, A. *J. Am. Chem. Soc.* **2001**, *123*, 10272.
- (7) Loeb, L. A.; Zakour, A. R. In *Nucleic Acid-Metal Ion Interactions*; Spiro, T. G., Ed.; John Wiley & Sons: New York, 1980; pp 115–144.
- (8) Aubagnac, J. L.; Devienne, T. M.; Combarieu, R.; Barascut, J. L.; Imbach, J. L.; Lazrek, H. B. *Org. Mass Spectrom.* **1983**, *18*, 361.
- (9) McCrery, D. A.; Gross, M. L. *Anal. Chim. Acta* **1985**, *178*, 91.
- (10) Tomer, K. B.; Gross, M. L. *Anal. Chem.* **1986**, *58*, 2527.
- (11) Hogg, A. M.; Kelland, J. G.; Vederas, C. J. *Helv. Chim. Acta* **1986**, *69*, 908.
- (12) Chiarelli, M. P.; Gross, M. L. *J. Phys. Chem.* **1989**, *93*, 3595.
- (13) Voyksner, R. D. *Org. Mass Spectrom.* **1987**, *22*, 513.
- (14) Madhusudanan, K. P.; Katti, S. B.; Hashmi, S. A. N. *Org. Mass Spectrom.* **1993**, *28*, 970.
- (15) Paul, G. J. C.; Theophanides, T.; Anastassopoulou, J.; Marcotte, I.; Bertrand, M. *Proceedings of the 43th ASMS Conference on Mass Spectrometry and Allied Topics*, Atlanta, GA, May 21–26, 1995; p 608.
- (16) Rodgers, M. T.; Armentrout, P. B. *Proceedings of the 44th ASMS Conference on Mass Spectrometry and Allied Topics*, Portland, OR, May 12–16, 1996; p 88.
- (17) Russo, N.; Toscano, M.; Grand, A. *J. Phys. Chem. A* **2003**, *107*, 11533.
- (18) Šponer, J.; Sabat, M.; Burda, J. V.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. B* **1999**, *103*, 2528.
- (19) Munõz, J.; Šponer, J.; Hobza, P.; Orozco, M.; Luque, J. *J. Phys. Chem. B* **2001**, *105*, 6051.
- (20) Burda, J. V.; Šponer, J.; Hobza, P. *J. Phys. Chem.* **1996**, *100*, 7250.
- (21) Šponer, J.; Burda, J. V.; Sabat, M.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. A* **1998**, *102*, 5951.
- (22) McFail-Isom, L.; Shui, X.; Williams, L. D. *Biochemistry* **1998**, *37*, 17105.
- (23) Russo, N.; Toscano, M.; Grand, A. *J. Mass. Spectrom.* **2003**, *38*, 265.
- (24) Marino, T.; Russo, N.; Toscano, M.; Grand, A. *Int. J. Quantum Chem.* **2004**, *98*, 347.
- (25) Russo, N.; Sicilia, E.; Toscano, M.; Grand, A. *Int. J. Quantum Chem.* **2002**, *90*, 903.
- (26) Kornilova, S. V.; Miskovsky, P.; Tomkova, A.; Kapinos, L. E.; Hackl, E. V.; Andrushchenko, V. V.; Grigoriev, D. N.; Blagoi, Yu. P. *J. Mol. Struct.* **1997**, *408/409*, 219.
- (27) Correia dos Santos, M. M.; Sousa, P. M. P.; Modesto, A. M.; Simoes, M.; Goncalves M. L. *Biochem. Biophys. Res. Commun.* **1998**, *245*, 267.
- (28) Hettich, R. H. *Int. J. Mass. Spectrom.* **2001**, *204*, 55.
- (29) Bianchi, E. M.; Ali, A.; Sajadi, S.; Song, B.; Sigel, H. *Chem. Eur. J.* **2003**, *9*, 881.
- (30) Deng, H.; Kebarle P. *J. Am. Chem. Soc.* **1998**, *120*, 2925.
- (31) Chaparro, A. L.; Vachet R. W. *J. Mass. Spectrom.* **2003**, *38*, 333.

- (32) Drouin, R.; Rodriguez, H.; Gao, S. W.; Gebreyes, Z.; O'Connor, T. R.; Holmquist, G. P.; Akman, S. A. *Free Radical Biol. Med.* **1996**, *21*, 261.
- (33) Bal, W.; Kasprzak, K. S. *Toxicol. Lett.* **2002**, *127*, 55.
- (34) Halliwell, B.; Aruoma, I. O. *FEBS Lett.* **1991**, *281*, 9.
- (35) Dizdaroglu, M. *Mutat. Res.* **1992**, *275*, 331.
- (36) Theophanides, T.; Anastassopoulou, J. *Crit. Rev. Oncol./Hematol.* **2002**, *42*, 57.
- (37) Burda, J. V.; Šponer, J.; Leszczynski, J.; Hobza, P. *J. Phys. Chem. B* **1997**, *101*, 9670.
- (38) Blades, A. T.; Jayaweera, P.; Ikononou, M. G.; Kebarle, P. *Int. J. Mass Spectrom. Ion Process.* **1990**, *101*, 325.
- (39) Walzer, N. R.; Stace, A. J.; Woodward, C. A. *Int. J. Mass Spectrom.* **1999**, *188*, 113.
- (40) Shvartsburg, A. A. *J. Am. Chem. Soc.* **2002**, *124*, 7910.
- (41) Shvartsburg, A. A. *J. Am. Chem. Soc.* **2002**, *124*, 12343.
- (42) Shvartsburg, A. A. *Chem. Phys. Lett.* **2002**, *360*, 479.
- (43) Shvartsburg, A. A.; Jones, R. C. *J. Am. Soc. Mass Spectrom.* **2004**, *15*, 406.
- (44) Cheng, Z. L.; Siu, K. W. M.; Guevremont, R.; Berman, S. S. *Org. Mass Spectrom.* **1992**, *27*, 1370.
- (45) Underwood, E. J. In *Trace Elements in Human and Animal Nutrition*; Academic Press: New York, 1977; p 430.
- (46) De Broe, M. E.; Coburn, J. W. In *Aluminium and Renal Failure*; Marcel Dekker: New York, 1990.
- (47) Perl, D. P.; Gajdusek, D. C.; Garruto, R. M.; Yanagikawa, R. T.; Gibbs, C. J. *Science* **1982**, *217*, 1053.
- (48) Foacin, J. F. *Nature* **1987**, *326*, 136.
- (49) Crapper, D. R.; Krishnan, S. S.; Kaehny, W. D. *Brain* **1976**, *99*, 67.
- (50) Wenk, G. L.; Stemmer, K. L. *Brain Res.* **1983**, *288*, 393.
- (51) Bishop, N. J.; Moreley, R.; Day, J. P.; Lucas, A. *New Engl. J. Med.* **1977**, *336*, 1557.
- (52) Birchall, J. D. *Chem. Br.* **1990**, 141.
- (53) Burgess, J. *Metal Ions in Solution*; John Wiley: London, 1978.
- (54) Kaim, W.; Schwederski, B. *Bioinorganic Chemistry: Inorganic Elements in the Chemistry of Life*; John Wiley: London, 1994.
- (55) Schmitt, W.; Jordan, P. A.; Henderson, R. K.; Moore, G. R.; Anson, C. E.; Powell, A. K. *Coord. Chem. Rev.* **2002**, *228*, 115.
- (56) Ganrot, P. O. *Environ. Health Perspect.* **1986**, *65*, 363.
- (57) Kiss, T.; Zatta, P.; Corain, B. *Coord. Chem. Rev.* **1996**, *149*, 329.
- (58) Martin Bruce, R. *Acc. Chem. Res.* **1994**, *27*, 204.
- (59) Zhang, R. Y.; Liu, Y.; Pang, D.-W.; Cai, R. X.; Qi, Y. P. *Anal. Sci.* **2002**, *18*, 761.
- (60) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03*, revision A.1; Gaussian, Inc.: Wallingford, CT, 2004.
- (61) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648.
- (62) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev. B* **1988**, *37*, 785.
- (63) Clark, T.; Chandrasekhar, J.; Schleyer, P. v. R. *J. Comput. Chem.* **1983**, *4*, 294 and references therein.
- (64) *NBO*, version 3.1; Glendening, E. D.; Reed, A. E.; Carpenter, J. E.; Weinhold, F.
- (65) Szczepaniak, K.; Kwaitkowski, J.; Kubulat K.; Person, W. B. *J. Am. Chem. Soc.* **1988**, *110*, 8319.
- (66) Nowak, M. I.; Lapinski, L.; Fullara, F. *Spectrochim. Acta* **1989**, *A45*, 229.
- (67) Brown, R. D.; Godfrey, P. D.; McNaughton, D.; Pierlot, A. P. *J. Am. Chem. Soc.* **1987**, *111*, 2308.
- (68) Dreyfus, M.; Bensaude, O.; Dodin, G.; Dubois, J. E. *J. Am. Chem. Soc.* **1976**, *98*, 2353.
- (69) Szczepaniak, K.; Szajda, W.; Person, W. B.; Leszczynski, J. *Can. J. Chem.* **1991**, *69*, 1718.
- (70) Russo, N.; Toscano, M.; Grand, A.; Jolibois, F. *J. Comput. Chem.* **1998**, *19*, 989.
- (71) Colominas, C.; Luque, F. J.; Orozco, M. *J. Am. Chem. Soc.* **1996**, *118*, 6811.
- (72) Paglieri, L.; Corongiu, G.; Estrin, D. A. *Int. J. Quantum. Chem.* **1995**, *56*, 615.
- (73) Roehrig, G. H.; Oyler, N. A.; Adamowicz, L. *J. Phys. Chem.* **1995**, *99*, 14285.
- (74) Holmen, A.; Broo, A. *Int. J. Quantum. Chem. (Quantum Biol. Symp.)* **1995**, *22*, 113.
- (75) Gould, I. R.; Burton, N. A.; Hall, R. J.; Hillier, I. H. *THEOCHEM* **1995**, *331*, 147.
- (76) Katritzki, A. R.; Karelson, M. *J. Am. Chem. Soc.* **1991**, *113*, 1561.
- (77) Fogarasi, G. *J. Mol. Struct.* **1997**, *413-414*, 271.
- (78) Kobayashi, R. *J. Phys. Chem.* **1998**, *102*, 10813.
- (79) Ha, T.-K.; Gunthard, H. H. *J. Mol. Struct.* **1993**, *300*, 619.
- (80) Rubini, P.; Lakatos, A.; Champmartin, D.; Kiss, T. *Coord. Chem. Rev.* **2002**, *228*, 137.
- (81) Pedersen, D. B.; Zgierski, M. Z.; Denomme, S.; Simard, B. *J. Am. Chem. Soc.* **2002**, *124*, 6686.
- (82) McFail-Isom, L.; Shui, X.; Williams, L. D. *Biochemistry* **1998**, *37*, 8741.
- (83) Munoz, J.; Šponer, J.; Hobza, P.; Orozco, M.; Javier Luque, F. *J. Phys. Chem. B* **2001**, *105*, 6051.
- (84) Ma, J.; Dougherty, D. *Chem. Rev.* **1997**, *97*, 1303.
- (85) Pedersen, D. B.; Zgierski, M. Z.; Denomme, S.; Simard, B. *J. Phys. Chem. A* **2003**, *107*, 6457.
- (86) Puškar, L.; Tomlins, K.; Duncombe, B.; Cox, H.; Stace, A. J. *J. Am. Chem. Soc.* **2005**, *127*, 7559.
- (87) Mercero, J. M.; Matxain, J. M.; Rezabal, E.; Lopez, X.; Ugalde, J. M. *Int. J. Quantum Chem.* **2004**, *98*, 409.
- (88) Tunega, D.; Haberhauer, G.; Gerzabek, M.; Lischka, K. *J. Phys. Chem. A* **2000**, *104*, 6824.
- (89) Rudolph, W. W.; Mason, R.; Pye, C. C. *Phys. Chem. Chem. Phys.* **2000**, *2*, 5030.

ARTICOLO 2

**Gas phase interaction of Zinc ion with purine and pyrimidine
DNA and RNA bases.**

Gas Phase Interaction of Zinc Ion With Purine and Pyrimidine DNA and RNA Bases

T. MARINO,¹ D. MAZZUCA,¹ M. TOSCANO,¹ N. RUSSO,¹
A. GRAND²

¹*Dipartimento di Chimica and Centro di Calcolo ad Alte Prestazioni per Elaborazioni Parallele e Distribuite-Centro d'Eccellenza MIUR, Università della Calabria, I-87030 Arcavacata di Rende (CS), Italy*

²*Département de Recherche Fondamentale sur la Matière Condensée, Service de Chimie Inorganique et Biologique, CEA-Grenoble, 17 Rue des Martyrs, 38054 Grenoble Cedex 9, France*

Received 24 October 2005; accepted 24 April 2006

Published online 31 August 2006 in Wiley InterScience (www.interscience.wiley.com).

DOI 10.1002/qua.21194

ABSTRACT: The interaction of uracil, thymine, cytosine, adenine, and guanine with zinc ion was studied at the density functional B3LYP/6-311+G(2df,2p) level. Different binding sites allowing both mono-metal and bi-metal coordination were considered for the different low-lying tautomers of nucleic acid bases. Zinc ion forms stable compounds with all nucleobases. Except for cytosine, mono-coordination appears to be less favored than bi-coordination in the other pyrimidines. Instead, the preferred sites in the case of adenine and guanine were found to be the N7 and O6 and N7 and N6 pairs of atoms, respectively. Zinc ion affinity was evaluated for all the complexes and compared with values previously obtained for other transition metal ions. In the present case, the following order of metal ion affinity (MIA) was found:

G>A>C>T>U. © 2006 Wiley Periodicals, Inc. *Int J Quantum Chem* 107: 311–317, 2007

Key words: zinc; density functional theory; purine and pyrimidine bases; metal ion affinity

Correspondence to: N. Russo; e-mail: nrusso@unical.it

Contract grant sponsor: Università degli Studi della Calabria.

Contract grant sponsor: Regione Calabria.

Contract grant number: 2000/2006 3.16 (PROSICA).

This article was presented at the 2nd International Theoretical Biophysics Symposium, June 28–July 1, 2005.

Introduction

The zinc ion is involved in different cellular processes, including proliferation, reproduction, immunofunction, and defense against free radicals [1–3]. Experimentally, it was found that about 25% of the total zinc present in rat liver lies in the nucleus [4], demonstrating the zinc ion can influence DNA as a component of chromatin structure, DNA replication, transcription, and repair [5]. Numerous enzymes employ zinc cation in the replication, transcription, and translation processes. Several proteins, such as zinc fingers, use the Zn^{2+} ion to stabilize the DNA structure [6]. A little more than a decade ago, it was proved that zinc ion controls conformational changes that occur within the DNA structure that can induce distortion of the Watson–Crick guanine–cytosine base pair [6, 7].

Recently, the effect of transition metal ions on oligodeoxyribonucleotide binding was investigated using matrix-assisted laser desorption/ionization Fourier transform ion cyclotron resonance (MALDI) (FTICR) mass spectrometry, revealing that bivalent metal ions, including Zn^{2+} , interact preferably with the nucleobase instead of the phosphate group exhibiting strong affinities [8–10].

While the interaction of alkali metal ions with isolated DNA and RNA bases was investigated widely in the gas phase at both the experimental [11–13] and theoretical [12, 14, 15] levels, no extensive study was devoted to the interaction of bare or hydrated zinc ion, with all nucleic acid bases. The investigation carried out by Burda et al. [16], concerning the interaction of zinc with purines, was performed at the Hartree–Fock (HF) level with second-order Møller–Plesset (MP2) single-point evaluation of energies.

The study of the interaction of bare zinc ion with the different low-lying tautomeric forms of nucleobases, presented in this work, can provide information regarding the electronic and structural properties of the complexes, that can be used in turn to clarify the role of Zn^{2+} in the biophysics of DNA and RNA molecules. The presence of solvent was not taken into account, although in some case, it may alter the metal ion binding properties.

The computational B3LYP/6-311+G(2df,2p) protocol adopted in this work was previously applied to the study of other metal ions interacting with DNA and RNA nucleobases [14, 15, 17–19]. It proved successful in the description of both structural and electronic properties.

Computational Methods

All the calculations were performed within the framework of density functional theory (DFT) using the B3LYP [20, 21] exchange–correlation functional, as implemented in the Gaussian 03 code [22]. Hessians for evaluation of zero-point vibrational effects (ZPVE) were calculated at the B3LYP/6-311G** level of theory on each optimized stationary point. Single-point computations were then carried out for a better estimate of energies using the all-electron 6-311+G(2df,2p) basis set for H, C, N, O, and Zn atoms.

The zinc ion affinity (MIA), assumed as the negative of the enthalpy variation (ΔH) for the metalation process, was calculated for all the Zn^{2+} -nucleic base adducts examined. Natural bond orbital analysis (NBO) was performed on the global minimum of each complex [23].

Results and Discussion

The gas phase tautomeric equilibria of the DNA and RNA free bases were extensively studied at theoretical [24–35] and experimental levels [36–40]. On the basis of these studies, it was shown that the formation of the noncanonical tautomeric forms for uracil, thymine, and adenine are improbable. Different situation occurs in the case of guanine and cytosine [32–35]. In fact, their enolic structures are very close in energy to those of ketoamine structures and were detected experimentally [32–35].

As in our other work [14, 15, 17–19], the most stable tautomeric forms of the free nucleobases and different attack sites (essentially oxygen and/or nitrogen atoms) for the zinc ion (II), were considered.

In particular, the noncanonical forms U2/T2 and U3/T3 obtained from the diketonic form U1/T1 after the N3 proton shift on O4 and O2, respectively, were considered for uracil and thymine.

The canonical (C1), hydroxyl O2–H (C2), hydroxyl O2–H rotamer (C3), and imino (C4) tautomers were selected for cytosine. The canonical (A1), the H–N7 bonded (A2), and the imino (A3) species are the three forms chosen for adenine.

Finally, the G1 and G2 canonical tautomers obtained changing the position of the hydrogen atom between N7 and N9, the G3 form obtained from G2 after a proton shift from N1 to O6, the G4 rotamer of G3, and the G5 tautomer arising from G1 after a

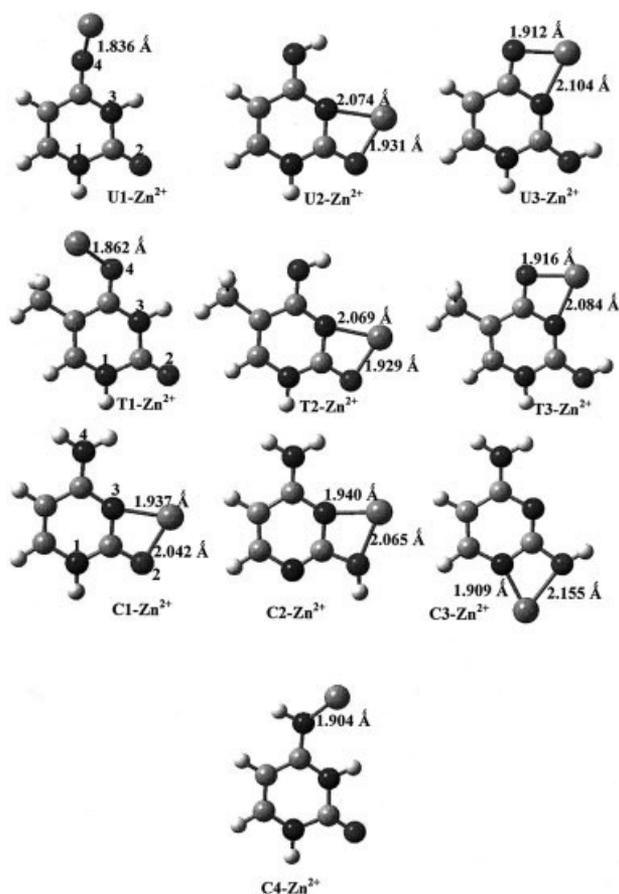


FIGURE 1. B3LYP/6-311+G(2df,2p) optimized structures of the complexes of uracil, thymine, and cytosine tautomers with Zn^{2+} ion. Distances are in Ångstroms (Å).

proton shift from N1 to O6, were selected for guanine.

The complexes obtained on the interaction of Zn^{2+} ion with the different tautomers of each nucleobase were depicted in Figures 1 and 2. Table I presents the relative energetic parameters.

In the case of uracil and thymine, the lowest-energy tautomers of free nucleobase, U1 and T1, show only two oxygen atoms as suitable coordination sites for zinc ion and give rise to the highest-energy adducts, U1- Zn^{2+} and T1- Zn^{2+} , that lie at 20.5 and 15.9 kcal/mol above the corresponding global minima U3- Zn^{2+} and T3- Zn^{2+} . For both nucleobases, this form energetically unfavorable represents the only case in which the metal ion binds in a mono-coordinated fashion. It is noteworthy that in the U1- Zn^{2+} complex, metal ion, lying out of the molecular plane of $\sim 100^\circ$, is 1.836 Å far

from O4 oxygen (see Fig. 1). In the T1- Zn^{2+} complex, the deviation from planarity of metal ion is less pronounced ($\sim 19.5^\circ$), and the distance Zn^{2+} -O4 is 1.862 Å.

In the complexes originating from the other tautomeric forms, U2, U3, T2, and T3, the metal ion is coordinated to two donor atoms (nitrogen and oxygen) of the free nucleic base. As previously mentioned, the U3- Zn^{2+} and T3- Zn^{2+} represent the global minima for uracil and thymine, respectively (see Table I and Fig. 1). Actually, the relative minima, U2- Zn^{2+} and T2- Zn^{2+} , are almost isoenergetic with the global ones, owing to the small energy difference between them (1.0 and 0.4 kcal/mol). It is precisely this very small energy difference that makes it difficult to explain why the U3- Zn^{2+} and T3- Zn^{2+} forms are slightly favored with respect to U2- Zn^{2+} and T2- Zn^{2+} ones. However, in the case of thymine, the larger stability of T3- Zn^{2+} can be due to the proximity of the O2 atom to the $-\text{CH}_3$

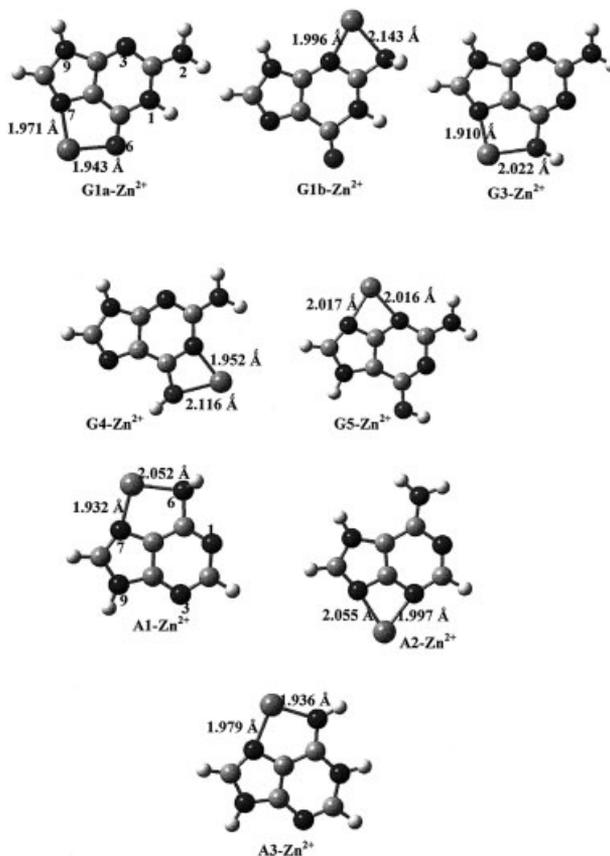


FIGURE 2. B3LYP/6-311+G(2df,2p) optimized structures of the complexes of guanine and adenine tautomers with Zn^{2+} ion. Distances are in Ångstroms (Å).

TABLE I
Absolute (E in a.u.) and relative (ΔE in kcal/mol) energies at 0 K of the Zn^{2+} -nucleobase complexes.

	E_{ZPE}	ΔE
U1— Zn^{2+}	-2193.49541	20.5
U2— Zn^{2+}	-2193.526489	1.0
U3— Zn^{2+}	-2193.528060	0.0
T1— Zn^{2+}	-2232.815716	15.9
T2— Zn^{2+}	-2232.840458	0.4
T3— Zn^{2+}	-2232.842921	0.0
C1— Zn^{2+}	-2173.67081	0.0
C2— Zn^{2+}	-2173.618137	33.0
C3— Zn^{2+}	-2173.637206	21.1
C4— Zn^{2+}	-2173.619397	32.3
G1a— Zn^{2+}	-2321.353784	0.0
G1b— Zn^{2+}	-2321.262749	57.1
G3— Zn^{2+}	-2321.299960	33.8
G4— Zn^{2+}	-2321.284313	43.6
G5— Zn^{2+}	-2321.322137	19.9
A1— Zn^{2+}	-2246.039660	27.5
A2— Zn^{2+}	-2246.054056	18.5
A3— Zn^{2+}	-2246.083517	0.0

donor group that enhances its negative charge favoring the interaction with cation.

It is interesting to underline that the free tautomers of uracil and thymine differ considerably in energy, but their complexes with metal ions often become energetically similar [17–19]. The distances in both bi-coordinated complexes of uracil and thymine are analogous, with the Zn–O bond shorter than the Zn–N one (see Fig. 1).

Cytosine, as in the interaction with other metal ions [14, 15, 17–19], showed a distinct behavior with respect to the other two pyrimidine bases. In fact, although the four tautomers of free base are very similar in energy, the corresponding complexes with Zn^{2+} have, instead, a very different stability that follows the trend: $\text{C1-Zn}^{2+} > \text{C3-Zn}^{2+} > \text{C4-Zn}^{2+} > \text{C2-Zn}^{2+}$. Except for C4-Zn^{2+} , all other complexes present a bi-coordinated metal ion (see Fig. 1).

As far as the geometrical parameters are concerned, it was observed that the cation–oxygen length is slightly longer than the cation–nitrogen length when the involved tautomer possesses the enolic form (see Fig. 1). Thus, the largest stability of C1-Zn^{2+} complex can be ascribed to the more symmetric disposition of metal cation with respect to oxygen and nitrogen atoms that allows two strong interactions.

The free molecule of guanine can have five different tautomeric forms whose maximum difference in energy is 4.4 kcal/mol. As in the case of cytosine, as well as for guanine, an infrared spectroscopy investigation pointed out the possible presence of more than one isomer in gas phase [40]. The low-lying energy tautomers of free nucleobase considered in this investigation, interacting with Zn^{2+} ion gave rise to five different species (see Fig. 2) with the following stability order: $\text{G1a-Zn}^{2+} > \text{G5-Zn}^{2+} > \text{G3-Zn}^{2+} > \text{G4-Zn}^{2+} > \text{G1b-Zn}^{2+}$. The G2-Zn^{2+} complex does not exist.

The G1a-Zn^{2+} and G1b-Zn^{2+} species derive from the same most stable canonical tautomer of the free base but they show significant energy differences, in part due to different size of the ring generated upon interaction with the metal ion. In G1a-Zn^{2+} global minimum complex, the ion is coordinated to the O6 and N7 atoms with bond lengths of 1.943 and 1.971 Å, respectively. In G1b-Zn^{2+} , lying at 57.1 kcal/mol above the global minimum, the cation approaches the N3 nitrogen atom and at the same time induces the pyramidalization of the amino group, so that it can also bind the N2 atom. The pyramidalization requires a considerable amount of energy, which is another reason for the minor stability of G1b-Zn^{2+} with respect G1a-Zn^{2+} (see Fig. 2).

In all the guanine complexes, the Zn^{2+} ion is bi-coordinated contrary to that occurs in the case of the interaction with the Cu^{2+} metal ion [18]. The same behavior, instead, was observed with the Mn^{2+} ion [17]. The simplest explanation for these findings lies in all likelihood in the different electronic nature of Zn^{2+} and Mn^{2+} (closed shells) with respect to that of Cu^{2+} (open shell) because, considerations about the size of the ions, in some respects not very different, should lead to opposite results. The second lowest-energy minimum, G5-Zn^{2+} has the Zn^{2+} ion bound to the N3 and N9 nitrogen atoms with the almost the same distance (2.016 and 2.017 Å, respectively). In the G4-Zn^{2+} structure, the Zn^{2+} ion establishes bonds with O6H (2.116 Å) and N1 (1.952 Å) atoms (see Fig. 2). This compounds appears to be topologically similar to the corresponding complex with Mn^{2+} ion [17], while, in the case of G4-Cu^{2+} complex [18], the Cu^{2+} ion results mono-coordinated to the N1 nitrogen atom.

The remaining adduct, G3-Zn^{2+} lies at 33.8 kcal/mol above the global minimum. Cation exhibits the same coordination mode encountered in the G1a-Zn^{2+} adduct although now the bond involves the

oxygen of an OH group, and not the O6 carbonyl oxygen (see Fig. 2). However, this observation can serve as a support to identify in the N7 and O6 atoms of free guanine, the preferred binding sites for Zn^{2+} ion.

As in the case of uracil and thymine, the metalation process of adenine occurs preferentially on the highest-energy tautomer of the isolated base (A3).

The interaction leads to complexes in which the zinc ion bi-coordinated gives rise to five- and four-membered cycles in $A1-Zn^{2+}$ and $A3-Zn^{2+}$ and $A2-Zn^{2+}$, respectively. The stability order of the various complexes can be rationalized on the basis of the size of the formed ring, taking into account that the five-membered rings are favored because of their minor annular strain. Moreover, the interaction with zinc ion is more stabilizing if carbonyl oxygen and imino nitrogen rather than hydroxyl and amino groups are involved.

$A1-Zn^{2+}$ and $A2-Zn^{2+}$ lie at 27.5 and 18.5 kcal/mol above the global minimum $A3-Zn^{2+}$, respectively (see Table I). In the most stable complex ($A3-Zn^{2+}$), the atoms involved in the metal ion coordination are N7 and N6-H with distances of 1.979 Å and 1.936 Å, respectively. In the $A1-Zn^{2+}$ compound, the cation interacts still with N6H2 and N7 couple of atoms but bond length with N6 (2.052 Å) reflects the minor reactivity of amino with respect to imino nitrogen. In the $A2-Zn^{2+}$ complex, the N9 and N3 atoms are involved in the Zn^{2+} coordination (see Fig. 2). The strain of four-membered cycle that originates on interaction with cation can be the reason of the scarce stability of this complex as also occur for $G1b-Zn^{2+}$.

For all global minimum complexes, the NBO analysis showed that no covalent bond between ion and ligand exists. The charge transfer of $\sim 0.4 |e|$ from ligand to cation underlines, a prevalent ionic contribution to the bond of Zn^{2+} with the coordinating atoms. Zn^{2+} ion, with its full-filled orbital configuration exhibits a behavior very similar to that observed in the case of the interaction between nucleobases and Mn^{2+} [17].

MIA VALUES

Absolute Zn^{2+} metal ion affinity (MIA) values, calculated at 298 K, are reported in Table II together with those previously obtained at the same level of theory for Mn^{2+} - and Cu^{2+} - DNA and RNA nucleic acid complexes [17, 18].

TABLE II
Zinc, manganese, and copper dications affinity values in kcal/mol for nucleobases.*

	MIA_{Zn}	MIA_{Mn}^a	MIA_{Cu}^b
U1— M^{2+}	175.6	—	238.2
U2— M^{2+}	207.9	173.8	246.8
U3— M^{2+}	214.4	181.1	248.5
T1— M^{2+}	185.9	—	250.9
T2— M^{2+}	219.5	180.4	268.9
T3— M^{2+}	221.1	186.6	265.8
C1— M^{2+}	226.1	191.4	264.4
C2— M^{2+}	193.8	159.3	253.9
C3— M^{2+}	207.0	172.5	262.2
C4— M^{2+}	195.8	157.8	258.3
G1a— M^{2+}	246.8	187.3	294.9
G1b— M^{2+}	189.4	—	262.9
G2— M^{2+}	—	207.7	290.8
G3— M^{2+}	213.8	169.0	293.5
G4— M^{2+}	204.3	167.1	281.2
G5— M^{2+}	230.4	192.8	280.8
A1— M^{2+}	212.9	170.5	250.1
A2— M^{2+}	229.9	192.7	287.4
A3— M^{2+}	259.1	216.2	294.9

* Values are reported at 298 K.

^a From Ref. [17].

^b From Ref. [18].

In principle, the MIA value should be referred to the most stable complex that the cation forms with the particular nucleobase. As we have observed in the previous section, the most stable complex can originate on interaction of Zn^{2+} metal ion with a tautomeric form that not always is the most stable one. This is not a novelty, as the same situation was found in other similar studies involving different transition metal ions [17–19].

The problem could be of little significance when tautomers fall in a narrow range of energy, as in the case of cytosine and guanine, since, as a general rule, their contemporary presence should guarantee the existence of all their complexes, so that the choice of the most reliable MIA values should depend only on the stability order of these complexes. In contrast, as far as uracil, thymine, and adenine are concerned, the high-energy difference between tautomers prevents, after all, the formation of adducts deriving from high-energy free forms irrespective of their intrinsic stability. That being the case, it is clear that, although $U3-Zn^{2+}$, $T3-Zn^{2+}$, and $A3-Zn^{2+}$ are much more stable than the other possible complexes of uracil, thymine and adenine

nucleobases, they cannot exist really, being the canonical U1, T1, and A1 the only tautomers present under any physical condition.

Thus, Table II shows that we can individuate as the most probable MIA for uracil a value of 175.6 kcal/mol. No comparison is possible with MIA of manganese cation, as it does not form stable complex with either uracil or thymine canonical tautomers. The metal affinity of Cu^{2+} for this nucleobase (238.2 kcal/mol) underlines the different bond strength in the two cases.

In correspondence with T1- Zn^{2+} complex we calculated a MIA value of 186.0 kcal/mol that is slightly higher than that of uracil toward the same metal ion and decidedly lower than that of thymine towards copper ion (250.9 kcal/mol).

For adenine, the A1- Zn^{2+} species defines a metal ion affinity value of 212.9 kcal/mol. Values for manganese and copper cations are 170.5 and 250.1 kcal/mol, respectively.

In the case of the cytosine-zinc complexes, the energy difference that separates the four obtained adducts is high, consequently the MIA values range from a maximum of 226.1 kcal/mol for C1- Zn^{2+} to a minimum of 193.8 kcal/mol for C2- Zn^{2+} . As previously discussed, all these complexes, except perhaps that originating from C4 form, are possible because of the coexistence of C1 and C2 tautomers and C3 rotamer. For this reason, the choice of the most probable MIA value depends on the stability order of complexes that, although with wider gaps, follows that of free tautomers. So, the value of 226.1 kcal/mol appears to be the most appropriate one in this case. The comparison with the other Mn^{2+} (191.4 kcal/mol) and Cu^{2+} (264.4 kcal/mol) ions proposes again the same trend of affinity in these metal ions series.

Finally, the guanine-zinc species present a situation very similar to that observed in the cytosine case. The five different complexes lie in a range of energy of ~ 57 kcal/mol. The highest MIA value, 246.8 kcal/mol, is found in correspondence of the G1a- Zn^{2+} complex and represents the most reliable one. In the case of the guanine- Mn^{2+} complexes, the lowest-energy minimum is different from that obtained in the Zn^{2+} and Cu^{2+} -guanine interaction. Nevertheless, comparing the MIA values of the three metal ions (246.8 for zinc, 187.3 for manganese and 294.9 kcal/mol for copper) for this nucleobase, it was found again that copper forms with DNA and RNA nucleic acids bases complexes more stable than zinc and manganese cations. The order

with which Zn^{2+} prefers to attack the nucleobases is $\text{G} > \text{C} > \text{A} > \text{T} > \text{U}$.

This trend, relatively to the guanine and adenine nucleobases, is the same found by Burda et al. [16] but, while the agreement with the MIA value for guanine is very good (only 1.3 kcal/mol of difference), the absolute value that we have obtained for adenine differs noticeably from that reported by these authors. In particular, the B3LYP/6-311+G(2df,2p) MIA value is 51.7 kcal/mol higher than that obtained by an MP2 single-point computation on a structure optimized at the HF level. The disagreement can be due in part to the different theoretical approaches, but mainly to the fact that the global minima do not coincide in the two cases.

Because of the lack of experimental information on this particular subject, no comparison is possible but, the reliability of the protocol used in this work was widely demonstrated in many previous similar studies [14, 15, 17–19].

Conclusions

On the basis of results obtained, some conclusions can be drawn:

1. Zinc ion appears to interact favorably with all nucleic acid bases giving rise to a number of stable complexes equal to that of their possible tautomeric forms.
2. The complexes that actually exist, however, derive from canonical tautomers of uracil, thymine and adenine and from the coexistent tautomers of cytosine and guanine.
3. The preferred attachment sites for the cation are limited to the species that have a real possibility to be present: the O4 atom of uracil and thymine for the mono-coordination, the O2 and N3 atoms of cytosine, the N7 and O6 atoms of guanine, and finally the N7 and N6 atoms of adenine for the bi-coordination cases.
4. The MIA values depend on the stability order of existing complexes other than on the coordination mode of cation to the base.
5. For uracil, thymine, and adenine, because of the predominance of canonical tautomer in any experimental condition, the MIA values were determined taking into account the complexes of this form. For cytosine and guanine, for whose many complexes are present, the

most stable ones were considered to determine the affinities of zinc cation.

6. The values we have proposed as the most reliable indicate that the nucleobase more subjected to the Zn^{2+} attack is the guanine, followed by cytosine, adenine, thymine and uracil. The present study can be useful for future experimental investigation on the zinc-nucleobases interaction; for the time being, it represents the first source of information in this area.

References

- Bray, T. M.; Bettger, W. J. *Free Radic Biol Med* 1990, 8, 281.
- Powell, S. R. *J Nutr* 2000, 130, 1447S.
- Ho, E. *J Nutr Biochem* 2004, 15, 572.
- Cousins, R. J. *Proc Nutr Soc* 1998, 57, 307.
- Falchuk, K. H. *Mol Cell Biochem* 1998, 188, 41.
- Vallee, B. L.; Falchuk, K. H. *Physiol Rev* 1993, 73, 79.
- Kang, S.; Wells, R. D. *J Biol Chem* 1994, 269, 9528.
- Sabat, M.; Lippert, B. In *Metal Ions in Biol Systems*; Vol 33; Sigel, A.; Sigel, H., Eds.; Marcel Dekker: New York, 1996; p 143.
- Sigel, H. *Chem Soc Rev* 1993, 18, 32.
- Hettich, R. H. *Int J Mass Spectrom* 2001, 204, 55.
- Cerda, B. A.; Wesdemiotis, C. *J Am Chem Soc* 1995, 117, 9734.
- Rodgers, M. T.; Armentrout, P. B. *J Am Chem Soc* 2000, 122, 8548.
- Cerda, B. A.; Wesdemiotis, C. *J Am Chem Soc* 1996, 118, 11884.
- Russo, N.; Toscano, M.; Grand, A. *J Phys Chem B* 2001, 105, 4735.
- Russo, N.; Toscano, M.; Grand, A. *J Am Chem Soc* 2001, 123, 10272.
- Burda, J. V.; Šponer, J.; Hobza, P. *J Phys Chem* 1996, 100, 7250.
- Russo, N.; Sicilia, E.; Toscano, M.; Grand, A. *Int J Quantum Chem* 2002, 90, 903.
- Marino, T.; Russo, N.; Toscano, M.; Grand, A. *Int J Quantum Chem* 2004, 98, 347.
- Russo, N.; Toscano, M.; Grand, A. *J Mass Spectrom* 2003, 38, 265.
- Becke, A. D. *J Chem Phys* 1993, 98, 5648.
- Lee, C.; Yang, W.; Parr, R. G. *Phys Rev B* 1988, 37, 785.
- Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Montgomery, J. A., Jr.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Petersson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Knox, J. E.; Hratchian, H. P.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.; Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03; Revision C.02*; Gaussian: Wallingford, CT, 2004.
- (a) Carpenter, J. E.; Weinhold, F. *J Mol Struct (THEOCHEM)* 1988, 169, 41; (b) Weinhold, F.; Carpenter, J. E. *Plenum* 1988, 227.
- Russo, N.; Toscano, M.; Grand, A.; Jolibois, F. *J Comput Chem* 1998, 19, 989.
- Colominas, C.; Luque, F. J.; Orozco, M. *J Am Chem Soc* 1996, 118, 6811.
- Paglieri, L.; Corongiu, G.; Estrin, D. A., *Int J Quantum Chem* 1995, 56, 615.
- Roehrig, G. H.; Oyler, N. A.; Adamowicz, L. *J Phys Chem* 1995, 99, 14285.
- Holmen, A.; Broo, A. *Int J Quantum Chem (Quantum Biol Symp)* 1995, 22, 113.
- Gould, I. R.; Burton, N. A.; Hall, R. J.; Hillier, I. H. *J Mol Struct (Theochem)* 1995, 331, 147.
- Katritzky, A. R.; Karelson, M. *J Am Chem Soc* 1991, 113, 1561.
- Kryachko, E. S.; Nguyen, M. T.; Zeegers-Huyskens, T. *J Phys Chem A* 2001, 105, 1288, 1934.
- Fogarasi, G. *J Mol Struct* 1997, 271, 413.
- Kobayashi, R. *J Phys Chem* 1998, 102, 10813.
- Ha, T.-K.; Gunthard, H. H. *J Mol Struct* 1993, 300, 619.
- Leszczynski, J. *J Phys Chem* 1998, 102, 2357.
- Szczepaniak, K.; Szczepaniak, K.; Kwaitkowski, J.; Kubulat, K.; Person, W. B. *J Am Chem Soc* 1988, 110, 8319.
- Nowak, M. I.; Lapinski, L.; Fullara, J. *Spectrochim Acta* 1989, A45, 229.
- Brown, R. D.; Godfrey, P. D.; McNaughton, D.; Pierlot, A. P. *J Am Chem Soc* 1987, 111, 2308.
- Dreyfus, M.; Bensaude, O.; Dodin, G.; Dubois, J. E. *J Am Chem Soc* 1976, 98, 2353.
- Szczepaniak, K.; Szczepaniak, M.; Szajda, W.; Person, W. B.; Leszczynski, J. *J Can Chem* 1991, 69, 1718.

ARTICOLO 3

A theoretical study on tautomerization processes of dehydrated and monohydrated cytosine.

A theoretical study on tautomerization processes of dehydrated and monohydrated cytosine

Donatella Mazzuca, Tiziana Marino, Nino Russo *, Marirosa Toscano

Dipartimento di Chimica and Centro di Calcolo ad Alte Prestazioni per Elaborazioni Parallele e Distribuite-Centro d'Eccellenza MURST, Università della Calabria, I-87030 Arcavacata di Rende (CS), Italy

Received 22 January 2007; accepted 16 February 2007
Available online 24 February 2007

Abstract

Some dehydrated and monohydrated tautomeric forms of cytosine enclosed in an energy range of about 20 kcal/mol were investigated at B3LYP/6-311+G(2df,2p) density functional level of theory with the aim to evaluate the interconversion barriers between them and to establish if it possible explain the coexistence, experimentally proven, of some keto-amino and enol-amino forms of free bases. Results indicated that, the clearing of the activation barriers involved in gas-phase is unlikely but, the presence of a coordination water molecule acts as a catalyst making quite easier the transformation processes. Limited to the two most stable keto-amino and enol-amino tautomers, interconversion was studied in the presence of two explicit water molecules. From the obtained data it appeared clear that the influence of the second water molecule as well as the introduction of solvent effects in reducing the height of the tautomerization barriers for these bihydrated systems is quite small.

© 2007 Elsevier B.V. All rights reserved.

Keywords: Tautomeric equilibria; Microhydration of DNA nucleic acid bases; Cytosine; Density functional computations

1. Introduction

Understanding the tautomeric equilibria of nucleic acid bases is fundamental as the hydrogen bond interactions of DNA and several models of spontaneous mutations of DNA are based on them.

The tautomerism phenomenon is important for all DNA and RNA bases but it has particular significance for cytosine and guanine owing to the existence for these last systems of several forms, whose presence was ascertained under thermal vaporization conditions, lying in a narrow range of energy. However, despite their comparable stability in gas phase, the interconversion of cytosine and guanine tautomers requires very large energy expense.

Cytosine tautomeric equilibria and those involved in some organic systems, have received all along great atten-

tion both from experimental [1–7] and theoretical [8–15] point of view.

Many solution techniques [4,5] were used successfully to determine the most stable tautomer of cytosine. The most part of these studies [1–5] indicated the canonical form of this base as the preferred one, while through a Resonance Enhanced Multiphoton Ionization (REMPI) [6] investigation it was established that both keto-amino and enol-amino tautomers coexist in the gas-phase. Matrix isolation infrared and Molecular beam microwave (MW) spectroscopies [7,16] suggested the contemporary presence of keto-amino, enol-amino and keto-imino tautomers.

Theoretical works [8–14] have taken into account several interesting aspects of this topic.

In particular, Gorb and Leszczynski [11,17] evaluated at MP4(SDQ)/MP2 level of computation, the gas phase energy barrier of the interconversion process between the two lowest energy keto-amino and enol-amino forms both in the absence and in the presence of a water molecule coordinated to the base. Results indicated that a lowering

* Corresponding author. Fax: +39 0984 492044.
E-mail address: nrusso@unical.it (N. Russo).

of 21.6 kcal/mol is obtained for the barrier in going from anhydrous to hydrated system.

B3LYP/6-31++G** and MP2 calculations in gas phase and in solvent were performed by Sambrano et al. [14] in order to determine the relative energies of six low lying tautomers of cytosine. At MP2 level, enol-amino form was found to be the most and less favoured one in gas phase and in solvent, respectively. According to B3LYP results, canonical tautomer was the predominant species both in gas phase and in water solution.

PCM/HF and PCM/AM1 solvation methods were used in the works of Alemán [12,18] to investigate the keto- and two enol- amino (rotamers) forms of cytosine in an environment of a discrete number of water molecules placed in a hydrophilic region of the base. In this situation, keto-amino tautomer remains the most abundant species but the energy gap that separates it from the two enol-amino rotamers becomes very large (16–19 kcal/mol with both the used solvent models).

Micro-hydration of cytosine by one and two water molecules was taken into account in a Molecular Dynamics and RI-MP2 [10] and MP2 (CCSD) [13] study, respectively. Different coordination sites for water were explored in both cases on three [13] and five [10] tautomeric forms of cytosine in order to determine the relative energies of these species. In the RI-MP2 [10] study, it was found that water destabilizes the keto-imino forms with respect to the canonical one but authors did not exclude their existence in biological conditions. The MP2 (CCSD) [13] investigation confirmed the major stability of monohydrated keto-amino tautomer with respect to the other forms because of its strong binding energy with water that generates a compensation effect for the energy difference in free monomers.

The unimolecular and bimolecular tautomerization of isolated and dimeric cytosine was investigated by Yang and Rodgers [8] at the MP2(full)/6-311+G(2d,2p)//MP2(full)/6-31G* and MP2(full)/6-311+G(2d,2p)//B3LYP/6-31G* levels of theory. In this study, the positive effects of dimers formation involving dual hydrogen bond interactions on the tautomerization kinetic of cytosine is suggested.

Russo et al. [9] studying the four most stable tautomers of cytosine in gas phase at B3LYP/6-311+G(2df,2p) density functional level, concluded that the interconversion barriers are too large to be overcome by thermal vaporization.

In most of studies present in the recent literature [10–14,17,19], particular emphasis was laid on the hydration of isolated nucleic acid bases because it was proven that explicit water molecules in their first solvation shell have a clear influence in the electronic structure of the solute and in the tautomerization processes.

For this reason we carried out a theoretical analysis in which the final goal was to establish the role of a water molecule coordinated to the most suitable sites on seven cytosine tautomers in the lowering of the barriers for all the possible interconversion processes between them.

Furthermore, the keto-amino \rightarrow enol-amino tautomerization was analyzed in more details with the aim to evaluate the influence of a second water coordination molecule on the barrier height of this process. Solvent effects on bihydrated systems tautomerization path were also taken into account.

2. Computational details

We performed a theoretical study based on density functional theory (DFT).

All computations were done using Gaussian 03 code [20] with the hybrid Becke exchange (B3)–Lee, Yang and Parr correlation (LYP) functional [21]. Geometry optimizations were performed with valence triple zeta 6-311+G(2df,2p) basis set which contains one set of diffuse and two sets of polarization functions on every atom.

Harmonic vibrational frequencies were computed at the same level of theory to verify the nature of minima and transition states of stationary points along the tautomerization path.

Transition states were confirmed applying the intrinsic reaction coordinated (IRC) procedure implemented in the Gaussian 03 package [22,23].

Zero point energy (ZPE) was enclosed in all energetic data.

Solvation energies were obtained as single point calculations using the self consistent reaction field polarisable continuum model (SCRF/PCM) procedure [24–26]. The surrounding water was modelled using the standard value of 78 for the dielectric constant.

3. Results and discussion

Despite the existence of some tautomers in DNA is hampered by the presence of the sugar moiety at N1 position, the study of these species and their tautomerization mechanisms can give useful information into the biological context provided that they are to be considered as simple model systems.

The seven isolated tautomers that we included in our study are sketched in Fig. 1.

Among these we have keto-amino (C1, C5), enol-amino (C2, C3), keto-imino (C4) and enol-imino (C6 and C7) forms.

C1, C2, C3, C4 and C5 are indicated as the most stable species by all literature sources although some controversial, concerning the relative stabilities of C1 and C2, exists at theoretical level.

Generally, B3LYP calculations [9] favours the C1 canonical over the C2 form contrary to what occurs with MPn or ab initio methods [8,10–14]. However, the energy separation between these two tautomers is computed to be a little more than 1 kcal/mol at all levels of theory, thus a possible error of the same order of magnitude could invalidate the interconversion barrier estimation.

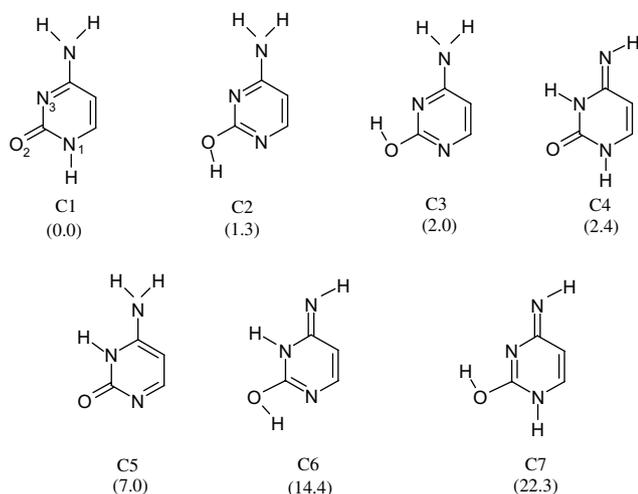


Fig. 1. Free cytosine tautomers. Numbers in parentheses represent relative energy values in kcal/mol.

C6 and C7 enol-imino rare tautomers lie at higher energy with respect to other selected species but they were considered on anyway to verify if also in these drastic cases it can exist some mechanism for generating them starting from the most common and stable forms.

The tautomerization path involving all these species is depicted in Fig. 2. Transition states in this figure, are referred to two types of processes that include proton transfers (TS_{1,2}, TS_{1,4}, TS_{4,6}, TS_{4,7}, TS_{3,5}) and rotations around C–OH single bond (TS_{2,3}). As reported in our previous [9] and in another theoretical work [8], all interconversion barriers between the first five most stable minima are insurmountable except for that promoting the C2–C3 transformation that is the only rotation. The C6 and C7

are in principle achievable from C4 but practically unreachable in gas phase (see Fig. 2).

On the basis of a study concerning the solvent-mediated tautomerization of purine bases [19] in which the authors suggest that water molecules may act as catalysts lowering significantly the barriers for the proton transfer mechanisms, and of the results obtained by Leszczynski et al. [11,17] in their investigation on the some interconversion process of cytosine tautomers, we retained interesting to explore the tautomerization path between the various species assisted by a water molecule.

Selecting the hydrophilic centres on cytosine ring we searched the most stable monohydrated form for all tautomers. The preferred microsolvated structures are reported in Fig. 3. Due to their comparable energy values, we included in this figure more than one hydrated form as far as the C1 and C4 tautomers are concerned. In fact, the C1H₂O and C1'H₂O differ by 0.7 kcal/mol, while C4H₂O, C4'H₂O and C4''H₂O fall in an energetic range of only 1.8 kcal/mol.

All the minima are characterized by two hydrogen bonds (see Fig. 3 for their length) involving the water and nitrogen and oxygen atoms of cytosine ring. The most stable situations are those in which the water molecule acts as hydrogen acceptor and donor simultaneously.

From the energetic point of view, hydrated system follows almost the same trend observed for dry tautomers. The only exception concerns the C4 and C3 whose stability is reversed into the two cases. Our trend is slightly different from that obtained at MP2 level of theory [10] in which the C2 tautomer is always favoured both as isolated than as hydrated system. In another MP_n ($n = 2, 4$) work [11], authors obtained different information depending on the used methods. In particular, the MP4 approach suggested

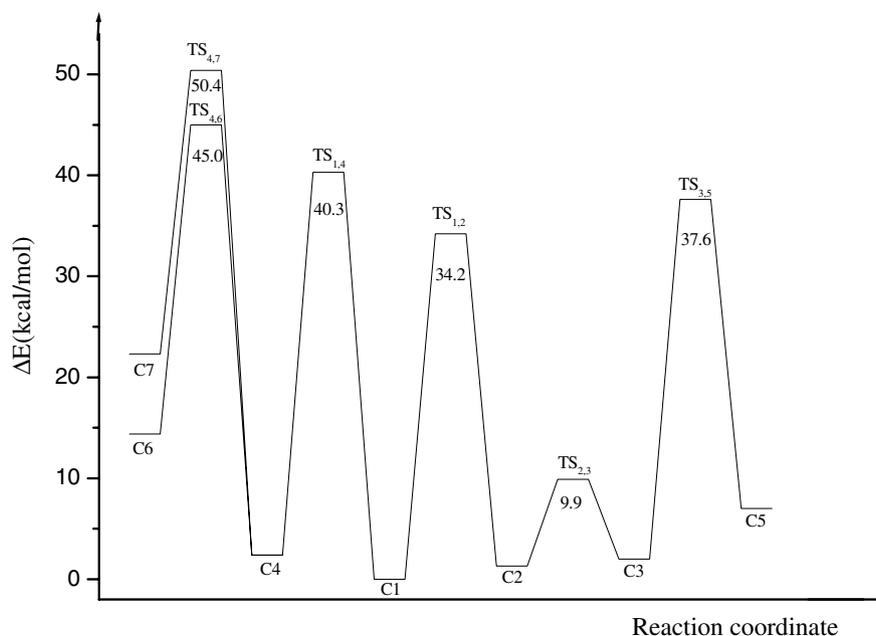


Fig. 2. Interconversion path of free cytosine tautomers. Transition states relative energy is given in kcal/mol.

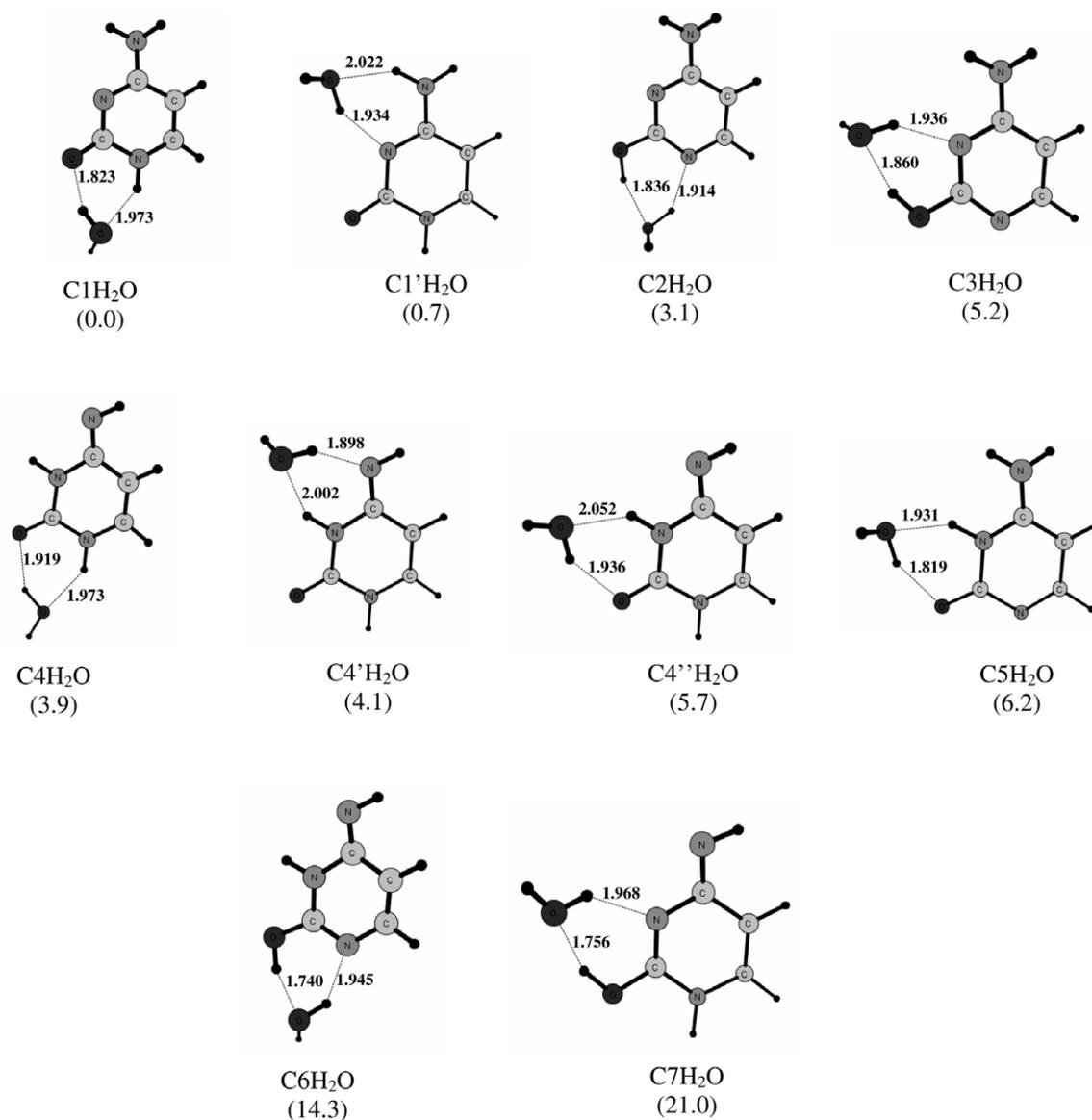


Fig. 3. Optimized geometries of monohydrated cytosine tautomers. Numbers in parentheses represent relative energy values in kcal/mol. Distances are in Å.

the predominance of the C1 keto-amino tautomer irrespective of the absence or the presence of water. At MP2 level C1 was found less stable than C2, while the stability of C1H₂O with respect to C2H₂O was found to be dependent on the used basis set.

The fact that the energetic range in which all hydrated tautomers fall is of the same width with respect to that of the forms lacking water, supports the probability that these species can exist and interconvert each in the other also when are monosolvated.

The water assisted tautomerization path is reported in Fig. 4.

All proton transfers involved in the transition states appear to be mediated by the water molecule. In the C1H₂O → C2H₂O interconversion, an hydrogen atom of the water molecule is transferred on the keto oxygen of the cytosine ring just as the water receives an hydrogen

from N1 atom. The imaginary frequency of 1487 cm⁻¹ corresponds to the double proton shift. The process requires 11.6 kcal/mol that is an energy amount far minor than that necessary without water mediation. Our result is substantially in agreement with the work of Gorb and Leszczynski [11] because these authors find barriers whose height is 13.4 and 15.8 kcal/mol at MP2 and MP4 levels, respectively.

If the water must have an active role in the tautomerization, the most stable hydrated form of the C4 (C4H₂O) cannot be obtained neither from C1H₂O nor from the C1'H₂O. In fact, in the first case, the water position is very far from the region involved in the tautomerization. The barrier for the C1H₂O → C4H₂O interconversion obtained without the water mediation retains almost the same value of that obtained in the C1 → C4 tautomerization (40.4 kcal/mol). Instead, as can be seen from the Fig. 3, the C1'H₂O system is not suitable to reach the C4H₂O

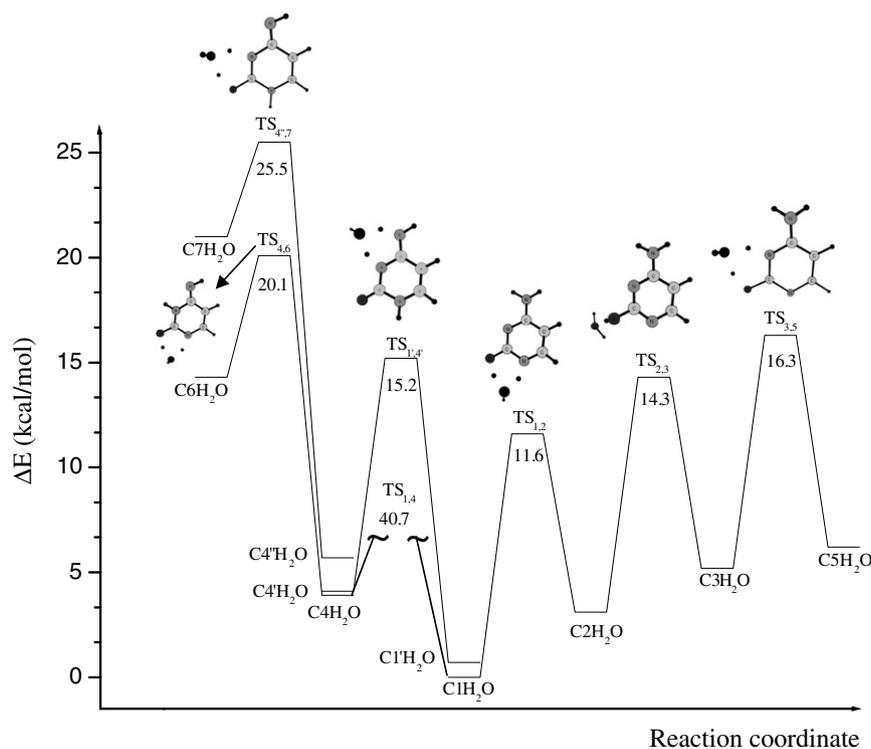


Fig. 4. Interconversion path of monohydrated cytosine tautomers. Transition states picture and their relative energy (in kcal/mol) are reported explicitly on the graph.

form because of the different placement of water in these two minima.

Starting from C1'H₂O it is possible obtain only the C4'H₂O system that however, is practically isoenergetic with C4H₂O (only 0.2 kcal/mol of energy difference).

The only interconversion occurring through a rotation, namely C2H₂O → C3H₂O, appears to be more difficult with respect to the same process between dehydrated tautomers. Transition state, characterized by the water molecule placed perpendicularly to the molecular plane, lies at 11.2 kcal/mol above C2H₂O. Imaginary frequency of 238 cm⁻¹ describes the out of plane motion of water molecule.

11.1 kcal/mol are required to convert C3H₂O into C5H₂O tautomer (22.5 kcal/mol less than the corresponding tautomerization between C3 and C5). The proton transfer is based on a like-shuttle water behaviour that removes a hydrogen from the hydroxyl group and gives it to the N3 atom. Also in this case, the vibrational imaginary frequency (1481 cm⁻¹) corresponds to the stretching of the O–H bond that is breaking coupled to the stretching of the N3–H bond that is forming.

The way to obtain, at least in principle, C6H₂O and C7H₂O species starts from C4H₂O and from C4'H₂O, respectively. Although the interconversion barriers are sensibly lower (about the half) than those involved in the C4 → C6 and C4 → C7 transformations, they remain still too high to cross.

The data obtained until now, underline that, except for the water assisted tautomerizations that involve C6H₂O

and C7H₂O, each individual process entails an energetic expense of about 11 kcal/mol. This, together with the fact that the proton transfers usually involve a quantum mechanical tunnelling that makes faster the process, suggests that the kinetic of the tautomerizations is strongly influenced by the presence of a water molecule.

In order to verify if the addition of a second water molecule coordinated to cytosine tautomers could have some further positive effects on the barriers height, we have studied the interconversion process between the bihydrated forms of C1 and C2.

Results were reported in Fig. 5.

The bihydrated form of C1(H₂O)₂ suitable for this interconversion was build up placing in the proximity of the tautomerization sites two water molecules linked by an hydrogen bond (see Fig. 5). However, this structure revealed as the most stable one among all other bi-hydrated forms of C1. The same result was also found by a previous theoretical RI-MP2 study [10]. The lowest energy C2(H₂O)₂ system shows an analogous disposition of the two water molecules.

When coordinated to a water dimer, the energy difference between the keto-amino and enol-amino tautomers increases with respect to that obtained in the cases of de- and mono-hydrated species. B3LYP computations suggest again a predominance of canonical over the enol-amino form ($\Delta E = 4.5$ kcal/mol) in agreement with the RI-MP2 findings [10].

As it can be observed, no significant improvement concerning the kinetic of the process occurs.

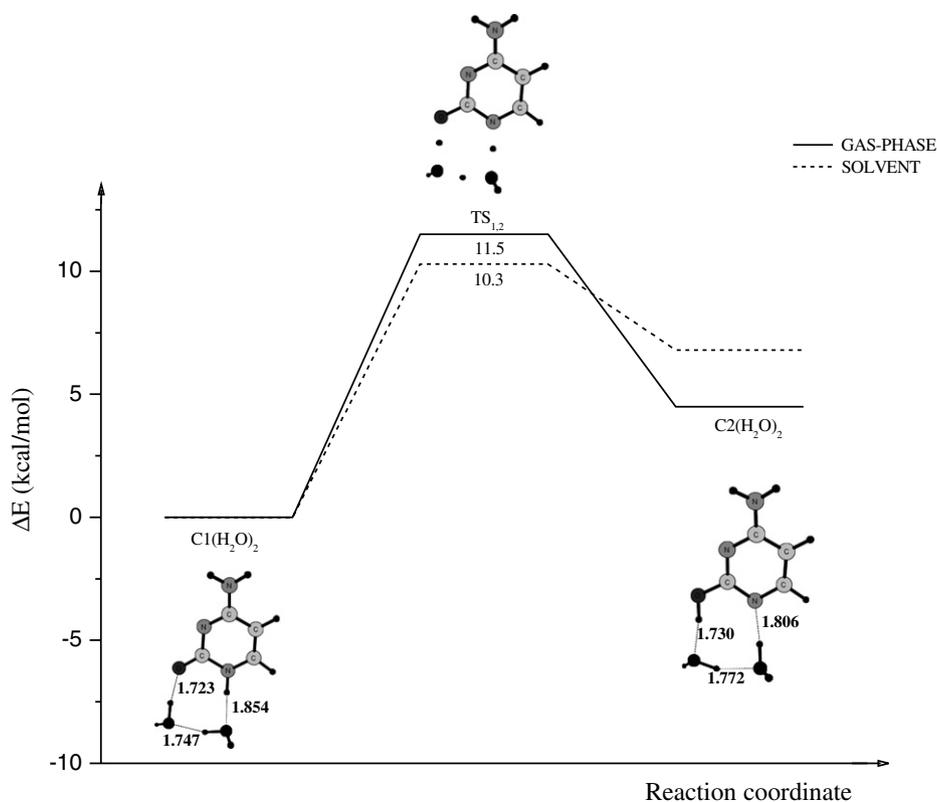


Fig. 5. Energetic profile for the interconversion between C1 and C2 bihydrated tautomers in gas phase (solid line) and in water (dashed line). Relative energy values are given in kcal/mol.

The barrier lowers of only 0.1 kcal/mol compared to that obtained previously.

In the study of Ahn et al. [19], it was demonstrated that the addition of an increasing number of water molecules in the coordination sphere of guanine, is able to lower gradually the barrier for the 9–7 and 9–3 tautomerizations through double and triple concerted proton transfers. The most relevant effect is obtained with the first explicit water but also the second one gives rise to substantial reduction of the barriers height. An opposite tendency can be observed as the third water molecule is considered.

Our results together with those obtained for guanine indicate that a precise number of water molecules depending on the structural features of nucleic acid bases and on the particular tautomerization process, exist that can make easier the interconversion between the various tautomeric forms.

Finally, the solvent effects as computed in a polarizable continuum were evaluated for the $C1(H_2O)_2 \rightarrow C2(H_2O)_2$ transformation. The energetic profile is reported in Fig. 5.

The solvation enhances the energy difference between bihydrated C1 and C2 forms (6.8 kcal/mol). This is in agreement both with the experimental observation that indicate the canonical as that far predominant in solution, and with previous theoretical determinations [12,14].

The barrier appears to be almost similar to that for bihydrated systems in gas phase (10.3 versus 11.5 kcal/mol) but the process is thermodynamically more unfavoured.

This last result proposes that the presence of explicit water molecules is in a certain sense defeated by the continuum electric field. This concept, transferred to a real solution, could mean that the water molecules directly linked to cytosine lose partially their ability as donor or acceptor of hydrogen atoms because of the presence of other water molecules in the outer coordination sphere.

4. Conclusions

The study in gas phase of seven cytosine tautomers as isolated, monohydrated and bihydrated species and of their interconversion processes was performed at B3LYP level of theory.

A combined explicit/polarisable continuum model calculation was used to evaluate the solvent effects on the $C1 \rightarrow C2$ transformation.

On the basis of obtained results, the following conclusion can be drawn:

- The tautomerization processes between isolated species cannot occur owing the high energetic barriers characterizing the involved proton shifts.
- Microhydration with one water molecule causes a great number of monohydrated systems and a more pronounced energy separation between some of these with respect to dehydrated tautomers. In the presence of monohydrated species, the complexity of the tautomer-

ization path increases especially as for as the keto-amino→keto-imino interconversion is concerned. The activation barriers appears to be reduced by a big amount of energy making all mutual transformations feasible at least at high temperature. A quantity of about 11.0 kcal/mol is required for each single tautomerization except for those that involve the C6 and C7 hydrated forms that continue to be unlikely also in these last conditions. The role played by the water molecule is that to mediate, thanks to its properties of donor and acceptor of hydrogen, the proton transfers that usually demand very high energy expenses.

- The introduction of a second water molecule in the coordination sphere of C1 and C2 tautomers results in a further destabilization of C2 with respect to C1 and, more importantly, does not affect the interconversion barrier between them.
- The solvent surrounding effects on the bihydrated C1 and C2 tautomeric forms is similar to that introduced by the second water molecule. The two systems are now still more different in energy with a net predominance of C1(H₂O)₂ species. The tautomerization process appears to be slightly more favoured with respect to the gas phase from the kinetic point of view but less probable from the thermodynamic one.
- From indications deriving from this study and from previous literature, the number of discrete water molecules to be coordinated to the base in order to have some benefits as far as the rate of tautomerization process is concerned, clearly depend on the structural characteristic of the base and on the type of transformation that occurs.

We think that, the lowering of the barriers observed in the tautomerization involving monohydrated species, together other considerations about their reciprocal concentrations and the occurrence of the quantum mechanical tunnelling that acts in the sense of increasing the rate of the processes can explain the coexistence of some tautomeric forms of cytosine under thermal vaporization conditions.

Acknowledgements

Financial support from the Università degli Studi della Calabria and Regione Calabria (POR Calabria 2000/2006, misura 3.16, progetto PROSICA) is gratefully acknowledged.

References

- [1] D.L. Barker, R.E. Marsh, *Acta Crystallogr.* 17 (1964) 1581.
- [2] R.J. McClure, B.M. Craven, *Acta Crystallogr., Sect. B* 29 (1973) 1234.
- [3] H.P. Weber, B.M. Craven, R.K. McMullan, *Acta Crystallogr., Sect. B* 36 (1980) 645.
- [4] T. Ueda, J.J. Fox, *J. Am. Chem. Soc.* 85 (1963) 4024.
- [5] M. Dreyfus, O. Bensaude, G. Dodin, J.E. Dubois, *J. Am. Chem. Soc.* 98 (1976) 6338.
- [6] E. Nir, M. Müller, L.I. Grace, M.S. de Vries, *Chem. Phys. Lett.* 317 (2000) 437.
- [7] M. Szczesniak, K. Szczepaniak, J.S. Kwiatkowski, K. Kubulat, W.B. Person, *J. Am. Chem. Soc.* 110 (1988) 8319.
- [8] Z. Yang, M.T. Rodgers, *Phys. Chem. Chem. Phys.* 6 (2004) 2749.
- [9] N. Russo, M. Toscano, A. Grand, *J. Am. Chem. Soc.* 123 (2001) 10272.
- [10] S.A. Trygubenko, T.V. Bogdan, M. Rueda, M. Orozco, F. Javier Luque, J. Šponer, P. Slavíček, P. Hobza, *Phys. Chem. Chem. Phys.* 4 (2002) 4192.
- [11] L. Gorb, J. Leszczynski, *Int. J. Quantum Chem.* 70 (1998) 855.
- [12] C. Alemán, *Chem. Phys.* 253 (2000) 13.
- [13] G. Fogarasi, P.G. Szalay, *Chem. Phys. Lett.* 356 (2000) 383.
- [14] J.R. Sambrano, A.R. de Souza, J.J. Queral, J. Andrés, *Chem. Phys. Lett.* 317 (2000) 437.
- [15] P.I. Nagy, W.M.F. Fabian, *J. Phys. Chem. B* 110 (2006) 25026.
- [16] R.D. Brown, P.D. Godfrey, D. McNaughton, A.P. Pierlot, *J. Am. Chem. Soc.* 111 (1989) 2308.
- [17] Y. Podolyan, L. Gorb, J. Leszczynski, *Int. J. Mol. Sci.* 4 (2003) 410.
- [18] C. Alemán, *Chem. Phys.* 244 (1999) 151.
- [19] D.-S. Ahn, S. Lee, B. Kim, *Chem. Phys. Lett.* 390 (2000) 384.
- [20] M.J. Frisch, G.W. Trucks, H.B. Schlegel, G.E. Scuseria, M.A. Robb, J.R. Cheeseman, J.A. Montgomery Jr., T. Vreven, K.N. Kudin, J.C. Burant, J.M. Millam, S.S. Iyengar, J. Tomasi, V. Barone, B. Mennucci, M. Cossi, G. Scalmani, N. Rega, G.A. Petersson, H. Nakatsuji, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, M. Klene, X. Li, J.E. Knox, H.P. Hratchian, J.B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R.E. Stratmann, O. Yazyev, A.J. Austin, R. Cammi, C. Pomelli, J.W. Ochterski, P.Y. Ayala, K. Morokuma, G.A. Voth, P. Salvador, J.J. Dannenberg, V.G. Zakrzewski, S. Dapprich, A.D. Daniels, M.C. Strain, O. Farkas, D.K. Malick, A.D. Rabuck, K. Raghavachari, J.B. Foresman, J.V. Ortiz, Q. Cui, A.G. Baboul, S. Clifford, J. Cioslowski, B.B. Stefanov, G. Liu, A. Liashenko, P. Piskorz, I. Komaromi, R.L. Martin, D.J. Fox, T. Keith, M.A. Al-Laham, C.Y. Peng, A. Nanayakkara, M. Challacombe, P.M.W. Gill, B. Johnson, W. Chen, M.W. Wong, C. Gonzalez, J.A. Pople, *Gaussian 03, Revision C.02*, Gaussian, Inc., Wallingford CT, 2004.
- [21] P.J. Stevens, F.J. Devlin, C.F. Chabrowski, M.J. Frisch, *J. Phys. Chem.* 98 (1994) 11623.
- [22] C. Gonzalez, H.B. Schlegel, *J. Chem. Phys.* 90 (1989) 2154.
- [23] C. Gonzalez, H.B. Schlegel, *J. Phys. Chem.* 94 (1990) 5523.
- [24] S. Miertos, E. Scrocco, J. Tomasi, *Chem. Phys.* 55 (1981) 117.
- [25] S. Miertos, J. Tomasi, *Chem. Phys.* 65 (1982) 239.
- [26] M. Cossi, V. Barone, R. Commi, J. Tomasi, *J. Chem. Phys.* 255 (1996) 327.