

Universidade de Brasília
Instituto de Química
Pós-Graduação em Química

Mestrado em Química

**Estudo teórico das contribuições energéticas envolvidas na
formação dos complexos: $[\text{Mg}(\text{H}_2\text{O})_n\text{-base nucleica}]^{2+}$**

Vytor Pinheiro Oliveira
Orientador: Prof. Dr. Edgardo Garcia

Brasília
2013

Universidade de Brasília
Instituto de Química
Pós-Graduação em Química

Mestrado em Química

**Estudo teórico das contribuições energéticas envolvidas na
formação dos complexos: $[\text{Mg}(\text{H}_2\text{O})_n\text{-base nucleica}]^{2+}$**

Vytor Pinheiro Oliveira
Orientador: Prof. Dr. Edgardo Garcia

Brasília
2013

*Aos meus pais,
aos meus irmãos
e à menina de ouro.*

AGRADECIMENTOS

Sobretudo, agradeço ao professor doutor Edgardo Garcia que me propôs este trabalho e teve paciência para me explicar repedidas vezes, inúmeras questões relacionadas ao projeto e à química computacional de modo geral.

Gostaria também de agradecer ao professor doutor João Batista por me acolher no Laboratório de Química Computacional (LQC) da UnB e a todos os meu colegas do LQC.

Também sou infinitamente grato às pessoas que me ajudaram de algum modo na escrita e na formatação dessa dissertação, em especial à minha mãe, à minha Irmã e à minha namorada Bruna Luana.

Por fim, agradeço à agência de fomento à pesquisa, CAPES, pela bolsa concedida durante a realização desse trabalho.

Resumo

Atualmente, apenas estudos teóricos de alto nível são capazes de tratar de maneira acurada interações específicas entre o Mg^{2+} e os ácidos nucleicos. Esses estudos são importantes para desvendar o papel do Mg^{2+} no dobramento e no mecanismo de reação de ribozimas. O presente trabalho tem como objetivo definir uma estratégia computacional capaz de descrever as contribuições energéticas envolvidas nas reações de complexação do Mg^{2+} com as bases nucleicas timina, uracila, guanina, citosina e adenina nos sistemas $[Mg(H_2O)_n\text{-base nucleica}]^{2+}$ ($n=6,5$ e 4) com a melhor razão possível entre acurácia e custo computacional. A combinação entre as energias eletrônicas de reação obtidas pelo método $MP2(FC)/6-31+G(2d,p)$ e as correções térmicas do cálculo de frequência $B3LYP/6-311G(2d,d,p)$ foram capazes de fornecer resultados semelhantes ao método de alto custo computacional CBS-QB3. As contribuições do solvente foram estimadas pelo modelo de contínuo CPCM; porém a definição da cavidade do solvente nos métodos PCM pode gerar desvios absolutos e relativos difíceis de serem estimados. Independentemente do modo de coordenação, o magnésio interagiu mais fortemente com os sítios O2 da citosina, O6 e N7 da guanina e mais fracamente com sítios da adenina, timina e uracila. Também observou-se que a aproximação de pares empregada por campos de força subestimou a energia de interação dos sistemas $[Mg(H_2O)_5\text{-base nucleica}]^{2+}$ em até 40% e mesmo no solvente contínuo os efeitos de polarização e transferência de cargas entre o metal e as bases nucleicas foram significativos. A estratégia computacional aqui definida poderá ser aplicada a sistemas análogos, assim como a sistemas maiores que envolvam interações similares. Os resultados apresentados poderão ser utilizados para parametrizar e validar outros métodos de menor custo computacional, como campos de força clássicos e métodos semi-empíricos.

Abstract

Currently, only high-level theoretical studies are able to treat specific interactions between Mg^{2+} and the nucleic acids accurately. These studies are important to unveil the role of Mg^{2+} in the folding and in the reaction mechanism of ribozymes. The present work is aimed at defining a computational strategy capable of describing the energies involved in the complexation reaction of Mg^{2+} with nucleic bases thymine, uracil, guanine, cytosine and adenine, in $[\text{Mg}(\text{H}_2\text{O})_n\text{-nucleobase}]^{2+}$ ($n=6,5$ and 4) systems with the best possible ratio between accuracy and computational cost. The combination between reaction electronic energies obtained at MP2(FC)/6-31+G(2d,p) level with thermal corrections calculated by B3LYP/6-311G(2d,d,p) frequencies were able to provide results similar to the high computational cost CBS-QB3. The solvent contributions were estimated using the CPCM. However, the electrostatic cavities definitions can produce absolute and relative errors in solution free energies difficult to be evaluated. Regardless of the coordination mode, magnesium interacted stronger at the cytosine O2, O6, and guanine N7 sites and weaker in adenine, thymine and uracil sites. It was also noted that the pairwise additive approach used by force fields underestimated the binding energy of $[\text{Mg}(\text{H}_2\text{O})_5\text{-nucleobase}]^{2+}$ systems by as much as 40% and even in the continuum solvent, polarization and charge transfer effects between metal and nucleic bases are significant. The method defined herein can be applied to study analog systems, as well as larger systems involving similar interactions. The presented results can be used to parameterize and validate other methods of lower computational cost, such as semi-empirical and classical force fields.

Sumário

Resumo.....	v
Abstract.....	vi
LISTA DE TABELAS.....	x
LISTA DE FIGURAS.....	xii
CAPÍTULO 1: Introdução ao tema.....	1
1 A descoberta do RNA catalítico.....	2
2 O cátion metálico e o mecanismo de ação das ribozimas.....	3
2.1. As funções do Mg^{2+}	4
3 Estudos experimentais e computacionais.....	5
4 Justificativa e objetivos.....	5
CAPÍTULO 2: Estudo dos complexos $[Mg(H_2O)_5\text{basenucleica}]^{2+}$	8
5 Metodologia computacional.....	9
6 Escolha do método.....	11
6.1. Cálculo da energia de interação.....	12
6.1.1 Procedimento counterpoise para corrigir o BSSE.....	13
6.2. Análise dos métodos.....	15
6.2.1 Estudo das energias de interação.....	15
6.2.2 Estudo da correlação eletrônica nos cálculos de ΔE e ΔE_{cp}	17
6.2.3 Aproximação Frozen Core.....	19
7 Escolha do conjunto de bases.....	20
7.1. Conjuntos de bases de Pople.....	22
7.2. Conjuntos de bases de Dunning.....	24
7.3. Análise dos conjuntos de bases.....	25
8 Otimização de geometria.....	29
9 Considerações finais.....	33
CAPÍTULO 3: Limitações dos campos de força clássicos na descrição dos complexos $[Mg(H_2O)_5\text{base nucleica}]^{2+}$	34

10	Campos de força clássicos	35
10.1.	Átomos ligados	35
10.2.	Átomos não ligados	37
11	Aproximação de pares e a energia de muitos corpos.....	38
11.1.1	ΔE_7 versus ΔE_{par}	41
11.1.2	BSSE e a correlação eletrônica:	42
12	Cargas atômicas parciais	43
12.1.	Influência do método e conjunto de bases nas cargas ESP	44
12.2.	Principais diferenças entre as cargas MP2 e HF	45
12.3.	Relação entre as cargas do Mg^{2+} e do sítio de coordenação	49
12.4.	Dipolos.....	51
12.5.	Transferência de cargas	53
13	Considerações finais	55
CAPÍTULO 4: Energias livres de formação dos complexos $[Mg(H_2O)_n \text{ base nucleica}]^{2+}$		56
14	Modos de coordenação.....	57
15	Contribuições térmicas a partir do cálculo de frequências	62
16	Métodos termoquímicos compostos CBS	63
16.1.	Definição de uma estratégia de cálculo acurada e prática	64
16.1.1	Contribuições eletrônicas.....	65
16.1.2	Contribuições térmicas.....	67
17	Energia livre de reação para o estado gasoso	69
18	Efeito do solvente:.....	72
18.1.	Modelos PCM	74
18.2.	Definições da cavidade.....	75
18.3.	Energia livre de solvatação.....	76
18.3.1	Cálculo da energia de hidratação do Mg^{2+}	78

18.4.	Energia livre de reação no solvente.....	81
18.4.1	Comparação entre $\Delta G_r(l)$ obtidos por diferentes métodos	84
19	Correlações com dados experimentais	87
20	Conclusões e perspectivas	88
	REFERÊNCIAS.....	91
	ANEXO 1 - Estudo sistemático dos conjuntos de bases	97
21	Adição de funções difusas nas bases de Pople	97
22	Adição de funções de polarização	98
23	Alterações nos conjuntos de bases de Dunning	101
	ANEXO 2 – Energias de interação.....	102
	ANEXO 3 – Trabalhos apresentados e artigos publicados.....	103

LISTA DE TABELAS

TABELA 1 ENERGIAS DE INTERAÇÃO (ΔE E ΔE_{CP}) E ESTIMATIVAS DOS BSSE OBTIDAS PELAS DIFERENÇAS ENTRE ΔE_{CP} E ΔE CALCULADAS A NÍVEL MP2(FC) E HF UTILIZANDO O CONJUNTO DE BASES AUG-CC-PVTZ,. TODOS OS VALORES ESTÃO EM KCAL/MOL.	16
TABELA 2 ENERGIA DE CORRELAÇÃO ΔE_{COR} E SEU VALOR PERCENTUAL ($\% \Delta E_{COR}$) EM RELAÇÃO A ENERGIA DE INTERAÇÃO (ΔE) SEM CORREÇÃO PARA O BSSE E COM A CORREÇÃO DE COUNTERPOISE (CP) CALCULADOS EM KCAL/MOL.	18
TABELA 3 ENERGIAS DE INTERAÇÃO (ΔE E ΔE_{CP}) OBTIDAS EM MP2(FULL)/AUG-CC-PCVTZ, MP2(FULL)/AUG-CC-PVTZ E MP2(FC)/AUG-CC-PVTZ, O PREFIXO AUG-CC- FOI ABREVIADO POR A- E O MÉTODO MP2 FOI OMITIDO, CP INDICA OS VALORES CORRIGIDOS COM COUNTERPOISE. TODOS OS VALORES ESTÃO EM KCAL/MOL.....	20
TABELA 4 ESQUEMA DE CONTRAÇÕES (CGTOS) DE ORBITAIS DO TIPO GAUSSIANA E ADIÇÕES DE FUNÇÕES DE POLARIZAÇÃO PARA OS ÁTOMOS DE LI-NE DOS CONJUNTOS DE BASES DE DUNNING.....	24
TABELA 5 COMPARAÇÃO ENTRE OS BSSE MÁXIMOS (MAX) E MÉDIOS (MÉDIA) E OS DESVIOS MÉDIOS E MÁXIMOS ABSOLUTOS (DMEDA E DMAXA) DE ΔE E ΔE_{CP} EM RELAÇÃO AOS VALORES MP2(FC)/AUG-CC-PVTZ PARA UMA SÉRIE DE CONJUNTOS DE BASES ORDENADO DECRESCENTEMENTE COM O NÚMERO DE FUNÇÕES DE BASES (F.B) DO CÁLCULO DO MAIOR COMPLEXO (GUA-N7).	26
TABELA 6 ENERGIAS DE INTERAÇÃO ΔE E ΔE_{CP} DOS COMPLEXOS DE ACORDO COM A NOTAÇÃO: GEOMETRIA//ENERGIA. O CONJUNTO DE BASES UTILIZADOS NOS CÁLCULOS FOI O 6-31+G(2D,P). TODOS OS VALORES ESTÃO EM KCAL/MOL.....	32
TABELA 7 ENERGIAS DE INTERAÇÃO ΔE_7 E ΔE_{PAR} , SEM A CORREÇÃO DE COUNTERPOISE E BSSE DADO PELA DIFERENÇA ENTRE OS VALORES CORRIGIDOS E NÃO CORRIGIDOS COM COUNTERPOISE. EMC SE REFERE À ENERGIA DE MUITOS CORPOS, (CORR) À CORRELAÇÃO ELETRÔNICA (ESTIMADA COMO MP2-HF) E CP À CORREÇÃO DE COUNTERPOISE. TODOS OS VALORES ESTÃO EM KCAL/MOL.	40
TABELA 8 COMPARAÇÃO ENTRE OS VALORES DAS MAIORES CARGAS ESP (MAX), DAS MENORES CARGAS (MIN) E A MÉDIA DOS VALORES ABSOLUTOS DAS CARGAS DE TODOS OS ÁTOMOS DOS COMPLEXOS (MEDABS).	44
TABELA 9 DESVIOS MÉDIOS ABSOLUTOS (DMEDA), DESVIOS MÁXIMOS (DMAX) E DESVIOS MÍNIMOS (DMIN) EM RELAÇÃO AOS VALORES OBTIDOS EM MP2/AUG-CC-PVTZ.	45
TABELA 10 DIPOLOS EM DEBYE CALCULADOS A PARTIR DAS CARGAS CHELPG PARA OS 8 COMPLEXOS UTILIZANDO DIFERENTES METODOLOGIAS. CPCM SE REFERE AO CÁLCULO NO SOLVENTE CONTÍNUO UTILIZANDO MP2/6-31+G(2D,P).	51
TABELA 11 VALOR DAS TRANSFERÊNCIAS DE CARGAS OBTIDAS NO SOLVENTE CONTÍNUO CPCM EM MP2/6-31+G(2D,P) E NO VÁCUO EM MP2/AUG-CC-PVTZ, HF/AUG-CC-PVTZ, MP2/6-31+G(2D,P) E HF/6-31G(D).	54
TABELA 12 VALORES DA ENTALPIA, ENTROPIA E ENERGIA LIVRE DA REAÇÃO 5 CALCULADOS UTILIZANDO O MÉTODO CBS-QB3 E COMPARADOS AOS VALORES EXPERIMENTAIS ⁸⁷ . OS VALORES ENTRE PARÊNTESES SE REFEREM AOS ERROS MÁXIMOS ESTIMADOS.	65
TABELA 13 DESVIOS MÉDIOS (DMED), DESVIOS MÉDIOS ABSOLUTOS (DMEDA) E DESVIOS MÁXIMOS (DMAX) EM RELAÇÃO ÀS ENERGIAS ELETRÔNICAS DAS REAÇÕES DE FORMAÇÃO DOS COMPLEXOS OBTIDOS EM CBS-QB3.	66
TABELA 14 DESVIOS MÉDIOS (DMEDA) E MÁXIMOS (DMAXA) ABSOLUTOS REFERENTES ÀS CONTRIBUIÇÕES TERMOQUÍMICAS OBTIDAS EM B3LYP/6-311G(2D,D,P) SEM FATOR DE ESCALA, COM EXCEÇÃO DOS DESVIOS SUBLINHADOS, QUE FORAM	

CALCULADOS EM RELAÇÃO AOS VALORES B3LYP/6-311G(2D,D,P) COM FATOR DE ESCALA DE 0,99 PARA O ZPE. TODOS OS VALORES ESTÃO EM KCAL/MOL.....	68
TABELA 15 ENERGIAS LIVRES DE REAÇÃO OBTIDAS PELO CBS-QB3 E PELA ESTRATÉGIA DE MENOR CUSTO. AS ENERGIAS LIVRES DESTE ÚLTIMO MÉTODO FORAM DECOMPOSTAS NAS CONTRIBUIÇÕES ELETRÔNICAS (ΔE_R), CORREÇÕES ENTÁLPICAS (ΔH_{CORR}), ENTALPIAS DE REAÇÃO (ΔH_R), CONTRIBUIÇÕES ENTRÓPICA ($-T\Delta S$). TODOS OS VALORES ESTÃO EM KCAL/MOL.	71
TABELA 16 ENERGIA DE HIDRATAÇÃO DO CÁTION METÁLICO OBTIDA CONSIDERANDO O CÁTION SEM A PRIMEIRA CAMADA DE SOLVATAÇÃO NO MEIO CONTÍNUO (MÉTODO I) E CONSIDERANDO EXPLICITAMENTE A PRIMEIRA CAMADA DE HIDRATAÇÃO (MÉTODO II).....	80
TABELA 17 ENERGIAS LIVRES E ENTALPIAS DE REAÇÃO EM FASE GASOSA E EM FASE AQUOSA, ADICIONANDO AS MÉDIAS DAS CORREÇÕES DO SOLVENTE ($\Delta \Delta E_{Med}$) OBTIDAS PELO MÉTODO CPCM COM RAIOS UAKS E NÍVEL PBE0/6-31G(D) E PELO CPCM, COM RAIOS BONDI À NÍVEL B3LYP/6-31G(D). TODOS OS VALORES ESTÃO EM KCAL/MOL.....	83
TABELA 18 ENTALPIAS E ENERGIAS LIVRES DA REAÇÃO DE FORMAÇÃO DOS COMPLEXOS NO SOLVENTE CONTÍNUO (A CONTRIBUIÇÃO DO SOLVENTE CONSIDERADA FOI A MÉDIA DOS VALORES OBTIDOS COM UAKS E BOND)I. OS VALORES ENTRE PARÊNTESES SE REFEREM ÀS ENERGIAS EM FASE GASOSA. TODOS OS VALORES ESTÃO EM KCAL/MOL.	86
TABELA 19 ESTUDO DA ADIÇÃO DE FUNÇÕES DIFUSAS NOS CONJUNTOS DE BASES DE POPLE A NÍVEL MP2, O CONJUNTO DE BASES UTILIZADO COMO REFERÊNCIA FOI O AUG-CC-PVTZ. OS VALORES ESTÃO EM KCAL/MOL.....	97
TABELA 20 ESTUDO DA ADIÇÃO DE FUNÇÕES DE POLARIZAÇÃO EM HIDROGÊNIOS NOS CONJUNTOS DE BASES DE POPLE DE VALÊNCIA TRIPLA À NÍVEL MP2. O CONJUNTO DE BASES UTILIZADO COMO REFERÊNCIA FOI O AUG-CC-PVTZ. OS VALORES ESTÃO EM KCAL/MOL.	98
TABELA 21 ESTUDO DA ADIÇÃO DE FUNÇÕES DE POLARIZAÇÃO EM ÁTOMOS PESADOS NOS CONJUNTOS DE BASES DE POPLE DE VALÊNCIA TRIPLA A NÍVEL MP2. O CONJUNTO DE BASES UTILIZADO COMO REFERÊNCIA FOI O AUG-CC-PVTZ. OS VALORES ESTÃO EM KCAL/MOL.	99
TABELA 22 ESTUDO DA ADIÇÃO DE FUNÇÕES DE POLARIZAÇÃO EM HIDROGÊNIOS NOS CONJUNTOS DE BASES DE POPLE DE VALÊNCIA DUPLA A NÍVEL MP2, O CONJUNTO DE BASES UTILIZADO COMO REFERÊNCIA FOI O AUG-CC-PVTZ. OS VALORES ESTÃO EM KCAL/MOL.	100
TABELA 23 ESTUDO DA ADIÇÃO DE FUNÇÕES DE POLARIZAÇÃO EM ÁTOMOS PESADOS NOS CONJUNTOS DE BASES DE POPLE DE VALÊNCIA DUPLA A NÍVEL MP2. O CONJUNTO DE BASES UTILIZADO COMO REFERÊNCIA FOI O AUG-CC-PVTZ. OS VALORES ESTÃO EM KCAL/MOL.	100
TABELA 24 ESTUDO DA ADIÇÃO DE FUNÇÕES DIFUSAS NOS CONJUNTOS DE BASES DE DUNNING A NÍVEL MP2. O CONJUNTO DE BASES UTILIZADO COMO REFERÊNCIA FOI O AUG-CC-PVTZ. OS VALORES ESTÃO EM KCAL/MOL.	101
TABELA 25 ENERGIAS DE INTERAÇÃO (ΔE E ΔE_{CP}) CALCULADAS MANTENDO AS GEOMETRIAS DOS COMPLEXOS OTIMIZADAS EM MP2/6-31+G(2D,P) FIXAS E VÁRIOS APENAS O MÉTODOS UTILIZADO PARA O CÁLCULO DAS ENERGIAS. TODOS OS VALORES ESTÃO EM KCAL/MOL.	102

LISTA DE FIGURAS

FIGURA 1 REPRESENTAÇÃO ESQUEMÁTICA DOS COMPLEXOS ESTUDADOS. OS RÓTULOS FORAM DADOS UTILIZANDO A ABREVIÇÃO DAS BASES NUCLEICAS (GUANINA GUA, CITOSINA CYT, ADENINA ADE, TIMINA THY, URACILA URA) SEGUIDO PELO SÍTIO AO QUAL O MAGNÉSIO SE COORDENA, OXIGÊNIO O OU NITROGÊNIO N E NÚMERO DE SUA POSIÇÃO NA ESTRUTURA DA BASE NUCLEICA. A LETRA A SE REFERE ÀS ÁGUAS DE HIDRATAÇÃO DO MAGNÉSIO.	12
FIGURA 2 TRÊS ESQUEMAS ILUSTRATIVOS DA INTERAÇÃO ENTRE DOIS MONÔMEROS DE ÁGUA; O PRIMEIRO MOSTRA COMO SERIA SE O CONJUNTO DE BASES FOSSE COMPLETO, O SEGUNDO MOSTRA COMO O BSSE E O BSIE SURGEM AO UTILIZAR UM CONJUNTO DE BASES INCOMPLETO E O TERCEIRO ESQUEMA MOSTRA COMO O PROCEDIMENTO COUNTERPOISE CORRIGE O BSSE.	14
FIGURA 3 PERFIL DAS ENERGIAS DE INTERAÇÃO (ΔE E ΔE_{CP}) CALCULADAS A NÍVEL MP2(FC) E HF UTILIZANDO O CONJUNTO DE BASES AUG-CC-PVTZ,. TODOS OS VALORES ESTÃO EM KCAL/MOL.....	17
FIGURA 4. ENERGIAS DE INTERAÇÃO ENTRE O MAGNÉSIO (II) PENTAHIDRATADO E AS BASES NUCLEICAS CALCULADAS UTILIZANDO A APROXIMAÇÃO FROZEN-CORE MP2(FC) VERSUS ENERGIAS DE INTERAÇÃO SEM A APROXIMAÇÃO MP2(FULL).....	19
FIGURA 5 ESQUEMA DE CONSTRUÇÃO DAS CONTRAÇÕES DE PRIMITIVAS DO CONJUNTO DE BASES 6-31G.....	22
FIGURA 6 COMPARAÇÃO ENTRE ΔE E ΔE_{CP} DE UMA SÉRIE DE CONJUNTOS DE BASES. OS VALORES SUPERIORES DE CADA BARRA SÃO REFERENTES AS ENERGIAS DE INTERAÇÃO CORRIGIDAS COM COUNTERPOISE (ΔE_{CP}) E OS VALORES INFERIORES SE REFEREM ÀS ENERGIAS NÃO CORRIGIDAS (ΔE), O COMPRIMENTO DAS BARRAS REPRESENTAM OS BSSE.	27
FIGURA 7: ESTIMATIVA DO FATOR MULTIPLICATIVO DE AUMENTO NO CUSTO COMPUTACIONAL PARA O CÁLCULO DA ENERGIA DE INTERAÇÃO EM MP2 DE UMA SÉRIE DE CONJUNTOS DE BASES, TENDO COMO PARÂMETRO DE COMPARAÇÃO O 6-31+G(D).	28
FIGURA 8 PRINCIPAIS DIFERENÇAS ENTRE AS GEOMETRIAS DOS COMPLEXOS OBTIDAS EM MP2(FC)/6-31+G(2D,P) EM RELAÇÃO ÀS GEOMETRIAS OBTIDAS EM HF/6-31+G(2D,P). VALORES EM ANGSTROMS (Å).	31
FIGURA 9 PERFIL DAS ENERGIAS DE INTERAÇÃO DOS 8 COMPLEXOS OBTIDOS DE ACORDO COM A NOTAÇÃO GEOMETRIA//ENERGIA. O CONJUNTO DE BASES UTILIZADOS NOS CÁLCULOS FOI O 6-31+G(2D,P). TODOS OS VALORES ESTÃO EM KCAL/MOL.	32
FIGURA 10 PERFIL DAS ENERGIAS DE INTERAÇÃO ΔE_7 E ΔE_{PAR} , PARA OS 8 COMPLEXOS. AS LEGENDAS COM (CP) SÃO DOS VALORES CORRIGIDOS COM COUNTERPOISE. OS CÁLCULOS FORAM FEITOS EM HF E MP2 UTILIZANDO AUG-CC-PVTZ. TODOS OS VALORES ESTÃO EM KCAL/MOL.	41
FIGURA 11 DIFERENÇA DAS CARGAS E DAS ISODENSIDADES (MP2-HF).....	48
FIGURA 12 CARGAS ATÔMICAS CHELPG DO MG E DO SÍTIO AO QUAL SE COORDENA, TANTO NO COMPLEXO COMO TAMBÉM A PARTE DO COMPLEXO. A PRIMEIRA COLUNA SE REFERE AO MAGNÉSIO (II) HEXAIDRATADO E À CARGA DO OXIGÊNIO DE UMA DE SUAS MOLÉCULAS DE H ₂ O.	50
FIGURA 13 MOMENTO DIPOLAR DOS 8 COMPLEXOS ESTUDADOS. CALCULADOS POR VÁRIOS MÉTODOS E ORDENADOS DECRESCENTEMENTE DE ACORDO COM OS VALORES DE MP2/6-31+G(2D,P) CPCM.	52
FIGURA 14 PERFIL DAS CARGAS TRANSFERIDAS DO [Mg(H ₂ O) ₅] ²⁺ PARA AS BASES NUCLEICAS NOS 8 COMPLEXOS ESTUDADOS. CPCM É REFERENTE ÀS TRANSFERÊNCIAS DE CARGAS NO SOLVENTE CONTÍNUO COM O MÉTODO MP2/6-31+G(2D,P). OS DEMAIS MÉTODOS FORAM CALCULADOS NO VÁCUO.....	54

FIGURA 15 REPRESENTAÇÃO ESQUEMÁTICA DOS COMPLEXOS ESTUDADOS. O PRIMEIRO NÚMERO INDICA QUANTAS MOLÉCULAS DE H ₂ O SE LIGAM AO Mg ²⁺ , SEGUIDO DA ABREVIÇÃO DAS BASES NUCLEICAS (GUANINA GUA, CITOSINA CYT, ADENINA ADE, TIMINA THY, URACILA URA) E DO SÍTIO AO QUAL O MAGNÉSIO SE COORDENA, OXIGÊNIO O OU NITROGÊNIO N E O NÚMERO DE SUA POSIÇÃO NA ESTRUTURA DAS BASES NUCLEICAS. A LETRA A SE REFERE ÀS ÁGUAS DE HIDRATAÇÃO DO MAGNÉSIO.	59
FIGURA 16 ENERGIAS ELETRÔNICAS DE INTERAÇÃO (ΔE) DE TODOS OS COMPLEXOS. A ABREVIÇÃO CP SE REFERE AOS VALORES CORRIGIDOS COM COUNTERPOISE.	60
FIGURA 17 MAPA DE POTENCIAL DAS BASES NUCLEICAS. OS VALORES EM VERMELHO SE REFEREM ÀS REGIÕES QUE O POTENCIAL ELETROSTÁTICO É NEGATIVO, ENQUANTO AS REGIÕES EM AZUL INDICAM POTENCIAL ELETROSTÁTICO POSITIVO.	61
FIGURA 18 PERFIL DAS ENERGIAS ELETRÔNICAS DE REAÇÃO, ENTALPIAS DE REAÇÃO E ENERGIAS LIVRES DE REAÇÃO EM FASE GASOSA A 298,15 K.	70
FIGURA 19 CICLO TERMODINÂMICO DAS REAÇÕES EM ANÁLISE BN = BASES NUCLEICAS E N=0,1,2.	76
FIGURA 20 ENERGIAS LIVRES DAS REAÇÕES DE FORMAÇÃO DOS 16 COMPLEXOS EM MEIO AQUOSO A 298,15K. AS CONTRIBUIÇÕES REFERENTES AO ESTADO GASOSO FORAM OBTIDAS PELOS MÉTODOS CBS-QB3 ($\Delta Gr(QB3)$) E PELA ESTRATÉGIA DE MENOR CUSTO COMPUTACIONAL (ΔGr) DEFINIDA NA SEÇÃO 16.1. O EFEITO DO SOLVENTE FOI CALCULADO UTILIZANDO O MODELO CPCM COM RAIOS UAKS E ENERGIA PBE0/6-31G(d) ($\Delta \Delta Esol (UAKS)$) E RAIOS BONDI COM B3LYP/6-31G(d) ($\Delta \Delta Esol (BONDI)$).	85

CAPÍTULO 1: Introdução ao tema

1 A descoberta do RNA catalítico

A maioria das reações químicas que ocorrem no interior das células dos seres vivos necessitam de catalisadores biológicos, chamados de enzimas, capazes de acelerar essas reações e garantir que elas ocorram de maneira altamente específica, viabilizando o devido funcionamento das células. ¹

Em 1926 James Batcheller Sumner cristalizou pela primeira vez uma enzima (urease) e descobriu que se tratava de uma proteína. Estudos subsequentes de diferentes enzimas constituídas de proteínas não só confirmaram a descoberta de Sumner, como também levaram a acreditar que a função de catálise biológica seria exclusiva das proteínas. ¹

O primeiro estudo a desafiar essa crença surgiu em 1978, quando Benjamin C. Stark, sob a orientação de Sidney Altman, descreveu que um fragmento de RNA seria essencial para que a enzima ribonuclease P da bactéria *Escherichia coli* (*E. coli*) apresentasse função enzimática. No entanto, Benjamin Stark não conseguiu definir se o RNA teria participação ativa na formação do sítio catalítico ou apenas estaria estabilizando um fragmento de proteína ligado a ele ².

Independentemente dos estudos de Altman, Thomas Cech, em 1982, ao estudar o RNA ribossômico do protozoário *Tetrahymena thermophila* descobriu que ele era capaz de catalisar a quebra e formação de ligações fosfodiéster na reação conhecida como *auto-splicing* mesmo na ausência de proteínas ³. Thomas Cech nomeou os RNA com propriedades enzimáticas de ribozima. Apesar de o trabalho ter sido bem aceito pela comunidade científica, restava saber se este era apenas um caso particular ou se existiriam outras ribozimas. ^{1,4}

No ano seguinte (1983), Guerrier-Takada, sob a orientação de Sidney Altman, e Norman Pace, demonstraram que a enzima ribonuclease P das bactérias *E. coli* e da *B. subtilis* também era uma ribozima. Além de descrever a primeira reação de catálise verdadeira (sem alteração do catalisador ao final da reação) envolvendo ribozimas, o trabalho de Guerrier-Takada *et al.* mostra que a atividade catalítica da ribonuclease P está diretamente relacionada às altas concentrações de MgCl₂ ^{4,5}.

Nos anos seguintes, centenas de ribozimas foram encontradas, demonstrando a grande significância do potencial catalítico de alguns RNA. Em

1989 Altman e Cech foram agraciados com o Prêmio Nobel devido à descoberta das ribozimas. ^{1,6}.

Atualmente, o RNA é a única molécula natural conhecida capaz de guardar informação genética e atuar como catalisador biológico, fortalecendo a hipótese de que o primeiro sistema capaz de autorreprodução seria formado por moléculas de RNA, hipótese chamada RNA World ^{1,7}. Com o passar do tempo, esse sistema foi evoluindo e, então, surgiram o DNA e as proteínas que compõem os sistemas complexos atuais, cujos DNA guardam informações genéticas necessárias para a síntese de proteínas, enquanto os RNA atuam como intermediários desse processo ⁸.

As propriedades catalíticas das moléculas de RNA representam uma potente ferramenta na produção de novas drogas baseadas em ribozimas sintéticas. Essas drogas atuam identificando e inibindo genes anormais ou de organismos patogênicos cuja sequência seja conhecida.^{9,10}

O desenvolvimento de drogas baseadas em ribozimas vêm se mostrando promissor para o tratamento de doenças virais como HIV, hepatite B e C, e também representa uma alternativa no tratamento de doenças como câncer e até mesmo o mal de Alzheimer ⁷. Atualmente já existem drogas de ribozimas em fase de teste em humanos para o tratamento de HIV ¹¹ e câncer renal ¹².

2 O cátion metálico e o mecanismo de ação das ribozimas

A notável necessidade de cátions metálicos bivalentes, em especial Mg^{2+} , para a atividade catalítica das primeiras ribozimas estudadas, serviram como indício de que estas ribozimas se tratavam de metaloenzimas, de modo que o Mg^{2+} atuaria não só estabilizando as cargas negativas do fosfato, mas também como cofator, ou seja, o cátion participaria ativamente no mecanismo catalítico das ribozimas.

Em 1993, Thomas Steitz propôs um mecanismo para as ribozimas similar ao das metaloproteínas DNA polimerase e fosfatase alcalina encontradas na bactéria *E. coli*, no qual dois cátions Mg^{2+} participam: um para ativar o grupo de ataque (nucleófilo), que seria uma molécula de água ou uma ribose, e o outro para estabilizar a carga negativa do grupo de saída; ambos agiriam como ácidos

de Lewis, estabilizando as cargas negativas do estado de transição ¹⁴. O papel do RNA seria apenas posicionar os cátions metálicos e orientar o substrato nos sítios específicos.

Embora haja evidências experimentais confirmando a participação ativa de dois cátions Mg^{2+} no mecanismo das ribozimas classificadas como introns do grupo I ¹⁵, esse mecanismo não pode ser generalizado. Afinal, já se sabe que algumas ribozimas pequenas (Hammerhead, Hairpin e VS) não necessitam de cátions metálicos bivalentes quando há altas concentrações de cátions monovalentes como Li^+ , Na^+ ou NH_4^+ para realizar a catálise de maneira eficiente¹⁶. Isso mostra que nem todas as ribozimas são necessariamente metaloenzimas e que tanto o cátion metálico como o RNA podem adotar papéis distintos em diferentes ribozimas ¹⁷.

Ainda hoje o papel dos cátions metálicos na estabilização e no mecanismo das ribozimas não foram totalmente elucidados ^{18,19}. Portanto está aberta a possibilidade para que modelos teóricos capazes de descrever questões energéticas e estruturais de difícil elucidação experimental, complementem os estudos atuais ²⁰.

2.1. As funções do Mg^{2+}

O Mg^{2+} é o cátion bivalente de maior concentração em meio fisiológico, possuindo funções essenciais na formação da estrutura tridimensional tanto do DNA como do RNA. Ao ocupar sítios específicos do DNA/RNA, o Mg^{2+} interage diretamente com átomos de oxigênio do fosfato e do açúcar ou átomos de oxigênio e nitrogênio das bases nucleicas, muitas vezes apresentando funções estruturais e catalíticas importantes ²¹⁻²³. Interações específicas indiretas formadas por ligações de hidrogênio entre as águas da primeira camada de hidratação do Mg^{2+} e as bases nucleicas ou fosfato também são importantes ²².

Ao interagir de maneira não específica ²⁴⁻²⁶, o Mg^{2+} participa da estabilização do DNA/RNA neutralizando as cargas negativas formadas ao longo do esqueleto de fosfato e diminuindo a repulsão eletrostática entre os segmentos de fosfato ou ocupando sítios em que haja alto potencial eletrostático negativo²².

Além de atuar na estabilização da estrutura terciária, o Mg^{2+} pode atuar como cofator em várias reações de catálise de ribozimas. Mesmo nos poucos casos em que ele pode ser substituído por outros cátions, isso raramente ocorre nas concentrações fisiológicas naturais das células ^{16,20,22}.

3 Estudos experimentais e computacionais

Nos estudos do dobramento de DNA e RNA ou do mecanismo de catálise das ribozimas, técnicas cristalográficas fornecem informações estruturais importantes, como, por exemplo, a posição dos cátions metálicos que ocupam sítios específicos no DNA/RNA. Porém, os dados cristalográficos não são capazes de informar exatamente como o Mg^{2+} se coordena ao DNA/RNA e, além disso, as condições de cristalização podem alterar os modos de ligação do cátion metálico ²⁷.

Cálculos teóricos de dinâmica molecular são úteis no monitoramento das mudanças conformacionais do RNA/DNA durante seu dobramento e formação da estrutura ativa. No entanto, esses métodos falham na descrição de sistemas que envolvem cátions metálicos bivalentes como Mg^{2+} devido ao tratamento simplificado dos campos de força, que em geral se baseiam na mecânica clássica. Estes métodos também desprezam ou tratam de forma simplificada as contribuições não eletrostáticas, tais como transferência de cargas e polarização, relevantes nesses sistemas ²⁸⁻³¹.

Em contrapartida, métodos quânticos de alto nível, apesar de terem seu uso restrito a pequenos complexos, são capazes de fornecer resultados satisfatórios para o estudo energético das interações entre o Mg^{2+} e pequenos fragmentos do DNA/RNA, como, por exemplo, as bases nucleicas e o fosfato ³¹⁻³⁶.

4 Justificativa e objetivos

Embora já existam vários trabalhos que utilizam métodos *ab initio* para descrever interações entre o magnésio(II) e pequenas partes do DNA/RNA ³⁷⁻⁴⁴, a construção de um conjunto bem organizado de resultados, obtidos por cálculos que utilizem um método computacional consistente, capaz de tratar de forma

acurada as principais contribuições energéticas dos sistemas, possibilitaria a extração de informações estatisticamente válidas.

Maria A. S. Prado, em sua tese de doutorado ³⁵, analisa os possíveis sítios de interação entre o Mg^{2+} e as bases nucleicas, utilizando métodos *ab initio*, semi-empírico, DFT e campos de força. Além de definir os sítios de cada base nucleica em que o Mg^{2+} interage mais fortemente, Maria A. S. Prado conclui que:

- Embora a correlação eletrônica seja relativamente pequena em sistemas Mg^{2+} bases nucleicas, esses efeitos tornam-se importantes ao adicionar as águas da primeira camada de hidratação do cátion e devem ser considerados no cálculo da energia de interação desses sistemas.
- O conjunto de bases 6-31+G(d) e os métodos HF ou B3LYP fornecem bons resultados em relação à otimização de geometria, porém cálculos de energia devem ser feitos utilizando-se um conjunto de bases maior tal como 6-311++g(2d,2p).
- O uso de bases de alto momento angular reduz o BSSE de modo que a correção de *counterpoise* se torna desnecessária.
- Interações do tipo base nucleica-(H_2O)_n e base nucleica-[$Mg(H_2O)_n$]²⁺ não são bem representadas por campos de força que utilizam a aproximação de pares.
- Cargas ChelpG refletem de maneira mais apropriada as variações no potencial eletrostático do Mg^{2+} do que as cargas de Mulliken.

O objetivo do presente trabalho é atualizar os resultados obtidos por Maria A. S. Prado, testando métodos computacionais que não existiam na época ou que eram inviáveis de serem feitas com as CPU de 32 bits. Também será avaliado o efeito do solvente e de contribuições termoquímicas, como a entropia e a entalpia, na energia de formação dos complexos [$Mg(H_2O)_n$ base nucleicas]²⁺ (n=4,5,6).

Ao final do trabalho busca-se a definição de uma estratégia baseada em cálculos *ab initio* que seja capaz de descrever as energias envolvidas na formação dos complexos $[\text{Mg}(\text{H}_2\text{O})_n \text{ base nucleicas}]^{2+}$ de maneira acurada, a um custo computacional que possibilite o emprego dessa estratégia em estudos futuros de sistemas maiores que apresente interações similares.

Espera-se também, que os dados apresentados neste trabalho sirvam como referência para parametrizar ou validar métodos de menor custo computacional, como semi-empíricos, DFT, campos de força clássicos, e que tais métodos possam ser aplicados em simulações de dinâmica molecular de ácidos nucléicos, dando resultados mais confiáveis que os métodos atuais.

CAPÍTULO 2: Estudo dos complexos

$[\text{Mg}(\text{H}_2\text{O})_5\text{basenucleica}]^{2+}$

Metodologia de cálculo.

5 Metodologia computacional

Métodos *ab initio* mecânico-quânticos são capazes de fornecer informações precisas sobre propriedades eletrônicas, estruturais e de estabilidade de sistemas moleculares, por considerar tanto interações de origem clássica (eletrostáticas) como também de origem quântica (orbital molecular). Sua maior vantagem em relação aos demais métodos é a ausência de parâmetros empíricos, fazendo uso apenas das equações da mecânica quântica e de algumas constantes experimentais ⁴⁵.

Apesar da aplicabilidade desses métodos estar restrita à pequenos sistemas moleculares devido ao alto custo computacional, os resultados obtidos pelos métodos *ab initio* podem ser utilizados na validação de funcionais de densidade e na parametrização de campos de força ou métodos semi-empíricos. Também podem ser combinados com mecânica molecular de forma que apenas uma pequena parte do sistema de interesse seja representado por métodos *ab initio* (estratégia conhecida como QM/MM), como ocorre nos estudos das interações entre fármacos e proteínas, nos quais apenas as interações envolvendo o sítio ativo, que necessitam de um tratamento mais acurado, são feitos com mecânica quântica, enquanto o restante do sistema é tratado por mecânica molecular.

Entre os métodos *ab initio*, baseados na função de onda, o mais simples é o Hartree-Fock (HF) que para resolver a equação de Schrödinger independente do tempo, admite várias aproximações. A de maior peso assume que cada elétron do sistema interage apenas com o campo médio dos demais elétrons, negligenciando a correlação eletrônica. Embora o método HF geralmente capture mais de 99% da energia total dos sistemas moleculares e apresente desvios em torno de 1% em distâncias de ligações, ele pode apresentar erros de 5% a 10% em propriedades como constantes de força e momentos de dipolo, bem como grandes erros na descrição de energias de atomização e energias de interação⁴⁶.

Para um tratamento mais rigoroso da correlação eletrônica, foram desenvolvidos os métodos conhecidos como pós-HF, que utilizam o HF como ponto de partida em seus cálculos. Entre esses métodos, o CCSD(T) é

considerado um referencial no cálculo de energias de interações intermoleculares, por ser capaz de reproduzir, com acurácia química, valores experimentais ⁴⁶, no entanto, o CCSD(T) possui elevado custo computacional.

Além do método, o cálculo da energia de um sistema molecular depende da descrição do seu orbital molecular, dada pela combinação linear de um conjunto de funções matemáticas chamada de conjunto de funções de basesⁱ. Geralmente, se utiliza uma combinação de funções de bases centradas nos átomos do sistema molecular, de formato aproximado ao dos orbitais monoelétrônicos do átomo de hidrogênio, para se capturar de maneira mais homogênea possível o orbital molecular. Pois, além do custo computacional depender diretamente do número de funções de bases, o orbital molecular se estende indefinidamente de modo que seria necessário um número infinito de funções de base para garantir sua descrição exata.

O custo computacional não só está diretamente relacionado ao tamanho do conjunto de bases, como também ao método teórico, de forma que enquanto o tempo computacional do método HF pode aumentar na proporção de até N^4 , sendo N o número de funções de bases empregado nos cálculos, o custo computacional do método MP2 aumenta aproximadamente na proporção de N^5 , o MP3 a uma proporção de N^6 , o MP4 e o CCSD(T) à N^7 . ⁴⁷ Desse modo, uma pequena redução no número de funções de base pode levar a uma considerável economia de tempo e recurso computacional, por isso a escolha do conjunto de bases deve ser guiada pela acurácia necessária ao estudo da propriedade de interesse do sistema ⁴⁸.

Na escolha do conjunto de bases, se deve considerar que o aumento do número de funções de bases, apesar de garantir a melhoria na descrição energética das moléculas, não garante uma descrição mais acurada de outras propriedades do sistema, como dipolos, cargas atômicas e energias de interação, de modo que um conjunto de bases grande pode dar resultados inferiores aos de um conjunto de bases menor. Vale ressaltar também que não há garantias de que um determinado conjunto de bases capaz de descrever bem certa propriedade do sistema dará bons resultados quando aplicado no cálculo de outras propriedades.

ⁱ Conjunto de funções de bases ou basis set, será abreviado para conjunto de bases.

Este capítulo tem como objetivo definir um método que apresente boa relação entre acurácia e custo computacional para descrever as interações eletrônicas entre o magnésio (II) pentahidratado e as bases nucleicas, adenina (Ade), guanina (Gua), citosina (Cyt), timina (Thy) e uracila (Ura), no vácuo, desconsiderando quaisquer contribuições vibracionais. Para isso, o capítulo foi dividido em duas seções; a primeira trata da escolha do nível teórico e de um conjunto de bases de alto custo para ser usado como referência, enquanto a segunda seção define um conjunto de bases que possa dar resultados semelhantes ao da primeira, a um menor custo computacional, com uma margem de erro aceitável.

6 Escolha do método

Por ser um cátion pequeno, o Mg^{2+} apresenta uma alta densidade de cargas positivas, logo, é esperado que as interações entre este cátion metálico bivalente e os átomos de nitrogênio e oxigênio das bases nucleicas sejam dominados por forças eletrostáticas de Coulomb, descritas suficientemente bem até mesmo pelo método HF. Entretanto, contribuições não eletrostáticas como dispersão, transferência de cargas e polarização também são significativas nessas interações ⁴⁹, principalmente ao incluir a primeira camada de hidratação do Mg^{2+} ³⁵, pois o cátion metálico polariza as moléculas de água coordenadas a ele, tornando-as capazes de fazer ligações de hidrogênio com as bases nucleicas mais fortes que as usuais ⁵⁰. Essas ligações de hidrogênio também envolvem contribuições não eletrostáticas diretamente relacionadas à correlação eletrônica ⁴⁵, como a dispersão, e não são tratadas devidamente pelo método Hartree-Fock (HF).

O método pós-HF de menor custo computacional, capaz de descrever satisfatoriamente bem interações dominadas pela componente eletrostática, mas que também apresente forças de dispersão considerável, é o método perturbacional Moller-Plesset de segunda ordem (MP2) ^{36,51}. Aliado a um conjunto de bases razoável, como o aug-cc-pVTZ ⁵², o MP2 é capaz de dar resultados mais acurados do que métodos de maior custo computacional, como o MP4 ou CCSD(T) combinados com funções de bases menores ^{46,50}

Os métodos baseados na teoria dos funcionais de densidade (DFT), também incluem o efeito da correlação eletrônica em seus cálculos e podem dar resultados tão bons, ou até melhores, do que aqueles obtidos pelo método MP2, a um menor custo computacional, porém esses métodos não descrevem bem forças de dispersão ⁴⁵.

6.1. Cálculo da energia de interação

Oito estruturas relacionadas como sendo os sistemas $[Mg(H_2O)_5\text{-base nucleica}]^{2+}$ mais estáveis (Figura 1), foram utilizadas como ponto de partida para o estudo das energias de interação entre o magnésio(II) pentahidratado e as bases nucleicas. Dentre essas oito estruturas, cinco delas foram retiradas da tese de doutorado de Maria A. S. Prado (Gua-O6, Cyt-N3, Ade-N1, Thy-O4 e Ura-O4) ³⁵, duas foram construídas com base nas estruturas descritas no trabalho de Ivan Sólt e colaboradores (CYT-O2, Ade-N7) ³⁴, e uma delas foi baseada em estudos de estruturas cristalinas que relatam a interação de cátions bivalente no sítio N7 da guanina, (Gua-N7) ²⁰. As valências livres referentes aos locais onde as bases nucleicas se ligam ao esqueleto do DNA/RNA, que ocorrem no átomo N1 das pirimidinas (citosina, uracila e timina) e no átomo N9 das purinas (guanina e adenina), foram preenchidas por hidrogênios.

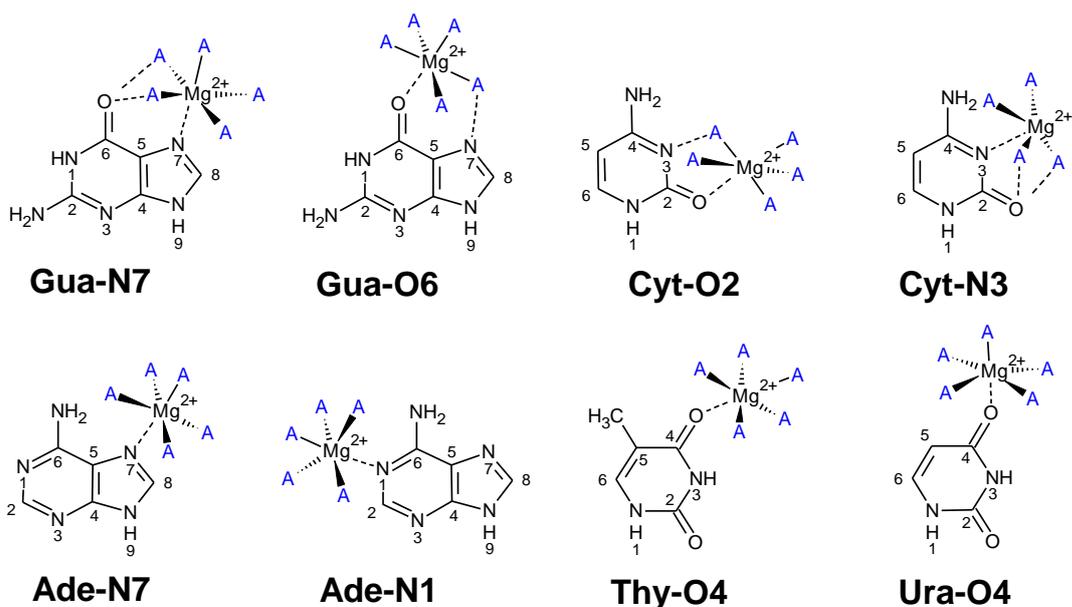


Figura 1 Representação esquemática dos complexos estudados. Os rótulos foram dados utilizando a abreviação das bases nucleicas (guanina Gua, citosina Cyt, Adenina Ade, Timina Thy, Uracila Ura) seguido pelo sítio ao qual o magnésio se coordena, oxigênio O ou nitrogênio N e número de sua posição na estrutura da base nucleica. A letra A se refere às águas de hidratação do Magnésio.

Todos os oito complexos tiveram as suas estruturas otimizadas no nível HF com o conjunto de base 6-31+G(d), em seguida essas estruturas foram reotimizadas, utilizando o método MP2(FC) e o conjunto de bases aug-cc-pVDZ. Os complexos foram então divididos em dois monômeros, mantendo os ângulos e as distâncias de ligação; o $[\text{Mg}(\text{H}_2\text{O})_5]^{2+}$ de cada complexo foi nomeado de monômero A e as bases nucleicas de monômero B, a energia de interação entre A e B (ΔE^{AB}) foi definida como a diferença entre a energia do complexo (E^{AB}) e a soma das energias de seus monômeros, E^A e E^B , calculados separadamente, conforme a equação 1.

$$\Delta E^{AB} = E^{AB} - (E^A + E^B) \quad [1]$$

6.1.1 Procedimento counterpoise para corrigir o BSSE

O cálculo da energia de interação (ΔE^{AB}) de acordo com a equação 1 só é consistente quando se aplica um conjunto de bases grande o suficiente para garantir que o orbital molecular dos monômeros seja descrito de maneira similar, tanto no complexo como separadamente, conforme o esquema 1 da Figura 2.

A inconsistência na equação 1, denominada erro de superposição de bases (BSSE), surge quando se utiliza um conjunto de bases incompleto. Pois, no complexo os átomos do monômero A mais próximos aos de B compartilham funções de bases⁵³, esse compartilhamento é benéfico e ajuda a suprir a defasagem do conjunto de bases incompleto dos monômeros, porém, quanto maior for a distância entre os monômeros menor é o compartilhamento, de modo que ao se dissociarem, cada monômero passa a ser descrito apenas por seu próprio conjunto de bases (esquema 2 da Figura 2). O menor número de funções de bases acessíveis aos monômeros quando estão separados faz com que suas energias E^A , E^B sejam superestimadasⁱⁱ devido à pior descrição matemática de seus orbitais moleculares em relação às que possuem no complexo. Como consequência disso, a energia ΔE^{AB} descrita na equação 1 é menor do que

ⁱⁱ Menos negativas.

deveria ser se o número de funções de bases acessíveis para cada monômero se mantivesse igual durante todo o cálculo, conforme o esquema 1 da Figura 2.

Uma forma simples de corrigir o BSSE no cálculo de ΔE^{AB} é substituir os conjuntos de bases dos monômeros (MCBS)ⁱⁱⁱ pelo conjunto de bases do complexo (DCBS)^{iv}, conforme o esquema 3 da Figura 2. De modo que para calcular a energia do monômero A, os átomos do monômero B passam a ser definidos como átomos fantasmas, posicionados na mesma geometria do complexo, porém sem elétrons ou núcleos. A energia de A passa a ser dada por $E^A(AB)$, na qual a notação entre parênteses (AB) se refere ao conjunto de bases do complexo (DCBS), esse procedimento é conhecido como *counterpoise* (cp)⁵⁴. A energia de interação entre o Magnésio (II) pentahidratado (A) e a base nucleica (B), corrigida com *counterpoise* (ΔE_{cp}^{AB}) passa a ser dada por:

$$\Delta E_{cp}^{AB} = E^{AB}(AB) - (E^A(AB) + E^B(AB)) \quad [2]$$

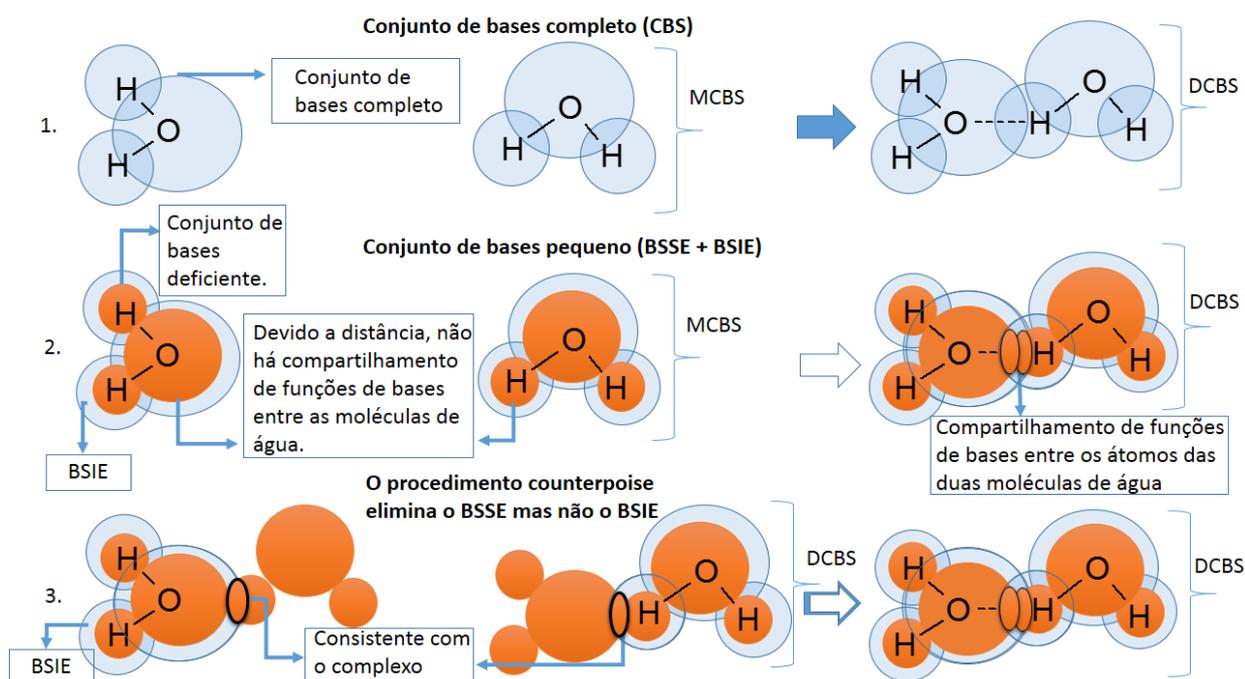


Figura 2 Três esquemas ilustrativos da interação entre dois monômeros de água; o primeiro mostra como seria se o conjunto de bases fosse completo, o segundo mostra como o BSSE e o BSIE surgem ao utilizar um conjunto de bases incompleto e o terceiro esquema mostra como o procedimento counterpoise corrige o BSSE.

ⁱⁱⁱ Quando o monômero tem acesso apenas ao seu conjunto de bases dá-se o nome de MCBS – Monomer Centered Basis Set.

^{iv} Quando o conjunto de bases de um dímero está acessível aos monômeros que o constituem dá se o nome de DCBS – Dimer Centered Basis Set.

Como é possível ver no esquema 3 da Figura 2, o procedimento *counterpoise* apesar de corrigir o BSSE, não corrige os erros devido ao uso de um conjunto de bases incompleto (BSIE)^v, presentes tanto no complexo como nos monômeros⁵⁵. Além disso, surge uma distorção na descrição do orbital molecular dos monômeros devido às diferenças na simetria do conjunto de bases do complexo (DCBS) em relação à simetria do conjunto de bases do monômero (DCBS). Apesar dessa distorção não ser significativa ao se empregar conjuntos de bases grandes, ela pode gerar erros consideráveis para conjuntos de bases muito pequenos⁵³.

Mesmo corrigindo o BSSE, o procedimento *counterpoise* faz com que a energia de interação ΔE^{AB} seja subestimada devido ao BSIE, enquanto os valores não corrigidos superestimam essas energias, sendo que à medida que se aumenta o conjunto de bases as energias de interação corrigidas e não corrigidas convergem a um mesmo valor.

Apesar dos valores corrigidos com *counterpoise* convergirem de maneira mais sistemática à medida que se aumenta o conjunto de bases, os valores não corrigidos podem apresentar menores desvios, principalmente ao se utilizar conjuntos de bases pequenos, devido ao cancelamento de erros entre o BSSE e o BSIE⁵⁰. Como o valor da energia de interação reside entre os valores corrigidos e não corrigidos, alguns grupos de pesquisa optam por utilizar a média entre esses valores⁵⁶. No presente trabalho optou-se por apresentar os valores corrigidos como margem superior dos resultados e os valores não corrigidos como margem inferior.

6.2. Análise dos métodos

6.2.1 Estudo das energias de interação

Para analisar as interações entre o $[\text{Mg}(\text{H}_2\text{O})_5]^{+2}$ e as bases nucleicas no vácuo, os valores de ΔE e ΔE_{cP} foram calculados conforme as equações 1 e 2, utilizando os métodos Hartree-Fock (HF) e MP2(FC). As geometrias dos

^v Basis Set Incompleteness Error

complexos, obtidas em MP2(FC)/aug-cc-pVDZ, foram mantidas fixas e o conjunto de bases utilizado para ambos os métodos foi o aug-cc-pVTZ. O BSSE foi estimado pela diferença entre ΔE e ΔE_{cp} .

Tabela 1 Energias de interação (ΔE e ΔE_{cp}) e estimativas dos BSSE obtidas pelas diferenças entre ΔE_{cp} e ΔE calculadas a nível MP2(FC) e HF utilizando o conjunto de bases aug-cc-pVTZ,. Todos os valores estão em kcal/mol.

	GUA-N7	GUA-O6	CYT-O2	CYT-N3	ADE-N7	ADE-N1	THY-O4	URA-O4
ΔE (HF)	-89,02	-91,91	-90,26	-79,09	-57,90	-56,04	-61,27	-61,62
ΔE_{cp} (HF)	-88,33	-91,29	-89,66	-78,37	-57,22	-55,39	-60,67	-61,09
BSSE (HF)	0,69	0,62	0,60	0,72	0,68	0,65	0,61	0,54
ΔE (MP2)	-95,74	-94,55	-92,19	-86,81	-67,87	-66,08	-62,45	-61,32
ΔE_{cp} (MP2)	-92,59	-91,90	-89,71	-83,50	-64,83	-63,22	-60,21	-59,49
BSSE (MP2)	3,15	2,65	2,48	3,30	3,04	2,86	2,24	1,83

Na Tabela 1, é possível notar que as interações entre os complexos de guanina (Gua-N7 e Gua-O6) e citosina (Cyt-O2 e Cyt-N3) foram em torno de 20 a 35 kcal/mol mais estáveis que as interações envolvendo os complexo de adenina (Ade-N7 e Ade-N1), timina (Thy-O4) e uracila (Ura-O4) em ambos os métodos.

A partir dos valores obtidos em MP2, não é possível notar uma relação direta entre o átomo ao qual o Mg^{2+} se coordena e a magnitude das interações, pois enquanto nos complexos de guanina o Mg^{2+} interagiu mais fortemente com o sítio de nitrogênio (N7), nos complexos de citosina o Mg^{2+} interagiu mais fortemente com o sítio de oxigênio (O2). Isso indica que outras contribuições, como distorções na geometria octaédrica do magnésio e formação de ligações de hidrogênio entre as águas coordenadas ao Mg^{2+} e as bases nucleicas, também são relevantes.

Em relação ao BSSE, o método HF apresentou valores inferiores a 1 kcal/mol, enquanto para o método MP2, o BSSE variou entre 1,8 a 3,3 kcal/mol. Em ambos os métodos o BSSE foi ligeiramente menor nos complexos em que o Mg^{2+} interage com um sítio de oxigênio da base nucleica.

Para uma comparação qualitativa entre os métodos HF e MP2, na Figura 3, os complexos foram ordenados decrescentemente em relação à magnitude de suas energias de interação, $\Delta E(MP2)$.

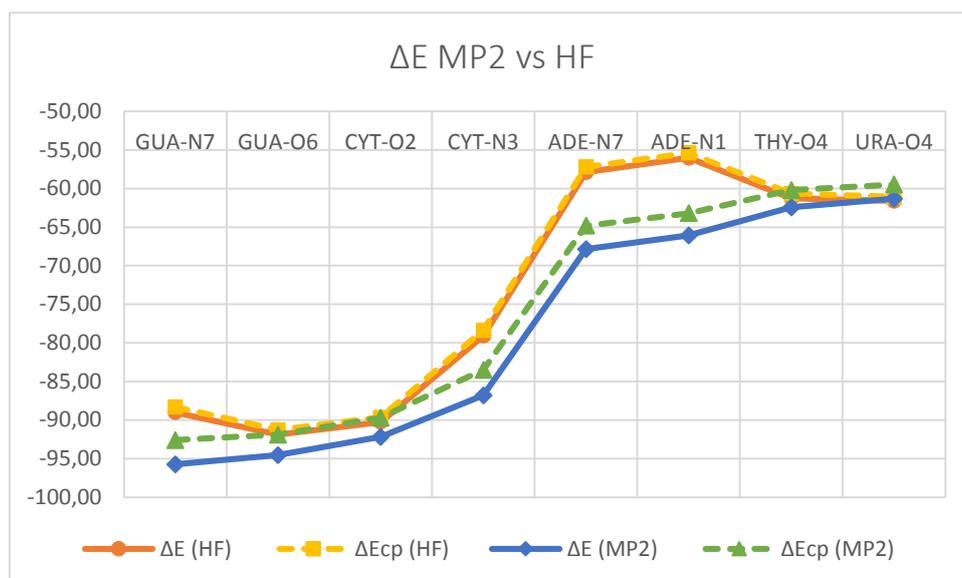


Figura 3 Perfil das energias de interação (ΔE e ΔE_{cp}) calculadas a nível MP2(FC) e HF utilizando o conjunto de bases aug-cc-pVTZ. Todos os valores estão em kcal/mol.

Observa-se, na Figura 3, que o método HF não descreveu a mesma ordem qualitativa que o MP2, pois subestimou as energias de interação do complexo Gua-N7 em relação a Gua-O6 e Cyt-O2, e dos complexos Ade-N7 e Ade-N1 em relação aos Thy-O4 e Ura-O4. As energias de interação corrigidas com *counterpoise* apresentaram o mesmo perfil dos valores não corrigidos em ambos os métodos.

6.2.2 Estudo da correlação eletrônica nos cálculos de ΔE e ΔE_{cp}

A princípio, a correlação eletrônica é a diferença da energia exata do sistema molecular e a energia HF calculada utilizando um conjunto de bases completo (CBS) ⁵⁷. Apesar de nem o conjunto de bases aug-cc-pVTZ ser completo, nem o método MP2 descrever a correlação eletrônica de maneira exata, geralmente, ao calcular diferenças de energia, há um certo cancelamento de erros entre o método de estrutura eletrônica MP2, que superestima a energia de correlação, e o conjunto de bases aug-cc-pVTZ, que por ser incompleto, faz com que apenas parte dessa energia seja capturada. De modo que a combinação MP2/aug-cc-pVTZ tenha boa concordância com alguns valores experimentais, como relatado para certas energias de atomização ^{46,50}. Por isso, utilizou-se as equações 3 e 4 e o conjunto de bases aug-cc-pVTZ para estimar

a magnitude da correlação eletrônica nos cálculos das energias de interação (ΔE e ΔE_{cp}):

$$\Delta E_{cor} = \Delta E(MP2) - \Delta E(HF) \quad [3]$$

$$\Delta E_{cor}(cp) = \Delta E_{cp}(MP2) - \Delta E_{cp}(HF) \quad [4]$$

Os valores de ΔE_{cor} e $\Delta E_{cor}(cp)$ foram comparados respectivamente com as energias $\Delta E(MP2)$ e $\Delta E_{cp}(MP2)$ para todos os complexos (Tabela 2).

Tabela 2 Energia de correlação ΔE_{cor} e seu valor percentual ($\% \Delta E_{cor}$) em relação a energia de interação (ΔE) sem correção para o BSSE e com a correção de counterpoise (cp) calculados em kcal/mol.

Complexos	Sem counterpoise			Com counterpoise		
	$\Delta E(MP2)$	ΔE_{cor}	$\% \Delta E_{cor}$	$\Delta E_{cp}(MP2)$	$\Delta E_{cor}(cp)$	$\% \Delta E_{cor}(cp)$
GUA-N7	-95,74	-6,73	7,03	-92,59	-4,27	4,61
GUA-O6	-94,55	-2,64	2,8	-91,90	-0,61	0,67
CYT-O2	-92,19	-1,93	2,09	-89,71	-0,05	0,05
CYT-N3	-86,81	-7,72	8,89	-83,50	-5,13	6,14
ADE-N7	-67,87	-9,97	14,69	-64,83	-7,60	11,73
ADE-N1	-66,08	-10,04	15,19	-63,22	-7,83	12,38
THY-O4	-62,45	-1,18	1,89	-60,21	0,46	-0,76
URA-O4	-61,32	0,30	-0,49	-59,49	1,60	-2,69

Na Tabela 2, é possível observar que a correlação eletrônica afetou de forma mais acentuada os valores de ΔE dos complexos nos quais o Mg^{2+} interage diretamente com o nitrogênio das bases nucleicas, Gua-N7, Cyt-N3, Ade-N7, Ade-N1. Nesses complexos $\% \Delta E_{cor}$ foi de 7% a 15% e $\% \Delta E_{cor}(cp)$ variou entre 5% a 12%. Em contrapartida, os valores de ΔE dos complexos com interações do tipo Mg-O (Gua-O6, Cyt-O2, Thy-O4 e Ura-O4) foram fracamente influenciados pela adição de correlação eletrônica, pois seus valores de $\% \Delta E_{cor}$ e $\% \Delta E_{cor}(cp)$ foram inferiores a 3%.

A comparação entre os resultados MP2 e HF confirmam a importância do efeito de correlação no estudo desses sistemas, principalmente para os complexos nos quais o magnésio(II) pentahidratado interage com o sítio de nitrogênio das bases nucleicas (Gua-N7, Cyt-N3, Ade-N7, Ade-N1). A maior correlação nas interações Mg-N ocorre devido a maior polarizabilidade do N em relação ao O, fazendo com que efeitos de dispersão sejam maiores nessas interações.

6.2.3 Aproximação Frozen Core

Como as interações interatômicas envolvem principalmente os elétrons da camada de valência ⁵⁸, uma aproximação muito comum é considerar que os orbitais atômicos de camadas internas permaneçam não perturbados, ou seja, tratados a nível Hartree-Fock. A essa aproximação dá-se o nome de frozen core (FC), logo, o método MP2(FC) inclui apenas perturbações nos elétrons de valência. No entanto, o cátion Mg^{2+} isolado, não possui elétrons de valência e sua energia calculada nos métodos MP2(FC) e HF são idênticas. Porém, os cálculos realizados neste trabalho define a carga 2+ para os complexos $[Mg(H_2O)_5\text{-base nucleica}]^{2+}$ e para o monômero $[Mg(H_2O)_5]^{2+}$ e não somente para o magnésio.

Para analisar a magnitude dessa aproximação, as energias de interação calculadas em MP2(FC) foram comparadas com as energias do método denominado MP2 completo (MP2(full)), que inclui a correlação eletrônica para todos os elétrons e conseqüentemente apresenta maior custo computacional.

Devido ao conjunto de bases aug-cc-pVTZ ter sido desenvolvido apenas para tratar efeitos de correlação eletrônica na camada de valência, utilizou-se também um conjunto de bases similar, o aug-cc-pCVTZ ⁵⁸, cuja única diferença é a adição de funções de polarização para elétrons de camada interna com a finalidade de se obter um tratamento mais rigoroso da correlação eletrônica envolvendo todos os elétrons ^{58,59} (Figura 4).

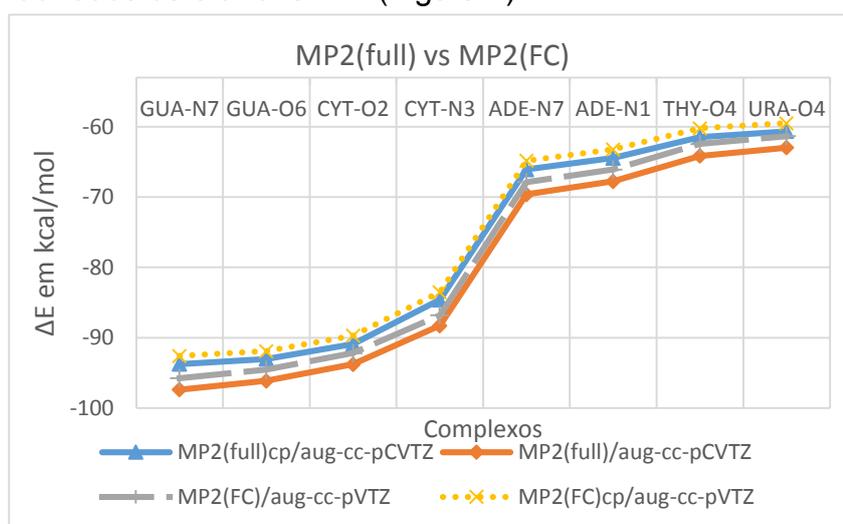


Figura 4. Energias de interação entre o magnésio (II) pentahidratado e as bases nucleicas calculadas utilizando a aproximação frozen-core MP2(FC) versus energias de interação sem a aproximação MP2(full).

Como todas as curvas da Figura 4 apresentaram o mesmo perfil, pode-se dizer que a inclusão de correlação eletrônica nos elétrons de camadas internas (MP2(full)) não alterou qualitativamente os resultados obtidos pela aproximação frozen-core MP2(FC).

A partir da Tabela 3, observa-se também que os valores corrigidos (cp) foram muito similares para os três métodos, pois seus valores diferiram menos de 3%. Em relação aos valores não corrigidos, é possível notar que o método MP2(full)/aug-cc-pVTZ subestimou, em aproximadamente 4 kcal/mol (6%), os valores de ΔE dados por MP2(full)/aug-cc-pCVTZ, ao passo que os valores não corrigidos com *counterpoise* (ΔE_{cp}) dados pelo método MP2(FC)/aug-cc-pVTZ, além de serem mais similares aos valores MP2(full)/aug-cc-pCVTZ, se situaram entre as energia de interação ΔE e ΔE_{cp} do método MP2(full)/aug-cc-pCVTZ, portanto, as energias ΔE obtidas através do método MP2(FC)/aug-cc-pVTZ, mesmo sem incluir a correção de *counterpoise*, aparentam ter uma ótima relação entre acurácia e custo computacional.

Tabela 3 Energias de interação (ΔE e ΔE_{cp}) obtidas em MP2(full)/aug-cc-pCVTZ, MP2(full)/aug-cc-pVTZ e MP2(FC)/aug-cc-pVTZ, o prefixo aug-cc- foi abreviado por a- e o método MP2 foi omitido, cp indica os valores corrigidos com *counterpoise*. Todos os valores estão em kcal/mol.

ΔE MP2	GUA-N7	GUA-O6	CYT-O2	CYT-N3	ADE-N7	ADE-N1	THY-O4	URA-O4
(FC)cp/a-pVTZ	-92,6	-91,9	-89,7	-83,5	-64,8	-63,2	-60,2	-59,5
(full)cp/a-pVTZ	-93,2	-92,5	-90,3	-84,1	-65,5	-63,9	-60,8	-60,1
(full)cp/a-pCVTZ	-93,8	-93,0	-90,9	-84,6	-66,1	-64,4	-61,5	-60,6
(FC)/a-pVTZ	-95,7	-94,6	-92,2	-86,8	-67,9	-66,1	-62,4	-61,3
(full)/a-pCVTZ	-97,4	-96,1	-93,8	-88,3	-69,6	-67,7	-64,1	-63,0
(full)/a-pVTZ	-101,9	-100,3	-97,8	-93,1	-74,3	-72,2	-68	-66,2

7 Escolha do conjunto de bases

Embora as funções de Slater descrevam os orbitais monoelétrônicos de forma precisa, suas integrais são de difícil resolução, tornando a aplicabilidade dos orbitais do tipo Slater (STO^{vi}) restrita apenas à sistemas monoatômicos e diatômicos. Para moléculas maiores, geralmente, utiliza-se um conjunto de

^{vi} STO - Slater Type Orbital

funções gaussianas (GTO^{vii}). Apesar de ser necessário combinar várias GTO para descrever um orbital atômico com formato aproximado ao de uma STO, as integrais de dois elétrons e vários centros calculadas para sistemas moleculares são feitas mais rapidamente por GTO. Isso ocorre devido à combinação de duas gaussianas formarem uma terceira gaussiana, o que simplifica as integrais.

Outra aproximação muito empregada na maioria dos conjunto de bases, para poupar recursos computacionais, é a combinação linear de gaussianas primitivas, a essa combinação dá-se o nome de contração de gaussianas^{viii} ou contração de orbitais do tipo gaussiana (CGTO^{ix}). Apesar dessa aproximação diminuir a flexibilidade das gaussianas primitivas, apenas as integrais das contrações (CGTO) são calculadas, outra vantagem é o menor espaço em disco necessário para o cálculo, pois apenas as contrações são armazenadas na memória do computador a cada ciclo SCF ⁶⁰.

A principal utilidade de CGTO está na descrição dos elétrons de camada interna. Pois eles possuem coeficientes menos difusos, logo, menor participação na formação de interações químicas, podendo ser tratados por apenas uma contração de várias gaussianas primitivas, geralmente 6 contrações são suficientes para uma boa descrição dos orbitais internos. Em contrapartida, orbitais de valência, por serem mais difusos e mais afetados pelo ambiente molecular, necessitam ser tratados por duas ou mais contrações.

A esses conjuntos de bases que utilizam apenas uma CGTO para orbitais de camada interna e várias CGTOS para descrever os orbitais de camada de valência, dá-se o nome de conjuntos de bases de valência dividida (Split-valence basis set), de modo que ao empregar duas CGTOs para descrever a valência dá-se o nome de duplo zeta (DZ), ao se utilizar três CGTOs dá-se o nome de triplo-zeta (TZ) e assim por diante.

Dentro dessa aproximação se destacam os conjuntos de bases 6-31G e 6-311G criados pelo grupo de Pople e os conjuntos de bases consistentes com métodos correlacionados desenvolvidos por Dunning e colaboradores.

^{vii} GTO - Gaussian Type Orbital

^{viii} Uma gaussiana primitiva é o mesmo que um orbital do tipo gaussiana (GTO)

^{ix} CGTO - Contracted Gaussian Type Orbital

7.1. Conjuntos de bases de Pople

-Valência dupla (DZ): esses conjuntos de bases são nomeados como n-ijG em que n é o número de gaussianas contraídas em uma única gaussiana, utilizadas para descrever os elétrons de camada interna e i,j são referentes à duas gaussianas utilizadas para descrever os elétron de valência de forma que os índices i e j dão o número de contrações de cada gaussiana.⁶¹

Conjunto de bases 6-31G⁶¹: Para cada átomo de Li-Ne na molécula, que possui os orbitais 1s de camada interna e os orbitais 2s, 2p de valência, o orbital s de camada interna é dado por uma CGTO formada por 6 gaussianas primitivas, e cada orbital do tipo s e p de valência é representado por duas CGTOs, a primeira formada por 3 primitivas e a segunda por apenas 1 primitiva, conforme ilustrado no esquema a baixo (Figura 5):

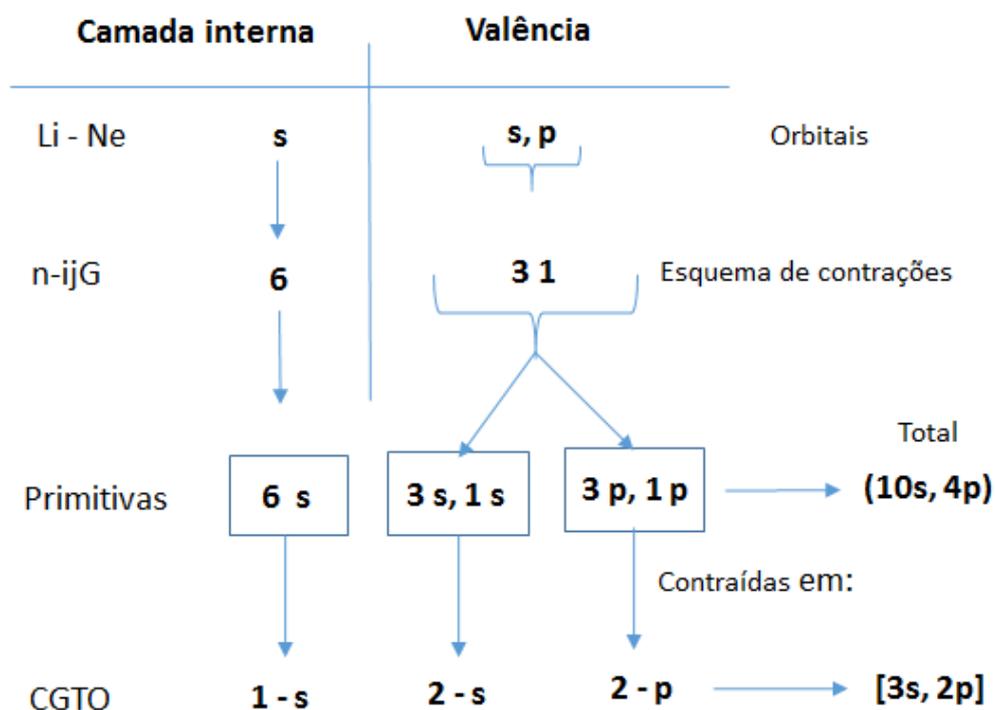


Figura 5 Esquema de construção das contrações de primitivas do conjunto de bases 6-31G

-Valência tripla (TZ), 6-311G⁶²: são semelhantes aos de valência dupla, porém tiveram seus coeficientes ajustados por métodos pós-HF (CISD^x) e contam com uma gaussiana de k contrações a mais para descrever os elétrons de valência, sua nomenclatura é n-ijkG^{xi}.

Outra diferença entre os conjuntos de valência dupla (DZ) e tripla (TZ) de Pople é que enquanto o primeiro utiliza 6 funções cartesianas para os orbitais do tipo d ($d_{x^2}, d_{y^2}, d_{z^2}, d_{xy}, d_{xz}, d_{yz}$), o segundo utiliza apenas 5 funções chamadas de funções de momento angular puro d ($d_{x^2-y^2}, d_{z^2}, d_{xy}, d_{xz}, d_{yz}$).

Para formar um conjunto de bases balanceado, não basta só aumentar as múltiplas contrações de gaussianas que descrevem a valência, pois formaria um grande número de funções s e p dando pouca flexibilidade ao sistema. Por exemplo, antes de aumentar a valência de dupla para tripla, é preferível adicionar funções de polarização, ou seja, funções de maior momento angular (d, f para Li – Ne e p, d para H e He) de forma que, para encontrar o conjunto de bases de Pople com melhor razão entre qualidade e custo computacional, é necessário testar várias combinações de funções de polarização.

Além de funções de polarização, no estudo de sistemas com interações fracas, como ligações de hidrogênio e interações do tipo π , geralmente é necessário adicionar gaussianas com expoentes pequenos, capazes de estender o alcance da função de onda. Essas gaussianas recebem o nome de funções difusas. A adição de funções difusas para os átomos pesados (todos com exceção de H e He) é simbolizada por + antes do G (n-ij+G), de modo que o conjunto de bases passa a acrescentar uma gaussiana difusa do tipo s e uma gaussiana difusa do tipo p tanto para átomos do tipo Li-Ne, como para Na-Ar. Também é possível adicionar funções difusas a todos os átomos, simbolizada por ++ antes do G (por exemplo 6-311++G), de forma que além das funções s e p difusas em átomos pesados, uma função s difusa é adicionada para cada hidrogênio.

^x CISD - Configuration Interactions with Single and Double excitations ou (HF +1 +2)

^{xi} Com exceção dos átomos Na-Ar que utilizam o esquema descrito por McLean.⁶²

7.2. Conjuntos de bases de Dunning

Os conjuntos de bases de valência dividida e consistentes com a correlação (cc), desenvolvido inicialmente por Dunning e colaboradores, conhecidos como cc-pVnZ^{xii} (cujas siglas pVnZ significam polarização de Valência e nZ = DZ, TZ e etc) foram desenvolvidos especificamente para serem empregados com métodos correlacionados ⁵², pois os expoentes das primitivas que formam as CGTOs de cada átomo foram obtidos utilizando métodos correlacionados (CISD), assim como os conjuntos de bases 6-311G de Pople.

Seu maior diferencial em relação aos demais conjuntos de bases é a convergência suave e monotônica^{xiii} da energia de correlação à medida que se aumenta o tamanho do conjunto de bases. Essa convergência gradual ocorre porque, diferentemente dos conjuntos de bases de Pople, nos de Dunning o aumento da multiplicidade das gaussianas de valência (DZ,TZ,QZ,5Z e etc) é acompanhado por um incremento sistemático de funções de polarização. A Tabela 4 mostra como as funções de polarização são adicionadas de acordo com o aumento do número de funções que descrevem os elétrons de valência para os átomos de Li-Ne.⁵²

Tabela 4 Esquema de contrações (CGTOs) de orbitais do tipo gaussiana e adições de funções de polarização para os átomos de Li-Ne dos conjuntos de bases de Dunning

	CGTO	F. de Polarização	F. adicionadas
Mínima	[2s,1p]		
cc-pVDZ	[3s,2p]	(1d)	[1s,1p,1d]
cc-pVTZ	[4s,3p]	(2d,1f)	[2s,2p,2d,1f]
cc-pVQZ	[5s,4p]	(3d,2f,1g)	[3s,3p,3d,2f,1g]

A vantagem dessa aproximação é o tratamento consistente da correlação eletrônica, de modo que as funções 1s,1p e 1d para os átomos Li-Ne adicionadas no conjunto de bases cc-pVDZ, captam uma quantidade da correlação eletrônica proporcional ao das funções 2s,2p,2d e 1f adicionadas no conjunto de bases cc-pVTZ, do mesmo modo as funções 3s,3p,3d,2f e 1g adicionadas ao conjunto cc-pVQZ captam uma porção comparável da correlação às funções adicionada ao

^{xii} cc-pVXZ deriva do inglês de forma que cc é a abreviação de correlated consistent; e pVXZ é X Zeta Valence Polarization cujo X pode ser double D, triple T, quadruple Q, quintuple 5, etc.

^{xiii} Funções monotônicas são funções invariavelmente crescentes ou decrescentes.

conjunto cc-pVTZ⁵². Esse tratamento permite que estimativas razoáveis dos valores esperados por um conjunto de bases completo (CBS) sejam feitas por meio de extrapolações, utilizando cálculos de apenas dois desses conjuntos de bases^{46,63}.

Assim como nos conjuntos de bases de Pople, os de Dunning também podem ser acrescidos de funções difusas, para tratar sistemas aniônicos ou para uma descrição mais acurada de interações que requerem um maior alcance da função de onda⁶⁴. Porém, nos conjuntos de Dunning ao invés de adicionar apenas uma função difusa, são adicionadas um conjunto de funções difusas simbolizado pelo prefixo (aug-)^{xiv}, que indica a adição de uma difusa por momento angular (s,p,f,g...) que o átomo possuir, por exemplo, para Li-Ne há o acréscimo de 3 funções difusas (1s, 1p e 1d). Existe também a possibilidade de inserir o dobro de funções difusas (d-aug-) e o triplo (t-aug-) para estudos mais acurados de interações de longo alcance, isto porém pode levar a sérias dificuldades de convergência na geometria de complexos.

Recentemente, Thrular e colaboradores, concluíram que em grande parte dos estudos computacionais a adição de funções difusas em hidrogênios e a adição de funções difusas de alto momento angular geram grandes aumentos no tempo computacional sem trazer melhorias significativas às propriedades estudadas⁶⁵. Eles desenvolveram então uma sequência de funções difusas parcialmente aumentadas para os conjuntos de bases de Dunning baseadas nas abreviações dos nomes dos meses em inglês, de modo que Jul- adiciona funções difusas da mesma forma que aug-, com exceção dos átomos de H e He que não são acrescidos de funções difusas. Jun- além de não utilizar funções difusas em hidrogênios, retira também a função difusa de maior momento angular, no caso de Li-Ne as difusas 1f. May- retira a maior função difusa subsequente (1d para Li-Ne) e assim sucessivamente⁶⁵.

7.3. Análise dos conjuntos de bases

Como o método de referência, MP2(FC)/aug-cc-pVTZ, escolhido na seção 5.2, possui um elevado custo computacional, sua aplicação seria proibitiva para

^{xiv} Aug - Augmented (aumentada)

complexos maiores, ou a uma extensa gama de sistemas. Com a finalidade de definir uma alternativa capaz de reproduzir os resultados desse método, a um menor custo computacional, com os menores desvios possíveis, optou-se por estudar o efeito da adição sistemática de funções difusas e de polarização nos conjuntos de bases de Pople e Dunning (Anexo 1). Os parâmetros levados em consideração por esse estudo foram, a magnitude do BSSE, a variação na energia de interação (ΔE) em relação aos valores obtidos com o conjunto de bases aug-cc-pVTZ e o número de funções de bases, conforme pode ser visto no Anexo 1.

A Tabela 5 apresenta os valores máximos (Max) e médios (Média) do BSSE e os desvios absolutos, máximos (DMaxA) e médios (DMedA), das energias de interação ΔE e ΔE_{cp} dos principais conjuntos de bases tratados no Anexo 1. O gráfico de barras (Figura 6), mostra uma comparação qualitativa entre esses conjuntos de bases. No gráfico da Figura 6, os valores superiores de cada conjunto de bases se referem a ΔE_{cp} , os valores inferiores à ΔE e o comprimento de cada barra, dado pela diferença entre esses dois valores representa a magnitude do BSSE.

Tabela 5 Comparação entre os BSSE máximos (max) e médios (Média) e os desvios médios e máximos absolutos (DMedA e DMaxA) de ΔE e ΔE_{cp} em relação aos valores MP2(FC)/aug-cc-pVTZ para uma série de conjuntos de bases ordenado decrescentemente com o número de funções de bases (F.B) do cálculo do maior complexo (Gua-N7).

Nível de cálculo MP2(FC)/ Conjunto de bases	ΔE (kcal/mol)		ΔE_{cp} (kcal/mol)		BSSE (kcal/mol)		F.B Gua-N7
	DMaxA	DMedA	DMaxA	DMedA	Max	Média	
Ref. aug-cc-pVTZ					3,30	2,70	1131
Jul-cc-pVTZ	0,51	0,26	0,12	0,08	2,96	2,52	996
Jun-cc-pVTZ	0,51	0,26	0,53	0,35	3,33	2,78	877
6-311+G (3d,p)	0,45	0,24	1,93	1,54	4,87	4,15	642
6-311+G (2d,p)	0,44	0,16	2,45	2,03	5,69	4,87	557
aug-cc-pVDZ	0,98	0,48	2,11	1,82	6,03	4,91	530
6-31+G (2d,p)	0,41	0,27	2,35	1,84	5,18	4,48	504
Jul-cc-pVDZ	0,82	0,37	2,38	2,03	5,18	4,39	470
6-31+ g (d)	2,53	1,01	4,93	3,69	8,93	7,13	357

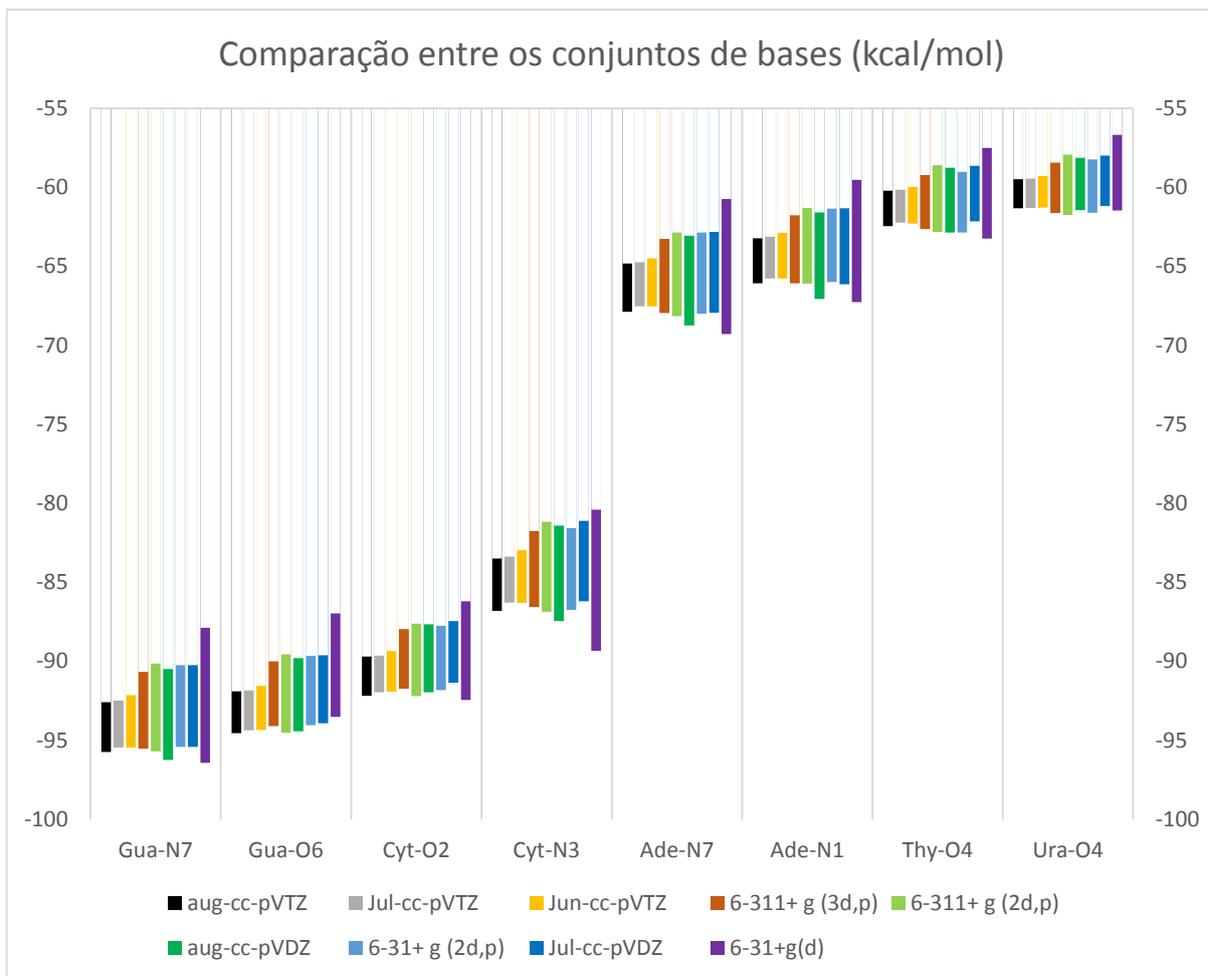


Figura 6 Comparação entre ΔE e ΔE_{cp} de uma série de conjuntos de bases. Os valores superiores de cada barra são referentes as energias de interação corrigidas com counterpoise (ΔE_{cp}) e os valores inferiores se referem às energias não corrigidas (ΔE), o comprimento das barras representam os BSSE.

O custo computacional relativo desses conjuntos de bases, a nível MP2, pode ser avaliado aplicando a equação 4. Essa equação fornece uma estimativa do fator multiplicativo de tempo de cálculo dos conjuntos de bases em relação a 6-31+G(d) para um determinado sistema.

$$Custo(MP2) = \left(\frac{N^{\circ} \text{ de funções de base do conj.de bases em análise}}{N^{\circ} \text{ de funções de base de } 6-31+g(d)} \right)^5 \quad [4]$$

Na Figura 7, a altura das barras corresponde ao fator multiplicativo de tempo de cálculo do complexo com o maior número de funções de bases (GUA-N7) em relação ao menor conjunto de bases 6-31+G(d).

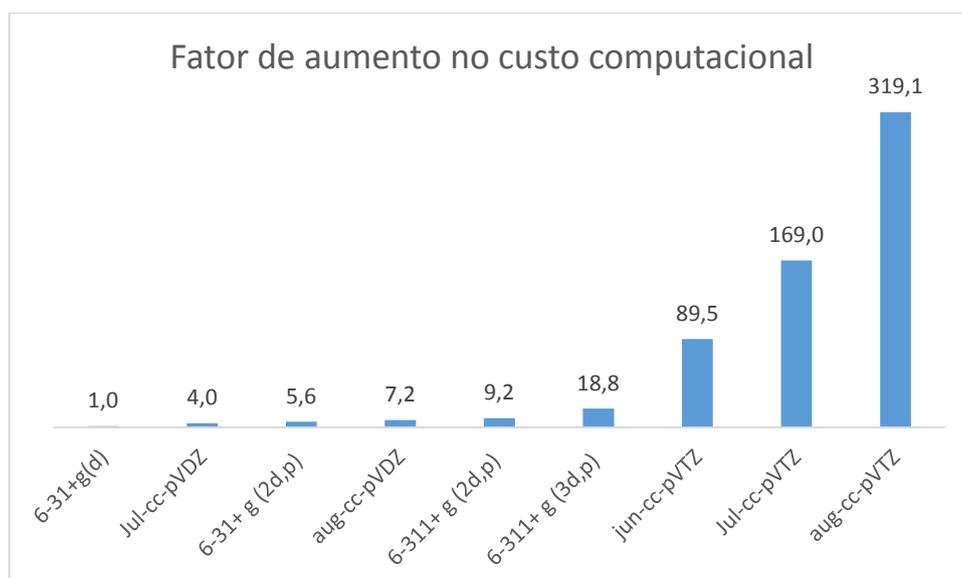


Figura 7: Estimativa do fator multiplicativo de aumento no custo computacional para o cálculo da energia de interação em MP2 de uma série de conjuntos de bases, tendo como parâmetro de comparação o 6-31+G(d).

Ao comparar os resultados apresentados na Tabela 5 e na Figura 7, observa-se que há uma melhora significativa ao passar do conjunto de bases 6-31+G(d) para o Jul-cc-pVDZ, pois o valor médio dos BSSE diminuiu 2,7 kcal/mol e os desvios médios absolutos de ΔE e ΔE_{cp} diminuiram 0,6 e 1,7 kcal/mol respectivamente, à um custo computacional aproximadamente quatro vezes maior.

O conjunto de bases subsequente, 6-31+G(2d,p) tem um fator multiplicativo ligeiramente maior (5,6x em relação a 6-31+G(d)), e apesar dos valor médio e máximo de seus BSSE, assim como os desvios de ΔE_{cp} , serem similares aos do conjunto de bases Jul-cc-pVDZ, o 6-31+G(2d,p) possui o menor desvio máximo de ΔE entre os conjuntos de bases analisados, 0,4 kcal/mol menor que o valor do Jul-cc-pVDZ.

Por sua vez, o conjunto de bases aug-cc-pVDZ, que foi utilizado para otimizar as geometrias, apresentou maiores valores de BSSE médio e máximo, 0,4 kcal/mol e 0,8 kcal/mol respectivamente, e maiores desvios médios e máximos absolutos de ΔE em relação a 6-31+G(2d,p), 0,6 e 0,2 kcal/mol respectivamente, a um maior custo computacional 1,6 vezes maior.

Os conjuntos de bases 6-311+G(2d,p) e 6-311+G(3d,p) não representam melhorias significativas em relação a 6-31+G(2d,p), pois nem seus BSSE médios e máximos, nem os desvios de ΔE e ΔE_{cp} chegam a 0,5 kcal/mol menores que os valores obtidos por 6-31+G(2d,p).

As energias de interação (ΔE e ΔE_{cp}) obtidas utilizando os conjuntos de bases Jun-cc-pVTZ e Jul-cc-pVTZ tiveram mínimos desvios em relação a referência (aug-cc-pVTZ) a um custo computacional significativamente menor, confirmando a proposição de Thrular de que tanto as funções de polarização em hidrogênio como as funções de polarização de maiores momentos angulares representam um grande aumento no custo computacional e não melhoram significativamente os resultados ⁶⁵.

Os conjunto de bases Jul-cc-pVDZ e 6-31+G(2d,p) representam uma boa opção para os cálculos de otimizações de geometria a nível MP2, principalmente quando estes cálculos não incluem as correções de *counterpoise*, devido aos seus pequenos desvios em relação aos valores de ΔE e os BSSE relativamente pequenos, à um custo computacional acessível. Já os conjuntos de bases Jul-cc-pVTZ e Jun-cc-pVTZ apresentaram uma performance similar ao do conjunto de bases aug-cc-pVTZ a um custo computacional significativamente menor, e podem ser utilizados para cálculos de energia (single point).

8 Otimização de geometria

O procedimento computacional chamado de otimização de geometria busca estruturas de mínima energia na hipersuperfície de energia potencial (SEP). Nesse procedimento as coordenadas dos átomos dos sistemas moleculares são ajustados de modo a minimizar o valor do gradiente de energia em função da posição dos átomos.

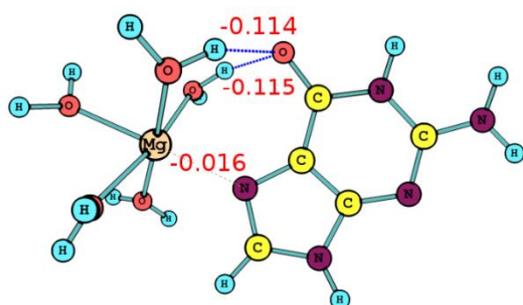
Como a otimização de geometria geralmente não é tão sensível ao método de cálculo e demanda um custo computacional maior que o cálculo da energia eletrônica do sistema, frequentemente se emprega um nível de teoria inferior para a otimização (nível X) em relação ao nível utilizado para o cálculo de energia (nível Y). Essa abordagem é abreviada como X//Y ^{45,48}.

Estudos de sistemas do tipo $M(H_2O)_n$ com $n=1,2,3,4$ e 6 e cátions metálicos, $M = Mg^{2+}$, Ca^{2+} , K^+ , concluíram que as estruturas otimizadas com o método correlacionado MP2 geram ligações M-O ligeiramente mais curtas quando comparadas às geometrias obtidas em HF. Entretanto, essas mudanças

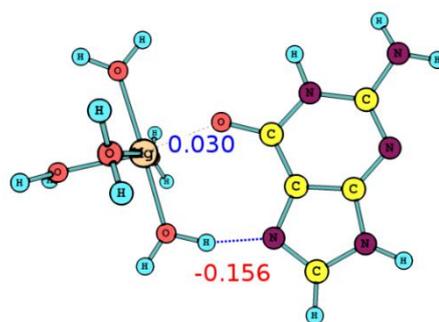
estruturais praticamente não alteraram a energia de interação entre o metal e as águas coordenadas a ele ⁶⁶⁻⁶⁸.

Nos complexos estudados, o magnésio interage não somente com os oxigênios das moléculas de água mas também com os sítios de nitrogênio e oxigênio das bases nucleicas, tais sítios são mais suscetíveis a efeitos de dispersão e transferência de carga, uma vez que a densidade eletrônica das bases nucleicas é mais facilmente polarizável que a das moléculas de água. Há também a formação de ligações de hidrogênio entre as bases nucleicas e as águas coordenadas ao magnésio, tais interações também envolvem forças de dispersão, que não são bem descritas a nível HF. Portanto, decidiu-se estudar as diferenças estruturais e energéticas ao adicionar o efeito da correlação eletrônica na otimização das geometrias.

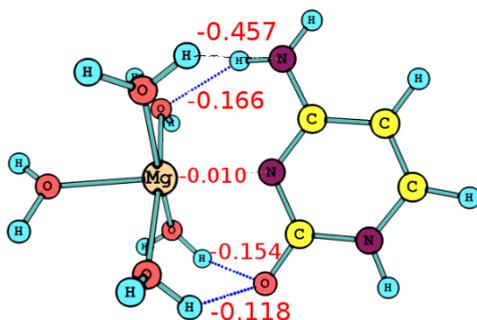
Partindo das estruturas otimizadas em HF/6-31+G(d), todos os complexos foram reotimizados em HF/6-31+G(2d,p) e MP2/6-31+G(2d,p), aplicando o algoritmo de Berny ⁶⁹, com a finalidade de analisar as diferenças entre as distâncias de ligação metal-base nucleica e os maiores desvios entre as estruturas otimizadas em MP2 em relação às estruturas obtidas em HF.



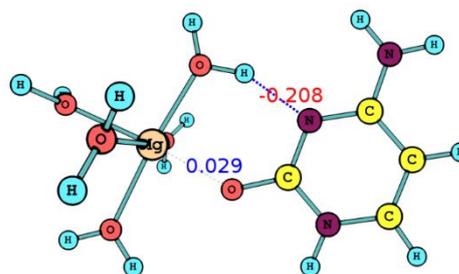
GUA-N7



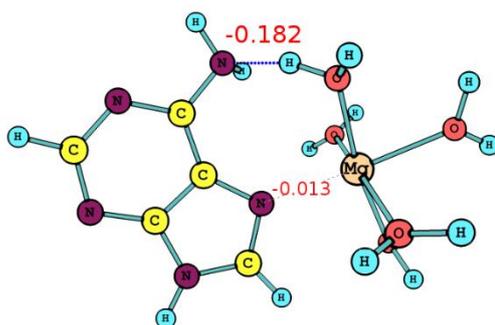
GUA-O6



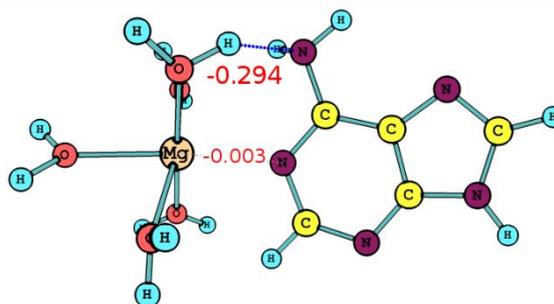
CYT-N3



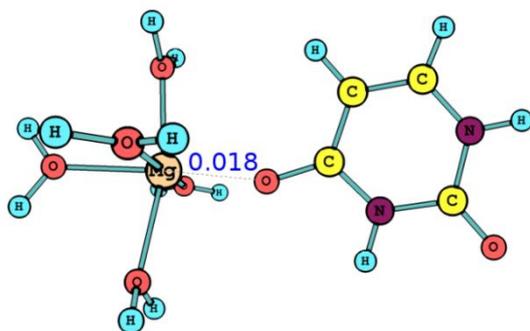
CYT-O2



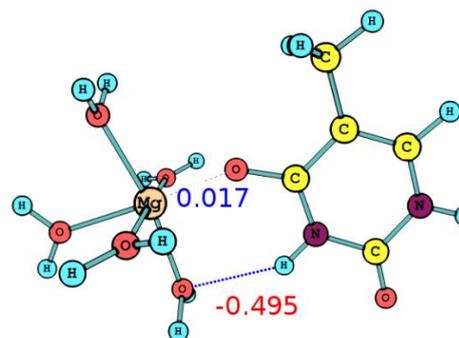
ADE-N7



ADE-N1



URA-O4



THY-O4

Figura 8 Principais diferenças entre as geometrias dos complexos obtidas em MP2(FC)/6-31+G(2d,p) em relação às geometrias obtidas em HF/6-31+G(2d,p). Valores em angstroms (Å).

Assim como mostrado na Figura 8, as maiores diferenças entre as distâncias de ligações obtidas nos nível MP2 e HF, com o conjunto de bases 6-31+G(2d,p), ocorreram entre as águas e os átomos de oxigênio e nitrogênio das bases nucleicas envolvidos em ligações de hidrogênio, que apresentaram encurtamentos de 0,11 Å (desvio de 6%) a 0,50 Å (desvio de 21%) nas geometrias MP2. Também foi possível observar um pequeno encurtamento das distâncias Mg-N e um alongamento entre Mg-O das bases nucleicas, porém essas variações foram inferiores a 2% das distâncias de ligação em MP2.

Em seguida, para analisar o efeito dessas mudanças nas energias de interação entre o magnésio (II) pentahidratado e as bases nucleicas, ΔE e ΔE_{cp} foram calculados a nível MP2/6-31+G(2d,p) para as geometrias obtidas em MP2 e HF (Tabela 6 e Figura 9).

Tabela 6 Energias de Interação ΔE e ΔE_{cp} dos complexos de acordo com a notação: geometria/energia. O conjunto de bases utilizados nos cálculos foi o 6-31+G(2d,p). Todos os valores estão em kcal/mol.

(geom./energ.) ΔE	Gua-N7	Gua-O6	Cyt-O2	Cyt-N3	Ade-N7	Ade-N1	Thy-O4	Ura-O4
MP2//MP2	-95,00	-93,65	-91,99	-86,57	-68,32	-65,98	-62,88	-61,63
HF//MP2	-92,40	-92,69	-90,78	-85,47	-66,09	-64,19	-61,39	-60,98
$\Delta(MP2-HF)//MP2$	-2,60	-0,96	-1,20	-1,11	-2,23	-1,79	-1,49	-0,65
% de MP2//MP2	2,7	1,0	1,3	1,3	3,3	2,7	2,4	1,1
MP2//MP2(cp)	-89,87	-89,23	-87,91	-81,46	-63,17	-61,34	-59,06	-58,26
HF//MP2(cp)	-87,57	-88,45	-86,88	-80,72	-61,32	-59,84	-57,80	-57,68
$\Delta(MP2-HF)//MP2(cp)$	-2,30	-0,78	-1,03	-0,75	-1,85	-1,50	-1,26	-0,58
% de MP2//MP2(cp)	2,6	0,9	1,2	0,9	2,9	2,5	2,1	1,0

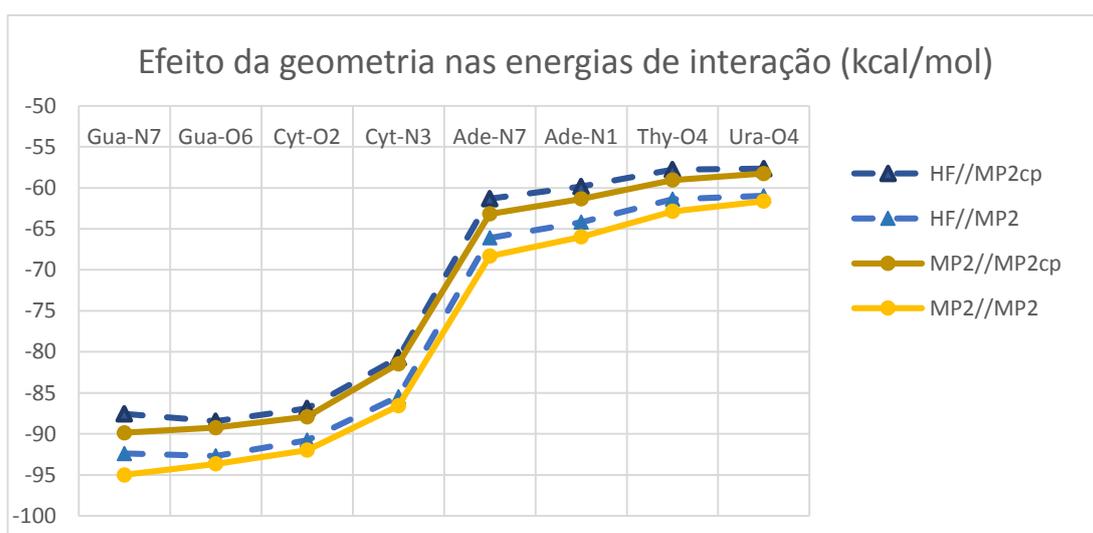


Figura 9 Perfil das energia de interação dos 8 complexos obtidos de acordo com a notação geometria/energia. O conjunto de bases utilizados nos cálculos foi o 6-31+G(2d,p). Todos os valores estão em kcal/mol.

As energias de interação calculadas com as geometrias HF foram maiores em relação às energias dadas utilizando as geometrias MP2, esse aumento foi de 1% a 3,3% para ΔE_{cp} e de 0,9% a 2,6% para ΔE . Apesar das diferenças entre as energias serem pequenas, há uma ligeira alteração qualitativa nos resultados, uma vez que o complexo GUA-N7 passou a ter uma energia de interação maior que a do complexo GUA-O6 nas geometrias HF (Figura 9).

9 Considerações finais

- A correlação eletrônica deve ser considerada nos cálculos da energia de interação (ΔE e ΔE_{cp}) entre o magnésio(II) pentahidratado e as bases nucleicas.

A aproximação *frozen core* MP2(FC) levou a pequenos desvios. Portanto, os cálculos dos capítulos 2 e 3 foram realizados utilizando essa aproximação.

- Como todos os cálculos deste trabalho foram feitos com o software Gaussian 09A optou-se por utilizar o conjunto de bases 6-31+G(2d,p), que está incluso no software, ao invés do Jul-cc-pVDZ.
- Apesar das geometrias obtidas em MP2 apresentarem diferenças significativas nas distâncias de ligações de hidrogênio obtidas em HF, essas mudanças pouco influenciaram nos cálculos das energias de interação.
- Alguns cálculos do capítulo 3 foram feitos em paralelo aos do capítulo 2. Nesses cálculos foram empregados as geometrias obtidas em MP2/aug-cc-pVDZ. Como a otimização de geometria é pouco sensível ao método e os conjuntos de bases aug-cc-pVDZ e 6-31+G(2d,p) empregam um número relativamente próximo de funções de bases, suas geometrias, em nível MP2, são similares.
- Optou-se por utilizar a geometria MP2(FC)/6-31+G(2d,p) no lugar da MP2(FC)/aug-cc-pVDZ, devido ao menor custo computacional, menores BSSE e por empregar um menor número de funções difusas, tais funções poderiam dificultar a convergência nas otimizações das geometrias de alguns complexos tratados no capítulo 4.

**CAPÍTULO 3: Limitações dos campos de força
clássicos na descrição dos complexos
[Mg(H₂O)₅ base nucleica]²⁺**

Energia de Muitos corpos e cargas ESP

10 Campos de força clássicos

Campos de força clássicos são empregados principalmente em simulações de sistemas que envolvem um grande número de átomos. Devido ao alto custo computacional, tais simulações são inviáveis de serem feitas utilizando métodos *ab initio*, DFT e até mesmo semi-empíricos. Dentre esses campos de força, o AMBER⁷⁰ foi desenvolvido para simulações de biomoléculas, como proteínas, pequenos RNA, e fragmentos de DNA/RNA. Os campos de força clássicos, assim como o AMBER, se baseiam em uma série de equações derivadas da mecânica clássica de Newton com termos parametrizados para reproduzir cálculos quânticos de alto nível e/ou dados experimentais.

A energia potencial do sistema molecular (E_{pot}) obtida por esses campos de força pode ser dividida entre termos que envolvem átomos ligados (E_{lig}) e termos que envolvem átomos não ligados ($E_{n\grave{a}o-lig}$):

$$E_{pot} = E_{lig} + E_{n\grave{a}o-lig} \quad [1]$$

10.1. Átomos ligados

Interações entre átomos considerados quimicamente ligados são tratadas basicamente por potenciais harmônicos. A energia dessas interações (E_{lig}) é calculada pelo somatório de três termos, conforme a equação 2:

$$E_{lig} = E_{est,r} + E_{ang,\theta} + E_{tor,\varphi} \quad [2]$$

Na qual,

- $E_{est,r}$ é a soma das energias potenciais de estiramento de todas as ligações químicas, formadas pelos N pares de átomos, calculada segundo a equação 3:

$$E_{est,r} = \sum_{n=1}^N K_{rn}(r_n - r_{eqn})^2 \quad [3]$$

K_{rn} é a constante de força da ligação química n;

r_n é o comprimento da ligação química n;

r_{eqn} é o comprimento de equilíbrio da ligação química n.

- $E_{ang,\theta}$ é a soma das energias potenciais de deformação dos N ângulos θ formados entre trios de átomos ligados, dada pela equação 4:

$$E_{ang,\theta} = \sum_{n=1}^N K_{\theta n}(\theta_n - \theta_{eqn})^2 \quad [4]$$

$K_{\theta n}$ é a constante de força de cada ângulo n;

θ_n é o valor de cada ângulo n;

θ_{eqn} é o valor de equilíbrio de cada ângulo n.

- $E_{tor,\varphi}$ é a energia potencial de torção, esse termo se refere a deformação dos ângulos φ de torção de todos os N diedros formados por quartetos de átomos ligados, dada pela equação 5

$$E_{tor,\varphi} = \sum_{n=1}^N \frac{Vm_n}{2} [1 + \cos(m_n\varphi_n - \gamma_n)] \quad [5]$$

Vm_n são as m_n constantes de torção de cada n diedro;

m_n se refere aos estados de mínima energia e só pode assumir valores inteiros;

φ_n são os ângulos diedros;

γ_n se refere aos ângulos de mudança de fase e pode assumir apenas os valores 0° e 180° .

10.2. Átomos não ligados

O termo não ligado da energia potencial ($E_{\text{não-lig}}$) é calculado entre os pares de átomos não ligados e geralmente emprega o potencial de Coulomb (E_{Col}) para descrever as forças eletrostáticas, e o potencial de Lennard-Jones^{xv} (E_{LJ}) para as forças de van der Waals e forças de dispersão de London, conforme a equação 6:

$$E_{\text{não-lig}} = E_{\text{Col}} + E_{\text{LJ}} \quad [6]$$

Na qual o potencial de Coulomb é dado por:

$$E_{\text{Col}} = \sum_{i < j}^{N'} \frac{q_i q_j}{\varepsilon R_{ij}} \quad [7]$$

E_{Col} é a soma das energias eletrostáticas entre os pares de átomos ij não ligados, calculada utilizando as cargas atômicas parciais q_i e q_j derivadas do potencial eletrostático de cálculos mecânico-quânticos de pequenos fragmentos. R_{ij} é a distância interatômica e ε é a constante dielétrica que leva em consideração o meio, quando este não está representado explicitamente ⁷¹.

O potencial de Lennard-Jones é descrito por:

$$E_{\text{LJ}} = \sum_{i < j}^{N'} \left(\frac{A_{ij}}{R_{ij}^{12}} - \frac{B_{ij}}{R_{ij}^6} \right) \quad [8]$$

E_{LJ} é a soma das energias entre os pares de átomos ij .

A_{ij} é o termo de van der Waals. B_{ij} é o termo de dispersão de London.

Interações intramoleculares não ligadas geralmente só são consideradas quando separadas por pelo menos três ligações ($j - i \geq 3$), sendo que

^{xv} Ou potencias similares aos de Lennard-Jones.

normalmente as interações entre terceiros vizinhos ($j - i = 3$) são multiplicadas por um fator de 0,2 a 0,5 ⁷¹.

Vários estudos indicam que até mesmo campos de força especificamente parametrizados para simular o DNA/RNA, como o AMBER, tratam interações envolvendo cátions bivalentes de maneira insatisfatória ^{30,72}. Isso ocorre principalmente porque nas interações entre os cátions bivalentes e o DNA/RNA, termos não eletrostáticos, como transferência de cargas, polarização e dispersão, são significativos ^{29,32,72}, e apesar do potencial de Lennard-Jones incluir uma fração desses termos, seus cálculos são feitos apenas como um somatório das interações entre pares de átomos, desconsiderando termos de maiores ordens. Dá-se, à esse somatório, o nome de aproximação de pares e a energia não contabilizada por esta aproximação se denomina energia de muitos corpos.

11 Aproximação de pares e a energia de muitos corpos

Quando um sistema molecular consiste de 3 ou mais monômeros (A,B,C,...,N), a energia de interação entre todos os monômeros pode ser escrita de duas formas:

- i) Como a diferença de energia entre o complexo e seus monômeros, equação 9:

$$\Delta E^{ABC\dots N} = E^{ABC\dots N} - (E^A + E^B + E^C + \dots + E^N) \quad [9]$$

- ii) Como a soma das interações entre todos os pares possíveis de serem formados mais o termo de muitos corpos (E^{MC}), conforme a equação 10:

$$\Delta E^{ABC\dots N} = \sum_{i < j}^n (\Delta E^{ij}) + E^{MC} \quad [10]$$

Na qual os monômeros (A,B,C,...,N) foram numerados como (1,2,3,...,n).

$$\Delta E^{ij} = E^{ij} - (E^i + E^j) \quad [11]$$

E ΔE^{ij} é a energia de interação dos pares (ou dímeros ij).

A energia de muitos corpos (E^{MC}) representa a fração da energia que não é descrita pelo somatório das interações de pares ⁴⁵ e é calculada pela diferença entre as equações 9 e 10:

$$E^{MC} = E^{ABC\dots N} - (E^A + E^B + E^C + \dots + E^N) - \sum_{i < j}^n (\Delta E^{ij}) \quad [12]$$

Na equação 10 existe a possibilidade de incluir termos de maior ordem, como a soma das interações entre todos os trios, quartetos e etc, porém sempre restará uma fração da energia de interação não representada por esses somatórios (E^{MC}) e embora, para se obter energias com acurácia quantitativa normalmente seja necessário incluir termos de maior ordem, a aproximação de pares geralmente é suficiente para uma descrição qualitativa da energia de sistemas compostos apenas por moléculas neutras ⁷³.

Devido a maior facilidade na parametrização e o menor custo computacional, a maioria dos campos de força clássicos utilizam a aproximação de pares e omitem o termo de muitos corpos nos cálculos das energias potenciais de átomos considerados como não ligados.

Com a finalidade de estudar a aplicabilidade da aproximação de pares e a magnitude do termo de muitos corpos, os complexos $[\text{Mg}(\text{H}_2\text{O})_5\text{-Base nucleica}]^{2+}$ foram divididos em 7 monômeros; 5 águas, Mg^{2+} e a base nucleica. Primeiramente se calculou a energia de interação entre os sete monômeros ($\Delta E^{ABC\dots N}$) de acordo com a equação 9. Para isso, as geometrias dos monômeros foram mantidas iguais às dos complexos, obtidas em nível MP2/aug-cc-pVDZ. As energias de cada complexo ($E^{ABC\dots N}$) e de seus monômeros ($E^A, E^B, E^C, \dots, E^N$) foram, então, calculadas em HF e MP2, ambas utilizando o conjunto de bases aug-cc-pVTZ. A essa energia foi dada o nome de $\Delta E7$.

As somas das energias das interações formadas pela combinação de todos os possíveis pares foram calculadas para se obter $\Delta E^{ABC\dots N}$, segundo a aproximação de pares (equações 10 e 11), desprezando o termo E^{MC} . As

geometrias dos pares foram mantidas iguais às que possuíam nos complexos, obtidas em MP2/aug-cc-pVDZ, e os cálculos das energias E^{ij} e E^i , E^j foram feitos utilizando os métodos HF e MP2 com o conjunto de bases aug-cc-pVTZ. A essa aproximação foi dada o nome de ΔE_{par} .

Os BSSE de $\Delta E7$ e ΔE_{par} foram estimados utilizando a correção de *counterpoise*. Para $\Delta E7$, essa correção foi feita calculando a energia de cada monômero utilizando o conjunto de bases do complexo e subtraindo estes valores pela energia eletrônica do complexo. Enquanto que para ΔE_{par} , a correção foi feita aos pares, utilizando o conjunto de bases dos dímeros ij para o cálculo das energias dos monômeros (E^i e E^j).

O efeito da correlação eletrônica foi estimado para $\Delta E7$ e ΔE_{par} como sendo a diferença entre as energias MP2/aug-cc-pVTZ e as energias HF/aug-cc-pVTZ. A energia de muitos corpos foi obtida pela diferença entre as energias $\Delta E7$ e ΔE_{par} em MP2/aug-cc-pVTZ e HF/aug-cc-pVTZ segundo a equação 13:

$$E^{MC} = \Delta E7 - \Delta E_{par} \quad [13]$$

Tabela 7 Energias de interação $\Delta E7$ e ΔE_{par} , sem a correção de *counterpoise* e BSSE dado pela diferença entre os valores corrigidos e não corrigidos com *counterpoise*. EMC se refere à energia de muitos corpos, (corr) à correlação eletrônica (estimada como MP2-HF) e cp à correção de *counterpoise*. Todos os valores estão em kcal/mol.

	GUA-O6	CYT-O2	GUA-N7	CYT-N3	ADE-N7	ADE-N1	THY-O4	URA-O4
$\Delta E7$ (MP2)	-371,43	-368,40	-366,55	-355,62	-343,30	-341,41	-337,84	-337,67
BSSE (MP2)	4,89	4,70	5,34	5,48	5,18	4,99	4,38	4,03
ΔE_{par} (MP2)	-506,43	-506,16	-500,27	-491,26	-477,45	-480,09	-475,99	-475,14
BSSE (MP2)	9,87	9,42	10,85	11,10	10,49	10,11	8,84	8,13
EMC(MP2)	135,00	137,76	133,72	135,64	134,15	138,68	138,15	137,47
%EMC(MP2)	-36,35	-37,39	-36,48	-38,14	-39,08	-40,62	-40,89	-40,71
EMC(MP2)cp	130,02	133,04	128,21	130,02	128,84	133,56	133,69	133,37
%EMC(MP2)cp	-35,47	-36,58	-35,49	-37,13	-38,10	-39,70	-40,09	-39,97
$\Delta E7$ (HF)	-371,75	-369,55	-362,62	-350,84	-336,55	-333,75	-339,74	-340,95
BSSE (HF)	0,94	0,92	0,97	0,99	0,95	0,92	0,90	0,86
ΔE_{par} (HF)	-498,13	-498,31	-488,85	-478,11	-464,04	-465,85	-469,70	-470,85
BSSE (HF)	0,99	0,96	1,03	1,04	1,01	0,98	0,95	0,91
EMC(HF)	126,38	128,76	126,23	127,27	127,49	132,10	129,96	129,90
%EMC(HF)	-34,00	-34,84	-34,81	-36,28	-37,88	-39,58	-38,25	-38,10
$\Delta E7$ (corr)	0,31	1,15	<u>-3,93</u>	<u>-4,78</u>	<u>-6,75</u>	<u>-7,66</u>	1,90	3,28
$\Delta E7$ (corr)cp	4,26	4,93	<u>0,44</u>	<u>-0,29</u>	<u>-2,52</u>	<u>-3,59</u>	5,38	6,45
ΔE_{par} (corr)	-8,30	-7,85	<u>-11,42</u>	<u>-13,14</u>	<u>-13,40</u>	<u>-14,24</u>	-6,29	-4,30
ΔE_{par} (corr)cp	0,57	0,61	<u>-1,60</u>	<u>-3,08</u>	<u>-3,92</u>	<u>-5,12</u>	1,61	2,92

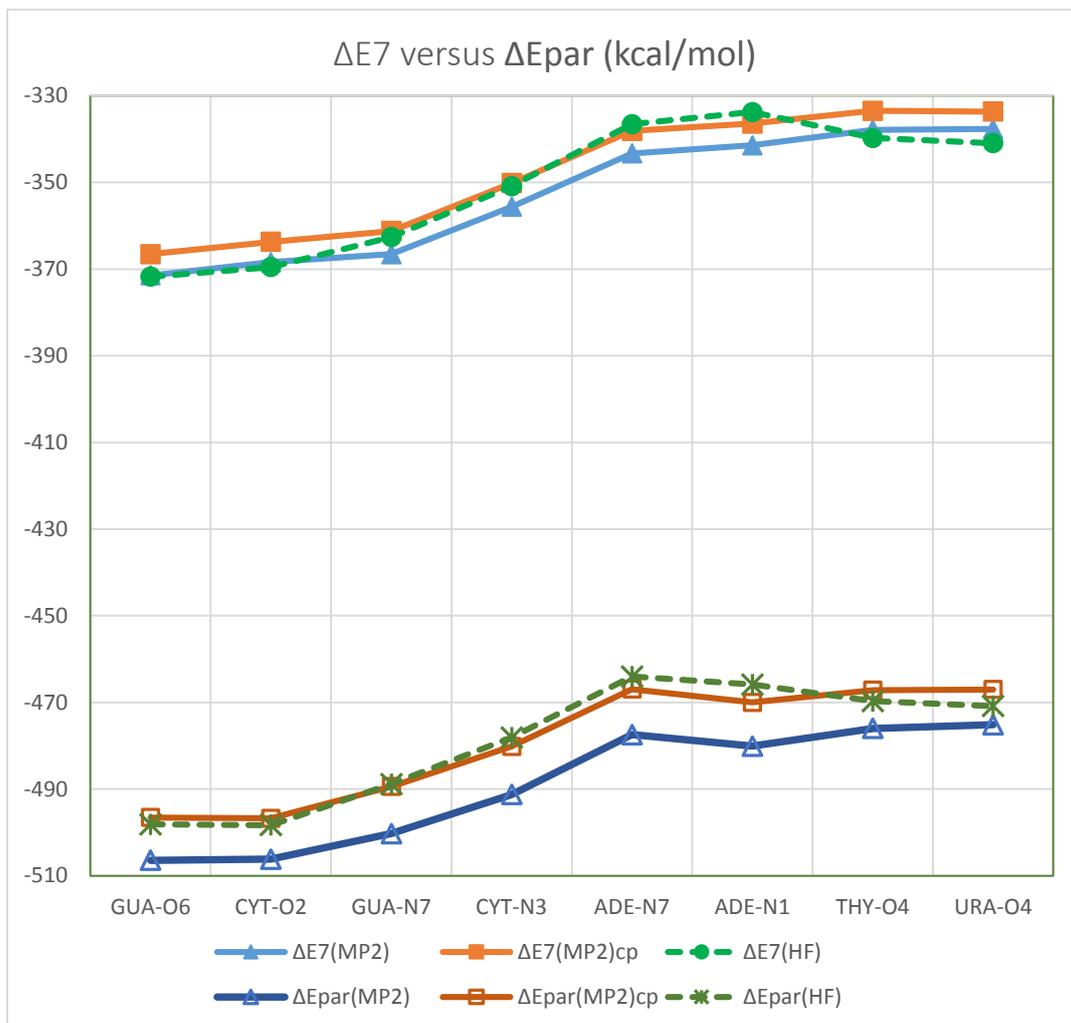


Figura 10 Perfil das energias de interação $\Delta E7$ e $\Delta Epar$, para os 8 complexos. As legendas com (cp) são dos valores corrigidos com counterpoise. Os cálculos foram feitos em HF e MP2 utilizando aug-cc-pVTZ. Todos os valores estão em kcal/mol.

11.1.1 $\Delta E7$ versus $\Delta Epar$

Apesar da correlação eletrônica ($\Delta E7_{corr}$ e $\Delta E7_{corr}(cp)$) representar valores inferiores a 2,5% de $\Delta E7(MP2)$ e $\Delta E7(MP2)_{cp}$ (Tabela 7), ela alterou o perfil de $\Delta E7(MP2)$ e $\Delta E7(MP2)_{cp}$ em relação à $\Delta E7(HF)$ e $\Delta E7(HF)_{cp}$, descrevendo menores energias de interação para os complexos ADE-N1 e ADE-N7 em relação aos complexos THY-O4 e URA-O4 (Figura 10).

As energias de muitos corpos, $EMC(MP2)$, $EMC(MP2)_{cp}$ e $EMC(HF)$, representaram mais de um terço de $\Delta E7(MP2)$ e $\Delta E7(MP2)_{cp}$ respectivamente, cerca de 35% a 40%, mostrando que a aproximação de pares é inapropriada para tratar as interações intermoleculares dos complexos estudados. Também é

possível notar que as energias de muitos corpos variaram menos de 6 kcal/mol entre os complexos (Tabela 7), o valor aproximadamente constante das energias de muitos corpos provavelmente se deve aos complexos serem pequenos e apresentarem estruturas e interações semelhantes.

Como consequência da pequena variação entre as energias de muitos corpos, os perfis das curvas $\Delta E_{\text{par}}(\text{MP2})$ e $\Delta E_{\text{par}}(\text{HF})$ apresentaram apenas pequenas diferenças em relação as curvas $\Delta E7(\text{MP2})$ e $\Delta E7(\text{HF})$ respectivamente (Figura 10). As principais diferenças foram que $\Delta E_{\text{par}}(\text{MP2})$ superestimou as energias de interação dos complexos ADE-N7 em relação a ADE-N1, THY-O4 e URA-O4 quando comparada a $\Delta E7(\text{MP2})$, ao passo que $\Delta E_{\text{par}}(\text{HF})$ superestimou a energia de interação dos complexos ADE-N1 e ADE-N7 em relação a THY-O4 e URA-O4 quando comparada a $\Delta E7(\text{HF})$ (Figura 10).

11.1.2 BSSE e a correlação eletrônica:

No nível HF, o erro de superposição de bases (BSSE) foi da ordem de 1 kcal/mol tanto para $\Delta E7(\text{HF})$ como para $\Delta E_{\text{par}}(\text{HF})$ (Tabela 7). Os pequenos BSSE indicam que o conjunto de bases aug-cc-pVTZ foi capaz de reproduzir energias de interação próximas ao limite do método HF. Os valores similares do BSSE para $\Delta E7(\text{HF})$ e $\Delta E_{\text{par}}(\text{HF})$ mostram que a contribuição do BSSE em HF/aug-cc-pVTZ, para o termo de muitos corpos, é pequena em HF.

A magnitude do BSSE para $\Delta E7(\text{MP2})$ foi de 4,0 a 5,5 kcal/mol, esses valores foram ainda maiores na aproximação de pares ($\Delta E_{\text{par}}(\text{MP2})$), variando de 8,1 a 11,1 kcal/mol, no entanto, a correção de *counterpoise* não interferiu na descrição qualitativa de $\Delta E7(\text{MP2})$, ou de $\Delta E_{\text{par}}(\text{MP2})$ (Figura 10).

A partir dos valores de $\Delta E7_{\text{corr}}$ e $\Delta E7_{\text{corr}}(\text{cp})$ na Tabela 7 é possível notar que a correlação eletrônica contribuiu para o aumento das energias de interação nos complexos em que o magnésio interage com um oxigênio da base nucleica (Mg-O) em relação aos complexos Mg-N. Esse comportamento foi qualitativamente reproduzido na aproximação de pares ($\Delta E_{\text{par}}_{\text{corr}}$ e $\Delta E_{\text{par}}_{\text{corr}}(\text{cp})$).

12 Cargas atômicas parciais

Cargas atômicas nada mais são do que uma forma de particionar a densidade eletrônica das moléculas entre seus átomos. Existem vários modos de se fazer essa partição ⁷⁴, seja utilizando diretamente a função de onda como as cargas de Mulliken, ou indiretamente, derivando as cargas através da análise de algum observável física calculada a partir da função de onda, como as cargas ESP.

Por reproduzirem bem propriedades como momento dipolares, multipolares e interações eletrostáticas, as cargas baseadas no potencial eletrostático (ESP) são as mais empregadas na parametrização de campos de forças clássicos, como o AMBER ^{45,74}.

As cargas ESP são ajustadas de modo que o potencial eletrostático $V_{ESP}(r)$ dado pela expansão de n monopolos de carga q centrados nos átomos (k) do sistema, reproduza o potencial eletrostático em torno da molécula (MEP) da melhor forma possível, conforme a equação 14:

$$V_{ESP}(r) = \sum_k^n \frac{q_k}{|r - r_k|} \quad [14]$$

Um dos algoritmos mais utilizados para o cálculo das cargas ESP é o ChelpG ⁷⁵ que constrói uma grade cúbica de pontos regularmente espaçados em torno da molécula, excluindo pontos dentro da superfície de van der Waals. Em seguida, o potencial eletrostático de cada ponto da grade é calculado e as cargas atômicas parciais centradas nos átomos são ajustadas de modo a reproduzir os valores do potencial eletrostático mantendo a carga total do sistema ⁷⁵. Assim como qualquer modelo que utilize monopolos centrados em átomos, as cargas ChelpG consideram os átomos como esfericamente simétricos e por isso não levam em conta possíveis distribuições anisotrópicas de cargas sobre os átomos ⁷⁵.

12.1. Influência do método e conjunto de bases nas cargas ESP

A descrição da densidade eletrônica de uma molécula depende tanto do nível de cálculo como também do conjunto de bases empregado. Logo, é de se esperar que as cargas atômicas derivadas do potencial eletrostático variem de acordo com o método adotado. Por isso foi feito um breve estudo sobre as influências do método e do conjunto de bases sobre as cargas ESP calculadas utilizando o algoritmo ChelpG.

Partindo das geometrias dos complexos otimizadas com MP2/6-31+G(2d,p), as cargas atômicas ESP foram calculadas em MP2/aug-cc-pVTZ e em HF/aug-cc-pVTZ, seus valores máximos (Max), mínimos (Min) e a média absoluta de suas cargas (Med Abs) podem ser vistos na Tabela 8.

Tabela 8 Comparação entre os valores das maiores cargas ESP (Max), das menores cargas (Min) e a média dos valores absolutos das cargas de todos os átomos dos complexos (MedAbs).

Complexos	MP2/aug-cc-pVTZ			HF/aug-cc-pVTZ		
	Max	Min	Med Abs	Max	Min	Med Abs
ADE-N1	1,56	-0,92	0,53	1,61	-0,94	0,56
ADE-N7	1,24	-0,87	0,49	1,30	-0,92	0,53
CYT-N3	1,16	-0,81	0,53	1,25	-0,85	0,57
CYT-O2	1,69	-0,98	0,62	1,75	-1,06	0,64
GUA-N7	1,15	-0,91	0,52	1,22	-0,97	0,57
GUA-O6	1,61	-0,92	0,56	1,69	-0,94	0,60
THY-O4	1,88	-1,00	0,56	1,88	-1,02	0,60
URA-O4	1,89	-0,99	0,64	1,96	-1,07	0,70

Na Tabela 8, é possível notar que em geral o método HF/aug-cc-pVTZ superestimou a magnitude tanto das cargas máximas como das mínimas, e suas cargas são em média 7% maiores do que as obtidas com o método MP2/aug-cc-pVTZ.

A fim de estimar a magnitude do efeito da inclusão de correlação eletrônica e a influência do conjunto de bases nas cargas parciais, foram calculados os desvios máximos (DMax), mínimos (DMin), médios absolutos (DMedA) e a percentagem de DMedA em relação as cargas ChelpG obtidas MP2/aug-cc-pVTZ para os métodos HF/aug-cc-pVTZ e MP2/6-31+G(2d,p), conforme a Tabela 9.

Tabela 9 Desvios médios absolutos (DMedA), desvios máximos (DMax) e desvios mínimos (DMin) em relação aos valores obtidos em MP2/aug-cc-pVTZ.

Complexos	HF/aug-cc-pVTZ				MP2/6-31+G(2d,p)			
	DMedA	%	DMax	DMin	DMedA	%	DMax	DMin
ADE-N1	0,03	6,6	0,12	-0,12	0,02	3,5	0,04	-0,06
ADE-N7	0,03	7,0	0,10	-0,10	0,02	3,8	0,05	-0,06
CYT-N3	0,06	12,1	0,22	-0,23	0,02	3,4	0,05	-0,04
CYT-O2	0,06	9,8	0,20	-0,24	0,02	3,6	0,07	-0,05
GUA-N7	0,05	9,0	0,12	-0,16	0,02	4,4	0,06	-0,06
GUA-O6	0,05	8,1	0,11	-0,15	0,02	4,1	0,06	-0,07
THY-O4	0,04	7,0	0,15	-0,16	0,02	3,2	0,04	-0,06
URA-O4	0,06	8,8	0,19	-0,21	0,02	3,1	0,04	-0,06

Ao analisar as médias dos desvios absolutos (Tabela 9), observa-se que os maiores desvios do método HF em relação ao MP2 ocorrem para os complexos formados pela citosina (CYT-N3 e CYT-O2). Também é possível notar que apesar do método HF ter superestimado as cargas, os desvios máximos e mínimos apresentaram magnitudes próximas.

Em relação à influência do conjunto de bases em nível MP2, observa-se (Tabela 9) que tanto os desvios máximos e mínimos como as médias dos desvios absolutos, foram pequenos e apresentaram valores similares para todos os complexos, ou seja, no nível MP2, o conjunto de bases 6-31+G(2d,p) é equivalente ao aug-cc-pVTZ para a descrição das cargas atômicas ESP.

12.2. Principais diferenças entre as cargas MP2 e HF

A diferença entre as cargas ESP MP2/aug-cc-pVTZ e HF/aug-cc-pVTZ de cada átomo (com exceção de hidrogênios) pode ser vista na primeira coluna da Figura 11. Os átomos de todos os complexos foram coloridos de acordo com a magnitude da diferença entre suas cargas em MP2 e HF, utilizando uma escala que varia de +0,236 em verde claro a -0,236 (vermelho claro).

Essas cargas foram, então, comparadas com as superfícies de isodensidade calculadas pelas diferenças entre as densidades MP2 e HF, conforme as imagens da segunda coluna da Figura 11. As partes em vermelho mostram as regiões onde a diferença entre as densidades eletrônicas foram positivas, essas regiões apresentaram maiores densidades eletrônicas em MP2

quando comparadas às densidades HF, ao passo que as partes azuis indicam as regiões cujas diferenças entre as densidades eletrônicas foram negativas, ou seja, a densidade eletrônica foi menor em MP2.

Deve-se ressaltar que diferentemente das cargas CHELPG, as densidades eletrônicas são observáveis mecânico-quânticos, e apesar das cargas CHELPG estarem diretamente relacionadas com essas densidades, por se tratarem de cargas pontuais ajustadas, suas diferenças (MP2-HF) não necessariamente são capazes de reproduzir os efeitos das diferenças entre as superfícies de isodensidade eletrônica MP2 e HF.

Δ Cargas (MP2-HF)	Δ Isodensidades (MP2-HF)
<p> $O(0.047)$ $C(-0.065)$ $N(0.028)$ $N(0.040)$ $Mg(-0.068)$ $O(0.038)$ $C(-0.095)$ $C(0.121)$ $O(0.042)$ $N(0.103)$ $O(0.025)$ $C(-0.146)$ $C(-0.165)$ $N(0.106)$ $O(0.095)$ $N(0.066)$ </p> <p>GUA-N7</p>	
<p> $O(0.043)$ $O(0.027)$ $O(0.042)$ $Mg(-0.077)$ $O(0.042)$ $O(0.049)$ $N(0.059)$ $C(-0.046)$ $C(0.098)$ $C(-0.149)$ $N(0.004)$ $C(-0.077)$ $O(0.072)$ $N(0.108)$ $N(0.102)$ $C(-0.154)$ $N(0.070)$ </p> <p>GUA-O6</p>	
<p> $N(0.099)$ $O(0.044)$ $O(0.032)$ $O(0.035)$ $Mg(-0.056)$ $O(0.035)$ $O(0.034)$ $N(0.197)$ $C(-0.236)$ $C(0.201)$ $C(-0.204)$ $O(0.083)$ $N(0.122)$ $C(-0.138)$ </p> <p>CYT-O2</p>	
<p> $O(0.035)$ $O(0.045)$ $O(0.049)$ $N(0.083)$ $Mg(-0.088)$ $O(0.035)$ $O(0.042)$ $C(-0.231)$ $N(0.182)$ $C(0.217)$ $C(-0.201)$ $C(-0.164)$ $O(0.095)$ $N(0.135)$ </p> <p>CYT-N3</p>	

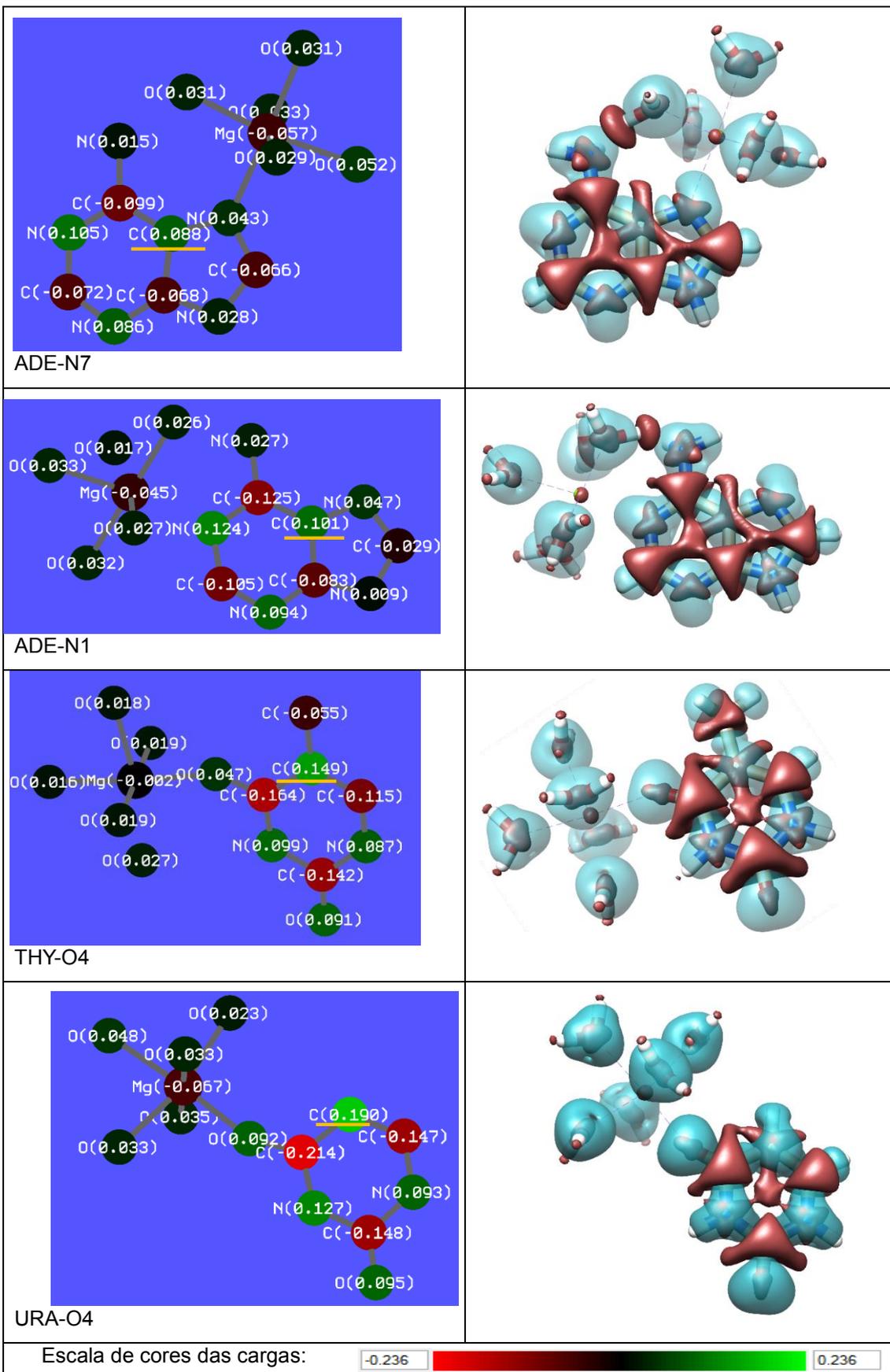


Figura 11 Diferença das Cargas e das isodensidades (MP2-HF).

Ao observar as isosuperfícies formadas pela diferença entre as densidades MP2 e HF (Figura 11), foi possível notar uma diminuição nas densidades eletrônicas MP2 em relação à HF (em azul) sobre os átomos de N, O e no carbono C5 (sublinhado em amarelo), enquanto nos demais átomos de carbono houve um aumento da densidade eletrônica em MP2 em relação à HF (superfície vermelha).

Essas mesmas variações também puderam ser notadas nas diferenças entre as cargas atômicas, de modo que os átomos cobertos pelas regiões cuja densidade eletrônica diminuiu apresentaram variações positivas em suas cargas (átomos verdes), enquanto átomos situados em regiões que apresentaram aumento na densidade eletrônica tiveram variações negativas em suas cargas (átomos vermelhos). Desta forma, as cargas CHELPG reproduziram satisfatoriamente os efeitos das variações de densidade eletrônica.

A partir dos valores da diferença entre as cargas MP2 e HF, é possível observar que as variações foram maiores nas cargas atômicas das bases nucleicas, principalmente às dos carbonos, e ocorrem porque o método HF não captura efeitos de dispersão apropriadamente. Por fim, é possível observar um aumento na densidade eletrônica próximo aos átomos de hidrogênio envolvidos em ligações de hidrogênio, tais interações também envolvem efeitos de dispersão eletrônica e não são bem descritos pelo método HF.

12.3. Relação entre as cargas do Mg^{2+} e do sítio de coordenação

Na Tabela 8, as maiores cargas atômicas (Max) foram as do magnésio. É possível notar que essas cargas são menores nos complexos em que o magnésio interage diretamente com o nitrogênio da base nucleica. Uma vez que o estudo das cargas ESP pode dar informações qualitativas sobre a distribuição da densidade eletrônica próxima a esses átomos, decidiu-se estudar a variação das cargas ESP do magnésio e do átomo das bases nucleicas ao qual ele se coordena.

Para isso, as cargas atômicas, tanto do magnésio como do sítio de nitrogênio ou oxigênio ao qual ele se coordena nas bases nucleicas, foram calculadas nos complexos e comparadas com as cargas obtidas mantendo a

geometria MP2/6-31+G(2d,p) fixa, porém separando o magnésio (II) pentahidratado das bases nucleicas. O mesmo estudo também foi feito com as cargas do magnésio e do oxigênio de uma das moléculas de H₂O do complexo [Mg(H₂O)₆]²⁺. Todas as cargas foram calculadas com o algoritmo ChelpG a nível MP2/aug-cc-pVTZ (Figura 12).

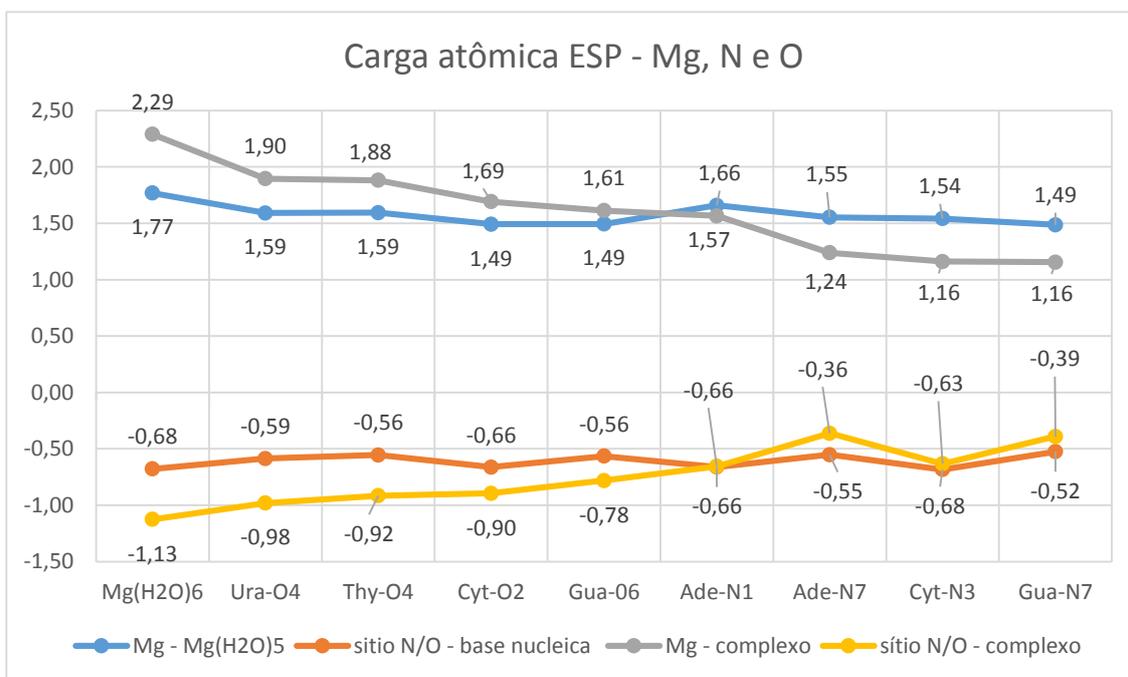


Figura 12 Cargas atômicas ChelpG do Mg e do sítio ao qual se coordena, tanto no complexo como também a parte do complexo. A primeira coluna se refere ao Magnésio (II) hexaidratado e à carga do oxigênio de uma de suas moléculas de H₂O.

Na Figura 12, quando o [Mg(H₂O)₅]²⁺ se coordena diretamente ao oxigênio, seja das bases nucleicas ou de uma sexta água, é possível notar que os átomos Mg e O se polarizam, pois as cargas dos oxigênios ficaram entre -0,22 a -0,45 mais negativas e as do magnésio aumentaram entre 0,12 a 0,52.

No entanto, quando o [Mg(H₂O)₅]²⁺ se coordena aos sítios de nitrogênio, o efeito mais notável é a transferência de cargas entre a base nucleica e o magnésio, pois as cargas do magnésio diminuiriam entre 0,09 e 0,33, ao passo que as cargas do N tiveram apenas um pequeno aumento, entre 0,01 a 0,19.

Todos os complexos [Mg(H₂O)₅ base nucleica]²⁺ aqui estudados têm suas energias de interação dominadas pela componente eletrostática (Coulomb), entretanto, a Figura 12 mostra que nos complexos em que o [Mg(H₂O)₅]²⁺ interage com o N das bases nucleicas a importância relativa dos termos de dispersão e transferência de cargas parece ser maior.

12.4. Dipolos

Uma das maiores dificuldades em simulações de sistemas biológicos, como o DNA, é a inclusão do efeito do solvente aplicando campos de força clássicos, devido à polarização entre o soluto e o solvente polar ser uma das principais contribuições que não podem ser divididas em um somatório de pares ⁴⁵. Para compensar parcialmente esse problema, alguns campos de força, como o AMBER, utilizam cargas derivadas do potencial eletrostático, calculadas na fase gasosa à nível HF com o conjunto de bases 6-31G(d) ⁷⁰, pois acredita-se que as cargas obtidas dessa maneira superestimam sistematicamente os dipolos das moléculas de modo a simular uma fração do efeito de polarização do solvente ⁷⁶.

Com a finalidade de estudar a validade dessa metodologia no cálculo dos dipolos dos complexos em análise, e estimar a diferença entre os dipolos obtidos pelos métodos MP2 e HF, e pelos conjuntos de bases aug-cc-pVTZ e 6-31+G(2d,p), calculou-se os dipolos de todos os complexos a partir das cargas ESP dadas pelo algoritmo ChelpG, mantendo as geometrias obtidas no vácuo em MP2/6-31+G(2d,p) fixas, com os métodos MP2/aug-cc-pVTZ, MP2/6-31+G(2d,p), HF/aug-cc-pVTZ e HF/6-31G(d). O efeito do solvente sob os dipolos foi estimado aplicando o método de contínuo CPCM aos cálculos MP2/6-31+G(2d,p). O parâmetro dos raios atômicos utilizados na construção das cavidades do solvente foi o Bondi ⁷⁷ multiplicado por um fator de 1,2. A constante dielétrica utilizada foi a da água ($\epsilon=78,35$).

Tabela 10 Dipolos em Debye calculados a partir das cargas ChelpG para os 8 complexos utilizando diferentes metodologias. CPCM se refere ao cálculo no solvente contínuo utilizando MP2/6-31+G(2d,p).

Métodos	ADE-N7	URA-O4	ADE-N1	THY-O4	GUA-O6	GUA-N7	CYT-N3	CYT-O2
MP2/aug-cc-pVTZ	13,50	12,07	11,32	11,68	7,92	7,52	5,75	3,76
MP2/6-31+G(2d,p)	13,60	12,12	11,38	11,71	7,97	7,57	5,76	3,72
HF/aug-cc-pVTZ	14,04	12,36	11,77	11,63	7,68	7,76	5,56	3,59
HF/6-31G(d)	14,98	12,55	12,77	12,22	8,91	8,73	6,26	4,34
CPCM	17,47	14,63	14,26	13,65	9,81	9,42	7,21	4,47

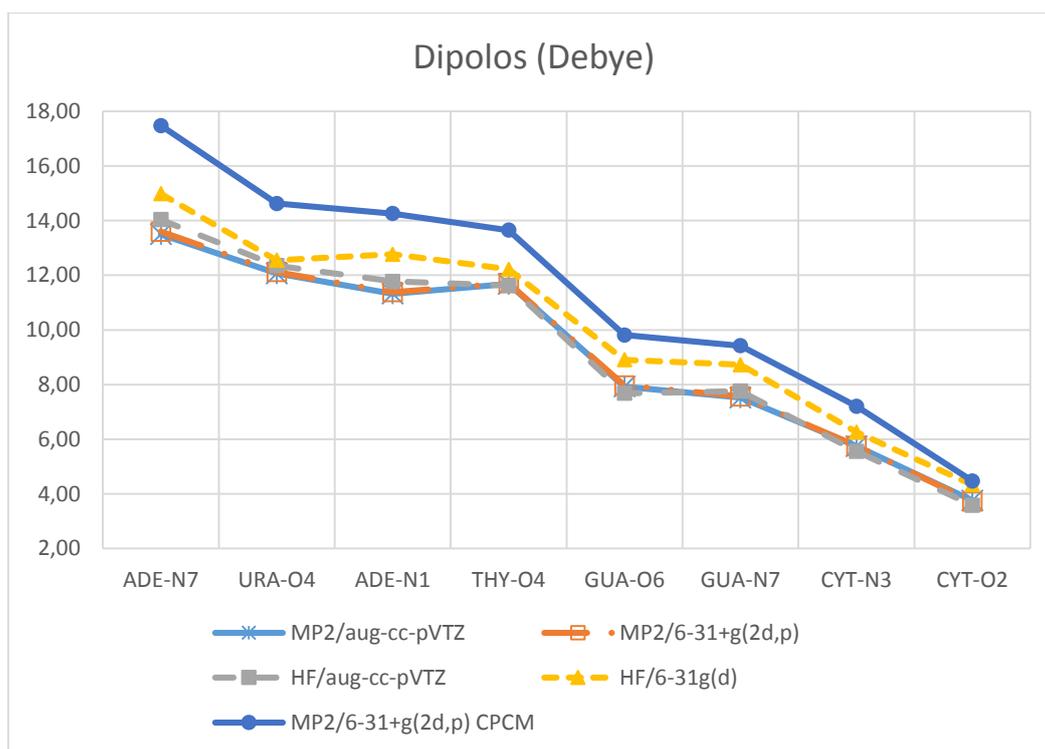


Figura 13 Momento dipolar dos 8 complexos estudados. Calculados por vários métodos e ordenados decrescentemente de acordo com os valores de MP2/6-31+G(2d,p) CPCM.

Na Tabela 10 e na Figura 13, é possível observar que o método HF/6-31G(d) (valores em **negrito**) superestimou os dipolos em relação ao demais métodos calculados no vácuo. Apesar das cargas HF/6-31G(d) gerarem dipolos mais próximos dos valores do solvente, isso não ocorre de maneira consistente, de modo que para os complexos CYT-O2, CYT-N3, GUA-N7 e GUA-O6 as diferenças entre os dipolos obtidos com HF/6-31G(d) e CPCM são inferiores a 1 D, enquanto que para os complexos THY-O4, ADE-N1, URO-O4 e ADE-N7, a diferença chega a 2,5 D.

Ao comparar os dipolos obtidos no vácuo pelos métodos MP2 e HF com o conjunto de bases aug-cc-pVTZ, nota-se que apesar de HF superestimar o valor das cargas, os dipolos dados pelos dois métodos foram similares, com diferenças inferiores a 5%.

O conjunto de bases 6-31+G(2d,p) é equivalente ao aug-cc-pVTZ na descrição tanto das cargas parciais ESP (Tabela 9) como também dos dipolos das moléculas estudadas, para o nível MP2 (Tabela 10).

12.5. Transferência de cargas

Interações entre as bases nucleicas e cátions bivalentes como Zn(II) e Mg(II) envolvem grandes transferências de cargas e polarização, esses dois fenômenos são considerados as principais contribuições para o termo de muitos corpos, ou seja, não são bem descritos pelo somatório das interações entre pares. Porém, ao incluir o efeito do solvente polar espera-se que, devido a sua alta constante dielétrica, o solvente ajude a estabilizar a carga do cátion metálico fazendo com que ocorra uma menor transferência de cargas entre o metal e as bases nucleicas.

No intuito de estudar a magnitude da influência do solvente na transferência de cargas entre o magnésio (II) pentahidratato e as bases nucleicas, utilizou-se o modelo de contínuo polarizável CPCM com a constante dielétrica $\epsilon=78,35$ para simular o meio aquoso. A transferência de cargas (TC) foi definida como a diferença entre a carga formal e a carga dos monômeros no complexo ⁷⁸, calculadas segundo as equações 15 e 16

$$TC = \sum \text{Cargas atômicas da base nucleica no complexo} \quad [15]$$

– Carga formal da base nucleica isolada

Como a carga formal das bases nucleicas é zero, temos;

$$TC = \sum \text{Cargas atômicas da base nucleica no complexo} \quad [16]$$

Também foram avaliados, os efeitos da correlação eletrônica e do conjunto de bases nas transferências de cargas. Para isso calculou-se o valor de TC utilizando as cargas ChelpG obtidas no vácuo em MP2/aug-cc-pVTZ, HF/aug-cc-pVTZ, MP2/6-31+G(2d,p) e HF/6-31+G(d), e no solvente implícito em MP2/6-31+G(2d,p). As geometrias dos complexos, obtidas no vácuo em MP2/6-31+G(2d,p), foram mantidas fixas para todos os cálculos. Utilizou-se os raios Bondi multiplicados por um fator de 1,2 na construção da cavidade para os cálculos no solvente.

Tabela 11 Valor das transferências de cargas obtidas no solvente contínuo CPCM em MP2/6-31+G(2d,p) e no vácuo em MP2/aug-cc-pVTZ, HF/aug-cc-pVTZ, MP2/6-31+G(2d,p) e HF/6-31G(d).

	ADE-N7	GUA-N7	CYT-N3	GUA-O6	ADE-N1	CYT-O2	THY-O4	URA-O4
MP2/aug-cc-pVTZ	0,44	0,37	0,31	0,28	0,27	0,17	0,05	0,03
HF/aug-cc-pVTZ	0,43	0,34	0,28	0,25	0,25	0,13	0,05	0,01
MP2/6-31+G(2d,p)	0,43	0,34	0,30	0,25	0,25	0,15	0,04	0,02
HF/6-31G(d)	0,39	0,30	0,24	0,24	0,23	0,12	0,05	0,02
CPCM	0,36	0,27	0,22	0,23	0,20	0,11	0,00	-0,01

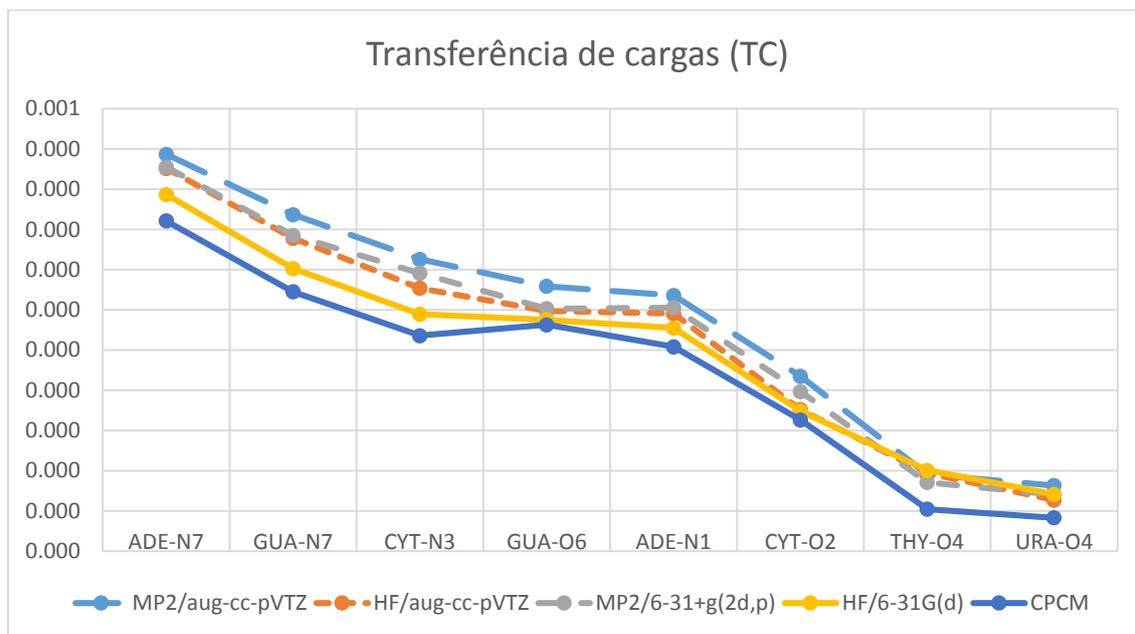


Figura 14 Perfil das cargas transferidas do $[Mg(H_2O)_5]^{2+}$ para as bases nucleicas nos 8 complexos estudados. CPCM é referente às transferências de cargas no solvente contínuo com o método MP2/6-31+G(2d,p). Os demais métodos foram calculados no vácuo.

A partir do perfil das transferências de cargas (TC) (Figura 14), nota-se que, em geral, esta é maior para os complexos nos quais o magnésio(II) pentahidratado interage com o nitrogênio das bases nucleicas, a única exceção foi o complexo Gua-O6 que apresentou transferência de cargas similar à do complexo Ade-N1.

Ao comparar os valores das TC obtidas no vácuo pelos métodos MP2/aug-cc-pVTZ, HF/aug-cc-pVTZ e MP2/6-31+G(2d,p) (Tabela 11) observa-se que a variação máxima entre esses três métodos foi de apenas 0,04 e que o perfil das curvas de TC na Figura 14 também foram semelhantes. Isso indica que o efeito da adição de correlação eletrônica foi pequeno na TC e que o conjunto de bases 6-31+G(2d,p) apresentou resultados próximos aos do aug-cc-pVTZ em MP2.

Também é possível observar que o método HF/6-31G(d) subestimou ligeiramente as transferências de cargas em relação aos demais métodos (desvio máximo de 0,07 em relação a MP2/aug-cc-pVTZ) e apresentou os valores de TC mais próximos aos obtidos pelo MP2/6-31+G(2d,p) com adição do efeito do solvente CPCM, com exceção dos complexos Ura-O4 e Thy-O4, nos quais o método HF/6-31+G(d) apresentou valores de TC mais similares aos demais cálculos realizados no vácuo.

13 Considerações finais

- A aproximação de pares fez com que as energias de interação intermoleculares entre o magnésio(II), as moléculas de água e as bases nucleicas fossem subestimadas em mais de um terço de seu valor. Campos de força que empregam essa aproximação dificilmente serão capazes de tratar de forma acurada as interações entre o Mg^{2+} e as bases nucleicas do DNA/RNA.
- Ao comparar os resultados obtidos com o método HF/aug-cc-pVTZ em relação aos valores em MP2/aug-cc-pVTZ, nota-se que a adição da correlação eletrônica pouco influenciou na descrição dos dipolos e nas transferências de cargas, entretanto, o método HF superestimou a magnitude das cargas atômicas ChelpG.
- O conjunto de bases 6-31+G(2d,p) apresentou valores equivalentes de cargas atômicas, dipolos e transferências de cargas em relação ao conjunto de bases aug-cc-pVTZ no nível MP2, a um menor custo computacional.
- Em concordância com a estratégia do AMBER, as moléculas cujas cargas foram calculadas no vácuo em HF/6-31G(d) apresentaram dipolos e transferências de cargas mais próximas às obtidas pelas moléculas solvatadas em água, porém de forma inconsistente.

CAPÍTULO 4: Energias livres de formação dos complexos $[\text{Mg}(\text{H}_2\text{O})_n \text{ base nucleica}]^{2+}$

Na fase gasosa e na fase aquosa

14 Modos de coordenação

Em meio aquoso, o cátion Mg^{2+} se apresenta em sua forma hidratada $[Mg(H_2O)_6]^{2+}$ de geometria octaédrica⁷⁹. Existem 3 modos principais de coordenação entre o cátion metálico hidratado e as bases nucleicas²², são eles:

- 1) **Modo bidentado:** o cátion libera duas moléculas de água, formando magnésio(II) tetrahidratado e interage diretamente com dois sítios das bases nucleicas.
- 2) **Modo monodentado:** o cátion libera uma molécula de água, formando o magnésio(II) pentahidratado e interage diretamente com um sítio de nitrogênio ou oxigênio das bases nucleicas, podendo também formar ligações de hidrogênio entre as águas coordenadas ao cátion metálico e os átomos de N e O das bases nucleicas.
- 3) **Modo Indireto:** o magnésio(II) hexahidratado interage apenas indiretamente com as bases nucleicas. O cátion polariza as moléculas de água coordenadas a ele, formando ligações de hidrogênio com os sítios de N e O das bases nucleicas.

Por se tratar de um cátion de pequeno raio iônico, 0,65 Å, o Mg^{2+} apresenta uma alta densidade de carga positiva⁴⁹, podendo ser classificado como um ácido duro. Essa característica faz com que, em meio biológico, sua primeira camada de hidratação seja bem definida, sua energia de hidratação seja alta e a velocidade de troca dessas águas sejam relativamente lentas²⁰.

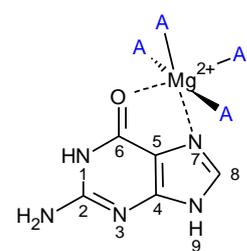
Estudos indicam que o magnésio tende a se manter hexacoordenado, substituindo as águas de hidratação, geralmente por átomos de oxigênio pertencentes ao DNA/RNA⁷⁹.

Os sítios mais comuns em que o magnésio se coordena são os oxigênios do esqueleto de fosfato do DNA/RNA, os sítios N7 da guanina e adenina, o sítio O6 da guanina, o N3 da citosina e o sítio O4 da timina e da uracila^{22,80}.

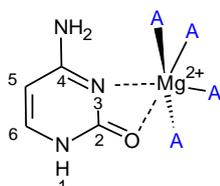
Com a finalidade de estudar a magnitude da energia de interação desses três modos de coordenação, 16 complexos descritos como sendo os mais estáveis foram selecionados; dentre eles 2 se ligam de maneira bidentada e suas geometrias iniciais foram retiradas da tese de doutorado de Maria A. S. Prado,³⁵ 8 complexos descritos no primeiro capítulo (Figura 1), se ligam de maneira monodentada e 6 complexos interagem indiretamente através das moléculas de água coordenadas ao Mg^{2+} , dos quais 5 geometrias iniciais foram retiradas da tese de Maria A. S. Prado³⁵ e uma estrutura do trabalho de Ivan Solt e colaboradores³⁴. A representação esquemática de todos os complexos pode ser vista na Figura 15.

Os complexos tiveram suas geometrias otimizadas em MP2/6-31+G(2d,p). Em seguida, as energias de interação corrigidas pelo procedimento *counterpoise* ΔE_{cp} e não corrigidas ΔE , assim como definidas nas equações 1 e 2 do capítulo 2, foram calculadas utilizando diferentes métodos *ab initio* (Figura 16):

- i) Em MP2/Jul-cc-pVTZ, esse método foi utilizado como parâmetro de referência por apresentar valores similares ao MP2/aug-cc-pVTZ, a um custo computacional menor.
- ii) Em HF/Jul-cc-pVTZ para que ao comparar com o método i se tenha uma noção do efeito da correlação eletrônica nas interações dos complexos que não foram abordados no capítulo 1.
- iii) Em MP2/6-31+G(2d,p) com a finalidade de verificar se o método que apresentou a melhor relação entre qualidade de resultado e custo computacional no capítulo 1, também fornece energias de interação próximas às obtidas pelo método de referência (MP2/Jul-cc-pVTZ) para os complexos nos quais o magnésio interage de forma bidentada ou de forma indireta com as bases nucleicas.

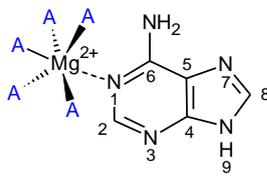


4-Gua-O6/N7

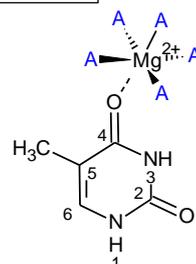


4-Cyt-O2/N3

MONODENTADOS

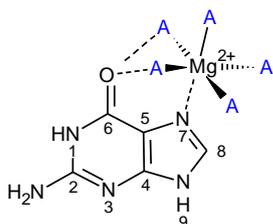


5-Ade-N1

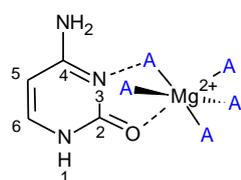


5-Thy-O4

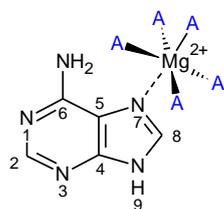
BIDENTADOS



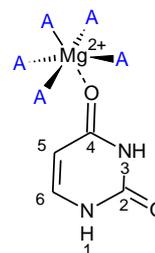
5-Gua-N7



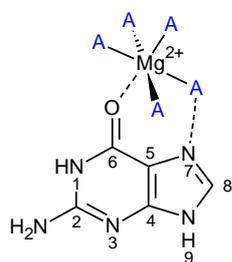
5-Cyt-O2



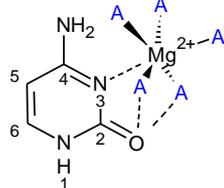
5-Ade-N7



5-Ura-O4

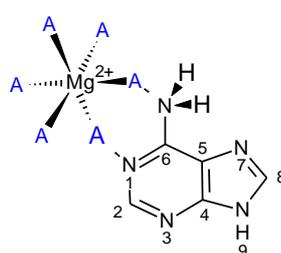


5-Gua-O6

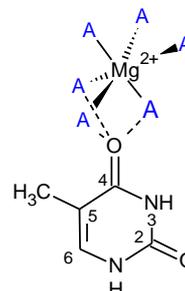


5-Cyt-N3

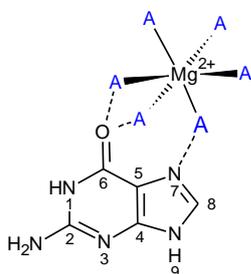
INDIRETAMENTE



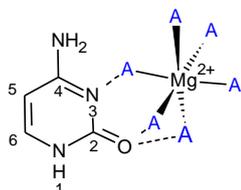
6-Ade-N1/N6



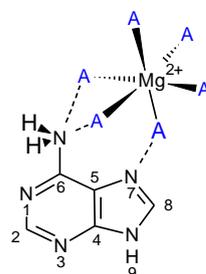
6-Thy-O4



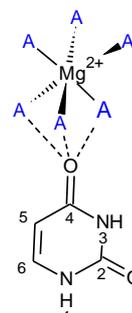
6-Gua-O6/N7



6-Cyt-O2N3



6-Ade-N6/N7



6-Ura-O4

Figura 15 Representação esquemática dos complexos estudados. O primeiro número indica quantas moléculas de H_2O se ligam ao Mg^{2+} , seguido da abreviação das bases nucleicas (guanina Gua, citosina Cyt, Adenina Ade, Timina Thy, Uracila Ura) e do sítio ao qual o magnésio se coordena, oxigênio O ou nitrogênio N e o número de sua posição na estrutura das bases nucleicas. A letra A se refere às águas de hidratação do Magnésio.

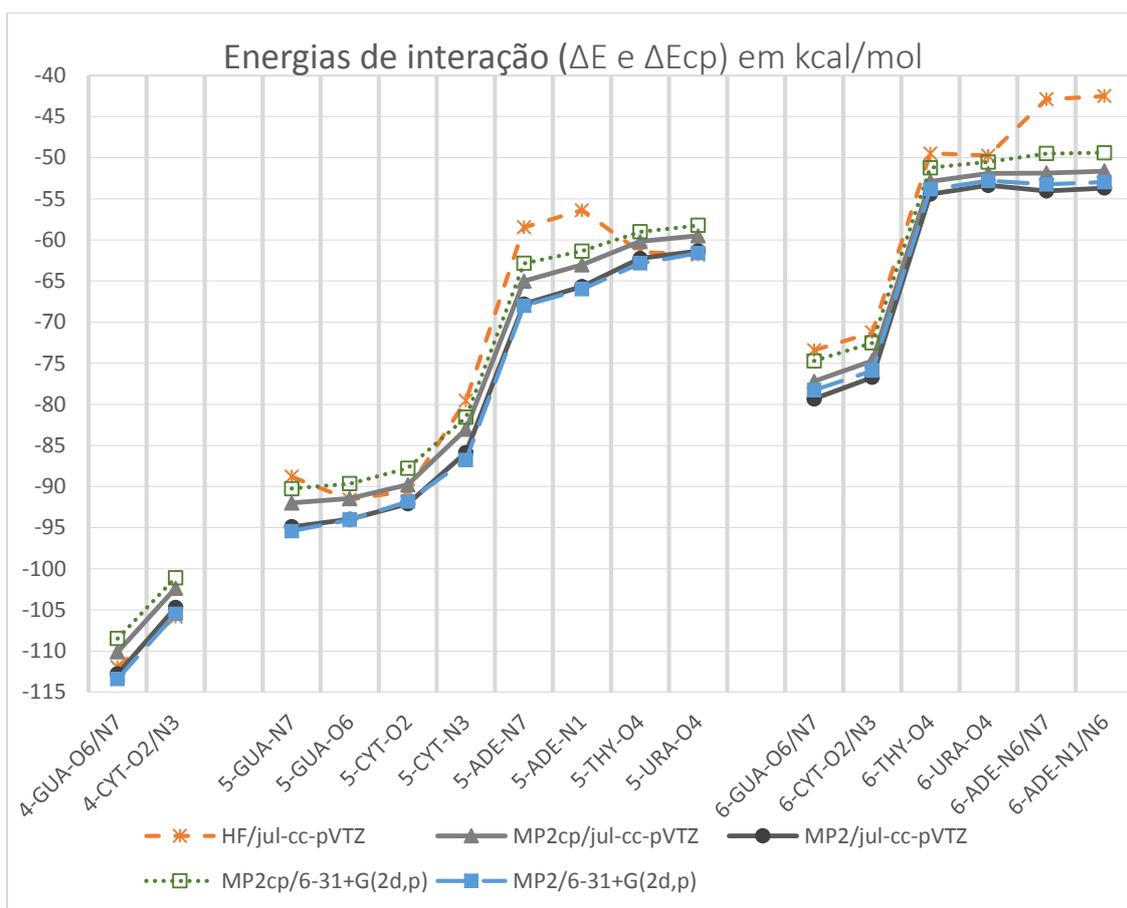


Figura 16 Energias eletrônicas de interação (ΔE) de todos os complexos. A abreviação cp se refere aos valores corrigidos com counterpoise.

Ao comparar o perfil das energias de interação obtidas em MP2/Jul-cc-pVTZ com o perfil de HF/Jul-cc-pVTZ na Figura 16, é possível notar que nos complexos cuja interação entre o cátion metálico e as bases nucleicas ocorre indiretamente, o efeito da adição de correlação eletrônica é maior para 6-Ade-N6/N7 e 6-Ade-N1/N6, representando cerca de 20% do ΔE desses complexos, enquanto para os demais complexos que interagem apenas indiretamente a correlação foi de aproximadamente 7% de ΔE . Nos complexos bidentados a correlação eletrônica não influenciou as energias de interação, pois seu valor foi inferior à 1% de ΔE .

Portanto deve-se empregar métodos correlacionados não só nos estudos das interações dos complexos monodentados em que o magnésio(II) pentahidratado interage diretamente com os sítios de nitrogênio das bases nucleicas^{xvi}, mas também para os complexos em que o magnésio(II) hexahidratado interage indiretamente com as bases nucleicas. Em ambos os

^{xvi} Capítulo 2, seção 6.2.2

modos de coordenação a correlação é maior quando o cátion interage com a adenina.

Na Figura 16, nota-se também que não só os perfis das interações ΔE e ΔE_{cp} calculados em MP2/Jul-cc-pVTZ foram similares como também seus valores diferiram em menos de 5% (1,4kcal/mol a 2,9kcal/mol) para todos os complexos. Isso mostra que o BSSE das interações não foi significativo no método de referência. Em relação ao método MP2/6-31+G(2d,p), suas energias de interação não corrigidas com *counterpoise* (ΔE) foram mais próximas às do método de referência do que suas energias corrigidas ΔE_{cp} . Consequentemente, a correção de *counterpoise* além de aumentar o custo computacional, levou a valores piores em relação aos da referência para as energias de interação aqui estudadas no nível MP2.

O cátion metálico Mg^{2+} interagiu mais fortemente com os sítios da guanina e citosina e preferencialmente de forma bidentada (Figura 16). Isso ocorreu devido à maior concentração de cargas negativas nos sítios O2/N3 da citosina e O6/N7 da guanina, como pode ser visto no mapa da superfície de potencial eletrostático da Figura 17. Já os sítios da adenina, timina e uracila, por terem uma menor densidade de cargas, resultam em interações de menor caráter eletrostático, portanto mais fracas com o Mg^{2+} .

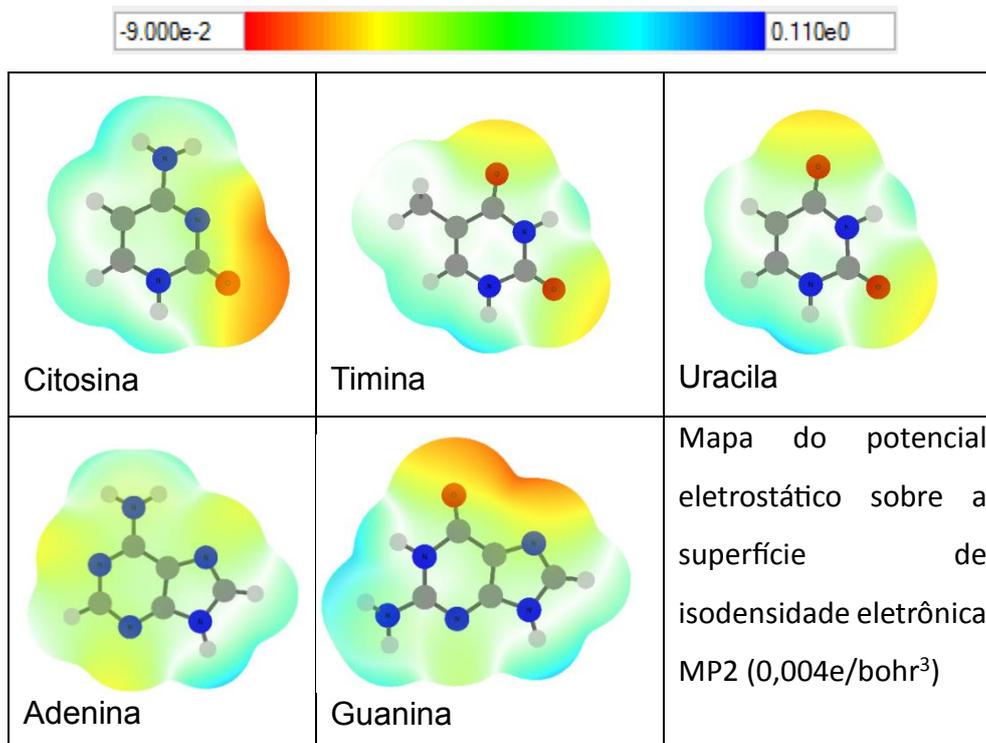
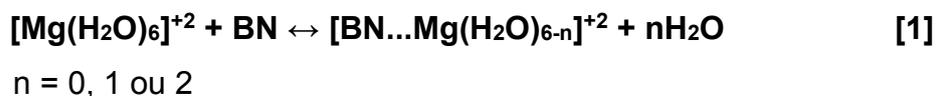


Figura 17 Mapa de potencial das bases nucleicas. Os valores em vermelho se referem às regiões que o potencial eletrostático é negativo, enquanto as regiões em azul indicam potencial eletrostático positivo.

Porém, as energias de interação (ΔE e ΔE_{cp}) tratadas na Figura 16, não levaram em consideração a perda energética na deformação das estruturas ao se complexarem, nem a energia envolvida na liberação de uma ou duas moléculas de água necessárias para que, a partir do magnésio(II) hexahidratado, sejam formados os complexos mono e bidentados respectivamente. Além do mais, não foram incluídas as correções térmicas, ou seja, os cálculos foram realizados considerando a temperatura de zero Kelvin, sem incluir a correção do ponto zero, e apenas a primeira camada de hidratação do cátion foi incluída, desconsiderando os efeitos do solvente em torno dos complexos.

A fim de contabilizar tais contribuições e correções nos cálculos computacionais, utilizou-se a equação química 1, na qual, se tem como reagentes as bases nucleicas (BN) e o magnésio(II) hexahidratado, formando os complexos da Figura 15 e liberando uma molécula de água para os complexos monodentados, e duas moléculas de água para os bidentados.



As contribuições térmicas foram adicionadas a partir dos cálculos das frequências vibracionais harmônicas de cada molécula e o efeito do solvente foi estimado por meio dos modelos de contínuo polarizável (PCM).

15 Contribuições térmicas a partir do cálculo de frequências

Enquanto o procedimento de otimização de geometrias calcula a derivada primeira da energia em relação à posição dos núcleos das moléculas, para se chegar a uma estrutura em um ponto estacionário na hipersuperfície de energia potencial (SEP), as frequências vibracionais são computadas a partir das derivadas segundas, que fornecem os valores das constantes harmônicas das forças vibracionais. A ausência de frequências imaginárias indica que a estrutura encontrada é um ponto de mínima energia na SEP.

Além de confirmar a natureza do ponto estacionário, a partir do cálculo das frequências é possível fazer correções térmicas na energia eletrônica, adicionando a ela as contribuições entálpicas e entrópicas. Para isso, admite-se que as moléculas se comportem como gases ideais e aplica-se as equações padrões da termodinâmica estatística de modo que as contribuições vibracionais, rotacionais e translacionais são tratadas separadamente conforme os modelos aproximados do oscilador harmônico, rotor rígido, e partícula na caixa para o *ensemble*^{xvii} canônico (pressão e temperatura constantes)⁸¹.

16 Métodos termoquímicos compostos CBS

Os métodos termoquímicos compostos, apesar de terem altíssimo custo computacional, são capazes de fornecer entalpias e energias livres com acurácia próxima ou até superior à experimental. Essas energias são obtidas a partir de uma série de cálculos e extrapolações feitas em diferentes níveis de teoria⁸². Entre os métodos compostos, o modelo CBS-QB3 é o mais empregado para cálculos termoquímicos devido à sua ótima relação entre qualidade de resultados e custo computacional⁸³. O CBS-QB3 foi desenvolvido por Petersson e colaboradores, seus cálculos envolvem 5 etapas⁸⁴:

- i) Otimização de geometria utilizando o funcional de densidade B3LYP e o conjunto de bases 6-311G(2d,d,p)^{xviii}.
- ii) Cálculo teórico das frequências em B3LYP/6-311G(2d,d,p), aplicando um fator de escala de 0,99 para o ZPE.
- iii) Cálculo da energia em CCSD(T)/6-31+G(d')^{xix}
- iv) Correções utilizando o cálculo de energia em MP4/CBSB4^{xx}.

^{xvii} Conjunto imaginário de sistemas idênticos.

^{xviii} O conjunto de bases 6-311G(2d,d,p) utiliza funções de polarização 2d para Na-Ar, d para Li-Ne e p para H e He (também chamado de CBSB7).

^{xix} 6-31+G(d') é uma modificação de 6-31+G(d) obtida combinando as funções sp de 6-31G com os exponenciais de polarização de 6-311G(d,p).⁸⁴

^{xx} CBSB4 ou 6-31+G(d(f),p) a notação se refere a adição de função de polarização d em Li-Si, df em P, S e Cl e uma função de polarização p no H e He.

v) Cálculo da energia em MP2/CBSB3^{xxi} e extrapolação da energia MP2 utilizando o esquema CBS⁸⁵.

A energia livre é então calculada da seguinte maneira:

$$G_{CBS-QB3} = E(MP2)_{CBS} + \Delta E(MP4) + \Delta E(CCSD(T)) + \Delta E_{termico_{B3LYP}} + \Delta E_{inter} + \Delta E_{spin} \quad [2]$$

Os termos ΔE_{inter} e ΔE_{spin} são determinados por parâmetros empíricos, o primeiro se refere ao coeficiente de sobreposição-interferência de dois elétrons e o segundo é uma correção ao problema de contaminação de spin⁸². Os termos $\Delta E(MP4)$ e $\Delta E(CCSD(T))$ são calculados conforme as equações 3 e 4⁸⁶:

$$\Delta E(MP4) = E_{MP4(SDQ)/CBSB4} - E_{MP2/CBSB4} \quad [3]$$

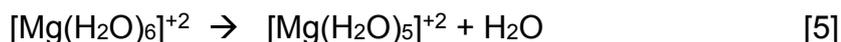
$$\Delta E(CCSD(T)) = E_{CCSD(T)/6-31+G(d')} - E_{MP4(SDQ)/6-31+G(d')} \quad [4]$$

Os erros médios absolutos do método CBS-QB3 para o conjunto de teste G2/97 é de apenas 1,10 kcal/mol. Esse conjunto de teste consiste de 302 energias obtidas experimentalmente, das quais 148 são calores de formação, 88 potenciais de ionização, 58 afinidades eletrônicas e 8 afinidades protônicas, envolvendo 148 moléculas neutras e 177 íons⁸³.

16.1. Definição de uma estratégia de cálculo acurada e prática

Para estimar a acurácia do método CBS-QB3, utilizou-se a reação em que o magnésio(II) hexahidratado libera uma molécula de água formando magnésio(II) pentahidratado (equação 5), essa reação foi escolhida por apresentar valores experimentais com pequenas margens de erros e envolver o cátion magnésio(II).⁸⁷:

^{xxi} CBSB3 ou 6-311+G(3d2f,2df,2p) que adiciona 3d2f em átomos Na-Ar, 2df para Li-Ne e 2p para H e He.



Primeiramente as estruturas dos reagentes e produtos foram otimizados em MP2/6-31+G(2d,p) em seguida, foram realizados os cálculos CBS-QB3 para se obter a entalpia, entropia e energia livre da reação (Tabela 12)

Tabela 12 Valores da entalpia, entropia e energia livre da reação 5 calculados utilizando o método CBS-QB3 e comparados aos valores experimentais ⁸⁷. Os valores entre parênteses se referem aos erros máximos estimados.

	Experimental ⁸⁷	CBS-QB3
$\Delta H(298\text{K})$ (kcal/mol)	24,6(1,0)	30,1
$\Delta S(298\text{K})$ (cal/K mol)	29,1(2,0)	36,2
$\Delta G(298\text{K})$ (kcal/mol)	16,0(0,5)	19,4

O método CBS-QB3 superestimou tanto o valor de ΔH como de ΔS obtendo um desvio de aproximadamente 3,4 kcal/mol para a energia livre de reação, esse valor será considerado o erro estimado do CBS-QB3 para os ΔG_r das reações de formação dos complexos (equação 1).

Devido ao alto custo computacional do método CBS-QB3, não foi possível calcular as energias livres em CBS-QB3 para 4 dos 16 complexos estudados (Figura 15), sendo eles: 6-Gua-O6/N7, 6-Ade-N1/N6, 6-Ade-N6/N7 e 6-Cyt-O2/N3. Portanto, para encontrar uma estratégia computacional capaz de reproduzir os resultados do CBS-QB3 a um menor custo computacional, as contribuições eletrônicas e térmicas das doze reações restantes foram calculadas em vários métodos e comparadas com os resultados obtidos pelo CBS-QB3.

16.1.1 Contribuições eletrônicas

As contribuições eletrônicas das reações de formação dos complexos (equação 1) foram calculadas pela diferença entre as energias eletrônicas dos reagentes e dos produtos (ΔE_r), obtidas no vácuo a zero Kelvin, conforme a equação 6:

$$\Delta E_r = E(\text{produtos}) - E(\text{reagentes}) \quad [6]$$

Primeiramente, otimizou-se a geometria de todas as moléculas dos reagentes e produtos a nível MP2/6-31+G(2d,p), em seguida, foram obtidas as energias eletrônicas dessas moléculas em MP2/6-31+G(2d,p), MP2/aug-cc-pVTZ e HF/aug-cc-pVTZ. As geometrias foram, então, reotimizadas em B3LYP/6-311G(2d,d,p), B3LYP/6-31G(d) e HF/6-31G(d) e as energias eletrônicas foram calculadas mantendo os mesmos níveis de teoria utilizados na reotimização. Os desvios desses métodos em relação às energias eletrônicas de reação calculadas em CBS-QB3 para 12 das 16 reações estudadas estão descritas na Tabela 13.

Tabela 13 Desvios médios (DMed), desvios médios absolutos (DMedA) e desvios máximos (DMax) em relação às energias eletrônicas das reações de formação dos complexos obtidos em CBS-QB3.

MÉTODOS	ΔE_r (kcal/mol)		
	DMed	DMedA	DMax
MP2/aug-cc-pVTZ//MP2/6-31+G(2d,p)	-0,92	0,92	1,85
HF/aug-cc-pVTZ//MP2/6-31+G(2d,p)	4,06	4,57	9,32
MP2/6-31+G(2d,p)	0,00	0,63	1,60
B3LYP/6-311G(2d,d,p) [CBSB7]	4,01	4,50	8,26
B3LYP/6-31G(d)	1,14	2,75	5,45
HF/6-31G(d)	4,29	4,29	7,45

A partir da Tabela 13 é possível notar que o método MP2/6-31+G(2d,p) apresentou desvios médios absolutos inferiores a 1kcal/mol e desvio máximo absoluto de apenas 1,6 kcal/mol em relação ao CBS-QB3. Esses desvios foram inferiores aos do método de maior custo computacional MP2/aug-cc-pVTZ//MP2/6-31+G(2d,p), que além de apresentar maiores desvios médios absolutos, em média subestimou o valor de ΔE_r em -0,92 kcal/mol.

As energias eletrônicas de reação obtidas em HF, mesmo com o conjunto de bases aug-cc-pVTZ, apresentaram altos desvios mostrando que os efeitos de correlação eletrônica foram significativos nos cálculos de ΔE_r . O funcional de densidade B3LYP também apresentou desvios significativamente maiores que os desvios obtidos em MP2.

Portanto, entre os métodos testados, o MP2/6-31+G(2d,p) foi considerado a opção com a melhor relação entre acurácia e custo computacional para substituir o CBS-QB3 nos cálculos das contribuições eletrônicas das reações estudadas.

16.1.2 Contribuições térmicas

O cálculo das contribuições termoquímicas a partir das frequências vibracionais é a etapa de menor custo computacional do CBS-QB3 e é feita independentemente dos cálculos de alto custo utilizados para as energias eletrônicas ⁸². Permitindo que o estudo das contribuições térmicas fosse realizado para todas as 16 reações.

No CBS-QB3, essas contribuições são obtidas pelas frequências harmônicas vibracionais calculadas em B3LYP com o conjunto de bases 6-311G(2d,d,p), utilizando um fator de escala de 0,99 para as energias de ponto zero ⁸². Como esse fator pouco alterou os valores das correções termoquímicas das reações estudadas, conforme mostra a Tabela 14, utilizou-se os cálculos feitos com as frequências não escalonadas de B3LYP/6-311G(2d,d,p) como valores de referência.

Os desvios máximos e médios absolutos em relação às correções termoquímicas do B3LYP/6-311G(2d,d,p), foram então calculados para MP2/6-31+G(2d,p), HF/6-31G(d) e B3LYP/6-31G(d) e podem ser vistos na Tabela 14.

O cálculo das frequências só fornece resultado significativo quando feito em pontos estacionários da hipersuperfície de energia, em pontos de máximo para estudar estados de transição e em pontos de mínimo para estudar estruturas de mínima energia. Portanto, todas as moléculas tiveram suas geometrias previamente otimizadas no mesmo nível de cálculo utilizado para obter as frequências. Não foram observadas frequências imaginárias em nenhum cálculo, confirmando que se tratavam de estruturas de mínima energia.

As contribuições térmicas analisadas foram:

1. A energia do ponto zero das reações (ΔZPE),

$$\Delta ZPE = ZPE(\text{produtos}) - ZPE(\text{reagentes}) \quad [7]$$

2. A correção entálpica das reações (ΔH_{corr});

$$\Delta H_{\text{corr}} = H_{\text{corr}}(\text{reagentes}) - H_{\text{corr}}(\text{produtos}) \quad [8]$$

Em que,

$$H_{\text{corr}} = ZPE + E_{\text{trans}} + E_{\text{vib}} + E_{\text{rot}} + K_B T \quad [9]$$

Etrans, Evib, Erot são as contribuições térmicas translacionais, vibracionais e rotacionais, respectivamente e $K_B T$ se refere à constante de Boltzmann multiplicada pela temperatura definida para os cálculos (298,15K), seu valor é de 0,59kcal/mol.

3. A correção da energia livre de reação ΔG_{corr} :

$$\Delta G_{corr} = G_{corr}(\text{reagentes}) - G_{corr}(\text{produtos}) \quad [10]$$

Em que,

$$G_{corr} = H_{corr} - TS \quad [11]$$

4. O termo entrópico $-T\Delta S_r$:

$$-T\Delta S_r = \Delta G_{corr} - \Delta H_{corr} \quad [12]$$

Em que,

$$S = S_{trans} + S_{vib} + S_{rot} + S_0 \quad [13]$$

$$\Delta S_r = S(\text{reagentes}) - S(\text{produtos}) \quad [14]$$

Strans, Svib, Srot e S_0 são as contribuições entrópicas, translacionais, vibracionais, rotacionais e do ponto zero, respectivamente.

Tabela 14 Desvios médios (DMedA) e máximos (DmaxA) absolutos referentes às contribuições termoquímicas obtidas em B3LYP/6-311G(2d,d,p) sem fator de escala, com exceção dos desvios sublinhados, que foram calculados em relação aos valores B3LYP/6-311G(2d,d,p) com fator de escala de 0,99 para o ZPE. Todos os valores estão em kcal/mol.

	ΔZPE		ΔH_{corr}		$-T\Delta S$		ΔG_{corr}	
	DMedA	DMaxA	DMedA	DMaxA	DMedA	DMaxA	DMedA	DMaxA
B3LYP/6-311G(2d,d,p)	<u>0,02</u>	<u>0,04</u>	<u>0,01</u>	<u>0,02</u>	<u>0,01</u>	<u>0,03</u>	<u>0,01</u>	<u>0,05</u>
B3LYP/6-31G(d)	0,52	0,91	0,41	0,78	0,99	1,55	1,22	2,12
MP2/6-31+G(2d,p)	0,20	0,49	0,11	0,31	0,56	1,47	0,53	1,42
HF/6-31G(d)	0,20	0,59	0,43	1,31	0,25	0,58	0,39	1,03

Apesar do B3LYP/6-31G(d) empregar o mesmo nível de cálculo da referência, seus desvios absolutos foram maiores para ΔZPE , $-T\Delta S$ e ΔG_{corr} do que o dos outros métodos testados. O método MP2/6-31+G(2d,p) teve boa concordância em relação aos valores de ΔZPE e ΔH_{corr} da referência, com desvios máximos inferiores a 0,5 kcal/mol, no entanto, esse método possui

elevado custo computacional para cálculo de frequências e seus desvios em relação a $-T\Delta S$ e ΔG_{corr} foram maiores que os do método de menor custo computacional HF/6-31G(d).

Entre os métodos testados, o HF/6-31G(d) apresentou os menores desvios médios e máximos absolutos para os valores de ΔG_{corr} . Devido ao menor custo computacional desse método em relação à referência, ele poderia ser aplicado futuramente no estudo de sistemas maiores similares aos estudados, poupando recursos computacionais, mas mantendo uma acurácia próxima à do B3LYP/6-31G(2d,d,p).

17 Energia livre de reação para o estado gasoso

A estratégia de menor custo computacional e maior acurácia em relação às entalpias (ΔH_r) e energias livres de reação (ΔG_r), em fase gasosa, obtidas pelo CBS-QB3 para as reações químicas descritas na equação 1, é composta de duas etapas:

- 1) Otimização das geometrias e cálculo das energias eletrônicas de reação (ΔE_r) em MP2/6-31+G(2d,p), conforme a equação 6.
- 2) Reotimização das geometrias seguida do cálculo de frequências em B3LYP/6-311G(2d,d,p) (CBSB7) para obter ΔH_{corr} , ΔG_{corr} e $-T\Delta S$ conforme as equações 8, 11 e 12.

As entalpias de reação foram então calculadas combinando as equações 6 e 8, conforme a equação 15:

$$\Delta H_r = \Delta E_r + \Delta H_{corr} \quad [15]$$

As energias livres de reação foram calculadas combinando as equações 6 e 10, conforme a equação 16:

$$\Delta G_r = \Delta E_r + \Delta G_{corr} \quad [16]$$

As principais contribuições de ΔH_r e ΔG_r calculadas para as 16 reações, assim como a comparação entre os valores de ΔG_r obtidos pela estratégia de cálculos de menor custo computacional e pelo CBS-QB3 para 12 das 16 reações estão na Tabela 15.

Para uma comparação qualitativa entre as contribuições eletrônicas, entrópicas e entálpicas das energias de reação, os valores de ΔE_r , ΔH_r e ΔG_r , foram agrupados conforme o modo de coordenação entre o Mg^{2+} e as bases nucleicas e organizados em ordem crescente de acordo com os valores de ΔG_r , como podem ser vistos na Figura 18.

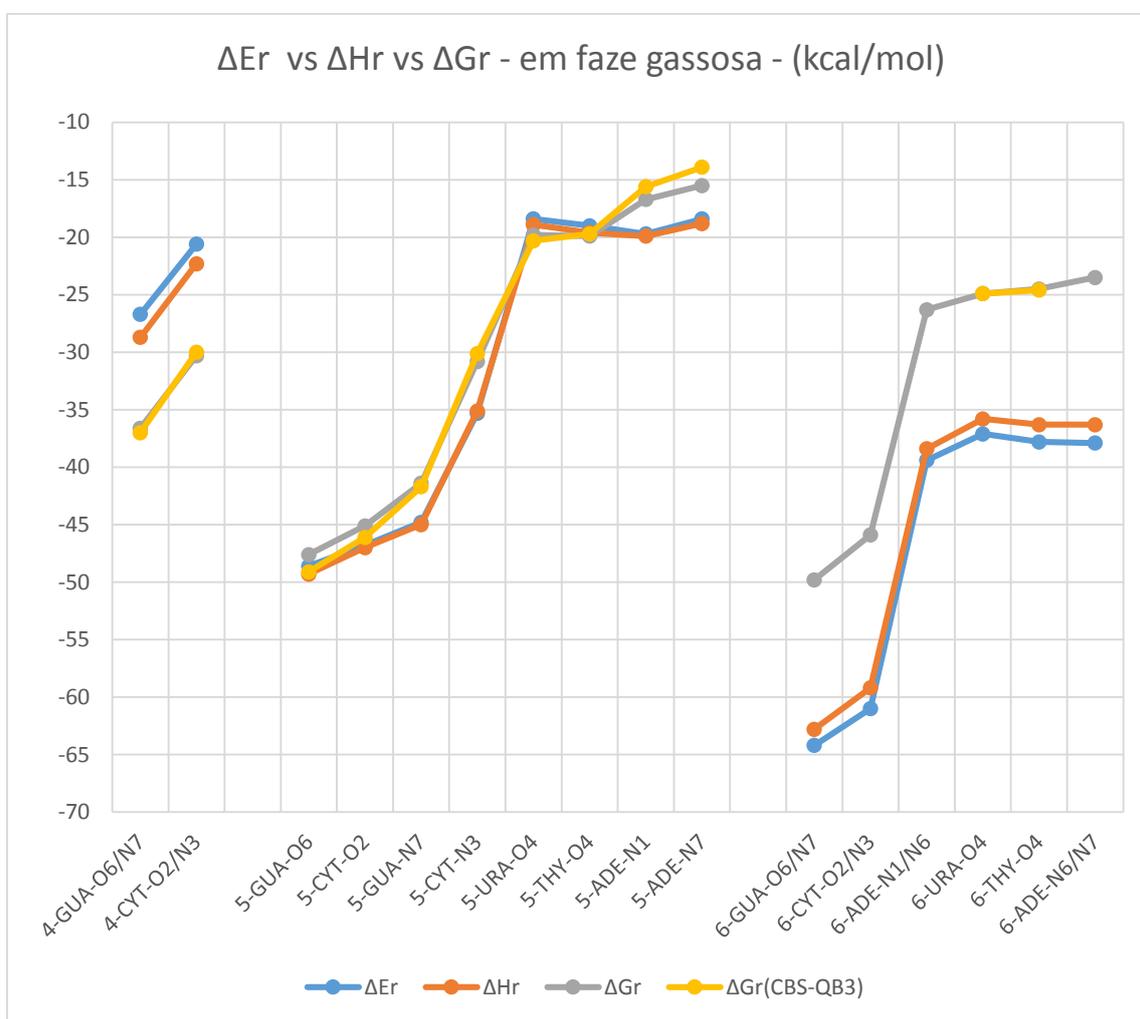


Figura 18 Perfil das energias eletrônicas de reação, entalpias de reação e energias livres de reação em fase gasosa a 298,15 K.

Tabela 15 Energias livres de reação obtidas pelo CBS-QB3 e pela estratégia de menor custo. As energias livres deste último método foram decompostas nas contribuições eletrônicas (ΔE_r), correções entálpicas (ΔH_{corr}), entalpias de reação (ΔH_r), contribuições entrópica ($-T\Delta S_r$). Todos os valores estão em kcal/mol.

Modos de interação	Reações (T=298,15K)	$\Delta Gr(gás)$ (kcal/mol)				CBS-QB3	
		$\Delta Hr(gás)$ (kcal/mol)		$-T\Delta Sr$	ΔGr	ΔGr	
		ΔE_r	Hcorr				ΔHr
Bidentados	4-GUA-O6/N7	-26,7	-2,0	-28,7	-7,9	-36,6	-37,0
	4-CYT-O2/N3	-20,6	-1,7	-22,3	-8,1	-30,3	-30,0
Monodentados	5-GUA-O6	-48,6	-0,7	-49,3	1,7	-47,6	-49,1
	5-CYT-O2	-46,8	-0,1	-47,0	1,9	-45,1	-46,1
	5-GUA-N7	-44,8	-0,2	-45,0	<u>3,6</u>	-41,4	-41,7
	5-CYT-N3	-35,3	0,2	-35,1	<u>4,3</u>	-30,8	-30,1
	5-URA-O4	-18,4	-0,4	-18,9	-0,9	-19,8	-20,3
	5-THY-O4	-19,0	-0,6	-19,6	-0,3	-19,9	-19,7
	5-ADE-N1	-19,7	-0,2	-19,9	<u>3,2</u>	-16,7	-15,6
	5-ADE-N7	-18,4	-0,4	-18,8	<u>3,3</u>	-15,5	-13,9
Interações indiretas	6-GUA-N7/O6	-64,2	1,4	-62,8	13,1	-49,8	
	6-CYT-O2/O6	-61,0	1,8	-59,2	13,4	-45,9	
	6-ADE-N1	-39,4	1,0	-38,4	12,2	-26,3	
	6-URA-O4	-37,1	1,4	-35,8	10,8	-24,9	-24,9
	6-THY-O4	-37,8	1,5	-36,3	11,8	-24,5	-24,6
	6-ADE-N7/N1	-37,9	1,6	-36,3	12,8	-23,5	

Apesar da metodologia de menor custo computacional apresentar desvios percentuais em relação aos valores de ΔGr do CBS-QB3 de até -11,5% para a 5-Ade-N7, a diferença máxima entre os ΔGr desses dois métodos foi de apenas 1,6 kcal/mol (também para a 5-Ade-N7) (Tabela 15), valor inferior à acurácia de $\pm 3,4$ kcal/mol^{xxii} CBS-QB3. O perfil das energias livres de reação (ΔGr) obtido pela metodologia de menor custo computacional foi muito similar ao do CBS-QB3 (Figura 18).

Ao comparar os valores de ΔH_{corr} da Tabela 15 e os perfis de ΔE_r e ΔHr é possível perceber que as contribuições entálpicas foram pequenas para todas as reações, pois ΔH_{corr} variou apenas entre -2,0 a 1,8 kcal/mol e não alterou o perfil de ΔE_r . Entretanto, ao comparar os valores ΔGr aos de ΔHr , nota-se que as contribuições entrópicas ($-T\Delta S_r$) não podem ser desconsideradas, pois seus valores variaram entre -8,1 a 13,4 kcal/mol, alterando significativamente o perfil de ΔGr em relação a ΔHr (Figura 18).

^{xxii} Estimada na seção 16.1

Nas reações de formação dos 2 complexos bidentados, a contribuição entrópica é negativa devido ao aumento dos graus de liberdade com a liberação das duas moléculas de água. O termo $-T\Delta S_r$ foi de aproximadamente -8 kcal/mol e representou de 22% a 27% de ΔG_r .

Nas reações de formação dos 8 complexos monodentados, a contribuição entrópica apresentou tanto valores positivos como negativos, variando de -0,9 a 4,3 kcal/mol. Para essas reações, o termo $-T\Delta S_r$ é maior nos casos em que o magnésio(II) pentahidratado interage diretamente com um sítio de nitrogênio, representando até 21% de ΔG_r (para a 5-Ade-N7).

Em relação às reações de formação dos 6 complexos nos quais o magnésio interage apenas indiretamente com as bases nucleicas, a contribuição entrópica é sempre positiva devido à diminuição dos graus de liberdade. Nessas reações os valores de $-T\Delta S_r$ foram similares, variando entre 10,8 kcal/mol a 13,4 kcal/mol. No entanto percentualmente as contribuições entrópicas representaram de -26% a -54% de ΔG_r para esses complexos.

Como as energias livres de reação (ΔG_r) foram todas inferiores a -10 kcal/mol, a formação de todos os complexos à 25°C pode ser considerada como termodinamicamente favorável em fase gasosa. As energias livres de todas as reações foram dominadas pelo termo entálpico (ΔH_r), porém os valores do termo entrópico ($-T\Delta S_r$) variaram entre -8,1 kcal/mol a 13,4 kcal/mol e representam de 36% a -35% do valor de ΔH_r .

18 Efeito do solvente:

Existem basicamente duas formas de inserir o efeito do solvente em simulações computacionais:

- Explicitamente: adicionando um aglomerado de moléculas de solvente em torno do fragmento estudado de forma que este esteja perfeitamente solvatado.
- Implicitamente: toda a informação sobre a estrutura explícita do solvente é perdida, o solvente é tratado como um contínuo dielétrico.

A forma mais realística de tratar o solvente é de maneira explícita, calculando as propriedades desejadas usando as médias estatísticas obtidas em várias configurações do sistema estudado. Porém, o tratamento explícito requer a utilização de uma grande quantidade de moléculas do solvente, geralmente limitando a descrição de tais moléculas aos métodos clássicos de campos de força que excluem ou tratam de forma imprecisa efeitos quânticos como a polarização que ocorre entre o soluto e as moléculas do solvente ⁸⁸. Já a aproximação do solvente por um contínuo dielétrico além de tratar da polarização soluto-solvente de maneira quântica elimina a necessidade de se utilizar um grande número de moléculas do solvente em várias configurações iniciais diferentes. Porém a aproximação do solvente contínuo possui várias limitações, as principais são:

- Não trata de forma precisa interações pontuais como ligações de hidrogênio entre o soluto e o solvente.
- Alguns termos necessários aos cálculos, como a cavidade eletrostática, não são observáveis e sua definição ideal é arbitrária, podendo variar de acordo com o tipo de molécula em análise ⁷⁴.

Uma forma de contornar o primeiro problema se dá pela adição de determinadas moléculas do solvente de forma explícita, como, por exemplo, a primeira camada de solvatação de cátions e ânions ⁸⁹. Quanto ao segundo problema, ainda não há uma forma de eliminá-lo, mas estudos de dinâmica molecular juntamente com a comparação de valores experimentais dos observáveis desejados em sistemas análogos aos estudados podem ajudar na validação da escolha desses termos ⁹⁰.

A interação eletrostática entre o soluto e o solvente implícito se dá de forma iterativa; as cargas do soluto polarizam o contínuo dielétrico, que em troca polariza o soluto, formando um campo de reação auto consistente (SCRF) que pode ser computado resolvendo a equação de Poisson. ⁷⁴

Porém a equação de Poisson tem solução analítica apenas para solutos pequenos como H₂O, NH₃, CH₄, cuja cavidade do solvente tem formato aproximadamente esférico ou elíptico. Moléculas maiores, como as bases nucleicas, necessitam de cavidades mais complexas que respeitem a forma do

soluto sem preencher espaços passíveis de serem ocupados pelo solvente evitando distorções na distribuição de cargas. Para resolver a equação de Poisson nesses casos faz-se necessário o uso de modelos numéricos aproximados, entre eles os mais utilizados são os PCM (*Polarizable Continuum Models*)⁹¹.

Os PCM transformam o problema de volume do contínuo dielétrico externo à cavidade do solvente em um problema de área, descrevendo o contínuo como cargas sobre a superfície da cavidade eletrostática definida em torno da molécula. Essa superfície é dividida em pequenos elementos de área, chamados *tesserae*, pelo algoritmo GePol (*Generating Polyhedra*) ou outro similar. Cargas pontuais capazes de descreverem aproximadamente as cargas de cada elemento de área são posicionadas no centro de cada *tessera*, simplificando a distribuição contínua de cargas sobre a superfície em cargas pontuais, conforme o método ASC (*Apparent Surface Charges*)⁹².

Para diminuir o problema causado por *tesserae* mal formados, muito pequenos ou de formatos oblongos, gerados nas reentrâncias da cavidade, utiliza-se o algoritmo de suavização desenvolvido por York e Karplus para o CPCM⁹³ e generalizado para os outros modelos PCM por Scalmani e Frisch⁹⁴ com o nome de CSC (*Continuous Surface Charge*). Esse algoritmo substitui as cargas pontuais dos *tesserae* situados nas reentrâncias das cavidades por funções gaussianas ponderadas pelas fração da área de cada *tesserae* internas às intersecções, diminuindo os problemas de descontinuidade que afetam a convergência dos cálculos SCRF, principalmente em procedimentos de otimização de geometria⁹⁰.

Existem pequenas variações em relação tanto aos modelos matemáticos para solucionar a equação de Poisson como também entre os parâmetros utilizados na definição da cavidade eletrostática do soluto.

18.1. Modelos PCM

IEF-PCM - Integral Equation Formalism PCM: é o modelo mais próximo de uma solução analítica para a equação de Poisson.⁹⁵ O formalismo matemático desse

método permite cálculos não só em líquidos isotrópicos, mas também em meios anisotrópicos como cristais líquidos e matriz sólida, além de soluções iônicas ⁹⁶.

C-PCM - *Conductor PCM*: em vez de tratar o solvente como um meio dielétrico, o trata como um meio condutor ⁹⁷. Apesar dessa aproximação levar a pequenos desvios ao simular solventes com baixa constante dielétrica, cálculos com solventes de alta constante dielétrica como a água dão valores muito similares aos do método IEF-PCM a um menor custo computacional ⁹⁸⁻¹⁰².

18.2. Definições da cavidade

Nos métodos IEF-PCM e CPCM as cavidades são construídas a partir de esferas de van der Waals sobrepostas e interligadas centradas nos átomos do soluto formando a superfície de van der Waals ⁹¹. Para se obter superfícies mais suaves, os raios de van der Waals são multiplicados por um fator de escala, geralmente entre 1,1 a 1,2.

Não existe consenso sobre qual o melhor valor para os raios dessas esferas, pois a cavidade eletrostática ideal pode variar com a carga da molécula e o ambiente químico de cada átomo. Os parâmetros mais utilizados para os raios de van der Waals são:

-BONDI ¹⁰³ Se baseia no banco de dados de estruturas cristalográficas de Cambridge^{91,104} e é o parâmetro com mais citações na literatura (mais de 10,000 citações) Por dar valores coerentes com dados experimentais ¹⁰⁵, seus valores são geralmente bem aceitos para uso geral ¹⁰². O fator de escala recomendado é de 1,17 até 1,20 ¹⁰¹.

-UFF (*Universal Force Field*) ¹⁰⁶: baseados no banco de dados do campo de força universal (UFF), sua maior vantagem é cobrir praticamente toda a tabela periódica incluindo raios de átomos que não possuem parâmetros Bondi. Como raios UFF são ligeiramente maiores ⁹⁴ o fator de escala recomendado é de 1,1.

-UATM (*United Atom Topology models*): desenvolvidos para cálculos de energia livre de hidratação. Nesses modelos os átomos de hidrogênio não possuem

esferas próprias, eles são incluídos nas esferas dos átomos em que estão ligados. Existem três variações que utilizam essa abordagem ¹⁰⁷:

- **UA0**: utiliza os raios UFF para átomos pesados. Tanto os parâmetros UA0 como os UFF, geralmente apresentam grandes desvios na descrição da energia de solvatação de íons ¹⁰⁸.
- **UAHF e UAKS**: em UAHF os raios foram parametrizados utilizando HF/6-31G(d) e em UAKS foram parametrizados utilizando o funcional PBE0/6-31G(d) ¹⁰⁷. Tais raios apresentam em média bons resultados para os cálculos de energias de solvatação ¹⁰⁸ quando combinados com métodos similares ao nível de cálculo para o qual foram parametrizados ¹⁰⁹.

18.3. Energia livre de solvatação

Utilizando o ciclo termodinâmico da Figura 19 o efeito do solvente pode ser adicionado como uma correção na energia livre da reação obtida na fase gasosa

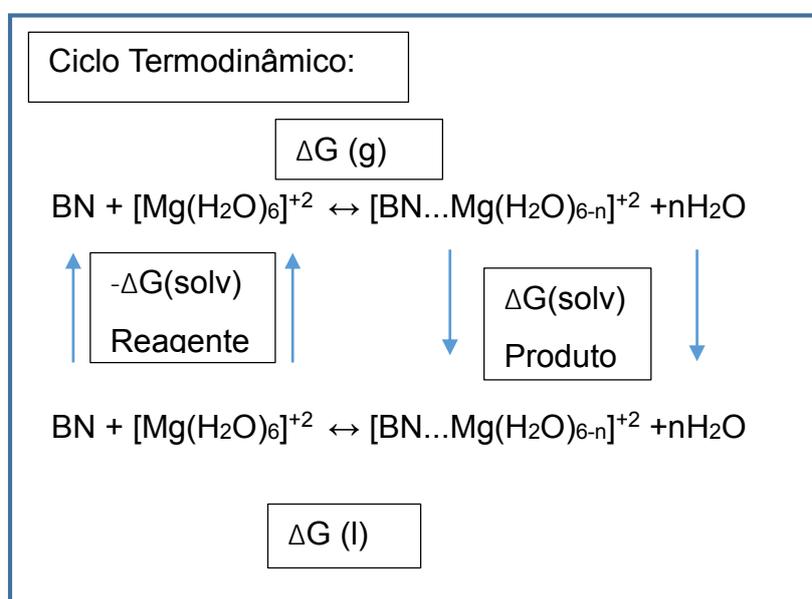


Figura 19 Ciclo termodinâmico das reações em análise BN = bases nucleicas e n=0,1,2.

$$\Delta G(aq) = \Delta G(gás) + \Delta \Delta G(solv) \quad [17]$$

Em que,

$$\Delta \Delta G(solv) = \Delta G(solv)Produtos - \Delta G(solv)Reagentes \quad [18]$$

A energia livre de solvatação $\Delta G(\text{solv})$ é a variação da energia livre necessária para trazer uma molécula do vácuo para dentro de um determinado solvente e pode ser escrita como ⁹²:

$$\Delta G_{\text{solv}} = W\left(\frac{M}{S}\right) + \Delta G_{\text{térmico}} \quad [19]$$

O primeiro termo da equação, $W(M/S)$, se refere ao trabalho de trazer uma molécula M inicialmente no vácuo para dentro do solvente S a zero Kelvin. O segundo termo se refere às variações nas contribuições térmicas (vibração, translação, rotação) e geralmente é mínima à temperatura e pressão ambiente para moléculas pequenas ¹⁰⁸. $W(M/S)$ pode ser dividido em quatro contribuições aproximadamente independentes entre si ⁹⁰, conforme a equação 20:

$$W\left(\frac{M}{S}\right) = \Delta E_{\text{ele}} + \Delta G_{\text{cav}} + \Delta G_{\text{dis}} + \Delta G_{\text{rep}} \quad [20]$$

ΔE_{ele} (eletrônico): diferença entre a energia eletrônica da molécula no solvente em relação a energia do vácuo, essa diferença se dá devido à polarização soluto-solvente e também à variação na energia eletrônica e nuclear gerada por mudanças na geometria do soluto ⁹¹.

ΔG_{cav} (cavitação): Energia necessária para criar um espaço vazio dentro do solvente, depende apenas de parâmetros geométricos do solvente e do soluto e da densidade do solvente, o PCM utiliza a formulação de Pierott estendida para cavidades de formato molecular por Claverie ⁹². Diferente dos outros termos a cavitação pode ser validada por meio de simulações computacionais em solventes explícitos ⁹⁰.

ΔG_{dis} (dispersão): termo mecânico-quântico relacionado à dispersão da nuvem eletrônica do soluto devido às interações com o solvente.

ΔG_{rep} (repulsão): termo positivo relacionado à troca eletrônica que ocorre em interações soluto-solvente à nível mecânico-quântico.

A dispersão e a repulsão são estimados pelos modelos PCM utilizando parâmetros do solvente e da superfície acessível ao solvente (SAS) conforme o formalismo de Floris e Tomasi ⁹².

18.3.1 Cálculo da energia de hidratação do Mg^{2+}

Para solutos neutros, as energias livres de solvatação variam entre 5 e -15 kcal/mol e podem ser obtidas experimentalmente com acurácia de $\pm 0,1$ kcal/mol, enquanto que solutos levemente carregados apresentam energias livres de solvatação maiores, entre -40 a -110 kcal/mol, e suas medidas experimentais podem ser obtidas com acurácia entre ± 2 a 5 kcal/mol ⁷⁴. Já as energias livres de solvatação, envolvendo espécies altamente carregadas, como o Mg^{2+} , são da ordem de centenas de kcal/mol, sendo que dados experimentais confiáveis para essas espécies são escassos ⁷⁴.

Nos modelos de solvente contínuo, devido as energias livres de solvatação de íons serem extremamente sensíveis ao tamanho e ao formato da cavidade eletrostática, tais espécies apresentam os maiores desvios em relação aos dados experimentais ^{74,108}.

Para definir o modelo de contínuo, o nível de cálculo e os parâmetros dos raios de van der Waals da cavidade que serão empregados nas reações químicas de interesse (equação 1), primeiramente foi feito um estudo comparativo entre diversas metodologias e os valores experimentais da energia de hidratação do Mg^{2+} .

Além de estudar várias definições de cavidade e nível de cálculo decidiu-se analisar a influência de se adicionar explicitamente a primeira camada de hidratação do Mg^{2+} . Para isso a energia de hidratação do Mg^{2+} (ΔG_{solv}) foi calculada de duas maneiras:

- i) Diretamente, adicionando o solvente implícito ao cátion:

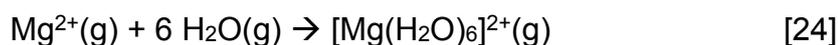


Por meio das equações 22 e 23:

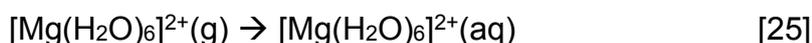
$$\Delta G_{\text{solv } 1} = E_{\text{ele}}Mg^{2+}(\text{aq}) - E_{\text{ele}}Mg^{2+}(\text{g}) \quad [22]$$

$$\Delta G_{\text{solv } 1}(\text{CDR}) = \Delta G_{\text{solv } 1} + \Delta G_{\text{cav}} + \Delta G_{\text{dis}} + \Delta G_{\text{rep}} \quad [23]$$

ii) Calculando a energia livre de formação da primeira camada de solvatação de maneira explícita ($\Delta G_{\text{r}}(\text{g})(\text{reação } 24)$), com o método CBS-QB3, no vácuo.



e em seguida adicionando o solvente implícito:



Por meio das equações 26 e 27:

$$\Delta G_{\text{solv } 2} = \Delta G_{\text{r}}(\text{g})(\text{reação } 24) + \Delta G_{\text{solv}}(\text{reação } 25) \quad [26]$$

$$\Delta G_{\text{solv } 2}(\text{CDR}) = \Delta G_{\text{r}}(\text{g}) + (\Delta G_{\text{solv}}(\text{reação } 25) + \Delta G_{\text{cav}} + \Delta G_{\text{dis}} + \Delta G_{\text{rep}}) \quad [27]$$

Diferentemente da contribuição eletrostática (E_{ele}), as energias de cavitação (ΔG_{cav}), dispersão (ΔG_{dis}), e repulsão (ΔG_{rep}) (CDR) são calculadas de modo empírico, em função da área da superfície acessível ao solvente, e não necessariamente fazem com que os valores das energias de solvatação sejam mais acurados, por isso as energia de hidratação do Mg^{2+} foram definidas tanto com essas energias (equações 22 e 26) como também sem elas (equações 23 e 27).

O cálculo da energia livre da formação da primeira camada de hidratação ($\Delta G_{\text{r}}(\text{g})(\text{reação } 24)$) foi feito utilizando o método CBS-QB3, sendo que as geometrias dos produto e reagentes foram previamente otimizadas em MP2/6-31+G(2d,p). As contribuições do solvente nas equações 22, 23, 26 e 27 foram calculadas por meio do método CPCM, utilizando a constante dielétrica da água ($\epsilon=78,35$) e empregando os raio de van der Waals, níveis de teoria e conjunto de

bases conforme descritos na Tabela 16. Variações nas geometrias e nas contribuições térmicas entre os cálculos feitos no vácuo e no solvente contínuo foram desconsideradas.

Tabela 16 Energia de hidratação do cátion metálico obtida considerando o cátion sem a primeira camada de solvatação no meio contínuo (método i) e considerando explicitamente a primeira camada de hidratação (método ii).

Raio e nível de cálculo	Método i (kcal/mol)		Método ii (kcal/mol)		ΔG_{solv} Literatura (kcal/mol)
	$\Delta G_{\text{solv1}}(\text{CDR})$	ΔG_{solv1}	$\Delta G_{\text{solv2}}(\text{CDR})$	ΔG_{solv2}	
BONDI B3LYP/6-31+G(2d,p)	-311,2	-315,8	-434,7	-446,4	Experimental
BONDI B3LYP/6-31G(d)	-311,2	-315,8	-432,9	-444,6	-455,5 ref. ¹¹⁰
BONDI HF/6-31+G(2d,p)	-311,2	-315,8	-434,2	-445,9	-441,2 ref. ¹¹¹
BONDI HF/6-31G(d)	-311,2	-315,8	-433,4	-445,1	Teórico
BONDI MP2/6-31+G(2d,p)	-311,2	-315,8	-435,8	-447,5	-433,3 ref. ¹¹¹
UAHF HF/6-31G(d)	-358,0	-361,7	-447,8	-452,8	-431,1 ref. ¹¹²
UAKS PBE0/6-31G(d)	-358,0	-361,7	-449,5	-454,5	

Comparando as diversas metodologias tanto em i como em ii é possível notar que o cálculo da energia de solvatação é mais sensível às variações nas definições dos raios de van der Waals do que às variações nos métodos ou conjuntos de bases.

Ao empregar o método i, observa-se que os valores obtidos estão bem abaixo do experimental. Uma maneira de corrigir este problema seria reparametrizar o raio de van der Waals do Mg^{2+} , variando o fator de escala até que ΔG_{solv} reproduzisse a energia de hidratação experimental. Entretanto, ao preencher a primeira camada de coordenação do metal explicitamente (método ii) as energias de hidratação se aproximaram dos valores experimentais. Isso ocorre porque as esferas das moléculas de água praticamente sobrepõem a esfera do pequeno cátion Mg^{2+} que passa a não apresentar uma participação significativa na formação da cavidade.

Esses resultados sugerem que a primeira camada de solvatação do cátion metálico não é reproduzida devidamente pelo modelo contínuo nos métodos testados. No entanto, incluir a primeira camada de solvatação explicitamente e simular o efeito das demais camadas de hidratação via modelo contínuo é razoável para a descrição da solvatação do cátion apresentando resultados comparáveis aos dados experimentais (Tabela 16).

Os métodos cujos resultados apresentaram menores desvios em relação aos dois valores experimentais descritos na literatura (-441,2 kcal/mol e -455,5 kcal/mol) foram o B3LYP/6-31G(d) com raios Bondi e o PBE0/6-31G(d) combinados com raios UAKS sem a adição das contribuições energéticas da cavitação, dispersão e repulsão (CDR).

18.4. Energia livre de reação no solvente

Equação 1 reação de formação dos complexos:



na qual BN se refere as bases nucleicas e $n = 0, 1$ e 2

Para se calcular a entalpia e a energia livre das reações de formação dos 16 complexos (Figura 15) em meio aquoso a 298,15 K, utilizou-se o ciclo termodinâmico da Figura 19 e as equações 17 e 18.

As energias de solvatação de todas as moléculas foram calculadas utilizando os dois métodos que deram os resultados mais próximos aos valores experimentais da energia de hidratação do Mg^{2+} :

1. **Método Bondi** – utilizou-se o método de contínuo CPCM, os parâmetros Bondi multiplicados por 1,2 na definição da cavidade eletrostática e o método B3LYP/6-31G(d) para calcular a diferença entre a energia eletrônica do vácuo e do solvente contínuo (termo ΔE_{ele} da equação 20).
2. **Método UAKS** – utilizou-se o método de contínuo CPCM, os parâmetros de cavidade UAKS multiplicados por 1,2 e o método PBE0/6-31G(d) para o cálculo do termo ΔE_{ele} .

Em ambos os métodos, as energias de cavitação, dispersão e repulsão não foram incluídas, devido as energias de hidratação do Mg^{2+} terem apresentados maiores erros ao contabilizar tais energias. Mudanças nas contribuições térmicas (vibracionais, translacionais e rotacionais) devido ao efeito do solvente

também não foram consideradas, pois para pequenas moléculas essas alterações são mínimas ¹⁰⁸.

As geometrias obtidas no vácuo em MP2/6-31+G(2d,p) foram mantidas fixas em todos os cálculos, desconsiderando possíveis alterações devido às interações entre o soluto e o solvente contínuo, tais alterações na geometria geralmente levam a erros desconsideráveis nas energias de solvatação de pequenas moléculas ⁸⁹, portanto, a energia de solvatação foi calculada apenas como a diferença entre as energias eletrônicas calculadas no vácuo e no contínuo ΔE_{solv} .

A contribuição do solvente foi calculada conforme a equação 28 e somada aos valores de ΔH_r e ΔG_r obtidos no estado gasoso para se obter $\Delta H_r(l)$ e $\Delta G_r(l)$

$$\Delta\Delta E_{solv} = \Delta E_{solv}(\text{produtos}) - \Delta E_{solv}(\text{reagentes}) \quad [28]$$

Em resumo os valores de $\Delta H_r(g)$ e $\Delta G_r(g)$ da fase gasosa foram calculados conforme a estratégia de menor custo computacional^{xxiii}:

1. Otimização das geometrias e cálculos das energias eletrônicas de reação em MP2/6-31+G(2d,p).
2. Reotimização das geometrias e cálculos dos parâmetros térmicos em B3LYP/6-311G(2d,d,p).

Esses valores foram então somados às médias aritméticas das contribuições do solvente ($\Delta\Delta E_{solv}$) calculadas com o método UAKS ($\Delta\Delta E_{solv1}$) e com o método Bondi ($\Delta\Delta E_{solv2}$) nas geometrias MP2/6-31+G(2d,p), conforme as equações 29 e 30:

$$\Delta H_r(l) = \Delta H_r(g) + \Delta\Delta E_{solv} \quad [29]$$

$$\Delta G_r(l) = \Delta G_r(g) + \Delta\Delta E_{solv} \quad [30]$$

^{xxiii} Definida na seção 17.

As entalpias (ΔH_r) e energias livres (ΔG_r) das 16 reações estudadas na fase gasosa e na fase aquosa ($\epsilon=78,35$), assim como os valores das contribuições do solvente calculados pelos métodos Bondi e UAKS estão na Tabela 17. Todos os cálculos foram realizados considerando a temperatura de 298,15 K.

Os erros esperados para os valores de $\Delta G_r(l)$ são de aproximadamente ± 5 kcal/mol, devido ao desvio de 3.4 kcal/mol do método CBS-QB3 em relação aos dados experimentais (Tabela 12 e Tabela 16) e aos desvios de aproximadamente 2 kcal/mol entre o CBS-QB3 e a estratégia de menor custo computacional adotada (Tabela 15).

Tabela 17 Energias livres e entalpias de reação em fase gasosa e em fase aquosa, adicionando as médias das correções do solvente ($\Delta\Delta E_{Med}$) obtidas pelo método CPCM com raios UAKS e nível PBE0/6-31G(d) e pelo CPCM, com raios Bondi à nível B3LYP/6-31G(d). Todos os valores estão em kcal/mol.

Reações	fase gasosa			contribuição do solvente			fase aquosa	
	$\Delta H_r(g)$	$-T\Delta S_r(g)$	$\Delta G_r(g)$	$\Delta\Delta E_{solv1}$ (UAKS)	$\Delta\Delta E_{solv2}$ (Bondi)	Média ($\Delta\Delta E_{Med}$)	$\Delta H_r(l)$	$\Delta G_r(l)$
4-GUA-O6/N7	-28,7	-7,9	-36,6	30,2	28,5	29,4	0,7	-7,2
4-CYT-O2/N3	-22,3	-8,1	-30,3	23,1	20,8	22,0	-0,3	-8,4
5-GUA-O6	-49,3	1,7	-47,6	41,0	39,5	40,3	-9,1	-7,3
5-CYT-O2	-47,0	1,9	-45,1	37,2	35,3	36,3	-10,8	-8,9
5-GUA-N7	-45,0	3,6	-41,4	37,7	35,2	36,5	-8,6	-5,0
5-CYT-N3	-35,1	4,3	-30,8	31,7	27,9	29,8	-5,2	-1,0
5-URA-O4	-19,0	-0,9	-20,0	18,5	17,6	18,1	-1,0	-1,9
5-THY-O4	-19,4	-0,3	-19,7	18,8	16,3	17,6	-1,9	-2,2
5-ADE-N1	-19,9	3,2	-16,7	21,7	19,4	20,6	0,7	3,8
5-ADE-N7	-18,8	3,3	-15,5	18,8	16,6	17,7	-1,2	2,3
6-GUA-O6/N7	-62,8	13,1	-49,8	50,9	48,7	49,8	-13,1	0,0
6-CYT-O2/N3	-59,2	13,4	-45,9	46,5	43,8	45,2	-14,1	-0,7
6-ADE-N1/N6	-38,4	12,2	-26,3	34,0	28,3	31,2	-7,3	5,0
6-URA-O4	-35,6	10,8	-24,7	31,4	29,3	30,4	-5,3	5,7
6-THY-O4	-36,3	11,8	-24,5	31,8	29,5	30,7	-5,7	6,2
6-ADE-N6/N7	-36,5	12,8	-23,7	31,9	26,1	29,0	-7,6	5,2

Pode-se observar, na Tabela 17, que as contribuições do solvente ($\Delta\Delta E_{solv}$) desfavorecem a formação dos complexos. Isso ocorre porque o solvente polar atenua o efeito das interações eletrostáticas, reduzindo assim as entalpias de complexação.

O valor de $\Delta\Delta E_{Med}$ foi de 17,3 kcal/mol a 49,8 kcal/mol e representou entre -76,3% a -103,3% do valor da entalpia de reação em fase gasosa ($\Delta H_r(g)$), conseqüentemente o valor de $\Delta H_r(l)$ diminui e a contribuição entrópica ($-T\Delta S_r$), que não foi alterada pelo solvente contínuo, passou a ser o termo dominante nas energias livres ($\Delta G_r(l)$) das reações de formação dos complexos bidentados e dos complexos em que o Mg^{2+} interage indiretamente. As únicas exceções foram para as reações dos complexos 6-Gua-O6/N7 e 6-Cyt-O2/N3, que apresentaram magnitudes semelhantes para $\Delta H_r(l)$ e $-T\Delta S_r$.

Dentro do erro estimado de ± 5 kcal/mol para as energia de $\Delta G_r(l)$, é possível afirmar apenas que a formação dos complexos bidentados 4-GuaO6/N7 e 4-Cyt-O2/N3 e dos complexos monodentados 5-Cyt-O2, 5-Gua-O6 e 5-Gua-N7 foi favorável termodinamicamente, enquanto que os 6-Ade-N1/N6, 6-Ura-O4, 6-Thy-O4 e 6-Ade-N6/N7 não devem ser formados em meio aquoso. A formação dos outros complexos não pôde ser definida, porém mesmo se a formação de tais complexos for possível eles serão mais lábeis do que os complexos formados por interações diretas nos sítios da citosina e da guanina.

18.4.1 *Comparação entre $\Delta G_r(l)$ obtidos por diferentes métodos*

As energias livres de reação em fase líquida ($\Delta G_r(l)$), calculadas combinando as contribuições do solvente $\Delta\Delta E_{solv}$ (UAKS) e $\Delta\Delta E_{solv}$ (Bondi) com as energias livres do estado gasoso ($\Delta G_r(g)$), obtidas tanto pelo método CBS-QB3 como pela estratégia de menor custo computacional, foram agrupadas segundo o modo de coordenação entre o Mg^{2+} e as bases nucleicas, conforme mostra a Figura 20.

As diferenças no perfil das curvas de energia livre de reação obtidas por $\Delta G_r + \Delta\Delta E_{solv}$ (UAKS) em relação ao perfil das energias $\Delta G_r + \Delta\Delta E_{solv}$ (Bondi) mostrados na Figura 20 ocorrem possivelmente devido ao raios UAKS não incluírem esferas explícitas para os átomos de hidrogênio, tal aproximação causa distorções na cavidade eletrostática que podem ser significativas nas descrições de ligações de hidrogênio.

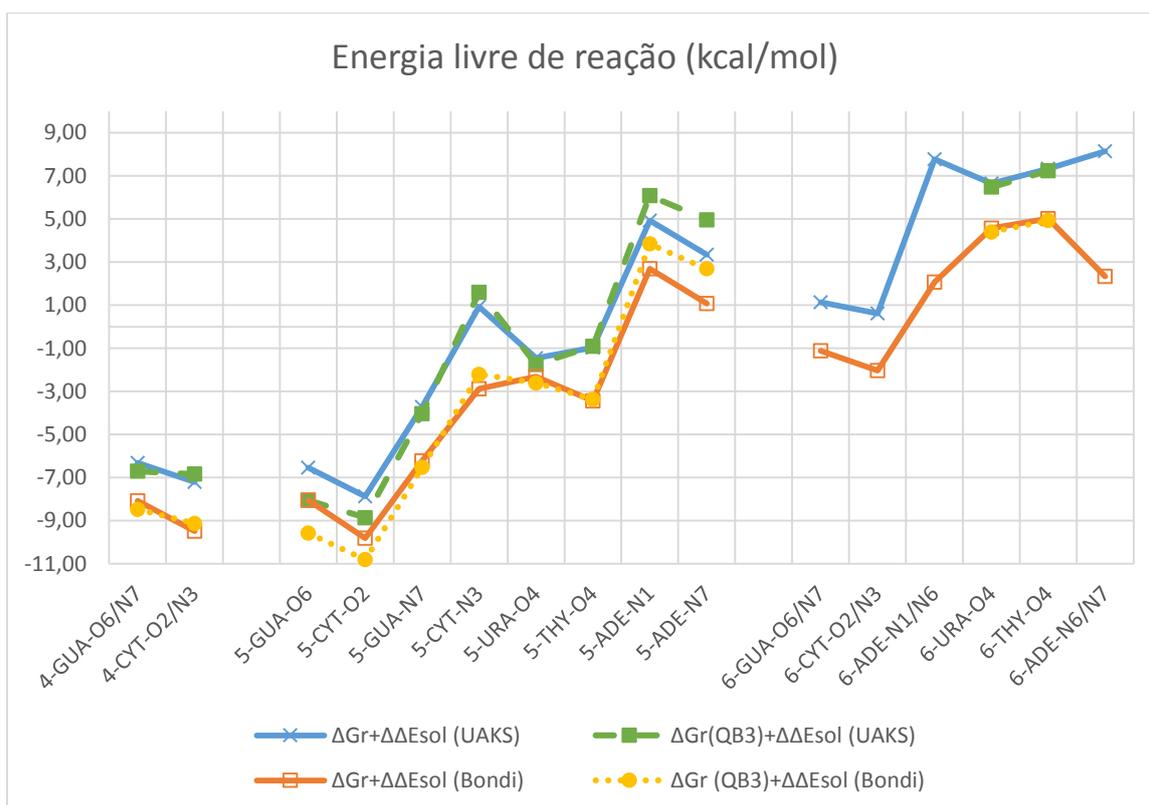


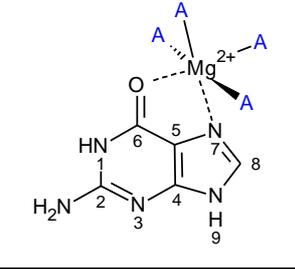
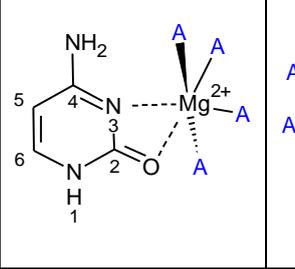
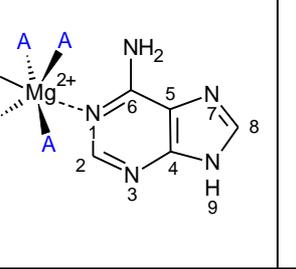
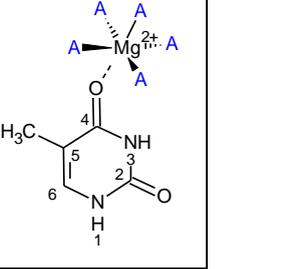
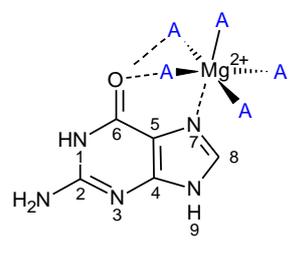
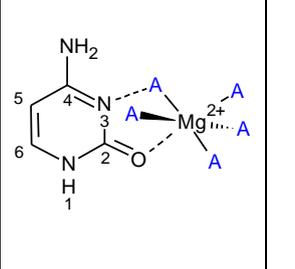
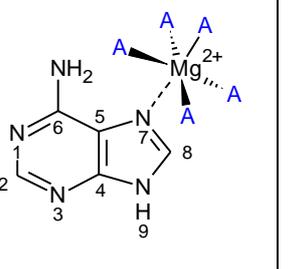
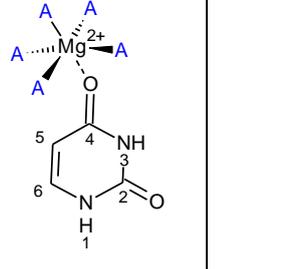
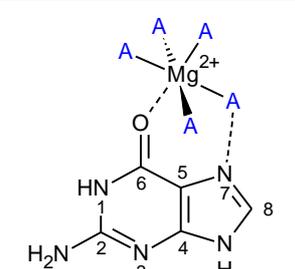
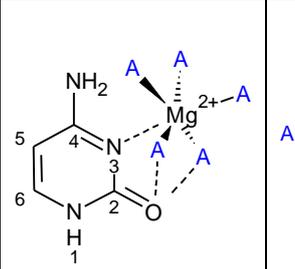
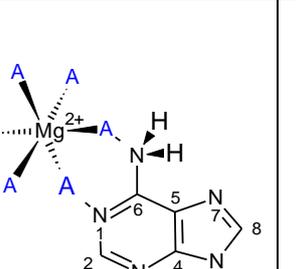
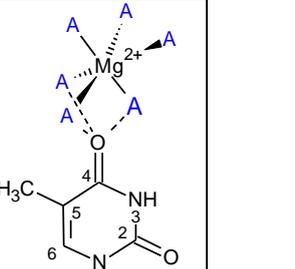
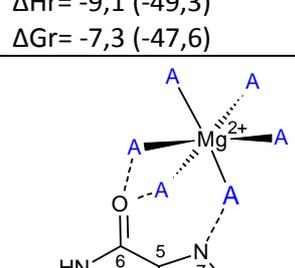
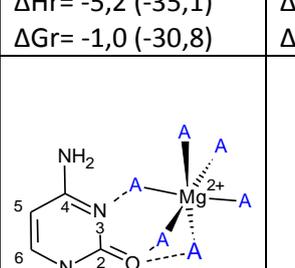
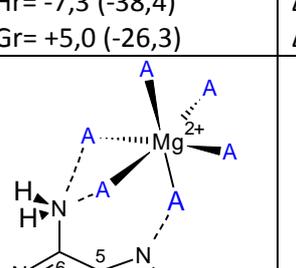
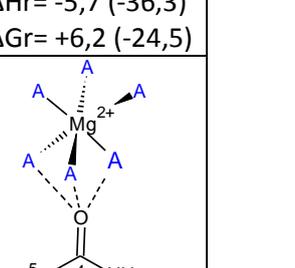
Figura 20 Energias livres das reações de formação dos 16 complexos em meio aquoso a 298,15K. As contribuições referentes ao estado gasoso foram obtidas pelos métodos CBS-QB3 ($\Delta Gr(QB3)$) e pela estratégia de menor custo computacional (ΔGr) definida na seção 16.1. O efeito do solvente foi calculado utilizando o modelo CPCM com raios UAKS e energia PBE0/6-31G(d) ($\Delta\Delta Esol$ (UAKS)) e raios Bondi com B3LYP/6-31G(d) ($\Delta\Delta Esol$ (Bondi)).

Os perfis das energias livres de reação ($\Delta Gr + \Delta\Delta E(UAKS)$ e $\Delta Gr + \Delta\Delta E(Bondi)$), obtidos combinando a estratégia de menor custo computacional para o cálculo de ΔGr^{xxiv} aos métodos de solvente, foram similares aos do CBS-QB3 ($\Delta Gr(QB3) + \Delta\Delta E(UAKS)$ e $\Delta Gr(QB3) + \Delta\Delta E(Bondi)$) apresentando desvios máximos inferiores a 2 kcal/mol a um custo computacional consideravelmente menor.

A estratégia de menor custo computacional para o cálculo das energias livres em fase gasosa combinada com raios Bondi provavelmente apresentam resultados mais acurados, pois ao incluir as esferas dos hidrogênios explicitamente os parâmetros Bondi reduzem possíveis distorções na cavidade eletrostática. Porém como não existem dados experimentais ou computacionais que validem estes resultados decidiu-se utilizar a média entre os valores $\Delta\Delta E(UAKS)$ e $\Delta\Delta E(Bondi)$.

^{xxiv} Definido na seção 17.

Tabela 18 Entalpias e energias livres da reação de formação dos complexos no solvente contínuo (a contribuição do solvente considerada foi a média dos valores obtidos com UAKS e Bond). Os valores entre parênteses se referem às energias em fase gasosa. Todos os valores estão em kcal/mol.

			
4-Gua-O6/N7 $\Delta H_r = +0,7$ (-28,7) $\Delta G_r = -7,2$ (-36,6)	4-Cyt-O2/N3 $\Delta H_r = -0,3$ (-22,3) $\Delta G_r = -8,4$ (-30,3)	5-Ade-N1 $\Delta H_r = +0,7$ (-19,9) $\Delta G_r = +3,8$ (-16,7)	5-Thy-O4 $\Delta H_r = -1,9$ (-19,4) $\Delta G_r = -2,2$ (-19,7)
			
5-Gua-N7 $\Delta H_r = -8,6$ (-45,0) $\Delta G_r = -5,0$ (-41,4)	5-Cyt-O2 $\Delta H_r = -10,8$ (-47,0) $\Delta G_r = -8,9$ (-45,1)	5-Ade-N7 $\Delta H_r = -1,2$ (-18,8) $\Delta G_r = +2,3$ (-15,5)	5-Ura-O4 $\Delta H_r = -1,0$ (-19,0) $\Delta G_r = -1,9$ (-20,0)
			
5-Gua-O6 $\Delta H_r = -9,1$ (-49,3) $\Delta G_r = -7,3$ (-47,6)	5-Cyt-N3 $\Delta H_r = -5,2$ (-35,1) $\Delta G_r = -1,0$ (-30,8)	6-Ade-N1/N6 $\Delta H_r = -7,3$ (-38,4) $\Delta G_r = +5,0$ (-26,3)	6-Thy-O4 $\Delta H_r = -5,7$ (-36,3) $\Delta G_r = +6,2$ (-24,5)
			
6-Gua-O6/N7 $\Delta H_r = -13,1$ (-62,8) $\Delta G_r = 0,0$ (-49,8)	6-Cyt-O2/N3 $\Delta H_r = -14,1$ (-59,2) $\Delta G_r = -0,7$ (-45,9)	5-Ade-N6/N7 $\Delta H_r = -7,6$ (-36,5) $\Delta G_r = +5,2$ (-23,7)	6-Ura-O4 $\Delta H_r = -5,3$ (-35,6) $\Delta G_r = +5,7$ (-24,7)

19 Correlações com dados experimentais

A subunidade ribossomal da bactéria *Haloarcula Marismortui*, composta por 3045 nucleotídeos, foi analisada a uma resolução de 2,4 Å para definir os sítios de ligação de íons metálicos^{113,114}, sendo identificados 116 íons Mg^{2+} nessa estrutura dos quais somente 9 interagem com o RNA apenas indiretamente por meio das águas coordenadas ao cátion. Isso indica que o magnésio(II) hexahidratado dificilmente forma interações indiretas com o RNA fortes o suficiente para ocupar sítios de coordenação bem definidos. Essa observação concorda com as energias livres de formação dos complexos $[Mg(H_2O)_6\text{-base nucleica}]^{2+}$ obtidas no presente trabalho serem próximas a zero ou positivas, sugerindo que, em meio aquoso, as interações indiretas entre as águas da primeira camada de hidratação do Mg^{2+} e as bases nucleicas são fracas e lábeis.

A partir de medidas de óptica não linear, Josef G. Holland e colaboradores²⁴ quantificaram os cátions Mg^{2+} que se ligavam às bases nucleicas de oligômeros formados por 21 nucleotídeos de adenina, timina, citosina ou guanina. Seus estudos mostraram que enquanto apenas três Mg^{2+} se ligaram nas bases nucleicas dos oligômeros de adenina e timina, nas bases nucleicas citosina e guanina, onze Mg^{2+} se ligaram aos seus respectivos oligômeros.

O maior número de cátions ligados às bases nucleicas citosina e guanina deve estar relacionado à possibilidade do magnésio(II) pentahidratado formar complexos mais estáveis com essas bases nucleicas. No entanto, os estudos de Josef G. Holland estimaram energias livres de ligação entre o Mg^{2+} e as bases nucleicas mais fortes para a timina do que para a citosina, tais energias não podem ser comparadas diretamente com as energias livres descritas no presente trabalho, pois os valores estimados por Holland não foram calculados para a formação de complexos específicos.

20 Conclusões e perspectivas

Quanto à metodologia:

A estratégia capaz de descrever a energia livre de reação em meio aquoso com melhor relação entre custo computacional e acurácia foi:

- 1) **Otimização de geometrias em MP2/6-31+G(2d,p) quando possível:** o método HF/6-31+G(2d,p) subestimou as distâncias das ligações de hidrogênio, porém as variações entre as geometrias obtidas por esses dois métodos influenciaram pouco nos cálculos da energia de interação entre o magnésio(II) pentahidratado e as bases nucleicas.
- 2) **Energia eletrônica em MP2/6-31+G(2d,p):** esse nível de cálculo é suficientemente acurado para o estudo das energias eletrônicas dos complexos aqui descritos, pois reproduz qualitativamente e quantitativamente bem, apresentando apenas pequenos desvios em relação aos valores das energias eletrônicas de reação do CBS-QB3 e às energias de interação calculadas com MP2/Jul-cc-pVTZ.
- 3) **Cálculo das contribuições térmicas em B3LYP/6-311G(2d,d,p):** juntamente com as energias eletrônicas MP2/6-31+G(2d,p). Esse método reproduz bem as energias obtidas por CBS-QB3. Para sistemas maiores, os parâmetros termoquímicos calculados em HF/6-31G(d) representam uma opção de menor custo computacional e resultados próximos aos do B3LYP/6-31G(2d,d,p).
- 4) **Cálculo da energia de solvatação utilizando o modelo CPCM:** devido à definição da cavidade eletrostática ser arbitrária, a energia de solvatação calculada por meio do solvente contínuo é a etapa mais passível de apresentar erros significativos, tanto quantitativos quanto qualitativos, nos cálculos das energias livres de reação. Os parâmetros UAKS e Bondi deram os melhores resultados para a energia de hidratação do Mg^{2+} .

Quanto à interação entre o Mg^{2+} e as bases nucleicas

O cátion metálico Mg^{2+} se complexa mais fortemente com os sítios de maior densidade eletrônica da guanina e da citosina independentemente do modo de interação. Enquanto na guanina o Mg^{2+} interage mais fortemente com o sítio N7 em relação ao O6, na citosina o Mg^{2+} interage mais fortemente com o sítio O2 em relação ao N3 (Figura 16). No entanto, ao estudar a energia eletrônica da reação de formação dos complexos, nota-se que o Mg^{2+} se complexa preferencialmente nos sítios de oxigênio (Figura 18).

Apesar das energias livres de formação dos sistemas $[Mg(H_2O)_n \text{ base nucleica}]^{2+}$, em fase gasosa, serem dominadas pela contribuição eletrônica, a contribuição entrópica é significativa e não pode ser desconsiderada como proposto por Iván Solt ³⁴, principalmente ao se incluir o efeito do solvente, pois em algumas reações o termo entrópico chega a ser o termo dominante na fase aquosa (Tabela 18).

Devido à lenta troca de água da primeira esfera de hidratação do cátion Mg^{2+} e aos complexos bidentados terem aproximadamente a mesma energia livre de formação que os complexos monodentados, estes últimos devem ocorrer com maior frequência. Os complexos que interagem indiretamente através das águas de hidratação com as bases nucleicas apresentaram as menores energias livres de formação e provavelmente não ocupam sítios de coordenação bem definidos no DNA/RNA, formando apenas ligações lábeis.

Perspectivas:

Estudos apontam para a possibilidade do cátion metálico Mg^{2+} ser capaz de estabilizar estruturas tautoméricas das bases nucleicas raras de serem formadas sem o cátion ¹¹⁵. Porém, estes estudos não incluíram nem mesmo a primeira camada de solvatação do Mg^{2+} . A aplicação do método descrito nesta dissertação em tais sistemas poderia trazer resultados de maior confiabilidade sobre a possível formação desses tautômeros e quais são os mais estáveis.

A estratégia de cálculo definida nesta dissertação também pode ser estendida para o estudo de outros cátions metálicos, sistemas maiores que incluam, por exemplo, não só a base nucleica, mas também o açúcar e o grupo fosfato ligados a ela. Esses resultados poderão então ser agrupados em um banco de dados contendo as principais interações entre cátions metálicos e os fragmentos do DNA/RNA com a finalidade de validar métodos DFT ou parametrizar campos de forças.

Atualmente já existem métodos como RI-MP2 capazes de poupar consideravelmente o tempo computacional e permitir o uso de funções de bases maiores ou sistemas maiores ¹¹⁶. Assim como esquemas de extrapolação para o conjunto de bases completo (CBS) a custos computacionais viáveis para os complexos descritos nesta dissertação ^{63,117}. A aplicabilidade desses métodos nos cálculos da interação entre cátions metálicos e fragmentos do DNA/RNA também devem ser investigadas.

Por fim, deve-se investigar alternativas menos arbitrárias para a definição da cavidade eletrostática empregada pelos solventes contínuos PCM, ou então empregar métodos contínuos parametrizados para reproduzir energias de solvatação experimentais como o SMD ¹¹⁸, no entanto tais métodos ainda não possuem parâmetros específicos para o Mg^{2+} .

REFERÊNCIAS

1. Cech, T. R. ;*Biosci. Rep.* **1989**, 24, 362.
2. Stark, B. C.; Kole, R.; Bowman, E. J.; Altman, S. ;*Proc. Natl. Acad. Sci. USA* **1978**, 75, 3717.
3. Kruger, K.; Grabowski, P. J.; Zaug, a J.; Sands, J.; Gottschling, D. E.; Cech, T. R. ;*Cell* **1982**, 31, 147.
4. Lewin, R. ;*Sci. - Res. News* **1984**, 223, 266.
5. Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. ;*Cell* **1983**, 35, 849.
6. Altman, S. ;*Biosci. Rep.* **1990**, 10, 317.
7. Walter, F.; Westhof, E. Catalytic RNA. *Encycl. Life Sci.* **2002**, 1–12.
8. Alberts, B.; Johnson, A.; Lewis, J.; Raff, M.; Roberts, K.; Walter., P. *Molecular Biology of the Cell*; 4th ed.; New York, Garland Science, 2002.
9. Khan, A. U. ;*Clin. Chim. Acta.* **2006**, 367, 20.
10. Mulhbacher, J.; St-Pierre, P.; Lafontaine, D. a ;*Curr. Opin. Pharmacol.* **2010**, 10, 551.
11. Mitsuyasu, R. T.; Merigan, T. C.; Carr, A.; Zack, J. a; Winters, M. a; Workman, C.; Bloch, M.; Lalezari, J.; Becker, S.; Thornton, L.; Akil, B.; Khanlou, H.; Finlayson, R.; McFarlane, R.; Smith, D. E.; Garsia, R.; Ma, D.; Law, M.; Murray, J. M.; von Kalle, C.; Ely, J. a; Patino, S. M.; Knop, A. E.; Wong, P.; Todd, A. V; Haughton, M.; Fuery, C.; Macpherson, J. L.; Symonds, G. P.; Evans, L. a; Pond, S. M.; Cooper, D. a ;*Nat. Med.* **2009**, 15, 285.
12. Weng, D. E.; Masci, P. a; Radka, S. F.; Jackson, T. E.; Weiss, P. a; Ganapathi, R.; Elson, P. J.; Capra, W. B.; Parker, V. P.; Lockridge, J. a; Cowens, J. W.; Usman, N.; Borden, E. C. ;*Mol. Cancer Ther.* **2005**, 4, 948.
13. Pyle, a M.; McSwiggen, J. a; Cech, T. R. ;*Proc. Natl. Acad. Sci. U. S. A.* **1990**, 87, 8187.
14. Steitz, T. a; Steitz, J. a ;*Proc. Natl. Acad. Sci. U. S. A.* **1993**, 90, 6498.
15. Stahley, M. R.; Strobel, S. a ;*Science* **2005**, 309, 1587.
16. Murray, J. B.; Seyhan, a a; Walter, N. G.; Burke, J. M.; Scott, W. G. ;*Chem. Biol.* **1998**, 5, 587.

17. Scott, W. G. ;*Curr. Opin. Struct. Biol.* **2007**, *17*, 280.
18. Hanna, R.; Doudna, J. a ;*Curr. Opin. Chem. Biol.* **2000**, *4*, 166.
19. Schnabl, J.; Sigel, R. K. O. ;*Curr. Opin. Chem. Biol.* **2010**, *14*, 269.
20. Johnson-Buck, A. E.; McDowell, S. E.; Walter, N. G. In *Metal Ions in Life Science, Volume 9*; Sigel, A.; Sigel, H.; Sigel, R. K. O., Eds.; Royal Society of Chemistry, 2011; Vol. 9, p. 175.
21. Anastassopoulou, J. ;*J. Mol. Struct.* **2003**, *653*, 19.
22. Pyle, A. M. ;*J. Biol. Inorg. Chem.* **2002**, *7*, 679.
23. Glasner, M. E.; Bergman, N. H.; Bartel, D. P. ;*Biochemistry* **2002**, *41*, 8103.
24. Holland, J.; Malin, J.; Jordan, D. S.; Geiger, F. M. ;*J. Am. Chem. Soc* **2011**, *133*, 2567.
25. Draper, D. E. ;*Rna* **2004**, *10*, 335.
26. Leipply, D.; Draper, D. E. ;*J. Am. Chem. Soc.* **2011**, *133*, 13397.
27. Feig, ndrew L.; Uhlenbeck, O. C. In *book*; 1999.
28. Petrov, A. S.; Lamm, G.; Pack, G. R. ;*Biopolymers* **2005**, *77*, 137.
29. Petrov, A. S.; Bowman, J. C.; Harvey, S. C.; Williams, L. D. ;*RNA* **2011**, *17*, 291.
30. Ditzler, M. a; Otyepka, M.; Sponer, J.; Walter, N. G. ;*Acc. Chem. Res.* **2010**, *43*, 40.
31. Petrov, A.; Pack, G.; Lamm, G. ;*J. Phys. Chem. ...* **2004**, 6072.
32. Petrov, A. S.; Lamm, G.; Pack, G. R. ;*J. Phys. Chem. B* **2002**, *106*, 3294.
33. Petrov, A. S.; Funseth-Smotzer, J.; Pack, G. R. ;*Int. J. Quantum Chem.* **2005**, *102*, 645.
34. Solt, I.; Simon, I.; Császár, A. G.; Fuxreiter, M. ;*J. Phys. Chem. B* **2007**, *111*, 6272.
35. Maria Aparecida da Silva Prado Estudo Teórico das Interações Mg 2 + e Bases Nucleotídicas, UnB: Brasília-DF, **2006**.
36. Prado, M. a. S.; Garcia, E.; Martins, J. B. L. ;*Chem. Phys. Lett.* **2006**, *418*, 264.
37. Rulisek, L.; Sponer, J. ;*J. Phys. Chem. B* **2003**, *107*, 1913.

38. Sponer, J.; Burda, J. V.; Sabat, M.; Leszczynski, J.; Hobza, P. ;*J. Phys. Chem. A* **1998**, *102*, 5951.
39. Russo, N.; Toscano, M.; Grand, A. ;*J. Phys. Chem. A* **2003**, *107*, 11533.
40. Burda, J. V.; Sponer, J.; Leszczynski, J.; Hobza, P. ;*J. Phys. Chem. B* **1997**, *101*, 9670.
41. Muñoz, J.; Sponer, J.; Hobza, P.; Orozco, M.; Luque, F. J. ;*J. Phys. Chem. B* **2001**, *105*, 6051.
42. Solt, I.; Simon, I.; Császár, A. G.; Fuxreiter, M. ;*J. Phys. Chem. B* **2007**, *111*, 6272.
43. Burda, J. V.; Leszczynski, J. In *book*; **2006**; pp. 389–410.
44. Hobza, P.; Leszczynski, J. ;*Theor. Comput. Chem.* **1999**, *8*, 85.
45. Sponer, J. P. J. P. H. In *Computational Studies of RNA and DNA Volume 2 Challenges and Advances in Computational Chemistry and Physics*; Lankaés, J. éSponer; J. Š. F., Ed.; Springer, **2006**; pp. 343–388.
46. Helgaker, T.; Ruden, T. a.; Jørgensen, P.; Olsen, J.; Klopper, W. ;*J. Phys. Org. Chem.* **2004**, *17*, 913.
47. Raghavachari, K.; Anderson, J. ;*J. Phys. Chem.* **1996**, *3654*, 12960.
48. Davidson, E. R.; Feller, D. ;*Chem. Rev.* **1986**, *86*, 681.
49. Bowman, J. C.; Lenz, T. K.; Hud, N. V; Williams, L. D. ;*Curr. Opin. Struct. Biol.* **2012**, *22*, 262.
50. Dunning, T. ;*J. Phys. Chem. A* **2000**, 9062.
51. Møller, C.; Plesset, M. ;*Phys. Rev.* **1934**, *46*, 618.
52. Dunning, T. H. ;*J. Chem. Phys.* **1989**, *90*, 1007.
53. Chalasinski, G.; Szczesniak, M. ;*Chem. Rev.* **1994**, 1723.
54. Bernardi, F.; Boys, S. F. *Mol. Phys.* **1970**, *19*, 553–566.
55. Van Duijneveldt, F. B.; van Duijneveldt-van de Rijdt, J. G. C. M.; van Lenthe, J. H. ;*Chem. Rev.* **1994**, *94*, 1873.
56. Kim, K. S.; Tarakeshwar, P.; Lee, J. Y. ;*Chem. Rev.* **2000**, *100*, 4145.
57. Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications, **1996**.

58. Woon, D. E.; Dunning, T. H. ;*J. Chem. Phys.* **1995**, *103*, 4572.
59. Estarellas, C.; Lucas, X.; Frontera, A.; Quiñero, D.; Deyà, P. M. ;*Chem. Phys. Lett.* **2010**, *489*, 254.
60. García, J. M. In *Introduction of Advanced Topics of Computational Chemistry*; Habana, E. de la U. de La, Ed.; Havana, **2003**; pp. 41–70.
61. Hehre, W. J. ;*J. Chem. Phys.* **1972**, *56*, 2257.
62. McLean, a. D.; Chandler, G. S. ;*J. Chem. Phys.* **1980**, *72*, 5639.
63. Park, Y.; Lee, J. ;*Bull. Chem. Soc.* **2007**, *28*, 386.
64. Kendall, R. a.; Dunning, T. H.; Harrison, R. J. ;*J. Chem. Phys.* **1992**, *96*, 6796.
65. Papajak, E.; Zheng, J.; Xu, X.; Leverentz, H. R.; Truhlar, D. G. ;*J. Chem. Theory Comput.* **2011**, *7*, 3027.
66. Kaupp, M.; Schleyer, P. R. ;*J. Phys. Chem* **1992**, *96*, 7316.
67. Glendening, E. ;*J. Phys. Chem.* **1996**, 4790.
68. Klobukowski, M. ;*Can. J. Chem.* **1992**, *70*, 589.
69. Schlegel, H. ;*J. Comput. Chem.* **1982**, *3*, 214.
70. Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. a. ;*J. Am. Chem. Soc.* **1995**, *117*, 5179.
71. Duan, Y.; Wu, C.; Chowdhury, S. ;*J. Comput. Chem.* **2003**, *24*, 1999.
72. Petrov, A. S.; Pack, G. R.; Lamm, G. ;*J. Phys. Chem. B* **2004**, *108*, 6072.
73. Dahlke, E. E.; Truhlar, D. G. ;*J. Phys. Chem. B* **2006**, *110*, 10595.
74. Cramer, C. J. *Essentials of Computational Chemistry: Theories and Models*; 2nd ed.; Wiley, **2005**.
75. Breneman, C. M.; Wiberg, K. B. ;*J. Comput. Chem.* **1990**, *11*, 361.
76. Zhang, W. E. I.; Yang, R.; Cieplak, P.; Luo, R. A. Y.; Lee, T.; Caldwell, J.; Wang, J.; Kollman, P. **2003**, 67168.
77. A. Bondi ;*J. Phys. Chem.* **1965**, *68*, 441.
78. Szefczyk, B.; Sokalski, W. A.; Leszczynski, J. ;*J. Chem. Phys.* **2002**, *117*, 6952.

79. Ikeda, T.; Boero, M.; Terakura, K. ;*J. Chem. Phys.* **2007**, *127*, 074503.
80. Pasternak, K.; Kocot, J.; Horecka, A. ;*J. Elementol* **2010**, *15*, 601.
81. Ochterski, J. ;*Gaussian Inc, Pittsburgh, PA* **2000**.
82. Montgomery, J. a.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. a. ;*J. Chem. Phys.* **1999**, *110*, 2822.
83. Montgomery, J. a.; Frisch, M. J.; Ochterski, J. W.; Petersson, G. a. ;*J. Chem. Phys.* **2000**, *112*, 6532.
84. Ochterski, J. W.; Petersson, G. a.; Montgomery, J. a. ;*J. Chem. Phys.* **1996**, *104*, 2598.
85. Montgomery, J. a.; Ochterski, J. W.; Petersson, G. a. ;*J. Chem. Phys.* **1994**, *101*, 5900.
86. Casanovas, R.; Frau, J.; Salva, A.; Donoso, J.; Muñoz, F. ;*Int. J. Quantum Chem.* **2010**, *110*, 323.
87. Peschke, M.; Blades, A. ;*J. Phys.* **1998**, 9978.
88. Cramer, C. J.; Truhlar, D. G. ;*Chem. Rev.* **1999**, *99*, 2161.
89. Ho, J.; Coote, M. L. ;*Theor. Chem. Acc.* **2009**, *125*, 3.
90. Mennucci, B.; Cammi, R. *Continuum solvation models in chemical physics: from theory to applications*; Mennucci, B.; Cammi, R., Eds.; John Wiley & Sons Ltd, **2007**.
91. Tomasi, J.; Mennucci, B.; Cammi, R. ;*Chem. Rev.* **2005**, *105*, 2999.
92. Tomasi, J.; Persico, M. ;*Chem. Rev.* **1994**, *94*, 2027.
93. York, D. M.; Karplus, M.; Biophysique, L. D. C.; Bel, I. Le; Pasteur, L.; Pascal, B. **1999**, 11060.
94. Scalmani, G.; Frisch, M. J. ;*J. Chem. Phys.* **2010**, *132*, 114110.
95. Cancès, E.; Mennucci, B. ;*J. Chem. Phys.* **2001**, *114*, 4744.
96. Mennucci, B.; Tomasi, J.; Industriale, C.; Chimica, D.; li, F. In *Introduction to Advanced Topic of Computational Chemistry*, **2003**; pp. 185–217.
97. Cossi, M.; Rega, N.; Scalmani, G.; Barone, V.; Chimica, D.; li, F.; Angelo, C. M. S. ;*J. Comput. Chem.* **2003**.
98. Tomasi, J.; Mennucci, B.; Cammi, R. ;*Chem. Rev.* **2005**, *105*, 2999.

99. Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. ;*J. Chem. Phys.* **2001**, *114*, 5691.
100. Cossi, M.; Rega, N. ;*J. Comput. Chem.* **2003**, *24*, 669.
101. Marenich, A. ;*J. Chem. Theory Comput.* **2008**, 877.
102. Klamt, A.; Eckert, F. ;*Fluid Phase Equilib.* **2000**, *172*, 43.
103. Bondi, A. ;*J. Phys. Chem.* **1964**, *68*, 441.
104. Rowland, R. S.; Taylor, R. ;*J. Phys. Chem.* **1996**, *3654*, 7384.
105. Marenich, A. V.; Olson, R. M.; Kelly, C. P.; Cramer, C. J.; Truhlar, D. G. ;*J. Chem. Theory Comput.* **2007**, *3*, 2011.
106. Rappe, A.; Casewit, C.; Colwell, W. A.; III, G.; Skiff, W. M. ;*J. Am. Chem. Soc.* **1992**, *2*, 10024.
107. Barone, V.; Cossi, M.; Tomasi, J. ;*J. Chem. Phys.* **1997**, *107*, 3210.
108. Takano, Y.; Houk, K. ;*J. Chem. Theory Comput.* **2005**, *1*, 70.
109. Ho, J.; Klamt, A.; Coote, M. L. ;*J. Phys. Chem. A* **2010**, *114*, 13442.
110. Larentzos, J. P.; Criscenti, L. J. ;*J. Phys. Chem. B* **2008**, *112*, 14243.
111. Asthagiri, D.; Pratt, L. R.; Paulaitis, M. E.; Rempe, S. B. ;*J. Am. Chem. Soc* **2004**, *126*, 1285.
112. Jiao, D.; King, C.; Grossfield, A. ;*J. Phys. Chem. B* **2006**.
113. Sigel, R.; Pyle, A. ;*Chem. Rev.* **2007**, *107*, 97.
114. Klein, D. J.; Moore, P. B.; Steitz, T. a ;*J. Mol. Biol.* **2004**, *340*, 141.
115. Kabelác, M.; Hobza, P. ;*J. Phys. Chem. B* **2006**, *110*, 14515.
116. Weigend, F.; Patzelt, H. ;*Chem. Phys. Lett.* **1998**.
117. Paizs, B.; Salvador, P.; Császár, A. G. ;*J. Comput. Chem.* **2001**, *22*, 196.
118. Marenich, A. V; Cramer, C. J.; Truhlar, D. G. ;*J. Phys. Chem. B* **2009**, *113*, 6378.

ANEXO 1 - Estudo sistemático dos conjuntos de bases

Para definir um conjunto de bases capaz de descrever as interações entre o Magnésio (II) pentahidratado e as bases nucleicas a nível MP2, levando em consideração o custo computacional e a acurácia dos resultados, foram realizados estudos sistemáticos do efeito da adição de funções difusas e de polarização nos conjuntos de bases de Pople e Dunning.

Os parâmetros comparativos utilizados nesse estudo foram a magnitude do BSSE e os desvios da energia de interação (ΔE e ΔE_{cp} definidos no capítulo 2 seção 6.1) em relação aos valores do maior conjunto de bases utilizado, o aug-cc-pVTZ. Em todos os cálculos utilizou-se as geometrias dos complexos obtidas pelo método MP2(FC)/aug-cc-pVDZ. Todos os cálculos foram realizados no nível de teoria MP2(FC).

As tabelas indicam os desvios máximos (DMaxA) e médios (DMedA) absolutos das interações não corrigidas (ΔE) e corrigidas com counterpoise (ΔE_{cp}) dos 8 complexos em relação aos valores MP2(FC)/aug-cc-pVTZ. As tabelas também incluem o maior valor (Max) e a média dos BSSE dos 8 complexos para cada conjunto de bases. Como parâmetro para estimar o custo computacional foi utilizado o número de funções de bases (F.B) do complexo GUA-N7.

21 Adição de funções difusas nas bases de Pople

Tabela 19 Estudo da adição de funções difusas nos conjuntos de bases de Pople a nível MP2, o conjunto de bases utilizado como referência foi o aug-cc-pVTZ. Os valores estão em kcal/mol.

Conjunto de bases	ΔE		ΔE		BSSE		F.B
	DMaxA	DMedA	DMaxA	DMedA	Max	Média	Gua-N7
Ref. aug-cc-pVTZ					3,30	2,70	1131
6-311++g (3df,3pd)	0,46	0,36	0,66	0,54	4,36	3,60	941
6-311+G (3df,3pd)	0,42	0,33	0,67	0,55	4,32	3,57	926
6-31++g (3df,3pd)	2,00	1,46	0,79	0,58	5,94	4,74	920
6-31+G (3df,3pd)	1,93	1,38	0,79	0,59	5,86	4,66	905
6-311G (3df,3pd)	4,68	4,07	0,63	0,34	7,91	6,45	858
6-31G (3df,3pd)	7,92	6,61	1,17	0,78	10,53	8,53	837

Na Tabela 19 nota-se que a adição de funções difusas nos hidrogênios (++) é desnecessária, pois tanto os valores dos desvios de ΔE , ΔE_{cp} como a magnitude do BSSE foram aproximadamente iguais aos dos conjunto de bases com funções de polarização apenas em átomos pesados (+). Porém ao retirar todas as funções difusas (6-31G(3df,3pd) e 6-311G(3df,3pd)) os desvios médios em relação aos valores sem correção de ΔE (aug-cc-pVTZ) aumentaram mais de 300% acarretando um acréscimo de aproximadamente 80% nos valores médios do BSSE.

22 Adição de funções de polarização

Nos conjuntos de bases TZ de Pople:

Tabela 20 Estudo da adição de funções de polarização em hidrogênios nos conjuntos de bases de Pople de valência tripla à nível MP2. O conjunto de bases utilizado como referência foi o aug-cc-pVTZ. Os valores estão em kcal/mol.

Conjunto de bases	ΔE		ΔE_{cp}		BSSE		F.B
	DMaxA	DMedA	DMaxA	DMedA	Max	Média	Gua-N7
Ref. aug-cc-pVTZ					3,30	2,70	1131
6-311+ g (3df,3pd)	0,42	0,33	0,67	0,55	4,32	3,57	926
6-311+ g (3df,3p)	0,62	0,43	0,88	0,71	4,62	3,84	851
6-311+ g (3df,2p)	0,49	0,22	0,90	0,72	4,29	3,63	806
6-311+ g (3df,p)	0,58	0,18	1,08	0,88	4,17	3,64	761
6-311+ g (3df)	1,22	0,67	1,95	1,28	5,31	4,65	716

A partir da Tabela 20 é possível notar que ao reduzir as funções de polarização nos hidrogênios de 3pd para apenas d em 6-311+G(3df,p), o aumento dos desvios máximos e médios de ΔE foi inferior a 0,5 kcal/mol, enquanto os BSSE se mantiveram praticamente iguais. Essa redução nas funções de polarização em hidrogênios diminuiu 165 funções de bases nos cálculos do complexo GUA-N7.

Já ao comparar o conjunto de bases 6-311+G(3df,p) com 6-311+G(3df) tanto os desvios médios e máximos aumentaram em torno de 0,5 kcal/mol, como também os valores médios e máximos dos BSSE aumentaram mais de 1 kcal/mol, sendo a diferença no número de funções de bases da GUA-N7 de

apenas 45, logo a adição de apenas uma função de polarização p nos hidrogênios foi considerada a escolha de melhor custo benefício.

Mantendo uma função de polarização p nos hidrogênios estudou-se o efeito da adição de funções de polarização d e f em átomos pesados, conforme a Tabela 21:

Tabela 21 Estudo da adição de funções de polarização em átomos pesados nos conjuntos de bases de Pople de valência tripla a nível MP2. O conjunto de bases utilizado como referência foi o aug-cc-pVTZ. Os valores estão em kcal/mol.

Conjunto de bases	ΔE		ΔE_{cp}		BSSE		F.B
	DMaxA	DMedA	DMaxA	DMedA	Max	Média	GUA-N7
Ref. aug-cc-pVTZ					3,30	2,70	1131
6-311+ g (3df,p)	0,58	0,18	1,08	0,88	4,17	3,64	761
6-311+ g (2df,p)	1,07	0,67	1,50	1,21	5,30	4,57	676
6-311+ g (3d,p)	0,45	0,24	1,93	1,54	4,87	4,15	642
6-311+ g (2d,p)	0,44	0,16	2,45	2,03	5,69	4,87	557
6-311+ g (d,p)	1,14	0,69	4,60	3,30	7,91	6,68	472

Observa-se, na Tabela 21, que ao retirar a função de maior momento angular f em 6-311+G(3d,p) há um menor acréscimo tanto nos desvios máximos e médios, como no BSSE, do que ao retirar uma função de polarização d e manter a f em 6-311+G(2df,p).

Já ao comparar o conjunto de bases 6-311+G(3d,p) com 6-311+G(2d,p) e 6-311+G(d,p) é possível observar que a diminuição de uma função d praticamente não alterou os desvios máximos e médios de ΔE , apenas aumentou os valores máximos e médios do BSSE. Porém ao retirar mais uma função d e formar o 6-311+G(d,p) os desvios médios e máximos de ΔE dobraram de valor, enquanto o acréscimo nos valores médios do BSSE, que foi de 0,72 para 6-311+G(2d,p), passou a ser de 2,53 em 6-311+G(d,p).

Logo, a combinação de funções de polarização com melhor razão entre custo computacional e qualidade dos resultados para os conjuntos de bases de valência tripla (TZ) foi 6-311+G(2d,p).

O mesmo estudo foi feito para os conjuntos de bases DZ de Pople:

Tabela 22 Estudo da adição de funções de polarização em hidrogênios nos conjuntos de bases de Pople de valência dupla a nível MP2, o conjunto de bases utilizado como referência foi o aug-cc-pVTZ. Os valores estão em kcal/mol.

Conjunto de bases	ΔE		ΔE_{cp}		BSSE		F.B
	DMaxA	DMedA	DMaxA	DMedA	Max	Média	Gua-N7
Ref. aug-cc-pVTZ					3,30	2,70	1131
6-31+ g (3df,3pd)	1,93	1,38	0,79	0,59	5,86	4,66	905
6-31+ g (3df,3p)	2,10	1,58	1,09	0,83	6,31	5,10	815
6-31+ g (3df,2p)	1,41	1,15	1,10	0,85	5,53	4,70	770
6-31+ g (3df,p)	1,14	0,84	1,36	1,06	5,22	4,59	725
6-31+ g (3df)	2,34	1,36	2,22	1,41	7,05	5,46	680

Assim como nos conjuntos de bases TZ a adição de apenas uma função p nos hidrogênios levou aos menores desvios nos valores não corrigidos de ΔE , como também, a menores valores de BSSE, se mostrando a melhor combinação.

Tabela 23 Estudo da adição de funções de polarização em átomos pesados nos conjuntos de bases de Pople de valência dupla a nível MP2. O conjunto de bases utilizado como referência foi o aug-cc-pVTZ. Os valores estão em kcal/mol.

Conjunto de bases	ΔE		ΔE_{cp}		BSSE		F.B
	DMaxA	DMedA	DMaxA	DMedA	Max	Média	GUA-N7
Ref. aug-cc-pVTZ					3,30	2,70	1131
6-31+ g (3df,p)	1,14	0,84	1,36	1,06	5,22	4,59	725
6-31+ g (2df,p)	1,20	0,90	1,41	1,08	5,37	4,68	623
6-31+ g (3d,p)	0,94	0,52	2,25	1,69	5,61	4,90	606
6-31+ g (2d,p)	0,41	0,27	2,35	1,84	5,18	4,48	504
6-31+ g (d,p)	1,33	0,52	4,19	3,14	7,59	6,21	402
6-31+ g (d)	2,53	1,01	4,93	3,69	8,93	7,13	357

Como pode ser visto na Tabela 23, o conjunto de bases 6-31+G(2d,p) se destacou em relação aos demais conjuntos de bases DZ de Pople por apresentar não só o menor desvio máximo absoluto em relação aos valores não corrigidos de ΔE , mas também pequenos valores médio e máximo do BSSE.

23 Alterações nos conjuntos de bases de Dunning

Tabela 24 Estudo da adição de funções difusas nos conjuntos de bases de Dunning a nível MP2. O conjunto de bases utilizado como referência foi o aug-cc-pVTZ. Os valores estão em kcal/mol.

Conjunto de bases	ΔE		ΔE_{cp}		BSSE		F.B
	DMaxA	DMedA	DMaxA	DMedA	Max	Média	GUA-N7
Ref. aug-cc-pVTZ					3,30	2,70	1131
Jul-cc-pVTZ	0,51	0,26	0,12	0,08	2,96	2,52	996
Jun-cc-pVTZ	0,51	0,26	0,53	0,35	3,33	2,78	877
May-cc-pVTZ	0,70	0,36	0,95	0,62	3,55	2,98	792
cc-pVTZ	1,87	1,52	0,53	0,35	5,58	4,56	724
aug-cc-pVDZ	0,98	0,48	2,11	1,82	6,03	4,91	530
Jul-cc-pVDZ	0,82	0,37	2,38	2,03	5,18	4,39	470

Ao retirar as funções difusas dos hidrogênios no conjunto de bases aug-cc-pVTZ formando Jul-cc-pVTZ nota-se pequenos desvios médio e máximo para os valores de ΔE_{cp} . E apesar dos desvios nos valores de ΔE terem chegado a 0,51 kcal/mol os valores médio e máximo do BSSE foram ligeiramente menores, com uma redução de 135 funções de bases nos cálculos do complexo GUA-N7.

Já ao retirar as funções difusas de maior momento angular do conjunto de bases Jul-cc-pVTZ formando Jun-cc-pVTZ, houve uma diminuição de 119 funções de bases, sendo que os valores não corrigidos de ΔE se mantiveram praticamente iguais, ocorrendo apenas um pequeno aumento nos desvios das energias corrigidas e consequentemente nos valores máximos e mínimos do BSSE.

Tanto os desvios médios absolutos como também o BSSE do conjunto de bases Jul-cc-PVDZ foram comparáveis ao conjunto de bases 6-31+G(2d,p) de Pople que se destacou por apresentar o menor desvio máximo absoluto em relação aos valores de ΔE da referência e ter BSSE relativamente pequenos a um baixo custo computacional.

ANEXO 2 – Energias de interação

Valores das energias de interação (ΔE e ΔE_{cp}) calculados no vácuo de todos os complexos estudados. As geometrias dos complexos foram otimizadas em MP2/6-31+G(2d,p) e apenas os cálculos das energias foram realizados em HF/Jul-cc-pVTZ, MP2/Jul-cc-pVTZ e MP2/6-31+G(2d,p).

Tabela 25 Energias de interação (ΔE e ΔE_{cp}) calculadas mantendo as geometrias dos complexos otimizadas em MP2/6-31+g(2d,p) fixas e vários apenas o métodos utilizado para o cálculo das energias. Todos os valores estão em kcal/mol.

Complexos	ΔE_{cp} HF Jul-cc-pVTZ	ΔE HF Jul-cc-pVTZ	ΔE_{cp} MP2 Jul-cc-pVTZ	ΔE MP2 Jul-cc-pVTZ	ΔE_{cp} MP2 6-31+G(2d,p)	ΔE MP2 6-31+G(2d,p)
4-GUA-O6/N7	-111,14	-111,95	-110,11	-112,79	-108,47	-113,39
4-CYT-O2/N3	-105,11	-105,83	-102,37	-104,71	-101,09	-105,49
5-GUA-N7	-88,06	-88,79	-91,99	-94,89	-90,24	-95,41
5-GUA-O6	-90,92	-91,59	-91,44	-93,97	-89,65	-94,04
5-CYT-O2	-89,89	-90,52	-89,77	-92,10	-87,76	-91,82
5-CYT-N3	-78,80	-79,53	-83,02	-85,86	-81,58	-86,76
5-ADE-N7	-57,77	-58,47	-65,01	-67,80	-62,86	-68,00
5-ADE-N1	-55,71	-56,39	-63,04	-65,66	-61,35	-66,01
5-THY-O4	-60,83	-61,46	-60,19	-62,26	-59,02	-62,86
5-URA-O4	-61,20	-61,79	-59,49	-61,35	-58,23	-61,61
6-GUA-O6/N7	-73,01	-73,42	-77,20	-79,33	-74,72	-78,27
6-CYT-O2/N3	-70,87	-71,24	-74,70	-76,71	-72,52	-75,90
6-THY-O4	-49,16	-49,48	-52,87	-54,43	-51,22	-53,86
6-URA-O4	-49,47	-49,76	-51,94	-53,37	-50,53	-52,82
6-ADE-N6/N7	-42,50	-42,90	-51,87	-54,03	-49,50	-53,25
6-ADE-N1/N6	-42,14	-42,51	-51,64	-53,69	-49,40	-52,97

ANEXO 3 – Trabalhos apresentados e artigos publicados

Trabalhos apresentados em eventos:

Continuum solvent effects on nucleobasis- $[\text{Mg}(\text{H}_2\text{O})_5]^{2+}$ complexation energies

Vytor P. Oliveira e Edgardo Garcia

XXXVIII Congresso de Químicos Teóricos de Expressão Latina, 2012, Natal-RN

Correlation effects on electron density and ESP charges in Nucleobasis $[\text{Mg}(\text{H}_2\text{O})_5]^{2+}$

Vytor P. Oliveira e Edgardo Garcia

IV Simpósio de estrutura eletrônica e dinâmica molecular, 2012, Pirenópolis-GO

Study of *ab initio* methods for a precise description of Nucleobasis- $[\text{Mg}(\text{H}_2\text{O})_5]^{2+}$ complexes

Vytor P. Oliveira e Edgardo Garcia

IV Simpósio de estrutura eletrônica e dinâmica molecular, 2012, Pirenópolis-GO

Nucleotide Bases and Pentahydrated Mg^{2+} Inner-Shell Interactions, an *ab initio* Approach

Vytor P. Oliveira e Edgardo Garcia

XLI Reuniao Anual da SBBq, 2012, Foz do Iguaçu-PR

An *ab initio* Study of Nucleotide bases and $[\text{Mg}(\text{H}_2\text{O})_5]^{2+}$ complexes

Vytor P. Oliveira e Edgardo Garcia

XVI Simpósio Brasileiro de Química Teórica, 2011, Ouro Preto-MG

Artigos publicados:

Estudo de Bases e Métodos em Complexos: $[\text{Mg}(\text{H}_2\text{O})_5]^{2+}$ Bases Nucleicas

Vytor P. Oliveira e Edgardo Garcia

Revista Processos Químicos (2012) 6:54-56

