



UNIVERSIDADE FEDERAL DE PERNAMBUCO
Centro de Ciências Exatas e da Natureza
Departamento de Química Fundamental
Programa de Pós-Graduação em Química

Dissertação de Mestrado

Modelo Sparkle:

Novas Estratégias de Parametrização

Agrinaldo Jacinto do Nascimento Junior

Recife-PE Brasil

Maio / 2009



UNIVERSIDADE FEDERAL DE PERNAMBUCO
CENTRO DE CIÊNCIAS EXATAS E DA NATUREZA
DEPARTAMENTO DE QUÍMICA FUNDAMENTAL
PROGRAMA DE PÓS-GRADUAÇÃO EM QUÍMICA

Modelo Sparkle:

Novas Estratégias de Parametrização

Agrinaldo Jacinto do Nascimento Junior*

Dissertação apresentada ao
Programa de Pós-
Graduação em Química da
UFPE como parte dos
requisitos para a obtenção
do título de Mestre em
Química.

Orientador: Prof. Dr. Alfredo Mayall Simas

Co-orientador: Prof. Dr. Ricardo Oliveira Freire

***Bolsista CNPq**

Nascimento Junior, Agrinaldo Jacinto do
Modelo sparkle: novas estratégias de parametrização / Agrinaldo Jacinto
do Nascimento Junior. - Recife : O Autor, 2009.
xxiv, 99 folhas: il. fig. tab.

Dissertação (mestrado) - Universidade Federal de Pernambuco. CCEN.
Química, 2009.

Inclui bibliografia e anexo.

1. Química computacional. 2. Metais alcalinos.
I. Título.

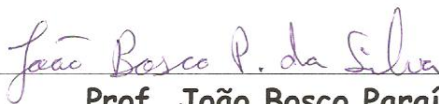
542 CDD (22.ed.) FQ 2010-009

Dissertação submetida ao Corpo Docente do Programa de Pós-Graduação em Química do Departamento de Química Fundamental da Universidade Federal de Pernambuco como parte dos requisitos necessários para a obtenção do Grau de Mestre em Química.

Aprovada:



Prof. Alfredo Mayall Simas
(Orientador)
Departamento de Química Fundamental
Universidade Federal de Pernambuco



Prof. João Bosco Paraíso da Silva
Departamento de Química Fundamental
Universidade Federal de Pernambuco



Prof. Gerd Bruno da Rocha
Departamento de Química
Universidade Federal da Paraíba

“Modelo Sparkle: Novas Estratégias de Parametrização”

Por

Agrinaldo Jacinto do Nascimento Junior

Departamento de Química Fundamental
Centro de Ciências Exatas e da Natureza
Universidade Federal de Pernambuco
Recife – PE Brasil
27 de maio de 2009

Dedico esta dissertação em primeiro lugar a Deus, o Criador e Mantenedor deste milagre chamado vida, ao meu pai Agrinaldo Jacinto do Nascimento e à minha mãe Maria do Carmo Silva do Nascimento que tanto amo.

Agradeço

À poderosa instituição chamada família de uma maneira geral. Em especial aos tios Zilda e Mardoqueu, pelo incentivo e por terem me acolhido quando aqui cheguei.

Ao grande amigo e irmão Dr. Rodrigo Meneghetti Pontes pelas conversas importantes de sabedoria e por ter me iniciado na Química Teórica.

Ao professor Dr. Ernani Abicht Basso, o Grupo ECO e todos os outros professores da Univesidade Estadual de Maringá muito importantes na minha formação como químico. Em especial ao Professor Expedito pelos muitos conselhos e amizade.

Ao meu orientador Professor Dr. Alfredo Mayall Simas por seus aconselhamentos e grande dedicação no desenvolvimento deste trabalho.

Ao meu co-orientador Professor Dr. Ricardo O. Freire pelos muitos e-mails, conversas via skype e suporte na maravilhosa cidade de Aracaju, crucial para que este trabalho fosse concluído.

Ao Professor Dr. Nivan B. da Costa Jr. pelo apoio ao dispor o laboratório de química computacional na UFS.

Ao professor Dr. Gerd B. da Rocha pelas explicações sobre o MOPAC.

Ao professor Dr. Oscar L. Malta por ter cedido o manuscrito de seu livro em preparação que foi importante para o meu entendimento de alguns dos conceitos aqui apresentados.

À minha namorada Ana Carolina Santos Albuquerque por sua dedicação e incentivo para que este trabalho fosse concluído.

Aos amigos Dimitri e Thiago pela disposição plena no trabalho, companheirismo e as importantes horas de discussões. Principalmente pelo auxílio através do desenvolvimento de programas importantes neste trabalho, a própria implementação do planetário, Gauss, Convent e outros.

Ao companheiro de Laboratório Dr. Gustavo L. C. de Moura pela disposição e discussões sobre métodos semi-empíricos.

À amiga Caroline pela amizade e troca de experiência em análises de Clusters.

Aos grandes estatísticos e companheiros de república Josimar Mendes, Jeremias Leão, Manoel Neto, Marcelo Bourguignon e ao economista Helton Saulo pelo apoio e disposição em esclarecer importantes conceitos sobre a estatística e o uso do programa R.

Ao amigo Josivandro do N. Silva pelas trocas de idéias que contribuíram para aperfeiçoar o trabalho.

Aos professores e funcionários do Departamento de Química Fundamental que auxiliaram no termino deste trabalho de alguma maneira.

À todos amigos distantes que me ajudaram direta ou indiretamente neste trabalho.

À RENAMI e ao Instituto do Milênio de Materiais Complexos, pelo apoio e ao CNPq pela bolsa concedida

Tout ce que nous voyons du monde n'est qu'un trait imperceptible dans l'ample sein de la nature. Nulle idée n'approche de l'étendue de ces espaces. Nous avons beau enfler nos conceptions, nous n'enfantons que des atomes au prix de la réalité des choses. C'est une sphère infinie dont le centre est partout, la ciconférence nulle part. Enfin c'est un des plus grands caractères sensibles de la toute puissance de Dieu, que notre imagination se perde dans cette pensée.

Blaise Pascal¹

SUMÁRIO

LISTA DE FIGURAS	X
LISTA DE SIMBOLOS.....	XVII
RESUMO	XX
ABSTRACT	XXII
1. INTRODUÇÃO	1
1.1- A PRIMEIRA VERSÃO DO MODELO SPARKLE (SMLC), 1994.....	2
1.2- A SEGUNDA VERSÃO DO MODELO SPARKLE (SMLCII), 1998	4
1.3- A TERCEIRA VERSÃO DO MODELO SPARKLE (SMLCIII), 2002	6
1.4- SPARKLE MODEL FOR THE CALCULATION OF ACTINIDE COMPLEXES, SMAC – UM MODELO PARA ACTNÍDEOS, 2002	7
1.5- A QUARTA E ÚLTIMA VERSÃO DO MODELO SPARKLE (SMLC/AM1), 2005-2006....	9
1.6- SPARKLE/PM3, 2006-2008.....	14
1.7- ANÁLISES UTILIZANDO MÉTODO AB-INITIO	14
2. A VALIDAÇÃO DO MODELO SPARKLE E A FUNÇÃO RESPOSTA: “PLANETÁRIO”	18
2.1- MODELO SMLC/AM1, UM MÉTODO COM BOA CAPACIDADE DE PREDIÇÃO DE GEOMETRIA	18
2.2- JUSTIFICATIVA PARA O USO DO MODELO SPARKLE.....	19
2.3- O ESTUDO DE UM CONJUNTO DE 45 CÉLULAS UNITÁRIAS DE HEXAAQUAMAGNÉSIO	20
2.4- TRATAMENTO ESTATÍSTICO.....	31
2.4.1- Testes de Hipótese Estatística	31
2.4.2- Definição para os Testes de Aleatoriedade e Especificidade das Distribuições	34
2.4.3- Distribuições Estatísticas Usadas	37
2.4.4- Método de Máxima Verossimilhança	39
2.5- ANÁLISE DOS RESULTADOS	41
2.5.1- Análise para as variáveis radiais	41
2.5.2- Análise para as variáveis latitudinais	43
2.5.3- Análise para as variáveis longitudinais	45
2.6- DEFINIÇÃO DA NOVA FUNÇÃO RESPOSTA: “PLANETÁRIO”	48
2.7- ANALISANDO GEOMETRIAS GERADAS ATRAVÉS DE MÉTODOS AB-INITIO	52
3. CONSTRUÇÃO DO BANCO DE PARAMETRIZAÇÃO.....	62
3.1- A ESCOLHA DOS COMPLEXOS A PARTIR DO CSD ¹⁵⁻¹⁷	62
3.2- DEFINIÇÃO DO GRUPO DE PARAMETRIZAÇÃO	67
3.2.1- Análise de Cluster DIANA.....	68
4. O PROCESSO DE PARAMETRIZAÇÃO	73
4.1- A PARAMETRIZAÇÃO	73
4.2- TÉCNICAS DE MINIMIZAÇÃO NÃO-LINEAR MULTIVARIADA.....	74
4.3- A ESCOLHA DO PARÂMETRO INICIAL PARA A PARAMETRIZAÇÃO	75
CONCLUSÕES	84
PERSPECTIVAS:	87

REFERÊNCIAS	88
ANEXO	94

LISTA DE FIGURAS

Figura 1 – Representação de bolas e bastões para o complexo cristalográfico [Eu(acac) ₃ o-fen] e o seu poliedro de coordenação.	3
Figura 2 – Representação esquemática em bolas e bastões do complexo cristalino [tris(Dipivaloilmetanato-O,O')(2,2':6',2''-terpiridina-N,N',N'')] de európio (III)]......	4
Figura 3 - Representação esquemática do ângulo formado entre dois átomos do poliedro de coordenação e o íon Lantanídeo (III).O átomo 1 representa o íon Lantanídeo (III) e os átomos 2 e 3 representam respectivamente dois átomos vizinhos no poliedro de coordenação.....	12
Figura 4 - Representação esquemática em bolas e bastões do complexo catiônico cristalino hexaaquamagnésio. Encontrado em 45 diferentes ambientes químicos no CSD.....	21
Figura 5 - Representação esquemática do poliedro de coordenação perfeito do hexaaquamagnésio. O sistema de coordenadas do poliedro será definido a partir do íon magnésio, sendo que o átomo 1 define a direção e o sentido positivo do eixo Z. A direção e o sentido positivo do eixo X é definido na direção do átomo 6. Por fim, a direção e o sentido positivo do eixo Y é definido na direção do átomo 3.....	24
Figura 6 - Representação de um sistema de coordenadas polares, onde d é a coordenada radial, θ a latitudinal e ϕ a coordenada longitudinal.	25
Figura 7 – Representação esquemática da regra da mão direita, que define o sistema de eixos perpendiculares únivoco centrado no metal do poliedro de coordenação. Dispondo o polegar da mão direita ao longo da ligação química de maior comprimento, define-se o eixo positivo Z. O sentido positivo do eixo X é definido pelo dedo indicador quando este está estendido. O último eixo a ser definido, o Y, tem seu sentido positivo indicado pelo dedo médio quando flexionado.....	25
Figura 8 – Representação esquemática num sistema de coordenadas polares, onde são aferidos os valores das coordenadas radiais, latitudinais e longitudinais para cada átomo coordenante do poliedro de coordenação do hexaaquamagnésio.....	26
Figura 9 - Gráfico da distribuição normal unicaudal à direita. Onde γ é o nível significância do teste e $1-\gamma$ é o nível de confiança do teste.	32
Figura 10 - Gráfico da distribuição normal unicaudal à esquerda.....	33
Figura 11 - Gráfico da distribuição normal bicaudal.	33
Figura 12 - Histograma do erro médio absoluto apresentado na Tabela 7 da variável radial para a ligação Mg-O do poliedro de coordenação do hexaaquamagnésio influenciado por 45 contra-íons e/ou moléculas apresentados na Tabela 3 e a sua correspondente distribuição “gama”, cujos parâmetros λ e β foram obtidos a partir método de máxima verossimilhança.	43
Figura 13 - Histograma para as medidas apresentadas na Tabela 8 do ângulo θ dos átomos coordenantes do hexaaquamagnésio influenciado por 45 contra-íons e/ou moléculas apresentados na Tabela 3, dividido por 180 e a sua correspondente distribuição beta, cujos parâmetros a e b foram obtidos a partir método de máxima verossimilhança.	45
Figura 14 - Histograma do erros apresentado na Tabela 9 para o ângulo ϕ dos átomos coordenantes do hexaaquamagnésio influenciado por 45 contra-íons e/ou moléculas apresentados na Tabela 3, e a sua correspondente distribuição von Mises, cujos parâmetros μ e κ foram obtidos a partir método de máxima verossimilhança.	48

Figura 15 - Gráfico de superfície dos erros médios absoluto absolutos do comprimento de ligação calculado a partir da função planetário para as diferentes combinações método/base.....	57
Figura 16 - Gráfico de superfície dos erros médios absolutos dos ângulos latitudinais dos oxigênios do poliedro de coordenação do hexaaquamagnésio calculados a partir da função planetário para as diferentes combinações método/base.	58
Figura 17 - Gráfico de superfície dos erros médios absoluto dos ângulos longitudinais dos oxigênios do poliedro de coordenação do hexaaquamagnésio calculados a partir da função planetário para as diferentes combinações método/base.	60
Figura 18 - Gráfico de superfície dos erros médios absoluto dos ângulos longitudinais dos oxigênios do poliedro de coordenação do hexaaquamagnésio calculados a partir da função planetário para as diferentes combinações método/base.	61
Figura 19 - Representação de Venn para a obtenção do conjunto universo de estruturas cristalográficas utilizadas no processo de parametrização. Onde; A = conjunto “original”; B = conjunto “Mg”; C = conjunto “Demais elementos”. A operação, $A \cap B \cap C$, resulta o conjunto universo de estruturas em estudo, que aqui chamamos grupo teste.	63
Figura 20 - Matriz representativa em coordenadas cartesianas (x,y,z) no formato .pdb, para o complexo cristalino hexaaquamagnésio, onde cada linha HETATM designa um átomo do complexo. A coluna A diz qual é o elemento, B determina um número que identifica o átomo que será utilizado nas linhas CONECT para identificar como os átomos estão conectados e C,D,E, indicam respectivamente as coordenadas X,Y,Z de cada elemento.....	65
Figura 21 - Representação esquemática de bolas e bastões para o complexo cristalino hexaaquamagnésio em conformidade com matriz da Figura 20.	66
Figura 22- Representação esquemática da análise de cluster, para escolha do grupo de parametrização.	67
Figura 23 - Representação esquemática do método de análise de desagrupamento DIANA, onde o cluster inicial contém cinco elementos (A, B, C, D, E). O primeiro passo consiste em dividi-lo em dois clusters ((C, D, E) e (A, B)) e assim por diante. Ao total, esta análise hierárquica é feita em quatro passos a partir do cluster inicial.	69
Figura 24 - Gráfico apresentado na Figura 23, apresentando o nível de cada divisão	70
Figura 25 - Dendograma simulado para um conjunto de 226 complexos considerando o erro médio das geometrias dos complexos. A altura do corte para a separação de novos clusters é feita na altura 50.	70
Figura 26 - Clusters obtidos a partir do corte na altura 50 do dendograma da Figura 25.	71
Figura 27 - Histograma das distâncias Mg-L, onde L poderá ser qualquer um dos átomos: C, O, N, Cl, Br e S. A maior quantidade de medidas é observada no intervalo de 2.0 Å a 2.2 Å.....	76
Figura 28 – (a) Histograma para o comprimento de ligação Mg-C, possuindo a maioria de suas medidas no intervalo de 2.3 Å a 2.4 Å. (b) Histograma para o comprimento de ligação Mg-S, que tem as suas medidas a partir de 2.55 Å.	77
Figura 30 - (a) Histograma para o comprimento de ligação Mg-Br, cujos valores se distribuem no intervalo de 2.45 Å a 2.70 Å. (b) Histograma para o comprimento de ligação Mg-Cl, que tem as suas medidas distribuídas no intervalo de aproximadamente 2.28 Å a 2.52 Å.....	78
Figura 31 - (a) Histograma para o comprimento de ligação Mg-N, onde a maioria das medidas se encontram no intervalo de 2.0 Å a 2.2 Å . (b) Histograma para o comprimento de	

	ligação Mg-O, cujo a maioria das medidas coincidem com o intervalo onde estão a maioria das medidas da ligação Mg-N variando de 2.0 Å a 2.2 Å.....	79
Figura 32 –	(a) Gráfico da energia de repulsão caroço-caroço em função das distâncias interatômica Mg-O para o modelo RM1. (b) Gráfico da energia de repulsão caroço-caroço em função das distâncias interatômica Mg-N para o modelo RM1.....	80

LISTA DE TABELAS

Tabela 1- Valores dos parâmetros para equação da energia de repulsão caroço-caroço do <i>sparkle</i> que representam o íon Th(IV).	9
Tabela 2- Número de complexos contidos em diferentes grupos de ligantes, definidos a partir das análises de cluster, para cada um dos íons lantanídeos trivalentes.	11
Tabela 3- Representação em duas dimensões das estruturas, com seus respectivos códigos advindos do CSD, que representam as moléculas e/ou contra íons presentes nas células unitárias de cada hexaaquamagnésio.....	21
Tabela 4- Comprimento das ligações químicas Mg-O para o poliedro de coordenação hexaaquamagnésio, onde a primeira linha designa os números que identificam os átomos representados no poliedro de coordenação da Figura 5. As últimas linhas indicam a média e a mediana de cada coluna.....	27
Tabela 5- Ângulos θ dos oxigênios no poliedro de coordenação do hexaaquamagnésio, onde a primeira linha designa os números que identificam os átomos representados no poliedro de coordenação da Figura 5. As últimas linhas indicam a média e a mediana de cada coluna.....	28
Tabela 6- Ângulos ϕ dos oxigênios no poliedro de coordenação hexaaquamagnésio, onde a primeira linha designa os números que identificam os átomos representados no poliedro de coordenação da Figura 5. As últimas linhas indicam a média e a mediana de cada coluna.	29
Tabela 7- Erro absoluto calculado para as distâncias Mg -O dos poliedros de coordenação do hexaaquamagnésio influenciado pelos contra-íons/ e ou moléculas a partir dos valores da Tabela 4.....	41
Tabela 8- Valores dos ângulos θ para os átomos de oxigênio nos poliedros de coordenação do hexaaquamagnésio para os vários contra-íons e /ou moléculas obtidos na Tabela 5, transformados para o intervalo de 0 a 1, que é o suporte da distribuição beta.	44
Tabela 9- Erros dos ângulos ϕ para os átomos de oxigênio nos poliedros de coordenação do hexaaquamagnésio influenciado pelos 45 contra-íons/ e ou moléculas apresentados na Tabela 3.	46
Tabela 10- Valores das distâncias radiais (Å) de cada átomo ligado ao magnésio, obtidos a partir das otimizações de geometria usando diferentes modelos <i>ab initio</i> . Cada coluna representa a posição do átomo no poliedro de coordenação segundo a Figura 5.....	53
Tabela 11- Valores dos ângulos latitudinais em graus para cada átomo ligado ao magnésio, obtidos a partir das otimizações de geometria usando diferentes modelos <i>ab-initio</i> . Cada coluna representa a posição do átomo no poliedro de coordenação segundo a Figura 5.	54
Tabela 12- Valores dos ângulos longitudinais em graus para cada átomo ligado ao magnésio, obtidos a partir das otimizações de geometria usando diferentes modelos <i>ab-initio</i> . Cada coluna representa a posição do átomo no poliedro de coordenação segundo a Figura 5.	55
Tabela 13- Valores calculados a partir da parcela radial da função planetário.	56
Tabela 14- Valores calculados a partir da parcela latitudinal do planetário.....	57
Tabela 15 - Valores calculados a partir da parcela longitudinal do planetário.	59
Tabela 16- Valores do erro médio total calculados pelo planetário.	60

Tabela 17- Valores dos parâmetros de partida ajustados para a equação da energia de repulsão caroço-caroço do <i>sparkle</i> que representam o íon Mg(II).	80
Tabela 18 - Média para os erros médios absolutos dos comprimentos de ligações Mg- L, onde L representa os átomos:O; N; C; S; Br; e Cl.....	82

LISTA DE SIGLAS

Ac(IV)	– Representação geral para todos os íons tetravalentes de actínídeos.
Acac	– Acetil acetonato.
AGNES	– <i>Agglomerative nesting</i> .
AM1	– <i>Austin model 1</i> .
CNDO	– <i>Complete neglect of differential overlap</i> .
CORE	– Variável que representa a carga para o elemento sparkle (++) no MOPAC.
CSD	– <i>Cambridge Structural Database</i> .
DIANA	– <i>Divisive analysis</i> .
DQF	– Departamento de Química Fundamental da UFPE.
ECP	– <i>Effective core potential</i> .
EMA	– Erro médio absoluto.
Eu(III)	– Íon trivalente de európio.
f.d.a	– Função de distribuição acumulada.
f.d.e	– Função de distribuição empírica.
f.d.p	– Função de densidade de probabilidade.
Gd(III)	– Íon trivalente de gadolínio.
GSA	– <i>Generalized simulated annealing</i> .
GUI	– <i>Graphical user interface</i> ou interface gráfica do usuário.
Heter.	– Coluna da planilha de classificação dos complexos que designa os tipos de átomos que compõem o poliedro de coordenação.
HF	– Hartree Fock.
INDO/S-CIS	– <i>Intermediate neglect of differential overlap/spectroscopic</i>
L	– Átomo coordenado ao <i>sparkle</i> .
Ln(III)	– Representação geral para todos os íons trivalentes de lantanídeos.

MNDO	– <i>Modified neglect of diatomic overlap.</i>
NAG	– <i>Numerical Algorithms Group.</i>
o-fen	– 1,10 – fenantrolina.
pdb	– <i>Protein data bank</i> , extensão dos arquivos obtidos a partir do CSD.
Pm(III)	– Íon trivalente promécio.
PM3	– <i>Parametric model 3.</i>
PM5	– <i>Parametric model 5.</i>
PM6	– <i>Parametric model 6.</i>
Refcode	– Coluna da planilha de classificação dos complexos que designa o código do cristal no CSD.
RHF e MP2(Full)	– Métodos <i>ab initio</i> utilizados no trabalho.
B3LYP	– Método de funcional da densidade usado no trabalho.
RM1	– <i>Recife model 1.</i>
RVN	– <i>Rank version of von Neumann's ratio.</i>
SA	– <i>Simulated annealing.</i>
Sm(III)	– Íon trivalente Sámario
SMAC	– <i>Sparkle model for actinide complexes.</i>
SMLC	– <i>Sparkle model for lanthanide complexes.</i>
SMLC/AM1	– Quarta e última versão para o modelo Sparkle utilizando o método AM1.
SMLC II	– Segunda versão do modelo Sparkle.
SMLC III	– Terceira versão do modelo Sparkle.
STO-3G, 3-21G, 6-31G, 6-31G*, 6-31+G	– Diversos conjuntos de bases constituídos por funções gaussianas usados neste trabalho.
Tb(III)	– Íon trivalente de térbio.
Th(IV)	– Íon tetravalente de tório.
UFPE	– Universidade Federal de Pernambuco.
u.m.a	– Unidade de massa atômica.

LISTA DE SIMBOLOS

$\text{\AA}$	– Unidade para distâncias interatômicas (Angström).
$ A $	– É a matriz que indica o número de objetos dentro do cluster A no DIANA.
a_k, b_k e c_k	– Parâmetros referentes à intensidade, à largura e à posição das funções gaussianas da expressão da energia de repulsão <i>core-core</i> , onde k varia de 1 a 3.
ALPAM1	– Parâmetro associado à dureza do <i>sparkle</i> no MOPAC.
AMAM1	– É a integral de dois elétrons envolvendo orbitais s expressa em E_h no MOPAC
B_q^k	– Parâmetros do campo ligante.
Beta(a,b)	– Distribuição beta. a e b, são os parâmetros da distribuição.
C_q^k	– Operadores de Racah.
d	– Coordenada radial.
d_{AB}	– Distância entre os átomos A e B
d_i^{exp}	– Distância internuclear metal–átomo coordenante de um complexo experimental.
d_i^{CSD}	– Distância internuclear metal–átomo coordenante de um complexo experimental advindo do CSD.
d_i^{calc}	– Distância internuclear predita pelo modelo Sparkle.
d_j	– é a distância entre o íon metálico e o íon metálico j.
e	– é a carga do elétron.
E_h	– Unidade de energia de Hartree.
EHEAT	– Variável que representa o calor de formação no MOPAC.
eV	– Elétron-volt, unidade de energia normalmente utilizada para potencial de ionização.
$E_N(A,B)$	– Energia de repulsão caroço-caroço.
E_n	– A nova energia do sistema, encontrado no GSA.
E_{ref}	– Energia de referência, utilizada no GSA.
$F(x)$	– Função de distribuição acumulada.
$f(x)$	– Função densidade de probabilidade.
f_{Resp}	– Função resposta utilizada na parametrização do modelo Sparkle.
F^{calc}	– Fator de estrutura simulado.
F^{obs}	– Fator da estrutura observada experimentalmente.
F_{resp}	– Função resposta planetário baseada nos erros quadráticos.
G_{resp}	– Função resposta planetário baseada nos erros absolutos.

$\text{Gama}(\lambda, \beta)$	– Distribuição gama, onde λ é o parâmetro de forma da distribuição, e β é o parâmetro de escala.
GSS	– Parâmetro referente às integrais monoatômicas de dois elétrons expressa em eV.
g_j	– É um fator de carga.
H_{CL}	– Operador de campo ligante.
H_0	– Hipótese nula.
H_1	– Hipótese alternativa.
h, k, l	– Coordenadas cartesianas para o cristal.
I_{corr}	– Intensidade experimental do feixe difratado de raio-x corrigido.
K	– Posto.
M	– Quantidade de complexos contidos num cluster.
Np	– Número total de complexos que participarão do conjunto de parametrização.
N	– Número de estruturas presentes no conjunto de parametrização.
NM	– Numerador de RVN.
O_h	– Grupo de simetria do hexaaquamagnésio.
R	– Fator residual.
R_i	– <i>Rank</i> (ordem) da i -ésima observação para o teste de Bartels.
$\langle r^k \rangle$	– Integrais radiais
T	– Sequência de observações no teste de Bartels.
U_τ	– Função escore.
v.a.	– Variável aleatória.
$VM(\mu, \kappa)$	– Distribuição von Mises. μ é o parâmetro de localização (média). κ é o parâmetro de concentração.
w	– Quantidade total de clusters.
$Y_{k,q}^*$	– Harmônicos esféricos dependentes de Ω_i .
Z_A	– Carga do núcleo A.
Z_B	– Carga do núcleo B.
α	– Parâmetro associado à dureza do <i>sparkle</i> no programa MOPAC.
γ	– Nível de significância dos teste estatístico.
δ_i	– Dissimilaridade média entre o objeto i e todos os outros no cluster A.
ΔE	– Diferença de E_n com E_{ref} .
ε	– Parâmetro que considera o deslocamento do baricentro da região de recobrimento.
ρ_j	– É a integral de recobrimento entre os orbitais 4f e a camada de valência do átomo ligante j coordenado ao íon lantanídeo.

θ	– Coordenada latitudinal.
θ_i^{calc}	– Ângulo da coordenada latitudinal para o poliedro de coordenação do complexo predito pelo modelo Sparkle.
θ_i^{CSD}	– Ângulo da coordenada latitudinal para o poliedro de coordenação do complexo predito pelo modelo Sparkle.
θ_i^{exp}	– Ângulo da coordenada latitudinal para o poliedro de coordenação do complexo experimental.
Θ	– Parâmetro equivalente a temperatura na termodinâmica estatística, utilizado no GSA.
σ_d	– Desvio-padrão dos erros absolutos das distâncias d.
σ_θ	– Desvio-padrão dos erros absolutos dos ângulos θ .
σ_φ	– Desvio-padrão dos erros dos ângulos φ .
σ_d^2	– Variância dos erros absolutos das distâncias d.
σ_θ^2	– Variância dos erros absolutos dos ângulos θ .
σ_φ^2	– Variância dos erros dos ângulos φ .
τ	– Vetores dos parâmetros da densidade de probabilidade de máxima verossimilhança.
$\hat{\tau}$	– Vetores dos estimadores da densidade de probabilidade de máxima verossimilhança.
φ	– Coordenada longitudinal.
φ_i^{calc}	– Ângulo da coordenada longitudinal para o poliedro de coordenação do complexo predito pelo modelo Sparkle.
φ_i^{CSD}	– Ângulo da coordenada longitudinal para o poliedro de coordenação do complexo.
φ_i^{exp}	– Ângulo da coordenada longitudinal para o poliedro de coordenação do complexo.
Ψ	– Porcentagem de estruturas em relação ao grupo teste que participarão do grupo de parametrização
Ω_i	– Coordenadas angulares para os i elétrons 4f.
ζ	– Um parâmetro utilizado para exemplificar o teste de hipótese.
ζ_0	– Pseudo valor utilizado para exemplificar o teste de hipótese.
$^\circ$	– Grau, unidade para ângulo.

RESUMO

O objetivo deste trabalho foi o de investigar a validade de cálculos *ab initio* ou obtidos pelo modelo Sparkle para compostos de coordenação considerados como espécies isoladas, embora possuam contra-íons e/ou outras moléculas em sua célula unitária. Para isto, estudamos o complexo hexaaquamagnésio, $[\text{Mg}(\text{H}_2\text{O})_6]^{+2}$, presente em 45 compostos diferentes. Com a finalidade de avaliar o quanto os contra-íons e/ou outras moléculas influenciam estes complexos, optamos por exprimir suas geometrias em coordenadas polares, colocando a origem no próprio íon magnésio. Mostramos através de testes estatísticos que os contra-íons e/ou outras moléculas que, no seu conjunto, produzem perturbações na geometria do complexo de forma pseudo-aleatória. Entretanto, os desvios médios absolutos das distâncias na geometria do complexo provocados pelos contra-íons e/ou outra moléculas (0.01Å), ficaram uma ordem de grandeza abaixo do erro médio normalmente observado para o modelo Sparkle/AM1 dos lantanídeos (0.2Å). Este resultado (0.01Å) dá validade ao cálculo de complexos metálicos como espécies isoladas sem a inclusão explícita de contra-íons e/ou outras moléculas, ou até mesmo da rede cristalina, tanto pelo modelo Sparkle, quanto por cálculos *ab-initio*. Este resultado mostra ainda que o modelo Sparkle pode ser melhorado para obter geometrias pelo menos uma ordem de grandeza mais exatas do que as conseguidas atualmente antes de a inclusão de contra-íons ou moléculas ter que ser feita de forma explícita. O resultado mostra ainda que cálculos *ab initio* de complexos isolados, por melhores que sejam, não poderão apresentar erros médios absolutos para as distâncias inferiores a 0.01Å, quando comparados a estruturas cristalográficas, sem a inclusão explícita de contra-íons e/ou moléculas, ou até mesmo do efeito do cristal como um todo. Ainda calculamos para o conjunto de estruturas cristalográficas dos complexos estudados, os desvios médios para as coordenadas latitudinal e longitudinal obtendo, respectivamente, os valores de 2.08° e 2.34°.

Assim sendo, tendo em vista que há espaço para melhorar a exatidão do modelo Sparkle, apresentamos duas novas ferramentas para o processo de parametrização. A primeira é uma nova função resposta, que chamamos planetário, que avalia as variáveis radiais, latitudinais e longitudinais de um poliedro de coordenação, de forma ajustada e equilibrada a partir dos desvios padrões encontrados a partir dos parâmetros das distribuições que mostramos que podem ser seguidas pelos dados: para a variável radial,

a distribuição gama; para a variável latitudinal, a distribuição beta; e para a variável longitudinal, a distribuição de von Mises. A segunda ferramenta é um novo critério para estabelecer o ponto inicial a ser utilizado na parametrização, baseado na observação direta do gráfico da energia de repulsão caroço-caroço em função das distâncias Metal-L, onde L é um átomo coordenante.

Também utilizamos a nova função resposta para avaliar a exatidão da geometria predita por cálculos *ab initio* para o mesmo complexo e por vários modelos. Assim sendo, otimizamos a geometria do hexaaquamagnésio isolado utilizando os métodos HF, B3LYP e MP2(Full) combinados às bases STO-3G, 3-21G, 6-31G, 6-31G*, 6-31+G e 6-31++G. A função resposta se mostrou de fato equilibrada para avaliar as geometrias do hexaaquamagnésio obtidas a partir destes modelos, sendo que a parcela radial mostrou ser a que mais variou e a mais difícil de ser predita. Destes modelos *ab initio*, o que apresentou o melhor resultado foi o MP2/3-21G.

O novo critério de escolha do ponto inicial de parametrização, também mostrou-se eficiente. De fato, o ponto de partida assim escolhido para os parâmetros do modelo Sparkle no RM1 para complexos do íon magnésio, sem qualquer otimização, já levou a erros médios absolutos para as ligações deste íon com átomos coordenantes oxigênio e nitrogênio (que estão presentes em mais de 80% dos complexos de magnésio do banco de dados CSD) menores que os apresentados pelo método PM3 e aproximadamente equivalentes aos apresentados pelo AM1 e PM6. Isto indica que, com a otimização subsequente, pode-se esperar um significativo aumento na exatidão do modelo Sparkle para o magnésio no âmbito do RM1. Este critério de escolha do ponto inicial de parametrização, revelou-se, assim, ser bastante eficaz.

Palavras-chaves: Química computacional, metais alcalinos.

ABSTRACT

The aim of this work is to investigate the validity of computations obtained by the Sparkle model for chemical systems which have counter-ions and/or other molecules in their unit cells. To do so, we studied the hexaaquamagnesium complex, $[\text{Mg}(\text{H}_2\text{O})_6]^{+2}$, which is present in 45 different compounds. Thus, in order to evaluate the amount of counter-ions and/or other molecules which might influence that complex, we chose to express their geometries in polar coordinates, placing the origin in the magnesium ion. We show, through statistical tests, that counter-ions and/or other molecules generate perturbations in the complex's geometry that, as a whole, behave in a pseudo-random manner. However, the mean absolute deviations in the distances of the complex's geometries, caused by counter-ions and/or other molecules (0.01Å), are one order of magnitude smaller than the mean error normally observed for the Sparkle/AM1 model for lanthanides (0.2Å). That result attaches validity to the computation of metallic complexes without the explicit inclusion of counter-ions and/or other molecules, or even of the whole crystalline network, for both Sparkle model and *ab initio* model chemistries. The above result still indicates that the Sparkle model can be further improved to obtain geometries one order of magnitude more accurate, before an explicit inclusion of counter-ions and/or other molecules becomes mandatory. In addition, the above result further demonstrates that *ab initio* calculations of isolated complexes cannot lead to mean absolute errors for distances less than 0.01Å without an explicit inclusion of counter-ions and/or other molecules, or even of the whole crystal network. Moreover, the mean absolute deviations in the latitudinal and longitudinal coordinates of the complexes' geometries studied was, 2.08° and 2.34°, respectively.

Therefore, being acutely conscious that there are enough space to further improve the Sparkle model's accuracy, we introduce two new tools for the parameterization process. The first one is a new response function, which we call "planetário", which uses, radial, latitudinal, and longitudinal coordinates of the edges of the coordination polyhedron. This response function is adjusted and balanced, taking into account standard deviations computed from the parameters of the distributions. For the radial variable, the gamma distribution; for the latitudinal variable, the beta distribution; and for the longitudinal distribution, the von Mises distribution. The second tool is a new criterium to establish a starting point to be used in the

parameterization, based on a direct observation of the core-core repulsion of energy plot as function of Metal-L distances, where L is a coordinating atom.

Finally, we used this new response function to evaluate the accuracy of coordination polyhedron geometries predicted by various ab initio model calculations for the same complex. Thus, we optimized the geometry of the isolated hexaaquamagnesium complex by means of HF, B3LYP, and MP2(Full) methods combined with STO-3G, 3-21G, 6-31G, 6-31G*, 6-31+G, and 6-31++G bases. The response function was indeed balanced to evaluate the hexaaquamagnesium's geometries obtained by that models, and the radial part presented greater variation and unpredictability. Considering ab initio models, the MP2/3-21G outperformed all the other models.

The new criterium for the choice of the starting point of parameterization, revealed itself to be quite efficient. In fact, the chosen starting point for the Sparkle model in RM1 for complexes of magnesium ion, without any further parameterization, provided mean absolute errors – for both oxygen and nitrogen coordinating atom bond distances (which are present in more than 80% of magnesium's complexes in the CSD data set) – already smaller than those of PM3, and approximately equivalent to those of AM1 and PM6. With subsequent optimization, it is reasonable to expect a significant increase in the Sparkle model's accuracy for the magnesium ion within RM1. Thus, the choice criterium for the starting point of parameterization seems to be indeed quite effective.

Words-keys: computational chemistry, alkali metals.

1. INTRODUÇÃO

Os métodos semi-empíricos têm sido amplamente utilizados, principalmente pelo fato de que o tempo de cálculo gasto nestes métodos é muito inferior aos métodos *ab initio*². Esta importante característica de fim prático permite hoje estudar grandes sistemas moleculares, antes impossíveis de serem calculados.

Desta maneira, a pesquisa na linha de desenvolvimento de métodos semi-empíricos se faz necessária para que possam existir os melhores parâmetros possíveis - dentro do erro natural do método - para representar da maneira mais geral possível os complexos cristalinos, para que então seja possível estudar e avaliar suas propriedades e características.

Em especial, nesta introdução apresentamos um breve histórico sobre a evolução do modelo Sparkle desenvolvido no Recife.

O Departamento de Química Fundamental da UFPE tem sido referência internacional no estudo da química de compostos de terras raras desde a sua fundação em 1982. A parceria entre o grupo de espectroscopia de terras raras e o grupo de química teórica e computacional, tem ajudado a projetar compostos de coordenação luminescentes de lantanídeos³ e a compreender suas propriedades.

Com o objetivo de desenvolver um modelo teórico para estudar e prever a geometria e propriedades espectroscópicas dos complexos de lantanídeos, é que surge em 1994⁴ a primeira versão do modelo Sparkle, na época intitulado *Sparkle Model for the Calculation of Lanthanide Complexes* (SMLC).

O modelo SMLC foi inspirado nos *sparkles*, - a escrita em itálico é usada para designar o modelo genérico apresentado no MOPAC⁵ e diferenciá-lo do desenvolvido em Recife - presentes no software MOPAC a partir da versão 6. O MOPAC é um software escrito pelo Prof. James Joseph Patrick Stewart para cálculos semi-empíricos de sistemas moleculares e que vem sendo aperfeiçoado ao longo dos anos. No MOPAC, *sparkle* é um termo utilizado para designar uma carga iônica pura, podendo esta variar de -3 a +3. Assim podem representar diferentes entidades químicas, como, por exemplo, Ca^{2+} , K^+ , Cs^+ , Cl^- , SO_4^{2-} , Al^{3+} , PO_4^{3-} e outros. Os raios iônicos destes *sparkles* possuem valores perto de 0.7 Å, possuem calor de atomização zero, não possuem orbitais e nem potencial de ionização. Por consequência, não podem doar nem receber elétrons. O

principal objetivo destes *sparkles* é o de simular contra-íons e dipolos para mimetizar certos tipos de interação em cálculos.

O modelo SMLC foi desenvolvido pela primeira vez para o íon Eu(III) conjugado ao método quântico semi-empírico AM1⁶(Austin Model 1). O modelo Sparkle considera que as ligações entre os íons lantanídeos e seus ligantes possuem caráter predominantemente eletrostático, já que os valores para as integrais de recobrimento dos orbitais 4f do lantanídeo com os orbitais dos outros átomos a ele coordenados são muito pequenos. Este enfoque permite estudar complexos de lantanídeos sem grande custo computacional, quando comparado aos métodos *ab initio*, o que é uma grande vantagem para o pesquisador.

Pode-se visualizar o *sparkle* como sendo uma carga pontual no centro de um potencial repulsivo esférico $\exp^{-\alpha d}$, onde α é um parâmetro que reflete a dureza do íon metálico e d é a distância entre o íon metálico e o ligante.

1.1- A Primeira Versão do Modelo Sparkle (SMLC), 1994

Na primeira versão do modelo Sparkle, SMLC, atribuiu-se a carga +3 ao *sparkle* para que pudesse representar o íon Eu(III). Atribuiu-se também o calor de formação do Eu(III) gasoso, +1005.3 Kcal. mol⁻¹ ao parâmetro EHEAT. Este valor foi obtido a partir da entalpia de atomização do európio, acrescida de seus três primeiros potenciais de ionização⁷. Por fim, otimizou-se os parâmetros ALPAM1 e AMPAM1. ALPAM1 está associado à dureza do *sparkle* e AMPAM1 (E_h) é a integral de dois elétrons envolvendo orbitais s, também representada como GSS (eV). Estes valores finais foram implementados no MOPAC 6.0.

Na época utilizou-se o algoritmo simplex básico, modificado e super modificado⁸ como procedimento de parametrização. A função resposta utilizada foi a seguinte:

$$f_{resp} = \sum_{i=1}^N (d_i^{exp} - d_i^{calc})^2 \quad (1)$$

Como podemos observar, esta função é o somatório do quadrado das diferenças entre as distâncias experimentais (d_i^{exp}) e calculadas (d_i^{calc}); distâncias estas entre o íon lantanídeo e um dos i átomos dos ligantes diretamente coordenados ao metal

e que definem os vértices do poliedro de coordenação. Esta função resposta avalia quão semelhantes as geometrias dos complexos cristalinos experimentais e calculadas são em termos das distâncias entre íon lantanídeo e os átomos coordenantes. Ou seja, a função deve tender a zero quando a estrutura calculada tender a coincidir com a estrutura experimental. Nesta primeira versão, uma única estrutura experimental foi utilizada: a do complexo de número de coordenação oito $[\text{Eu}(\text{acac})_3 \text{ o-fen}]^9$ representado na Figura 1, onde acac significa acetilacetonato e o-fen significa 1,10-fenantrolina.

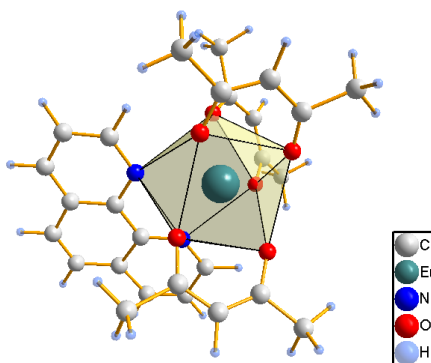


Figura 1 – Representação de bolas e bastões para o complexo cristalográfico $[\text{Eu}(\text{acac})_3 \text{ o-fen}]$ e o seu poliedro de coordenação.

Apenas 24 das 36 distâncias interatômicas possíveis do poliedro de coordenação foram tomadas para a realização desta parametrização.

Os parâmetros obtidos para esta versão SMLC foram os seguintes: ALPAM1 igual a $2.0 \text{ angström}^{-1}$; AMAM1 igual a $2 E_h$, onde E_h é a unidade de energia atômica dada por Hartree, ou 54.42 eV ; carga igual a $+3e$; e por último EHEAT igual a $+1005.3 \text{ kcal. mol}^{-1}$.

Para testar esses parâmetros, em 1994 Andrade, e outros¹⁰ otimizaram a geometria do poliedro de coordenação do $[\text{tris}(\text{Dipivaloilmetanato-O,O'}) (2,2':6',2''\text{-terpiridina-N,N',N''}) \text{ de európio (III)}]$ representado na Figura 2, que é um complexo com número de coordenação nove, encontrando um valor de $2.83 \text{ \AA}$ para a distância média interatômica e um valor de desvio médio absoluto igual a $0.12 \text{ \AA}$.

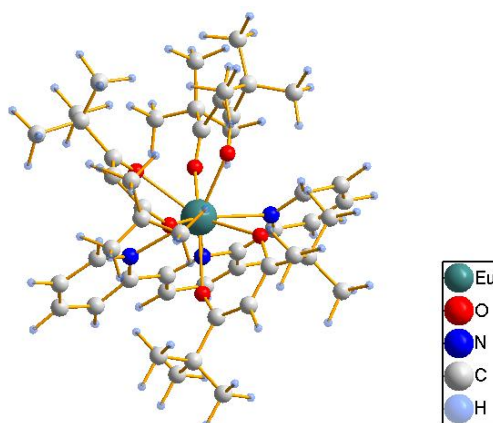


Figura 2 – Representação esquemática em bolas e bastões do complexo cristalino [tris(Dipivaloilmetanato-O,O')(2,2':6',2'')-terpiridina-N,N',N'') de európio (III)].

Para os ângulos de ligação, o desvio médio absoluto encontrado foi de 5°. Estes resultados reforçaram a idéia de que a ligação química nos complexos de lantanídeos é essencialmente eletrostática, aumentando a confiança no modelo Sparkle.

1.2- A Segunda Versão do Modelo Sparkle (SMLCII), 1998

A segunda versão do modelo SMLC^{11,12}, chamada SMLCII, foi concluída em 1998 e publicada em 2004. Esta nova versão foi uma tentativa de melhorar o modelo de várias maneiras, como enumeraremos a seguir: (i) inclusão da massa atômica do európio; (ii) reparametrização do modelo com o método AM1 incluindo todas as 36 distâncias do poliedro de coordenação; (iii) implementação do modelo no software MOPAC 93r2¹³, que é mais robusto que sua versão anterior MOPAC 6.0; (iv) e a inclusão de funções gaussianas esféricas no cálculo da energia de repulsão caroço-carço.

A inclusão da massa do európio como sendo a massa do *sparkle* foi muito importante, pois permitiu o cálculo do momento de inércia do complexo e das suas propriedades espectroscópicas na região do infravermelho: intensidades, frequências e modos normais de vibração. Permitiu também calcular as funções de partição rotacional e vibracional e conseqüentemente, propriedades termodinâmicas, como capacidade calorífica, entropia, energia livre e entalpia. Também passaram a poder ser calculadas propriedades cinéticas, como constantes de velocidade, energias de ativação e mecanismos de reação através de coordenadas intrínsecas. Também é possível

determinar estados de transição e energias de ativação envolvidas em uma reação química. Enfim, uma gama de cálculos passa a ser realizável.

A reparametrização continuou sendo feita unicamente com o mesmo complexo (Eu(acac)₃ o-fen) utilizado na versão anterior, até porque os recursos computacionais no laboratório eram muito escassos na época e os softwares utilizados durante o processo de parametrização ainda eram muito limitados em termos de automação.

Por fim, a inclusão de funções gaussianas esféricas tinha o objetivo de descrever melhor as distâncias entre o íon Eu(III) e os seus ligantes. Os autores realizaram testes e concluíram que 2 funções gaussianas eram suficientes para bem representar as ligações químicas envolvidas. A seguir, apresentamos a Equação 2 que representa a energia de repulsão caroço-caroço, utilizada no SMLC II:

$$E_N(A, B) = Z_A Z_B \langle s_A s_A | s_B s_B \rangle (1 + e^{-\alpha_A d_{AB}} + e^{-\alpha_B d_{AB}}) + \left(\frac{Z_A Z_B}{d_{AB}} \right) \left(\sum_{k=1}^{ng_A} a_{k_A} e^{-b_{k_A} (d_{AB} - c_{k_A})^2} + \sum_{k=1}^{ng_B} a_{k_B} e^{-b_{k_B} (d_{AB} - c_{k_B})^2} \right), \quad (2)$$

onde $E_N(A, B)$ é a energia de repulsão caroço-caroço; Z_A e Z_B representam as cargas do núcleo A e do núcleo B, respectivamente; $\langle s_A s_A | s_B s_B \rangle$ são os parâmetros GSS que estão associados à integral de 2-elétrons envolvendo orbitais s; α_A e α_B são os parâmetros que estão relacionados às dureza das espécies A e B, respectivamente; ng_A e ng_B são os números de funções gaussianas centradas no átomo A e átomo B, d_{AB} é a distância entre os átomos A e B; a_{k_A} e a_{k_B} são os parâmetros referentes à intensidade das funções gaussianas, b_{k_A} e b_{k_B} são os parâmetros referentes às larguras das funções gaussianas; e, por último, c_{k_A} e c_{k_B} são os parâmetros referentes à posição das funções gaussianas.

Desta maneira, na segunda versão o número de parâmetros aumentou de dois para oito parâmetros ajustáveis, já que nesta temos os novos termos a_{k_A} , a_{k_B} , b_{k_A} , b_{k_B} , c_{k_A} e c_{k_B} . Estas modificações resultaram em um modelo mais eficiente e capaz de apresentar resultados quantitativos bem mais significativos. Assim, para poder testar os parâmetros, otimizou-se a geometria de um grupo de 96 estruturas utilizando a seguinte equação para calcular o erro médio absoluto (EMA) das distâncias dentro do poliedro de coordenação:

$$EMA = \frac{1}{N} \sum_{i=1}^N |d_{exp_i} - d_{calc_i}|. \quad (3)$$

Os valores do erro médio absoluto das distâncias dentro do poliedro de coordenação foram de 0.1923 Å para a primeira versão (SMLC) e de 0.1506 Å para a nova versão (SMLC II), o que significou uma melhora de 21.68%.

Além destas inovações acima citadas, o outro diferencial foi a realização de uma análise de agrupamentos com o software Statistica 6.0 utilizando como critério de distância a euclidiana. As variáveis escolhidas para a análise foram EMAs e distâncias por tipos de ligantes classificadas em 5 grupos: β- dicetona, enxofre, tridentados, nitratos, amina e macrocíclicos. Este estudo possibilitou avaliar o erro em diferentes grupos de ligantes, quando pôde-se perceber que o grupo que apresentou os menores erros foi o da β- dicetona, que era justamente o grupo do complexo cristalino que participou da parametrização. Este resultado tinha total sentido, pois a parametrização foi realizada sobre aquele tipo de ligante. Sendo assim, os parâmetros tinham de fato um poder de predição quantitativo da geometria do estado fundamental para este grupo de complexos.

Em suma, os resultados desta segunda versão foram melhores que a primeira e a principal razão que poderia justificar isso, seria justamente a inclusão e otimização das gaussianas esféricas no modelo. Nesta versão, além da parametrização do método AM1, foram testados os métodos *Modified Neglect of Diatomic Overlap*, MNDO, e o *Parametric Model 3*, PM3. Destes três, o melhor resultado foi o AM1 com duas gaussianas.

1.3- A Terceira versão do Modelo Sparkle (SMLCIII), 2002

A terceira versão do modelo Sparkle, SMLCIII¹⁴ tem como diferença de seus predecessores dois itens apenas: (i) o banco de parametrização foi aumentado; e (ii) O modelo foi estendido aos actínídeos através da parametrização inédita do Th (IV).

Foi implementado no banco de parametrização além do [tris(acetilacetato) (1,10 fenantrolina) Eu(III)], dois novos complexos advindos do CSD¹⁵⁻¹⁷, o [tris(2,6-dimetil-3,5-heptanodionato) (trifenilfosfina) Eu(III)] e o [tris(tenoiltrifluoroacetato)

(terpiridina) Eu(III)] estes possuíam respectivamente números de coordenação oito, sete e nove.

O aumento do conjunto de parametrização de um para três complexos objetivava uma maior representatividade de números de coordenação e de diferentes ligantes, já que a segunda versão descrevia muito bem apenas os complexos com ligantes β -dicetona.

A função resposta para esta versão sofreu uma pequena modificação como podemos observar na Equação (4):

$$f_{resp} = \sum_{i=1}^N \frac{(d_i^{\text{exp}} - d_i^{\text{calc}})^2}{d_i^{\text{exp}}} . \quad (4)$$

A única diferença da Equação 2 para a Equação 4 é que no somatório aparece o termo d_i^{exp} no dividendo. Esta mudança contribuiu para que a função resposta considerasse as distâncias, todas com o mesmo peso.

Os algoritmos utilizados para realizar a parametrização (simplex básico, modificado e super modificado⁸), não sofreram modificações. Imaginava-se que apenas o incremento destas duas novas estruturas ao banco de dados resultaria numa melhora significativa dos parâmetros, já que o universo de complexos foi aumentado. Porém, quando se comparou estatisticamente o SMLC II com SMLC III, percebeu-se que não havia grandes diferenças entre os dois.

1.4- *Sparkle Model for the Calculation of Actinide Complexes*, SMAC – Um Modelo para Actínídeos, 2002

A segunda novidade desta versão, a implementação do modelo para os actínídeos, *Sparkle Model for the Calculation of Actinide Complexes*, SMAC, em particular para o Th(IV), abre uma reflexão sobre a aplicação deste modelo. Pois diferentemente dos lantanídeos, os elétrons f dos actínídeos são mais externos quando comparados aos elétrons f dos lantanídeos.

A questão principal neste momento era verificar se o carácter covalente dos elementos do bloco 5f seria tão forte a ponto de o modelo Sparkle não poder representá-lo. Observou-se claramente esta tendência covalente dos actínídeos nos complexos

presentes no CSD que apresentavam grande número de coordenação como o UF_5 . Porém, para o Th(IV) , muitos complexos tinham forte caráter eletrostático, principalmente quando se tratava das ligações Th(IV)-O e Th(IV)-N , enquanto as seguintes ligações: Th(IV)-S ; Th(IV)-P ; Th(IV)-C ; e por último Th(IV)=O apresentavam um maior caráter covalente. A principal motivação para desenvolver o SMAC vem justamente do fato de o CSD apresentar uma quantidade significativa de complexos cristalinos de Th(IV) com caráter eletrostático. Desta maneira, foi entendido que valeria a pena desenvolver um modelo para este universo de estruturas.

A metodologia de parametrização do Th(IV) , não diferiu da aplicada ao Eu(III) . A única diferença estava nos valores característicos do *sparkle* que haveria de ser parametrizado, neste caso para o íon Th(IV) . O valor implementado para a carga foi de $+4e$, o calor de formação que preenche o parâmetro EHEAT foi $1674.64 \text{ kcal.mol}^{-1}$ (Este valor foi obtido a partir da entalpia de atomização do tório, acrescida de seus quatro primeiros potenciais de ionização⁷) e, por último, a massa atômica do tório de 232.04 u.m.a.

O conjunto de parametrização para o SMAC do Th(IV) veio do CSD e era constituído por duas estruturas: bis(etilenodiamônio) tetrakis(mimetilmalonato) aqua- Th(IV) , que tinha o número de coordenação oito e o bis(etilenodiamônio) tetrakis(malonato) aqua- Th(IV) , que tinha o número de coordenação nove. Foi inserida ainda nesta versão as mesmas funções gaussianas esféricas para representar a ligação Th(IV)-L (onde L representa qualquer ligante). Sendo assim, a quantidade de parâmetros a serem encontradas era igual às do SMLCII e SMLCIII. Os mesmos algoritmos de minimização foram utilizados para encontrar os parâmetros assim como a mesma função resposta do SMLC(III).

Para esta primeira parametrização do SMAC, foram encontrados os valores apresentados na Tabela 1.

Tabela 1- Valores dos parâmetros para equação da energia de repulsão caroço-caroço do *sparkle* que representam o íon Th(IV).

<i>GSS</i>	55.5533997779
<i>ALP</i>	2.1616053239
<i>a₁</i>	0.4411098727
<i>b₁</i>	7.1471674779
<i>c₁</i>	1.7519078812
<i>a₂</i>	0.0214177229
<i>b₂</i>	7.8489816588
<i>c₂</i>	2.6498997377

E para validar estes parâmetros foi escolhido, a partir do CSD, um conjunto com mais de 40 complexos.

1.5- A Quarta e Última Versão do Modelo Sparkle (SMLC/AM1), 2005-2006

A última versão do modelo Sparkle usando o método AM1 culminou na criação do modelo agora chamado SMLC/AM1¹⁸⁻²⁹ com o objetivo de que se constituísse em um método ainda mais robusto. A outra razão para justificar esta última versão do modelo Sparkle para o método AM1 era porque o grupo havia desenvolvido uma reparametrização do mesmo, conhecida agora por método RM1³⁰, que inclusive já se encontra implementado em vários pacotes de softwares de química quântica como o Spartan'06³¹, Hyperchem 8³² e MOPAC 2007⁵, todos mundialmente distribuídos e utilizados. Desta maneira, a motivação do grupo para o desenvolvimento de novos modelos Sparkle passa a visar o desenvolvimento de um novo modelo Sparkle bem mais exato para o método RM1. Como etapa intermediária, foi desenvolvida a versão SMLC/AM1.

A versão SMLC/AM1 sofreu alguns melhoramentos no processo de parametrização dos *sparkles* e passou a representar, além do íon Eu(III), os demais íons lantanídeos trivalentes. Os parâmetros finais dos *sparkles* que representam cada íon lantanídeo (III) foram publicados em uma série de artigos: lantânio (III) e lutécio (III)²⁷; cério (III) e érbio (III)²⁵; praseodímio (III)²²; itérbio (III)²³; neodímio(III)¹⁸, promécio

(III) e samário (III)²⁸; európio (III), gadolínio (III) e térbio (III)²¹, disprósio (III)²⁰, hólmio (III)¹⁹, túlio (III)²⁶.

Não foi somente o número de *sparkles* representados no modelo que aumentou. Para cada parametrização houve um aumento significativo do banco de dados, onde estas estruturas eram advindas do CSD e todas possuíam um fator $R < 0.05$ Å. Por exemplo, o íon Eu (III) na primeira versão (SMLC) tinha apenas um complexo cristalino em seu banco de dados enquanto nessa última possuía 96 estruturas. Este aumento no banco de dados de teste foi muito importante para avaliar se os parâmetros encontrados estavam representando bem as geometrias do estado fundamental de todo o universo de estruturas.

Um grupo de parametrização com uma centena ou mais de complexos seria um grande problema, pois para cada conjunto de parâmetros encontrado no programa de parametrização, uma centena ou mais de cálculos de otimização de geometrias deveriam ser feitos com o objetivo de avaliar o quanto as geometrias geradas a partir destes parâmetros se assemelhariam com as experimentais – o que inviabilizaria o processo.

Considerando que uma boa parametrização precisa de aproximadamente 500 pontos de simplex, um banco deste tamanho gastaria em torno de três dias para que todas as suas estruturas fossem calculadas. Logo, estimamos que tal parametrização viesse a durar cerca de 1500 dias, o que representa cerca de quatro anos.

Desta maneira, a saída encontrada para essa problemática foi fazer uma parametrização com o menor grupo de complexos cristalinos possíveis que pudesse corretamente representar o universo dos complexos.

Com esta finalidade, uma análise de agrupamentos baseada nos valores dos erros do método SMLCII foi aplicada, onde a métrica de dissimilaridade utilizada foi a distância euclidiana. O resultado desta análise mostrou que os conjuntos universo das estruturas de cada íon lantanídeo trivalente poderiam ser divididos em sete conjuntos de acordo com seus ligantes: β -dicetonas; nitratos; monodentados; bidentados; tridentados; polidentados; dilantanídeos. Podemos observar esta divisão e quantas estruturas haviam para cada íon lantanídeo (III) na Tabela 2.

Tabela 2- Número de complexos contidos em diferentes grupos de ligantes, definidos a partir das análises de cluster, para cada um dos íons lantanídeos trivalentes.

Tipos de Ligantes		Número de Estruturas													
Grupos	La	Ce	Pr	Nd	Pm	Sm	Eu	Gd	Tb	Dy	Ho	Er	Tm	Yb	Lu
β- dicetonas	1	4	7	2	2	5	13	5	4	1	2	2	1	1	0
Nitratos	32	7	17	18	2	7	17	8	4	5	4	11	4	8	10
Monodentados	6	4	8	9	2	7	5	10	12	5	7	6	3	9	6
Bidentados	2	1	1	6	2	4	7	3	2	4	2	6	0	5	2
Tridentados	3	7	3	7	2	5	13	2	3	0	0	0	0	3	0
Polidentados	19	10	4	12	4	8	17	30	8	11	9	9	4	15	7
Dilantanídeos	12	4	8	3	1	6	24	12	9	4	3	5	3	6	5
TOTAL	75	37	48	57	15	42	96	70	42	30	27	39	15	47	30

De uma maneira geral, os autores convencionaram que 15 estruturas eram suficientes para representar adequadamente o universo de complexos de cada íon lantanídeo (III) existente no CSD até aquela época.

Para cada íon lantanídeo foram selecionados 15 complexos que participariam do grupo de parametrização. A escolha destes 15 complexos foi feita a partir de uma amostragem sobre os grupos apresentados na Tabela 2.

Por exemplo, para o íon Eu (III), a quantidade de complexos retirados de cada grupo foi a seguinte: três do grupo β- dicetona; dois do grupo nitrato; quatro do grupo tridentado; três do grupo polidentado; e por último três do grupo dieurópio. O requisito tomado para selecionar as estruturas dentro de cada um dos grupos para participar do grupo de parametrização foi o tempo de cálculo gasto para a otimização das geometrias de suas respectivas estruturas. Ou seja, as estruturas que demandaram o menor tempo de cálculo foram as selecionadas a participar do conjunto de parametrização. Esta seleção, feita desta forma, é importante, pois o processo de parametrização é iterativo: para cada novo parâmetro encontrado, as geometrias das estruturas do grupo de parametrização deverão ser reotimizadas, com este processo sendo repetido até que os parâmetros cumpram os requisitos estabelecidos pelo parametrizador.

Outra importante mudança que aconteceu nesta última versão do modelo Sparkle foi o desenvolvimento de uma nova função resposta onde, além dos erros nas distâncias, foram também incluídos os erros nos ângulos formados entre dois átomos adjacentes do poliedro de coordenação e o íon lantanídeo, tal como indicado na Figura 3 abaixo.

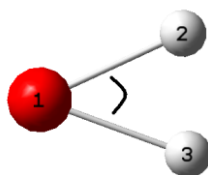


Figura 3 - Representação esquemática do ângulo formado entre dois átomos do poliedro de coordenação e o íon Lantanídeo (III). O átomo 1 representa o íon Lantanídeo (III) e os átomos 2 e 3 representam respectivamente dois átomos vizinhos no poliedro de coordenação.

Desta maneira, os ângulos do poliedro de coordenação puderam ser melhores avaliados como podemos observar na nova função resposta utilizada:

$$f_{resp} = \sum_{i=1} \left\{ \sum_{j=1} [100 \cdot (d_{i,j}^{CSD} - d_{i,j}^{calc})^2] + \sum_{k=1} \left[\frac{2}{3} \cdot (\theta_{i,k}^{CSD} - \theta_{i,k}^{calc})^2 \right] \right\}, \quad (5)$$

onde i representa o número de diferentes complexos; j representa todas as distâncias do poliedro de coordenação; $100 \text{ \AA}^{-2}$, e $\frac{2}{3} \text{ grau}^{-2}$, são, respectivamente, os fatores peso de Dewar³³ para o comprimento de ligação e ângulo de ligação; d é a distancia entre o íon e o ligante; θ são todos os ângulos entre o íon e dois átomos vizinhos no poliedro de coordenação como representado na Fig.1.; e as siglas CSD e calc, referem-se às estruturas experimental e calculada, respectivamente.

Esta mudança na função resposta foi importante porque os autores perceberam que a Equação 4 representava bem somente as distâncias Ln(III)-L do complexo cristalino, enquanto as distâncias L-L no poliedro de coordenação não eram bem representadas.

Outro importante desenvolvimento neste momento foi a automatização dos programas de parametrização. Esta realização representou um importante avanço na parametrização de todos os íons lantanídeos trivalentes.

Em geral, o protocolo de parametrização foi idêntico para os íons lantanídeos, com exceção do íon Pm(III), pois o seu banco de estruturas não foi advindo do CSD e sim construído a partir de estruturas com geometrias geradas por métodos *ab-initio/ECP* utilizando somente o ECP quase relativístico desenvolvido por Dolg, M. e

colaboradores³⁴, implementado no Gaussian 98 como MWB50 ECP³⁵. Este ECP inclui 46+4fⁿ elétrons no caroço permitindo que até 11 elétrons sejam tratados explicitamente.

Na parametrização, foi utilizada ainda uma combinação de algoritmos de minimização não linear da função resposta: o simplex básico utilizado nas primeiras versões e o Newton-Raphson com o objetivo de encontrar o melhor parâmetro para representar de maneira geral o maior número de geometrias do estado eletrônico fundamental dos complexos cristalinos.

Sobre a escolha dos parâmetros iniciais para cada *sparkle* Ln(III), foi adotada a estratégia de utilizar os parâmetros do átomo vizinho mais próximo na tabela periódica. Desta maneira, como o primeiro *sparkle* a ser parametrizado foi o que representava o íon Eu(III), seus parâmetros finais foram utilizados para iniciar a parametrização do Sm(III) e Gd(III). Da mesma maneira, os parâmetros finais do Sm(III) serviram como parâmetros de partida para o Pm(III) e assim por diante.

Para validar os parâmetros, foi utilizado o grupo teste de cada Ln(III), que seria justamente o conjunto universo de cada íon.

Outro ponto importante que deve ser ressaltado é que nas versões anteriores do modelo Sparkle os ligantes do tipo nitrato não tinham a geometria de seus poliedros de coordenação bem representada, pois estes coordenavam-se como monodentados. Nesta última versão este problema foi corrigido.

O modelo SMLC/AM1, tornou-se uma boa alternativa frente ao método *Ab initio* já que teve um excelente avanço na reprodução de geometrias do estado eletrônico fundamental dos complexos de lantanídeos.

Quando analisamos a evolução do modelo Sparkle^{4,11}, detectamos que hoje, na versão SMLC/AM1¹⁸⁻²⁹, este modelo apresenta um erro médio absoluto de 0.09 Å em relação às distâncias Eu(III) – L, onde L representa oxigênio ou nitrogênio, enquanto na primeira versão este mesmo erro era de 0.68 Å. Desta maneira, observamos uma melhora de mais de 86% no modelo em termos da exatidão naquelas distâncias. Com relação ao erro médio absoluto de todas as distâncias no poliedro de coordenação, observamos que na primeira versão o valor era de 0.437 Å e nesta última versão SMLC/AM1 o valor caiu para 0.190 Å, o que representa uma melhora de mais de 56%.

Também, uma nova versão do modelo SMAC para o Th(IV) foi preparada. As diferenças desta versão para a sua anterior apresentada por Rocha, G. B. e colaboradores¹⁴ foram: (i) o aumento do banco de dados para o íon Th(IV); (ii) aumento do grupo de parametrização para 15 estruturas; (iii) escolha do grupo de parametrização

a partir de análise de cluster semelhante às demais seleções do modelo SMLC/AM1; (iv) e por último a utilização da nova função resposta dada na Equação 5.

Esta soma de inovações resultou em uma melhora significativa do modelo SMAC, pois o erro médio de todas as distâncias do poliedro de coordenação sofreu uma redução de 0.239 Å para 0.170 Å, o que representou um grande avanço. Tanto que em alguns casos particulares como o da ligação Th(IV)-O, o erro foi reduzido em 40%. Desta maneira, os autores concluíram que tanto o modelo SMLC/AM1, quanto o SMAC se tornaram opções de metodologias alternativas com qualidade igual ou superior, em alguns casos, aos métodos *ab initio*.

1.6- Sparkle/PM3, 2006-2008

O método PM3^{36,37} foi desenvolvido pelo professor James Stewart em 1989, ou seja, em 2009 está completando 20 anos. Este método é mundialmente utilizado e bem conhecido principalmente por apresentar um erro médio de entalpia de formação para a maioria dos compostos menor que o método AM1. Isto motivou a criação do modelo Sparkle/PM3³⁸⁻⁴², cujos parâmetros para os íons Ln(III) estão publicados em artigos para lantânio e lutécio³⁸; praseodímio e itérbio³⁹; neodímio, promécio, samário⁴⁰; túlio⁴¹; disprósio, hólmio e érbio⁴².

A metodologia utilizada para esta parametrização foi a mesma utilizada no modelo SMLC/AM1, inclusive os conjuntos de 15 complexos cristalinos utilizados nas parametrizações foram exatamente os mesmos.

1.7- Análises Utilizando Método *Ab-Initio*

Atualmente, apenas duas abordagens de química quântica são utilizadas para o cálculo de geometrias de complexos com lantanídeos: (i) Modelos Sparkles/(AM1, PM3 e PM6⁴³); (ii) e por último o potencial efetivo de caroço ("*Effective Core Potential*, ECP") que pertence a classe *ab-initio*.

Nenhuma comparação entre os métodos *ab initio*/ECP e SMLC/AM1 havia sido realizada para que estes pudessem ser comparados. Desta maneira, Freire, R.O. e outros⁴⁴ realizaram cálculos *ab initio*/ECP com a finalidade de testar este método e se

surpreenderam com a exatidão e a viabilidade computacional do mesmo quando comparado ao modelo SMLC/AM1 utilizado pelo grupo naquele momento.

Cálculos com boa precisão utilizando métodos ECPs para otimização de geometria de complexos de lantanídeos, só poderão ser obtidos mediante a inclusão explícita de efeitos relativísticos.

O modelo Sparkle assume que os elétrons dos orbitais 4f não contribuem para a ligação química devido ao pequeno valor para a integral de recobrimento dos orbitais 4f do metal com os orbitais dos átomos coordenantes. Desta maneira, assume-se que a natureza de sua ligação química é puramente eletrostática. Já um ECP é a soma de funções potenciais que substituem os elétrons do caroço e os orbitais do caroço. Existem vários ECPs^{34,45-50} disponíveis para lantanídeos.

Estudos mostraram que cálculos utilizando funcional de densidade reproduzem aceitavelmente as geometrias para complexos de lantanídeos e actnídeos⁵¹⁻⁵⁷. De posse dessa informação Freire R. O. e outros⁵⁸ escreveram um trabalho comparando a precisão das medidas do poliedro de coordenação para diferentes complexos de lantanídeos utilizando o método ECP quase relativístico desenvolvido por Dolg, M. e colaboradores³⁴. Os autores combinaram os diferentes métodos de cálculo RHF, B3LYP, e MP2(*full*) com as seguintes bases: STO-3G, 3-21G, 6-31G, 6-31G* e 6-31+G. Esta série de cálculo foi realizada para um grupo de complexos contendo diferentes lantanídeos. A Equação 4 mais uma vez foi utilizada para avaliar as geometrias dos complexos calculados a partir de comparações com suas geometrias presentes no CSD.

O resultado encontrado neste estudo foi que o RHF/STO-3G/ECP apresentou-se como o mais eficiente para representar a geometria cristalográfica dos poliedros de coordenação dos complexos isolados de íons lantanídeos. Além disso, um outro resultado interessante é que o aumento do conjunto base ou a inclusão de correção eletrônica só aumentaram o desvio e agravaram a qualidade da predição das geometrias dos íons lantanídeos. Concluíram ainda que a recomendação de Cosentino⁵⁹ para se usar o método RHF/3-21G/ECP é justificada. Embora não seja necessariamente o mais exato dos métodos para prever geometrias experimentais.

1.8- Objetivo

O objetivo desta dissertação, é investigar se o emprego de métodos de estrutura eletrônica para a otimização de geometrias de complexos metálicos isolados é justificado para representar as geometrias de estruturas experimentais. Com esta finalidade, estudaremos um conjunto de 45 compostos contendo o íon $[\text{Mg}(\text{H}_2\text{O})_6]^{2+}$ com o objetivo de verificar se as variações sofridas nas coordenadas radial, latitudinal e longitudinal devido aos efeitos de contra-íons, outras moléculas, ou rede cristalina, são, ou significativas, ou sistemáticas, ou pseudo-aleatórias, em torno de uma média para cada tipo de coordenada. Caso se revelem como pseudo-aleatórias, a estrutura representada por este conjunto de valores médios é a que melhor representará este universo de 45 complexos.

Ainda determinaremos, caso os dados sejam pseudo-aleatórios nas coordenadas radial, latitudinal e longitudinal, os limites empíricos de exatidão para métodos semi-empíricos que busquem prever as geometrias experimentais a partir de cálculos nos complexos isolados utilizando métodos robustos de estatística.

Ainda, com a finalidade de avaliar a prática comum de se utilizar métodos *ab-initio* para o cálculo de geometrias de complexos de íons metálicos isolados, calcularemos a geometria do íon $[\text{Mg}(\text{H}_2\text{O})_6]^{2+}$ escolhendo um conjunto destes métodos e verificaremos seus erros em relação à estrutura que contenha os valores médios em torno das medidas pseudo-aleatórias das coordenadas r , θ e ϕ .

Uma vez conhecido o erro limite empírico, sugeriremos duas ferramentas que poderão auxiliar o parametrizador de métodos semi-empíricos a alcançar estes limites. A primeira será duas funções resposta equilibradas nas três coordenadas esféricas utilizando os desvios e variâncias obtidas a partir do estudo dos 45 compostos contendo hexaaquamagnésio. A segunda, uma nova técnica iterativa de se obter parâmetros iniciais baseada na observação direta do gráfico de energia de repulsão caroço-caroço em função da distância de ligações, alterando os valores dos parâmetros das gaussianas da equação. Com a finalidade de testar esta nova técnica observaremos em um banco de dados (contendo, além do íon magnésio, somente átomos parametrizados no modelo RM1) as distâncias entre o íon magnésio e os átomos coordenantes. A partir dos parâmetros preliminares obtidos com esta técnica para o método RM1, calcularemos as geometrias dos complexos contidos no banco de dados. Utilizaremos ainda os métodos

AM1, PM3 e PM6 calculando ao final o erro absoluto em termos das distâncias de ligações em relação ao banco de dados para cada método.

2. A VALIDAÇÃO DO MODELO SPARKLE E A FUNÇÃO RESPOSTA: “PLANETÁRIO”

2.1- Modelo SMLC/AM1, um Método com Boa Capacidade de Predição de Geometria

A grande diminuição nos erros médios absolutos do modelo Sparkle resultou num método com boa capacidade de predição da geometria do poliedro de coordenação do estado fundamental eletrônico de complexos cristalinos. Tanto que a exatidão deste modelo é suficiente para justificar sua aplicação à teoria do campo ligante para o estudo dos complexos de lantanídeos. Com o mesmo, pode-se calcular e prever propriedades espectroscópicas de complexos de íons lantanídeos, tais como, posição de níveis excitados singletos e tripletos e seus espectros eletrônicos. Isto permite que o pesquisador consiga construir equações de taxas que envolvem mecanismos de transferências de energia, o que permite determinar rendimentos quânticos de emissão teórico e a eficiência da luminescência dos complexos. No programa ZINDO⁶⁰ está implementado o método semi-empírico INDO/S^{61,62} (*Intermediate Neglect of Differential Overlap/Spectroscopic*) que permite realizar estes tipos de cálculos. Para compreendermos como uma boa geometria influencia na aplicação do modelo de campo ligante, apresentamos o operador de campo ligante que é representado pela Equação 6:

$$H_{CL} = \sum_{k,q,i} B_q^k C_q^k(\Omega_i), \quad (6)$$

onde o posto k que interessa aos lantanídeos deve ser par e menor ou igual a 6; $q = (-k, -k+1, \dots, 0, \dots, k-1, k)$; i representa cada um dos elétrons 4f; B_q^k são os parâmetros de campo ligante; C_q^k são os operadores de Racah; e Ω_i são as coordenadas angulares para os i elétrons 4f.

Para facilitar a aplicação deste operador, Malta introduziu o modelo de recobrimento simples em 1982^{63,64} com os seguintes pressupostos: a energia potencial dos elétrons da subcamada 4f depende da interação das cargas distribuídas uniformemente em pequenas regiões em torno da meia distância entre o íon central e o ligante; a carga desta região é definida a partir de $-g_j e \rho_j$, onde g_j é um fator de carga, e é a carga do elétron e ρ_j é a integral de recobrimento entre os orbitais 4f e a camada de valência do átomo ligante j coordenado ao íon lantanídeo. Neste modelo, os parâmetros de campo ligante são dados pela equação 7:

$$B_q^k = e^2 \langle r^k \rangle \left(\frac{4\pi}{2k+1} \right)^{\frac{1}{2}} \sum_j g_j \rho_j (2\varepsilon_j)^{k+1} \frac{Y_{k,q}^*(\Omega_j)}{d_j^{k+1}}, \quad (7)$$

onde $\langle r^k \rangle$ são as integrais radiais; ε é um parâmetro que considera o deslocamento do baricentro da região de recobrimento; $Y_{k,q}^*$ são os harmônicos esféricos dependentes de $\Omega \equiv \{\theta, \phi\}$; finalmente, d_j é a distância entre o íon metálico e o átomo coordenante j. O parâmetro ε pode ser aproximado pela expressão $\varepsilon \cong \frac{1}{1 \mp \rho_j}$. Assim, se ε estiver mais deslocado para o íon lantanídeo, utilizamos a operação de subtração e se estiver mais deslocado para o ligante utilizamos a operação de adição. Normalmente, ε é próximo de um, pois ρ_j para íons lantanídeos são muito pequenos.

Notamos que os operadores tensoriais de Racah, C_q^k na Equação 6 estão unicamente em função de Ω_j , enquanto os parâmetros do campo ligante na Equação 7 são diretamente proporcionais às coordenadas angulares Ω_j e inversamente proporcionais a d_j que está elevado à potência $k+1$. Desta maneira, a Equação 7, é sensível a variações em d_j pois havendo erro, este será propagado na ordem $d^{-(k+1)}$. Justamente pela boa exatidão do modelo SMLC/AM1^{18-23,25-29} em termos das distâncias obtidas entre o metal e os átomos coordenantes é que se têm obtido bons resultados na predição das propriedades espectroscópicas.

2.2- Justificativa para o Uso do Modelo Sparkle

Normalmente o modelo Sparkle é utilizado para calcular a geometria dos complexos de lantanídeos de forma isolada, não levando em consideração nem a

presença dos contra-íons ou de outras moléculas que possam existir na célula unitária, nem da estrutura cristalográfica como um todo. Nesta seção avaliaremos se isso é válido.

Em um banco de dados com estruturas cristalográficas contendo um mesmo íon metálico poderá haver: (i) complexos com diferentes números de coordenação; (ii) complexos com números de coordenação iguais, porém com poliedros de coordenação diferentes; (iii) complexos com o mesmo poliedro de coordenação formado por diferentes ligantes; (iv) complexos com os mesmos ligantes, porém dispostos espacialmente de formas diversas; (V) mesmos complexos, porém com diferentes contra-íons e/ou moléculas neutras na célula unitária do cristal.

Daremos nessa seção ênfase ao último grupo de complexos citado, pois gostaríamos de avaliar o quanto as diferentes moléculas neutras e/ou contra-íons contidos nas células unitárias dos cristais afetam a geometria do poliedro de coordenação como um todo, mais especificamente as distâncias do íon metálico central aos vértices do poliedro de coordenação, bem como seus ângulos. Estas variáveis geométricas descrevem a região cuja geometria o modelo Sparkle busca prever.

Nosso objetivo aqui não é o de descobrir como descrever o efeito do contra-íon, mas ter uma noção da extensão deste efeito na geometria do complexo.

Sabemos que o modelo Sparkle não foi parametrizado incluindo de forma explícita o efeito individual destes contra-íons e/ou moléculas neutras no complexo. Desta forma, os parâmetros do modelo Sparkle foram obtidos para levar a uma única estrutura de um dado complexo para todos os que possuem diferentes moléculas neutras e/ou contra-íons na célula unitária do cristal.

2.3- O Estudo de um Conjunto de 45 Células Unitárias de Hexaaquamagnésio

Com o objetivo de validar o modelo Sparkle, propomos o estudo inédito do caso de todos os complexos catiônicos hexaaquamagnésio contidos no CSD, para inferir como e com que intensidade o contra-íon influencia o complexo cristalino.

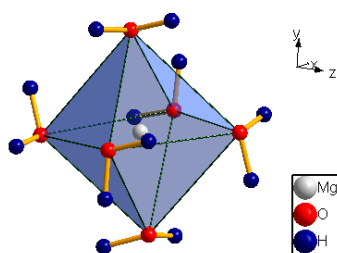


Figura 4 - Representação esquemática em bolas e bastões do complexo catiônico cristalino hexaaquamagnésio. Encontrado em 45 diferentes ambientes químicos no CSD.

Para esta análise, encontramos um conjunto de células unitárias de 45 diferentes cristais contendo o mesmo complexo catiônico, hexaaquamagnésio, porém com diferentes moléculas neutras e/ ou contra-íons, tal como apresentado na Tabela 3. O objetivo deste estudo é determinar o quanto e como estes complexos são assim afetados pela presença destes em termos da geometria do complexo como um todo.

Tabela 3- Representação em duas dimensões das estruturas, com seus respectivos códigos advindos do CSD, que representam as moléculas e/ou contra íons presentes nas células unitárias de cada hexaaquamagnésio.

AQMEDA	BIHLIO	BIHNEM	BIKPUG
CIRVAA01	CIRVAA	DAMZAT	EXUBOO

Tabela 3- Continuação

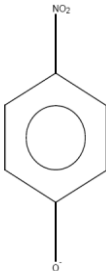
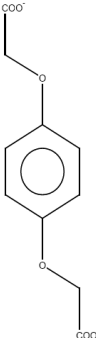
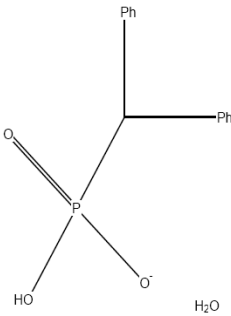
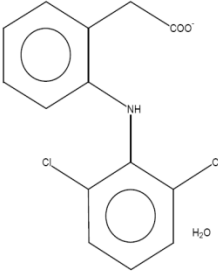
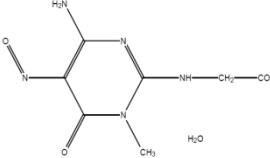
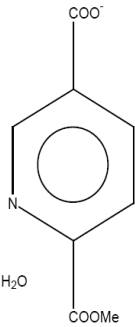
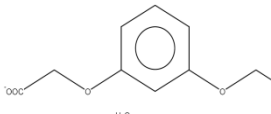
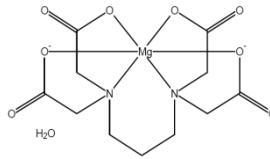
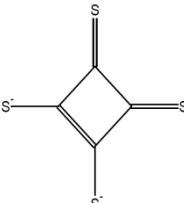
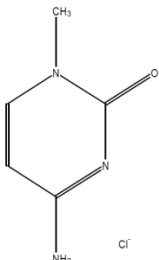
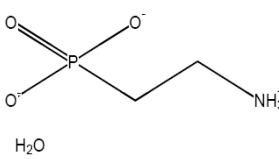
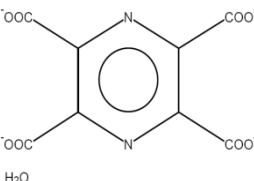
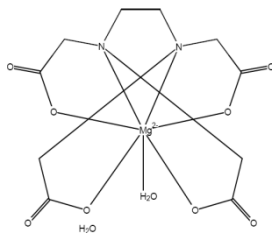
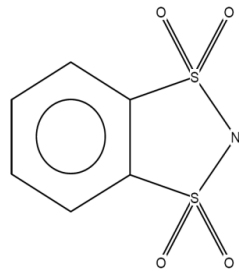
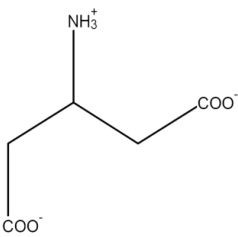
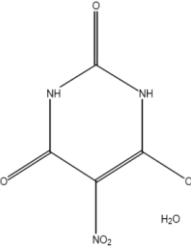
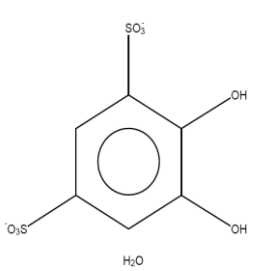
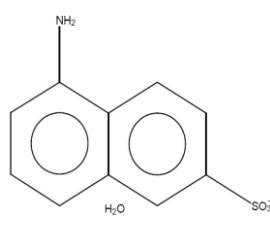
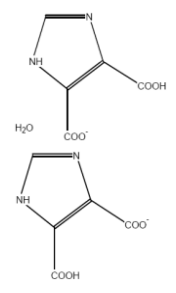
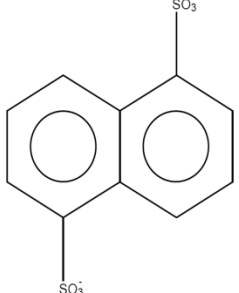
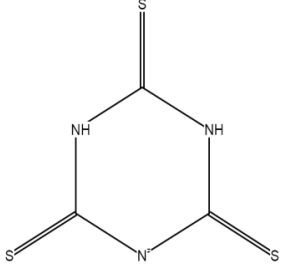
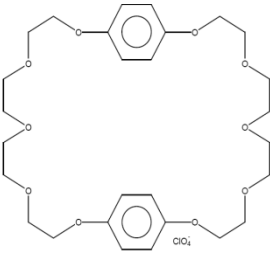
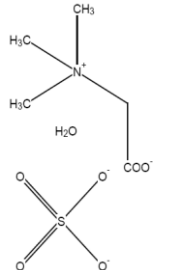
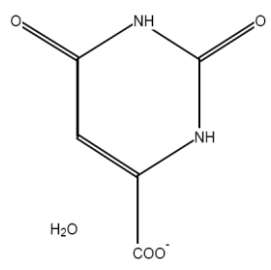
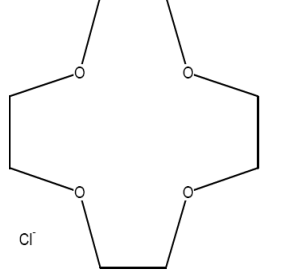
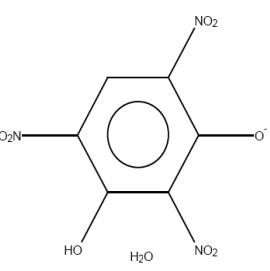
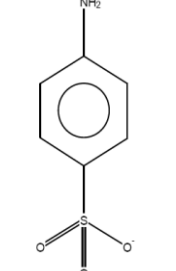
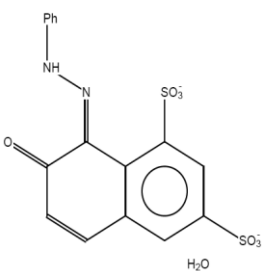
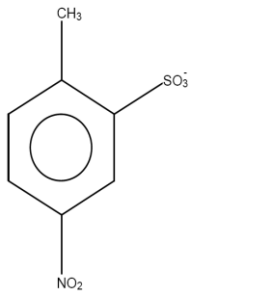
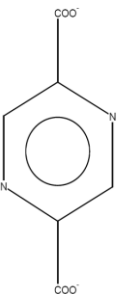
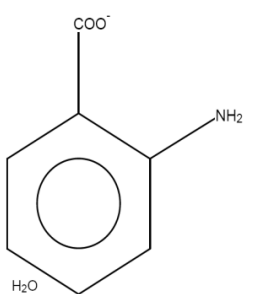
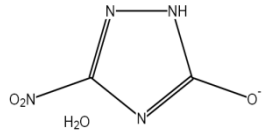
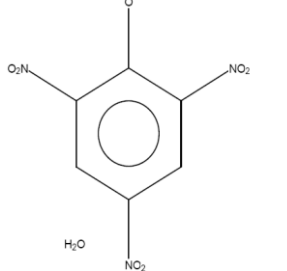
			
H ₂ O	H ₂ O		H ₂ O
FEDQUA01	FEDQUA	FETHAO	GATLUI
			
	H ₂ O		H ₂ O
GOLPIG	GUHHOC	HETZUB	ILUXAO
			
H ₂ O	H ₂ O		Cl ⁻
IWOYEV	JAMQAQ	JEYCEW	KIDLUF
			
H ₂ O	H ₂ O	H ₂ O	
KIMNID	MAVDOD	MGEDTA01	MGNTSP
			
	COO ⁻	H ₂ O	H ₂ O
MINKOJ	MISXER	MOYTAV	NOGLAW

Tabela 3- Continuação

			
OMARID	POHGIC	QERJIH	RAKYOR
			
RAMXOS	RATRIN	REFTEC	SIMZUJ
			
TOXDMG	TUXYOA	VOPCEI	WIKXAP
			
XEZCIP	XECJIY	XITDUZ	XUXLUX
			
ZARMEK			

Porém, para que este estudo possa ser realizado, devemos antes de tudo definir um sistema de coordenadas atômicas padrão para todos os hexaaquamagnésio. Pois as geometrias dos complexos advindas do CSD estão em coordenadas cartesianas e os dados que queremos devem estar em um sistema de coordenadas polares. Para estabelecer o mesmo eixo de coordenadas cartesianas para todos os 45 diferentes complexos hexaaquamagnésio para então poder efetuar as comparações, convencionamos que o eixo de cada estrutura química deve primeiramente ter por origem as coordenadas do íon magnésio. Em seguida, numeramos cada átomo de oxigênio segundo a sua posição no poliedro, como observamos na Figura 5:

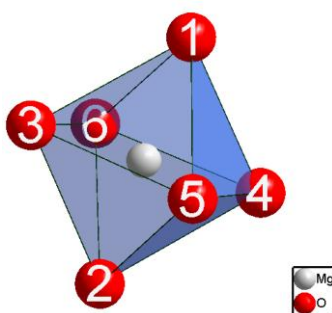


Figura 5 - Representação esquemática do poliedro de coordenação perfeito do hexaaquamagnésio. O sistema de coordenadas do poliedro será definido a partir do íon magnésio, sendo que o átomo 1 define a direção e o sentido positivo do eixo Z. A direção e o sentido positivo do eixo X é definido na direção do átomo 6. Por fim, a direção e o sentido positivo do eixo Y é definido na direção do átomo 3.

Vamos, portanto, exprimir as coordenadas dos vértices do poliedro de coordenação em um sistema de coordenadas esférico-polares em cuja origem situa-se o íon metálico. Assim, as coordenadas de cada vértice são definidas em função das variáveis d , θ e φ , como indicado na Figura 6.

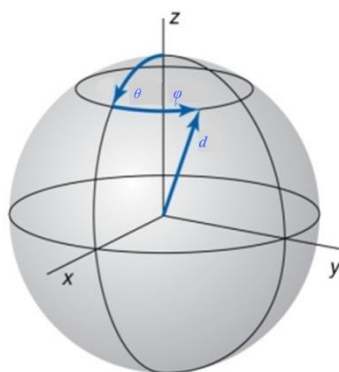


Figura 6 - Representação de um sistema de coordenadas polares, onde d é a coordenada radial, θ a latitudinal e φ a coordenada longitudinal.

Os intervalos onde as coordenadas d , θ e φ estão definidas são:

$$0 \leq d \leq \infty \quad 0 \leq \theta \leq \pi \quad 0 \leq \varphi \leq 2\pi. \quad (8)$$

Na Figura 5, os átomos 1 e 2 são os átomos axiais do poliedro. Definimos a direção e sentido positivo do eixo Z a partir da ligação Mg-O₁. O eixo X deve estar no plano que contém os átomos Mg, O₁ e O₆, este deve ser ortogonal ao eixo Z definido em Mg-O₁ e o sentido positivo é definido a partir da direção do vetor que representar a ligação Mg-O₆. Por último determinamos o eixo Y utilizando a regra da mão direita.



Figura 7 – Representação esquemática da regra da mão direita, que define o sistema de eixos perpendiculares único centrado no metal do poliedro de coordenação. Dispondo o polegar da mão direita ao longo da ligação química de maior comprimento, define-se o eixo positivo Z. O sentido positivo do eixo X é definido pelo dedo indicador quando este está estendido. O último eixo a ser definido, o Y, tem seu sentido positivo indicado pelo dedo médio quando flexionado.

Para encontrar estes eixos ortogonais, aplicamos o método de Gram-Schmidt. Definido o novo sistema de coordenadas cartesianas, encontramos os valores de d , θ e φ no sistema de coordenadas esféricas para cada átomo de oxigênio em cada vértice do poliedro de coordenação tal como aparece na Figura 8, abaixo.

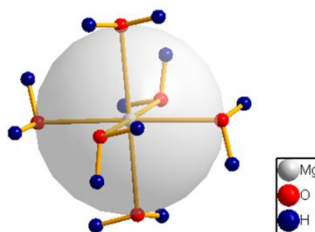


Figura 8 – Representação esquemática num sistema de coordenadas polares, onde são aferidos os valores das coordenadas radiais, latitudinais e longitudinais para cada átomo coordenante do poliedro de coordenação do hexaaquamagnésio.

Para encontrar estes valores de d , θ e φ para os átomos nos vértices do poliedro, utilizamos as seguintes equações:

$$d = \sqrt{x^2 + y^2 + z^2}; \quad \theta = \tan^{-1} \left(\frac{\sqrt{x^2 + y^2}}{z} \right); \quad \varphi = \tan^{-1} \left(\frac{y}{x} \right) \quad (9)$$

Nas próximas linhas apresentaremos os valores aferidos para as coordenadas radiais latitudinais e longitudinais respectivamente para os complexos hexaaquamagnésio contido nas 45 células unitárias dos contra-íons e/ou moléculas apresentadas nas Tabelas 4.

Abaixo estão indicados todos os valores de d , que representam a distância entre o Mg e o oxigênio calculadas considerando que o centro de origem das coordenadas polares esteja sobre o Magnésio.

Tabela 4- Comprimento das ligações químicas Mg-O para o poliedro de coordenação hexaaquamagnésio, onde a primeira linha designa os números que identificam os átomos representados no poliedro de coordenação da Figura 5. As últimas linhas indicam a média e a mediana de cada coluna.

Estrutura	1	2	3	4	5	6
AQMEDA	2.08729	2.08729	2.08670	2.08670	2.05311	2.05311
BIHLIO	2.06751	2.06751	2.06038	2.06038	2.04463	2.04463
BIHNEM	2.07211	2.07211	2.05253	2.05253	2.05155	2.05155
BIKPUG	2.08444	2.08413	2.07909	2.08000	2.06549	2.06522
CIRVAA01	2.07080	2.07080	2.06781	2.06781	2.05794	2.05794
CIRVAA	2.05905	2.05905	2.05153	2.05153	2.03583	2.03583
DAMZAT	2.11711	2.11711	2.06483	2.06483	2.02967	2.02967
EXUBOO	2.08935	2.08794	2.06289	2.06277	2.05168	2.05119
FEDQUA01	2.11242	2.11242	2.08943	2.08943	2.02003	2.02003
FEDQUA	2.11188	2.11188	2.08863	2.08805	2.02043	2.01983
FETHAO	2.12763	2.12756	2.06737	2.06737	2.04338	2.04307
GATLUI	2.07888	2.07849	2.07080	2.07080	2.04746	2.04712
GOLPIG	2.06451	2.06451	2.05656	2.05656	2.04810	2.04810
GUHHOG	2.09535	2.09475	2.06275	2.06275	2.04341	2.04325
HETZUB	2.08581	2.08581	2.07521	2.07521	2.06096	2.06096
ILUXAO	2.08886	2.08886	2.06391	2.06391	2.02645	2.02645
IWOYEE	2.11531	2.11448	2.05122	2.05877	2.05420	2.04867
JAMQAQ	2.07474	2.06741	2.06760	2.07388	2.04963	2.04892
JEYCEW	2.08404	2.05813	2.08332	2.05813	2.04312	2.04293
KIDLUF	2.11581	2.11581	2.09113	2.09113	2.06855	2.06855
KIMNID	2.07147	2.07147	2.06719	2.06719	2.04110	2.04110
MAVDOD	2.11323	2.03581	2.04354	2.06581	2.07469	2.05194
MGEDTA01	2.08702	2.08702	2.05132	2.05132	2.04692	2.04692
MGNTSP	2.07988	2.07988	2.07680	2.07680	2.02231	2.02231
MINKOJ	2.08585	2.08562	2.06276	2.06256	2.04461	2.04419
MISXER	2.09109	2.09109	2.07446	2.07420	1.99973	1.99884
MOYTAV	2.06215	2.06215	2.05367	2.05367	2.03902	2.03902
NOGLAW	2.06240	2.06176	2.05860	2.05927	2.04678	2.04578
OMARID	2.08837	2.05233	2.07511	2.05373	2.08007	2.02768
POHGIC	2.06626	2.06614	2.05233	2.05332	2.03970	2.03834
QERJIH	2.12298	2.11705	2.04519	2.03966	2.03723	2.03247
RAKYOR	2.10355	2.10288	2.06008	2.06054	2.03662	2.03662
RAMXOS	2.10050	2.06591	2.08783	2.04860	2.03579	2.03009
RATRIN	2.10237	2.06530	2.05127	2.10114	2.05091	2.01501
REFTEC	2.07862	2.07384	2.05615	2.06694	2.04730	2.04232
SIMZUJ	2.09461	2.09461	2.08258	2.08258	2.02478	2.02478
TOXDMG	2.08118	2.08118	2.05826	2.05910	2.05800	2.05400
TUXYOA	2.09563	2.09554	2.05490	2.05391	2.02595	2.02595
VOPCEI	2.08813	2.08813	2.05959	2.05959	2.02517	2.02517
WIKXAP	2.11223	2.11195	2.03833	2.03853	2.02482	2.02393
XECJIY	2.11223	2.11195	2.03833	2.03853	2.02482	2.02393
XEZCIP	2.08199	2.08082	2.06428	2.06274	2.03628	2.03541

Tabela 4- Continuação

XITDUZ	2.09629	2.09629	2.05491	2.05424	2.05321	2.05299
XUXLUX	2.08780	2.08694	2.05669	2.05669	2.04558	2.04558
ZARMEK	2.07682	2.05595	2.04904	2.06110	2.06412	2.04296
Média	2.08990	2.08417	2.06371	2.06410	2.04314	2.03943
Mediana	2.08780	2.08562	2.06275	2.06256	2.04461	2.04293

Para θ encontramos os seguintes valores para os oxigênios no poliedro de coordenação de acordo com a numeração dada na Figura 5:

Tabela 5- Ângulos θ dos oxigênios no poliedro de coordenação do hexaaquamagnésio, onde a primeira linha designa os números que identificam os átomos representados no poliedro de coordenação da Figura 5. As últimas linhas indicam a média e a mediana de cada coluna.

Estrutura	1	2	3	4	5	6
AQMEDA	0	180.000	93.164	86.836	91.450	88.550
BIHLIO	0	180.000	90.226	89.774	87.705	92.295
BIHNEM	0	180.000	89.596	90.404	86.931	93.069
BIKPUG	0	179.974	88.821	91.170	91.736	88.255
CIRVAA01	0	180.000	88.303	91.697	90.689	89.311
CIRVAA	0	180.000	91.063	88.937	89.581	90.419
DAMZAT	0	180.000	94.368	85.632	91.411	88.589
EXUBOO	0	179.997	89.137	90.824	93.099	86.862
FEDQUA01	0	180.000	90.501	89.499	89.702	90.298
FEDQUA	0	180.000	90.400	89.568	89.610	90.357
FETHAO	0	179.962	86.917	93.083	87.967	92.050
GATLUI	0	179.975	90.112	89.888	90.911	89.078
GOLPIG	0	180.000	90.649	89.351	88.928	91.072
GUHHOG	0	179.978	93.322	86.678	87.070	92.913
HETZUB	0	180.000	86.009	93.991	89.051	90.949
ILUXAO	0	180.000	87.410	92.590	88.013	91.987
IWOYEV	0	178.389	91.186	89.062	90.511	89.755
JAMQAQ	0	177.498	88.390	89.197	90.993	91.850
JEYCEW	0	175.516	92.644	90.124	87.272	88.137
KIDLUF	0	180.000	87.534	92.466	91.281	88.719
KIMNID	0	180.000	92.248	87.752	89.271	90.729
MAVDOD	0	173.139	89.933	90.650	91.118	84.339
MGEDTA01	0	180.000	87.418	92.582	90.448	89.552
MGNTSP	0	180.000	90.524	89.476	87.079	92.921
MINKOJ	0	179.962	89.251	90.775	90.312	89.718
MISXER	0	180.000	90.297	89.697	95.088	84.903
MOYTAV	0	180.000	90.804	89.196	88.971	91.029
NOGLAW	0	179.965	88.264	91.718	87.212	92.788
OMARID	0	177.177	86.501	88.685	87.974	89.571

Tabela 5- Continuação

POHGIC	0	179.961	88.879	91.133	88.867	91.146
QERJIH	0	170.360	87.353	90.330	84.491	97.000
RAKYOR	0	179.980	90.131	89.850	91.834	88.166
RAMXOS	0	172.592	92.183	85.867	87.014	92.159
RATRIN	0	172.641	89.667	89.803	94.363	85.717
REFTEC	0	178.713	89.439	90.325	88.478	89.733
SIMZUJ	0	180.000	89.366	90.634	89.613	90.387
TOXDMG	0	177.742	91.748	88.228	88.871	91.129
TUXYOA	0	179.973	92.613	87.383	87.646	92.354
VOPCEI	0	180.000	91.076	88.924	88.970	91.030
WIKXAP	0	179.974	88.138	91.854	88.663	91.347
XECJIY	0	179.974	88.138	91.854	88.663	91.347
XEZCIP	0	179.978	92.550	87.428	87.524	92.466
XITDUZ	0	180.000	92.021	87.960	90.974	89.007
XUXLUX	0	179.986	88.045	91.955	92.952	87.048
ZARMEK	0	178.702	90.764	89.120	88.828	88.285
Média	0	178.936	89.936	89.866	89.536	90.186
Mediana	0	179.980	90.112	89.803	89.051	90.387

Sobre a coluna 1 encontramos 0 porque o átomo O_1 define a origem de θ , enquanto o O_2 é o outro átomo axial e sendo assim as medidas de seus ângulos deveriam coincidir com o limite de θ que é 180° , ou estar muito próximo deste. Esta variação dos ângulos é justamente devido à presença do contra íon e /ou molécula.

A última tabela que apresentaremos segue abaixo e nela estão descritos os valores aferidos para as coordenadas longitudinais em cada átomo de O.

Tabela 6- Ângulos φ dos oxigênios no poliedro de coordenação hexaaquamagnésio, onde a primeira linha designa os números que identificam os átomos representados no poliedro de coordenação da Figura 5. As últimas linhas indicam a média e a mediana de cada coluna.

Estrutura	3	4	5	6
AQMEDA	89.5809	269.5809	180.0000	0
BIHLIO	89.0794	269.0794	180.0000	0
BIHNEM	90.7568	270.7568	180.0000	0
BIKPUG	87.0181	267.0117	180.0252	0
CIRVAA01	89.3130	269.3130	180.0000	0
CIRVAA	89.6366	269.6366	180.0000	0
DAMZAT	87.9605	267.9605	180.0000	0
EXUBOO	92.0171	272.0176	179.9745	0
FEDQUA01	93.2257	273.2257	180.0000	0
FEDQUA	93.2499	273.2347	179.9842	0
FETHAO	92.3891	272.3891	180.0203	0

Tabela 6- Continuação

GATLUI	90.6016	270.6016	179.9761	0
GOLPIG	90.1690	270.1690	180.0000	0
GUHHOG	91.2746	271.2746	180.0223	0
HETZUB	87.1639	267.1639	180.0000	0
ILUXAO	86.7466	266.7466	180.0000	0
IWOYEV	92.9127	271.2898	178.1633	0
JAMQAQ	88.2725	268.9879	177.1470	0
JEYCEW	87.2053	270.7155	175.2263	0
KIDLUF	89.1355	269.1355	180.0000	0
KIMNID	91.2168	271.2168	180.0000	0
MAVDOD	94.3587	268.9723	182.2518	0
MGEDTA01	89.5487	269.5487	180.0000	0
MGNTSP	92.2696	272.2696	180.0000	0
MINKOJ	89.2649	269.2745	179.9767	0
MISXER	85.7967	265.7708	179.9907	0
MOYTAV	91.3375	271.3375	180.0000	0
NOGLAW	89.6393	269.6099	179.9969	0
OMARID	89.5992	268.3965	175.9919	0
POHGIC	89.2204	269.1825	180.0273	0
QERJIH	88.5866	275.9458	183.0946	0
RAKYOR	91.6099	271.6264	180.0000	0
RAMXOS	91.5461	269.9138	180.2917	0
RATRIN	90.3990	268.7417	178.4262	0
REFTEC	90.2231	270.4101	178.4534	0
SIMZUJ	86.5103	266.5103	180.0000	0
TOXDMG	90.2573	269.8122	180.0000	0
TUXYOA	90.8218	270.8212	180.0000	0
VOPCEI	89.0466	269.0466	180.0000	0
WIKXAP	92.1513	272.1250	180.0293	0
XECJIY	92.1513	272.1250	180.0293	0
XEZCIP	92.0582	272.0535	179.9902	0
XITDUZ	88.6644	268.6577	179.9804	0
XUXLUX	88.5019	268.5019	180.0000	0
ZARMEK	87.9445	270.0243	181.0273	0
Média	90.0096	270.0486	179.7799	0
Mediana	89.6393	269.8122	180.0000	0

No sistema de coordenadas convencionado para o estudo deste universo de complexos, as medidas de φ para os O_1 e O_2 são indeterminadas, pois o oxigênio 1 se encontra sobre o eixo Z, enquanto o oxigênio 2 está a 180° , ou muito próximo deste. Sendo assim uma pequena variação neste ângulo poderá causar uma grande diferença

nas medidas destas coordenadas longitudinais. Assim, as medidas de O_2 para φ serão também consideradas como indeterminadas neste estudo.

Sendo o sentido positivo do eixo X aquele que define a origem das medidas de φ , o O_7 necessariamente será igual a zero, pois o vetor que representa a sua ligação com o magnésio é coplanar ao plano X-Y.

O vetor que representa a ligação Mg- O_5 - como observamos na Tabela 6 - é oposto ao O_6 , já que o hexaaquamagnésio é um complexo altamente simétrico e a medida de 180° se repete muitas vezes. Por isto só avaliamos os φ dos átomos 3 e 4 que não estão envolvidos na definição das coordenadas.

Desta maneira, observamos que esta última tabela possui menos colunas para serem avaliadas devido à própria definição do sistema de coordenadas para um complexo que, se perfeito, possuiria simetria O_h .

Para avaliar os dados destas três últimas tabelas utilizaremos algumas ferramentas de estatística que estarão descritas na próxima seção.

2.4- Tratamento Estatístico

Se o modelo Sparkle é válido, então a estrutura que o mesmo propõe para o complexo hexaaquamagnésio deverá ser uma boa representação de todas as 45 estruturas anteriormente citadas. Assim, a influência dos contra-íons e/ou moléculas nas respectivas células unitárias deveria ou ser não significativa e menor do que o erro do modelo, ou então se comportar de forma pseudo aleatória – o que daria validade ao uso de uma estrutura média para representar o conjunto das 45 estruturas como um todo. Para quantificar isto, precisamos de técnicas estatísticas específicas para o nosso caso.

2.4.1- Testes de Hipótese Estatística

A seguir descreveremos brevemente os testes de hipóteses que utilizamos em nosso estudo para verificar a aleatoriedade dos dados e as distribuições aplicáveis em cada caso respectivamente, todos disponíveis no software livre R ⁶⁵ versão 2.8.1.

Definimos o teste de hipóteses como sendo uma regra de decisão usada para rejeitar ou aceitar hipóteses estatísticas com base em elementos amostrais. As hipóteses sempre deverão ser duas: (i) H_0 , que é conhecida como hipótese nula; (ii) H_1 , que é a

hipótese alternativa. A hipótese nula (H_0) é aquela que tem a suposição que o pesquisador deseja provar e a hipótese alternativa é aquela que rejeita a hipótese nula. Desta maneira quando aceitamos uma hipótese automaticamente rejeitamos a outra.

Nesta dissertação trabalharemos com três distribuições: gama, beta e von Mises. Para exemplificar como funciona um teste de hipóteses, escolhemos a distribuição normal, pelo fato de que o teorema central do limite⁶⁶ garante que toda distribuição que possui os dois primeiros momentos (média e variância) é bem aproximada pela distribuição normal. Apresentamos abaixo os seguintes testes de hipóteses comentados: (i) unilateral à direita; (ii) unilateral à esquerda; (iii) ou bicaudais.

1 – Teste unilateral à direita:

$$H_0: \zeta = \zeta_0$$

$$H_1: \zeta > \zeta_0$$

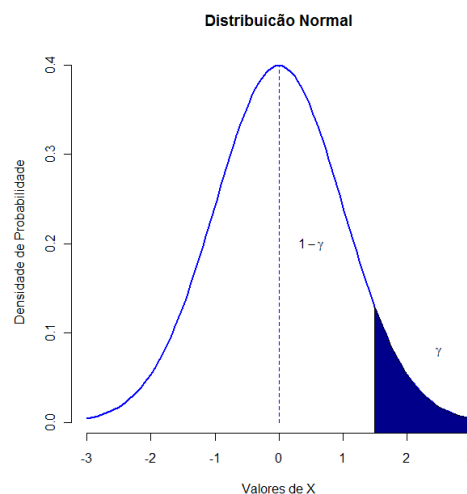


Figura 9 - Gráfico da distribuição normal unicaudal à direita. Onde γ é o nível significância do teste e $1-\gamma$ é o nível de confiança do teste.

No gráfico acima, γ é a região crítica ou região de rejeição (da hipótese nula) e $(1-\gamma)$ é a região não-crítica ou região de aceitação. O sinal de maior no teste indica que os valores de x , que são as observações, devem unicamente ser positivos. A fronteira entre essas regiões será dada por um valor tabelado, de acordo com a distribuição da estatística de teste sob a hipótese nula. Desse modo, se o valor calculado for maior que o valor tabelado, rejeitamos $H_0: \zeta = \zeta_0$, onde ζ é um parâmetro qualquer e ζ_0 é um pseudo valor.

2 – Teste unilateral à esquerda:

$$H_0: \zeta = \zeta_0$$

$$H_1: \zeta < \zeta_0$$

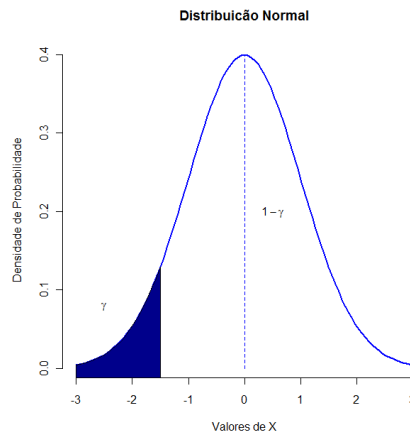


Figura 10 - Gráfico da distribuição normal unicaudal à esquerda.

Para o caso unilateral à esquerda temos a mesma sistemática do caso anterior, entretanto a região crítica agora se posiciona à esquerda do gráfico. A fronteira entre essas regiões será dada por um valor tabelado, de acordo com a distribuição da estatística de teste sob a hipótese nula. Desse modo, se o valor calculado for maior que o valor tabelado, rejeitamos $H_0: \zeta = \zeta_0$.

3 – Teste bilateral:

$$H_0: \zeta = \zeta_0$$

$$H_1: \zeta \neq \zeta_0$$

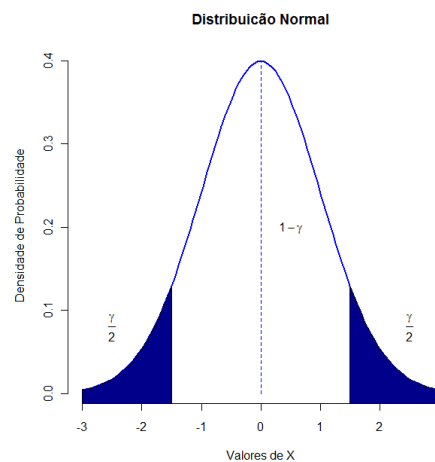


Figura 11 - Gráfico da distribuição normal bicaudal.

No caso bilateral a região crítica será dada por $\frac{\gamma}{2}$ em ambas as caudas, pois aqui temos um sinal de diferente ($\neq$) na hipótese alternativa. Note que na hipótese nula sempre temos uma igualdade ($=$) e na hipótese alternativa uma desigualdade ($<$, $>$ ou $\neq$).

Na presente dissertação serão usados três tipos de testes de hipóteses: o primeiro é o teste de Bartels⁶⁷ para testar a aleatoriedade dos dados; o segundo é o teste de Kolmogorov-Smirnov⁶⁸ para testar a especificidade da função de distribuição dos dados e o último é o teste de Watson⁶⁹ para testar a especificidade da distribuição em dados circulares. Apresentaremos na próxima seção a definição de cada um destes.

De maneira geral, escolhemos o valor do nível de significância para os testes aqui realizados como sendo igual a 5%. Desse modo, a verificabilidade da aceitação ou não da hipótese nula será dada pelo p – *valor*, ou seja, valores de $p < 0.05$ indicam a rejeição da hipótese nula.

2.4.2- Definição para os Testes de Aleatoriedade e Especificidade das Distribuições

- **Teste de Bartels**

O problema de testes de aleatoriedade numa seqüência de observações é de extrema importância na área de estatística. Neste trabalho, utilizamos o teste de Bartels para verificar a aleatoriedade dos valores de **d** e **θ** para as ligações químicas do poliedro de coordenação.

O teste de Bartels é baseado em métodos não-paramétricos, ou seja, é um teste que não faz suposições prévias sobre a distribuição real, utilizando ao invés os dados disponíveis para gerar uma distribuição empírica, que representa uma aproximação à distribuição real. O teste de Bartels é baseado nas seguintes hipóteses:

H₀ - A amostra é aleatória

H₁ - A amostra não é aleatória

Considere R_i o rank (ordem) da i -ésima observação numa seqüência de T observações. Assim, o teste de Bartels é baseado na seguinte estatística de teste:

$$RVN = \frac{\sum_{i=1}^{T-1} (R_i - R_{i+1})^2}{\sum_{i=1}^T (R_i - \bar{R})^2} \quad (10)$$

Sob condições de aleatoriedade (sob $\mathbf{H}_0$) o denominador de RVN (*rank version of von Neumann's ratio*) é uma constante igual a $T(T^2 - 1)/12$. Desta maneira, temos o seguinte numerador (NM) de RVN,

$$NM = \sum_{i=1}^{T-1} (R_i - R_{i+1})^2, \quad (11)$$

que é equivalente à estatística de teste.

Para a realização do teste, usamos a função *bartels.test* do software estatístico R 2.8.1, onde esta função realiza o teste de aleatoriedade Bartels, que é baseado na RVN⁶⁷. Os pontos críticos da distribuição empírica são tabelados em Bartels⁶⁷.

- **Teste de Kolmogorov-Smirnov**

Este teste é usado para testar a hipótese de se um conjunto de dados provém de uma distribuição especificada. Suponha que tenhamos uma amostra $X_1, \dots, X_n$ de uma população P , sobre a qual estamos considerando uma variável aleatória (v.a) X . Designamos por $f(x)$ a função densidade e por $F(x)$ a função de distribuição acumulada (f.d.a) de X . Estimar $f(x)$ é equivalente a estimar $F(x)$, onde o objetivo é testar se a amostra observada veio de uma distribuição de probabilidade especificada. A seguir temos as seguintes hipóteses:

$$\mathbf{H}_0: F(x) = F_0(x), \text{ para todo } x$$

$$\mathbf{H}_1: F(x) \neq F_0(x)$$

Em vez de considerar o histograma como estimador de $f(x)$, considera-se a *função de distribuição empírica* (f.d.e), $F_e(x)$, definida abaixo, como estimador de $F(x)$, para todo valor de $x \in \mathfrak{R}$.

Dadas n observações de uma variável quantitativa e um número x real qualquer, indicar-se-á por $N(x)$ o número de observações menores ou iguais a x , e chamar-se-á de função de distribuição empírica (*f.d.e*) a função $F_n(x)$ ou $F_e(x)$.

$$F_e(x) = F_n(x) = \frac{N(x)}{n} \quad (12)$$

Os probabilistas Kolmogorov e Smirnov propuseram uma estatística para o teste obtida tomando o máximo dos valores absolutos das diferenças $F(x) - F_e(x)$, $i=1, \dots, n$. Nessas diferenças, calculadas nos valores amostrais, $F(x_i)$ é o valor calculado sob a hipótese nula H_0 , ou seja, é o valor que a *f.d.a* hipotetizada toma no ponto x_i . Formalmente, a estatística a ser usada no teste é

$$D = \max_{1 \leq i \leq n} |F(x_i) - F_e(x_i)| \quad (13)$$

O valor encontrado deve ser comparado com um valor crítico, obtido na distribuição X , fixado o nível de significância do teste (γ). Se D é maior que o valor tabelado, rejeitamos H_0 .

- **Teste de Watson**

Realiza o teste de homogeneidade entre duas amostras de dados circulares (dados circulares são dados que pertencem a intervalos de comprimento 2π). Valores críticos para a estatística de teste são obtidos utilizando distribuição assintótica da estatística de teste. É recomendado utilizar valores críticos obtidos junto com seus respectivos intervalos de confiança para amostras superiores a 17. Os quadros estão disponíveis para pequenas amostras e podem ser encontrados em Mardia⁶⁹.

2.4.3- Distribuições Estatísticas Usadas

- **Distribuição Gama**

A distribuição Gama tem como suporte o intervalo de zero a infinito ($[0; \infty)$). A distribuição Gama, já foi utilizada no modelo Sparkle/PM3³⁸⁻⁴², para avaliar a qualidade dos parâmetros. Temos que a função de densidade de probabilidade da distribuição Gama é dada por

$$f(x; \lambda, \beta) = \frac{\beta^\lambda}{\Gamma(\lambda)} x^{\lambda-1} e^{-\beta x}, \quad x \geq 0, \quad \lambda, \beta > 0 \quad (14)$$

Nesta equação a variável aleatória x com limite $[0; \infty)$, designa os valores dos EMAs. O termo λ é o parâmetro de forma da distribuição, β é o parâmetro de escala, ambos devem ser positivos. Temos para a densidade Gama que seu valor esperado é dado por, $E(X) = \frac{\lambda}{\beta}$, e a variância, $Var(X) = \frac{\lambda}{\beta^2}$. E ainda temos que função Gama, é definida pela integral:

$$\Gamma(\lambda) = \int_0^\infty t^{\lambda-1} e^{-t} dt \quad (15)$$

A estimação dos parâmetros da distribuição gama não tem forma fechada, desse modo temos que usar métodos de estimação não linear, por meio das técnicas de Newton Raphson, entre outros. A distribuição é muito utilizada em dados médicos para modelar tempo de vida. Uma das maiores dificuldades encontradas em trabalhos que envolvem a distribuição gama é a estimação de seus parâmetros λ e β , devido à complexidade e extensão dos cálculos envolvidos. São muitos os métodos que podem ser usados como: o método dos quadrados mínimos; o método dos momentos; e o da máxima verossimilhança, que é o método que utilizaremos neste trabalho. De uma maneira geral, todos possuem suas limitações, seja por razões matemáticas ou por produzirem estimativas ineficientes. O método dos quadrados mínimos apresenta várias dificuldades quando utilizado na distribuição gama e por isso não é recomendado. Normalmente, os métodos dos momentos são os mais comumente utilizados, porém

segundo Thom⁷⁰, deve-se preferir o método da máxima verossimilhança devido às suas melhores propriedades.

- **Distribuição Beta**

A distribuição de probabilidade do tipo Beta é da família das distribuições contínuas e possui dois parâmetros de forma: a e b . A seguir apresentamos a equação da função de densidade da probabilidade dada por:

$$f(x, a, b) = \frac{1}{B(a, b)} x^{a-1} (1-x)^{b-1}, 0 < x < 1, a, b > 0, \quad (16)$$

cujo valor esperado é dado por, $E(X) = \frac{a}{a+b}$, e a variância é dada por

$Var(X) = \frac{ab}{(a+b+1)(a+b)^2}$, a função $Beta(a, b)$ é dada por:

$$B(a, b) = \int_0^1 x^{a-1} (1-x)^{b-1} dx \quad (17)$$

A função $Beta(a, b)$ aparece como uma constante de normalização para assegurar que a integral da densidade $f(x, a, b)$ seja igual a um. A distribuição Beta é geralmente usada para dados entre zero e um, por exemplo taxas e proporções.

- **Distribuição von Mises**

A distribuição de von Mises também conhecida como distribuição normal circular, é uma distribuição de probabilidade contínua no círculo; seu suporte é dado em qualquer intervalo de comprimento 2π . Pode ser entendida como o análogo circular à distribuição normal. Seu uso é justificado em estatística direcional.

Considerando x uma variável angular aleatória, é também comum ver a distribuição de von Mises como uma distribuição de números complexos $z = e^{ix}$ ao invés de uma distribuição de números reais x .

A função de densidade de probabilidade da distribuição von Mises, $f(x; \mu, \kappa)$ é dada por:

$$f(x; \mu, \kappa) = \frac{e^{\kappa \cos(x-\mu)}}{2\pi I_0(\kappa)}, \quad (18)$$

onde $I_0(x)$ é a função de Bessel modificada de ordem 0; κ é o parâmetro de concentração. O valor esperado da distribuição de von Mises é simplesmente μ , que é o parâmetro de localização (média). A variância de $z = e^{ix}$, onde o ângulo x segue a distribuição von Mises é também denominada por variância circular de x , que é dada por:

$$var(z) = 1 - \frac{I_1(\kappa)^2}{I_0(\kappa)^2}, \quad (19)$$

e $I_1(x)$ é a função de Bessel modificada de ordem 1.

2.4.4- Método de Máxima Verossimilhança

O método de máxima verossimilhança⁷¹, assim como o método de mínimos quadrados, permite a estimação dos parâmetros de modelos econométricos e a realização de testes de hipóteses relativos a restrições lineares e não lineares ao vetor de parâmetros. Embora seja bastante antigo⁷², foi apenas a partir dos anos oitenta, em função do desenvolvimento dos computadores pessoais de grande potência, que o método de máxima verossimilhança começou a ser utilizado extensivamente em diversas áreas do conhecimento, quando se desejava obter os parâmetros de um dado modelo.

Este método consiste em obter estimadores pontuais a partir da maximização da função de verossimilhança. Segundo Cordeiro⁷³, as estimativas de máxima verossimilhança apresentam propriedades importantes como: consistência, invariância e eficiência assintótica.

Uma amostra aleatória $(y_1, \dots, y_n)$, retirada de uma população com uma função de densidade de probabilidade $f(y, \tau)$, a qual depende do vetor de parâmetros τ , tem uma função de densidade de probabilidade (*f. d. p*) conjunta dada por

$$\prod_{i=1}^n f(y_i, \tau). \quad (20)$$

Isto é, a função de densidade de probabilidade conjunta é simplesmente o produto das densidades de cada uma das observações, $f(y_1, \tau) \times \dots \times f(y_n, \tau)$, onde τ é um vetor de parâmetros e y_i é uma variável aleatória. Note que, antes da retirada da amostra, cada observação é uma variável aleatória cuja função de densidade de probabilidade é igual à função de densidade de probabilidade da população. A média e a variância de cada observação a ser retirada são iguais à média e variância da população em questão.

Dessa forma, para a resolução do problema de estimação deve-se escolher o vetor $\hat{\tau}$ que maximize a probabilidade de obtenção da amostra específica que se tem em mãos. Em outras palavras, queremos o vetor $\hat{\tau}$ que faz a probabilidade de obter-se a amostra já obtida a maior possível, ou seja, temos que achar o $\hat{\tau}$ que maximize a função de verossimilhança.

Temos, portanto a função de verossimilhança $L(\tau, y)$, onde y é fixo e τ é a variável, e o problema consiste em obter-se o vetor $\hat{\tau}$ que maximiza esta função. O estimador de máxima verossimilhança $\hat{\tau}$ é o vetor que faz $L(\hat{\tau}', y) > L(\bar{\tau}', y)$, em que $\bar{\tau}'$ é outro estimador de τ .

Consideremos agora as seguintes definições,

- Função escore (U_τ): $\frac{\partial \ln L(y, \tau)}{\partial \tau}$, ou seja, a função escore é dada pela derivada do logaritmo da função de verossimilhança.

- Matriz de Informação: $E \left(- \frac{\partial^2 \ln(L(y, \tau))}{\partial \tau \partial \tau'} \right) = I(\tau)$

Note que o estimador de máxima verossimilhança, $\hat{\tau}$, vai ser a solução do conjunto de equações $U_\tau = 0$. Mais ainda, dadas algumas condições bem gerais, é possível mostrar-se que $\hat{\tau}$ é consistente, assintoticamente normalmente distribuído e tem variância $[I(\tau)]^{-1}$. Este valor, $[I(\tau)]^{-1}$, é conhecido como o limite inferior de Cramer-Rao⁷⁴, pois não existe outro estimador consistente do vetor τ que tenha variância menor.

Neste sentido, o estimador de máxima verossimilhança $\hat{\tau}$ é também eficiente assintoticamente.

2.5- Análise dos Resultados

2.5.1- Análise para as variáveis radiais

Com a finalidade de mostrar que a variação da geometria do poliedro de coordenação sob a influência dos contra íons e/ou moléculas se comporta como se fosse aleatória, vamos tratar primeiro as variáveis radiais. Para isso calculamos a média de cada coluna da Tabela 4, onde estas individualmente representam os átomos ligados ao poliedro de coordenação, como apresentado na Figura 5.

Em alguns casos os complexos, por se apresentarem simétricos, podem exibir dentro de um mesmo poliedro de coordenação valores para o comprimentos de ligação Mg-O iguais, o que indicaria em uma análise estatística sob todos estes valores, uma pseudo não aleatoriedade devido às repetições de valores. Embora, os valores sejam iguais, eles representam medidas diferentes quando comparados aos erros em torno da média de sua respectiva coluna. Desta maneira, a forma que escolhemos para resolver o problema das repetições numéricas foi justamente o cálculo do erro absoluto para todos os valores da Tabela 4. Pois assim representamos de maneira unívoca a variação que o contra-íon e/ou molécula causa em cada caso. Escolhemos calcular o erro médio absoluto para que os dados pudessem estar no suporte da distribuição gama. Apresentamos estes novos valores na Tabela 7.

Tabela 7- Erro absoluto calculado para as distâncias Mg -O dos poliedros de coordenação do hexaaquamagnésio influenciado pelos contra-íons/ e ou moléculas a partir dos valores da Tabela 4.

Estruturas	1	2	3	4	5	6
AQMEDA	0.00261	0.00312	0.02300	0.02261	0.00997	0.01368
BIHLIO	0.02239	0.01666	0.00333	0.00372	0.00149	0.00520
BIHNEM	0.01779	0.01206	0.01118	0.01157	0.00841	0.01212
BIKPUG	0.00546	0.00004	0.01538	0.01591	0.02235	0.02579
CIRVAA01	0.01910	0.01337	0.00410	0.00371	0.01480	0.01851
CIRVAA	0.03085	0.02512	0.01218	0.01256	0.00731	0.00360
DAMZAT	0.02721	0.03294	0.00112	0.00073	0.01346	0.00976
EXUBOO	0.00055	0.00377	0.00082	0.00132	0.00855	0.01176

Tabela 7- Continuação

FEDQUA01	0.02251	0.02825	0.02572	0.02533	0.02310	0.01940
FEDQUA	0.02198	0.02771	0.02492	0.02396	0.02271	0.01960
FETHAO	0.03772	0.04339	0.00366	0.00328	0.00025	0.00364
GATLUI	0.01102	0.00568	0.00709	0.00670	0.00433	0.00769
GOLPIG	0.02539	0.01966	0.00715	0.00754	0.00497	0.00867
GUHHOG	0.00545	0.01058	0.00096	0.00135	0.00027	0.00382
HETZUB	0.00409	0.00164	0.01150	0.01111	0.01782	0.02153
ILUXAO	0.00104	0.00469	0.00020	0.00018	0.01668	0.01298
IWOYEE	0.02541	0.03031	0.01249	0.00532	0.01106	0.00924
JAMQAQ	0.01516	0.01676	0.00389	0.00978	0.00649	0.00949
JEYCEW	0.00587	0.02604	0.01961	0.00597	0.00002	0.00350
KIDLUF	0.02591	0.03164	0.02742	0.02704	0.02541	0.02912
KIMNID	0.01843	0.01270	0.00348	0.00310	0.00204	0.00167
MAVDOD	0.02333	0.04836	0.02017	0.00171	0.03155	0.01251
MGEDTA01	0.00288	0.00285	0.01239	0.01278	0.00378	0.00749
MGNTSP	0.01002	0.00429	0.01309	0.01271	0.02083	0.01712
MINKOJ	0.00405	0.00145	0.00095	0.00154	0.00148	0.00476
MISXER	0.00119	0.00692	0.01075	0.01011	0.04340	0.04059
MOYTAV	0.02775	0.02202	0.01004	0.01042	0.00411	0.00041
NOGLAW	0.02750	0.02241	0.00511	0.00483	0.00364	0.00635
OMARID	0.00153	0.03184	0.01140	0.01037	0.03693	0.01175
POHGIC	0.02364	0.01803	0.01138	0.01077	0.00343	0.00109
QERJIH	0.03308	0.03288	0.01852	0.02444	0.00591	0.00696
RAKYOR	0.01364	0.01871	0.00363	0.00356	0.00651	0.00281
RAMXOS	0.01060	0.01827	0.02412	0.01550	0.00734	0.00934
RATRIN	0.01247	0.01887	0.01244	0.03705	0.00778	0.02442
REFTEC	0.01128	0.01033	0.00756	0.00285	0.00416	0.00289
SIMZUJ	0.00471	0.01044	0.01887	0.01848	0.01836	0.01465
TOXDMG	0.00873	0.00299	0.00544	0.00500	0.01486	0.01457
TUXYOA	0.00573	0.01137	0.00881	0.01018	0.01719	0.01348
VOPCEI	0.00177	0.00396	0.00412	0.00451	0.01797	0.01426
WIKXAP	0.02233	0.02778	0.02538	0.02556	0.01831	0.01550
XECJIY	0.02233	0.02778	0.02538	0.02556	0.01831	0.01550
XEZCIP	0.00791	0.00335	0.00057	0.00136	0.00685	0.00402
XITDUZ	0.00639	0.01212	0.00880	0.00986	0.01007	0.01356
XUXLUX	0.00210	0.00277	0.00702	0.00741	0.00244	0.00615
ZARMEK	0.01308	0.02822	0.01467	0.00300	0.02098	0.00353

Primeiramente realizamos o teste de Bartels, para as variáveis radiais. Este teste é usado para verificar a não aleatoriedade dos dados, onde obtivemos que o teste foi não significativo ao nível de significância adotado, $\gamma = 0.05$. Concluimos isso a partir do $p - valor$ (0.2147) obtido. Como vemos, o $p - valor$ é maior que 0.05, daí aceitamos a hipótese nula de aleatoriedade dos dados. Saliente-se que o nível de significância de 0.05 será usado para todos os testes apresentados nesse trabalho. Prosseguindo a

análise, vamos verificar se podemos utilizar a distribuição Gama (λ, β) para estes dados. Faremos esta verificação usando o teste, não-paramétrico, de Kolmogorov-Smirnov. Como o p-valor deste teste é igual a 0.2272, não rejeitamos a hipótese de nulidade, ou seja, os dados seguem uma distribuição *Gama*(λ, β). Assim, temos que os dados para as variáveis radiais se comportam como se fossem aleatórios e podem ser estudados por essa distribuição. Em seguida, obtivemos os estimadores de máxima verossimilhança dos parâmetros: $\hat{\lambda} = 1.268$ e $\hat{\beta} = 99.74$. Dessa forma, temos que a média é igual a 0.01271 e a variância é igual a 0.0001274. Abaixo apresentamos um gráfico a partir dos dados acima apresentados:

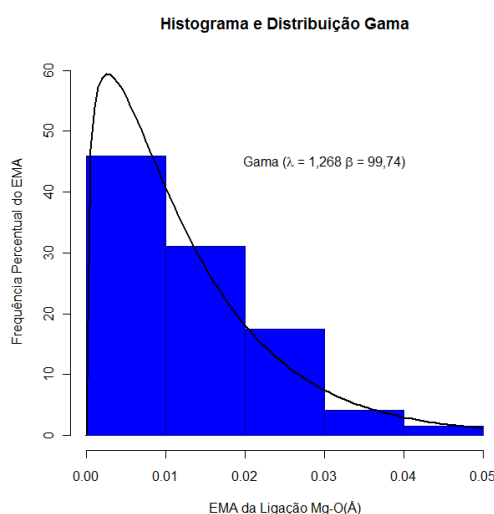


Figura 12 - Histograma do erro médio absoluto apresentado na Tabela 7 da variável radial para a ligação Mg-O do poliedro de coordenação do hexaaquamagnésio influenciado por 45 contra-íons e/ou moléculas apresentados na Tabela 3 e a sua correspondente distribuição “gama”, cujos parâmetros λ e β foram obtidos a partir método de máxima verossimilhança.

2.5.2- Análise para as variáveis latitudinais

Para a obtenção dos dados das variáveis latitudinais, dividimos os valores obtidos na Tabela 5 por 180, uma vez que o suporte da distribuição gama varia de zero a um. Desta maneira, transformamos os dados para poder prosseguir com o estudo. Os valores obtidos estão apresentados na Tabela 8.

Tabela 8- Valores dos ângulos θ para os átomos de oxigênio nos poliedros de coordenação do hexaaquamagnésio para os vários contra-íons e /ou moléculas obtidos na Tabela 5, transformados para o intervalo de 0 a 1, que é o suporte da distribuição beta.

Estruturas	3	4	5	6
AQMEDA	0.48242	0.51758	0.50805	0.49195
BIHLIO	0.49875	0.50125	0.48725	0.51275
BIHNEM	0.50225	0.49775	0.48295	0.51705
BIKPUG	0.50650	0.49345	0.50965	0.49031
CIRVAA01	0.50943	0.49057	0.50383	0.49617
CIRVAA	0.49409	0.50591	0.49767	0.50233
DAMZAT	0.47573	0.52427	0.50784	0.49216
EXUBOO	0.50458	0.49521	0.51722	0.48257
FEDQUA01	0.49722	0.50278	0.49834	0.50166
FEDQUA	0.49760	0.50222	0.49783	0.50198
FETHAO	0.51713	0.48287	0.48870	0.51139
GATLUI	0.49938	0.50062	0.50506	0.49488
GOLPIG	0.49640	0.50360	0.49405	0.50595
GUHHOG	0.48155	0.51845	0.48372	0.51619
HETZUB	0.52217	0.47783	0.49473	0.50527
ILUXAO	0.51439	0.48561	0.48896	0.51104
IWOYEY	0.49479	0.50659	0.50284	0.49864
JAMQAQ	0.49554	0.49106	0.50552	0.51028
JEYCEW	0.50069	0.51469	0.48484	0.48965
KIDLUF	0.51370	0.48630	0.50712	0.49288
KIMNID	0.48751	0.51249	0.49595	0.50405
MAVDOD	0.50361	0.49963	0.50621	0.46855
MGEDTA01	0.51434	0.48566	0.50249	0.49751
MGNTSP	0.49709	0.50291	0.48377	0.51623
MINKOJ	0.50430	0.49584	0.50173	0.49843
MISXER	0.49832	0.50165	0.52827	0.47168
MOYTAV	0.49553	0.50447	0.49428	0.50572
NOGLAW	0.50955	0.49035	0.48451	0.51549
OMARID	0.49269	0.48056	0.48874	0.49762
POHGIC	0.50629	0.49377	0.49370	0.50637
QERJIH	0.50183	0.48530	0.46939	0.53889
RAKYOR	0.49917	0.50073	0.51019	0.48981
RAMXOS	0.47704	0.51213	0.48341	0.51199
RATRIN	0.49891	0.49815	0.52424	0.47621
REFTEC	0.50181	0.49688	0.49154	0.49852
SIMZUJ	0.50352	0.49648	0.49785	0.50215
TOXDMG	0.49016	0.50971	0.49373	0.50627
TUXYOA	0.48546	0.51452	0.48692	0.51308
VOPCEI	0.49402	0.50598	0.49428	0.50572
WIKXAP	0.51030	0.48966	0.49257	0.50748
XECJIY	0.51030	0.48966	0.49257	0.50748
XEZCIP	0.48571	0.51417	0.48625	0.51370
XITDUZ	0.48867	0.51123	0.50541	0.49448
XUXLUX	0.51086	0.48914	0.51640	0.48360
ZARMEK	0.49511	0.50424	0.49349	0.49047

A análise será de forma semelhante à apresentada para as variáveis radiais. As variáveis em estudo nesse tópico são as latitudinais. O teste de Bartels apresentou um p -valor igual a 0.7070, onde concluímos que não rejeitamos a hipótese nula ao nível de significância de 0.05, ou seja, os dados apresentam-se de forma aleatória. Os parâmetros a e b da distribuição beta foram obtidos através do método de máxima verossimilhança e seus valores foram respectivamente 931.1 e 933.6. No próximo teste, Kolmogorov-Smirnov, obtivemos um p -valor igual a 0.9981. Daí chegamos à conclusão de que aceitamos a hipótese nula de que os dados podem ser estudados através de uma distribuição Beta (a,b). Sabendo que os dados são aleatórios e que seguem uma distribuição beta, temos então o valor de 0.0001340 para a variância. Esta variância foi obtida a partir dos parâmetros obtidos pelo método da máxima verossimilhança. Podemos observar na Figura 13 o histograma das medidas para o erro médio absoluto de θ e a distribuição beta.

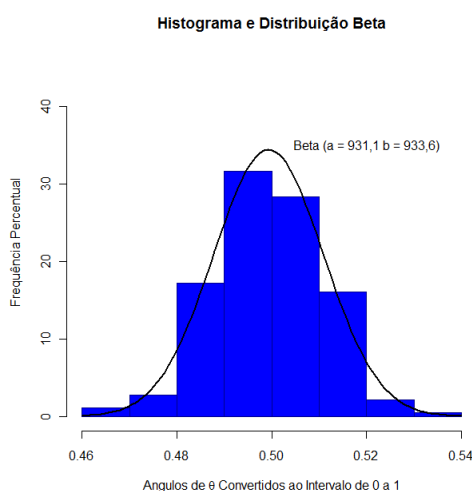


Figura 13 - Histograma para as medidas apresentadas na Tabela 8 do ângulo θ dos átomos coordenantes do hexaaquamagnésio influenciado por 45 contra-íons e/ou moléculas apresentados na Tabela 3, dividido por 180 e a sua correspondente distribuição beta, cujos parâmetros a e b foram obtidos a partir método de máxima verossimilhança.

2.5.3- Análise para as variáveis longitudinais

Por fim, para a última variável em estudo transformamos as medidas dos ângulos longitudinais dos átomos de oxigênio do hexaaquamagnésio apresentados na Tabela 6

de graus para radianos. Desta maneira, estes valores passam a estar dentro do suporte da distribuição von Mises, estando contidos em intervalo de comprimento 2π .

Com a mesma argumentação aplicada ao caso das variáveis radiais, calculamos a média de cada coluna, que representa um átomo de oxigênio do poliedro de coordenação do hexaaquamagnésio, obtendo em seguida os erros para cada medida a partir da média de sua coluna. Nesta distribuição, diferente do caso da distribuição gama, os erros não precisam ser absolutos, pois a distribuição von Mises só exige que os valores estejam contidos num intervalo circular de comprimento 2π . Estes novos dados são apresentados na Tabela 9.

Tabela 9- Erros dos ângulos φ para os átomos de oxigênio nos poliedros de coordenação do hexaaquamagnésio influenciado pelos 45 contra-íons/ e ou moléculas apresentados na Tabela 3.

Estruturas	3	4
AQMEDA	-0.46763	-0.42869
BIHLIO	-0.96912	-0.93018
BIHNEM	0.70829	0.74722
BIKPUG	-3.03690	-2.99151
CIRVAA01	-0.73553	-0.69659
CIRVAA	-0.41197	-0.37303
DAMZAT	-2.08805	-2.04911
EXUBOO	1.96905	2.00747
FEDQUA01	3.17714	3.21608
FEDQUA	3.18613	3.24028
FETHAO	2.34059	2.37952
GATLUI	0.55305	0.59198
GOLPIG	0.12047	0.15940
GUHHOG	1.22607	1.26501
HETZUB	-2.88466	-2.84572
ILUXAO	-3.30198	-3.26305
IWOYEY	1.24129	2.90309
JAMQAAQ	-1.06066	-1.73712
JEYCEW	0.66692	-2.80428
KIDLUF	-0.91305	-0.87411
KIMNID	1.16827	1.20720
MAVDOD	-1.07623	4.34906
MGEDTA01	-0.49987	-0.46093
MGNTSP	2.22103	2.25996
MINKOJ	-0.77404	-0.74472
MISXER	-4.27778	-4.21292
MOYTAV	1.28898	1.32792
NOGLAW	-0.43870	-0.37035
OMARID	-1.65207	-0.41047

Tabela 9- Continuação

POHGIC	-0.86605	-0.78927
QERJIH	5.89723	-1.42300
RAKYOR	1.57783	1.60030
RAMXOS	-0.13475	1.53643
RATRIN	-1.30688	0.38934
REFTEC	0.36150	0.21351
SIMZUJ	-3.53829	-3.49936
TOXDMG	-0.23639	0.24767
TUXYOA	0.77266	0.81216
VOPCEI	-1.00195	-0.96302
WIKXAP	2.07644	2.14169
XECJIY	2.07644	2.14169
XEZCIP	2.00498	2.04858
XITDUZ	-1.39087	-1.34524
XUXLUX	-1.54670	-1.50777
ZARMEK	-0.02423	-2.06511

Para verificar se os dados se comportam como se fossem aleatórios, novamente aplicamos o teste de Bartels obtendo um p -valor de 0.2243, que confirma a hipótese de que este conjunto de valores se comporta como se os mesmos fossem aleatórios. Procederemos neste instante da mesma maneira que nos casos anteriores, já que utilizaremos o estimador de máxima verossimilhança para obter os parâmetros da distribuição que, diferentemente das variáveis anteriores, possui dados circulares. Os parâmetros estimados foram $\mu = 8.922e^{-7}$ e $\kappa = 847.3$. Para avaliar se estes parâmetros seguem a distribuição para dados circulares von Mises utilizamos o teste de Watson (Watson's Two-Sample Test of Homogeneity), em que obtivemos um p -valor igual a 0.1000. Desta maneira, os dados estão dispostos de uma forma aleatória e podem ser analisados a partir da distribuição von Mises. A seguir, apresentamos o gráfico da distribuição:

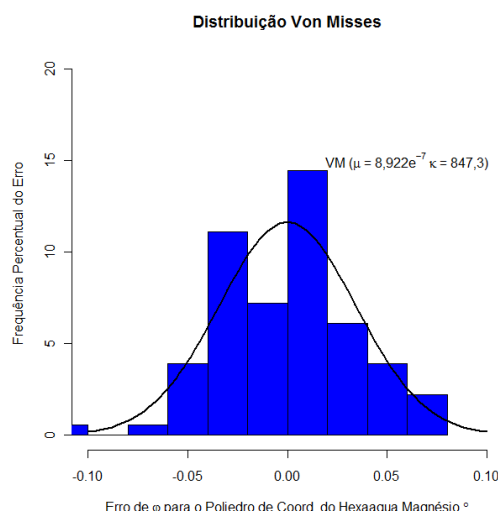


Figura 14 - Histograma do erros apresentado na Tabela 9 para o ângulo ϕ dos átomos coordenantes do hexaaquamagnésio influenciado por 45 contra-íons e/ou moléculas apresentados na Tabela 3, e a sua correspondente distribuição von Mises, cujos parâmetros μ e κ foram obtidos a partir método de máxima verossimilhança.

2.6- Definição da nova função resposta: “Planetário”

O estudo feito neste capítulo mostrou empiricamente que o contra-íon e/ou molécula afetam significativamente, porém de forma pseudo-aleatória, a geometria do poliedro de coordenação do complexo cristalino. Detectamos também que num conjunto de quarenta e cinco células de cristais contendo hexaaquamagnésio advindos do CSD, os dados para as variáveis radial, latitudinal e longitudinal se comportam como se fossem aleatórias e portanto podem ser descritas por três distribuições respectivamente: distribuição gama; distribuição beta; e distribuição von Mises. Este resultado valida o modelo Sparkle, pois nesta ótica os melhores parâmetros que representam o *sparkle* são aqueles que geram através do processo de otimização de geometria um complexo que tenha suas coordenadas d , θ e ϕ equivalendo aos respectivos valores médios de um conjunto de células unitárias contendo o mesmo complexo perturbado por diferentes contra-íons e/ou moléculas.

Os desvios médios absolutos das distâncias na geometria do complexo provocados pelos contra-íons e/ou moléculas (0.01Å) ficaram uma ordem de grandeza abaixo do erro médio normalmente observado para o modelo Sparkle/AM1 dos

lantânídeos (0.2Å). Este resultado valida a realização de cálculos de otimização de geometria para complexos metálicos sem a inclusão explícita de contra-íons e/ou moléculas, ou até mesmo da rede cristalina. pelo modelo Sparkle. Este resultado ainda mostra que os parâmetros do modelo Sparkle podem ser melhorados para que seja possível obter geometrias de poliedros de coordenação dos complexos pelo menos uma ordem de grandeza mais exata do que as conseguidas atualmente, antes de a inclusão de contra-íons ou moléculas ter que ser feita de forma explícita.

A função resposta é uma ferramenta que ajudará a encontrar os parâmetros que se aproximam deste limite teórico. Esta compara a estrutura do banco de parametrização com sua análoga gerada a partir do novo parâmetro encontrado. Desta maneira, os melhores parâmetros deverão ser aqueles nos quais o valor da função resposta se aproxime do zero. Um valor zero significa que as duas estruturas deverão ser idênticas, tendo desta maneira em igualdade todas suas distâncias, ângulos e diedros. Com a finalidade de obter melhores parâmetros com os erros médios absolutos equilibrados entre distâncias e ângulos, apresentaremos uma nova proposta de função resposta para a parametrização do modelo Sparkle baseada nos resultados discutidos neste capítulo.

Pretendemos que todas as três variáveis, d , θ e ϕ , tenham pesos equivalentes na função resposta. Chamaremos neste trabalho esta nova função resposta de planetário, pois se assemelha muito ao sistema elaborado na antiguidade para localizar as estrelas fixas, onde a Terra seria o metal e os átomos coordenantes do poliedro de coordenação as estrelas. Desta maneira, este sistema é regido por coordenadas esféricas, que de longa data é muito bem descrito na literatura.

Propomos duas formas para a função resposta: uma, G_{resp} , baseada nos erros absolutos e outra alternativa, F_{resp} , baseada nos erros quadráticos. Em G_{resp} os desvios são ponderados pelo desvio padrão e em F_{resp} , os desvios são ponderados pela variância. Os valores do desvio padrão e da variância vêm do estudo realizado para os quarenta e cinco complexos hexaaquamagnésio. Apresentamos a seguir as duas propostas:

$$G_{resp} = \frac{1}{n_i \sigma_d} \sum_{i=1}^{n_i} |d_i^{CSD} - d_i^{Calc}| + \frac{1}{n_j \sigma_\theta} \sum_{j=1}^{n_j} |\theta_j^{CSD} - \theta_j^{Calc}| + \frac{1}{n_k \sigma_\phi} \sum_{k=1}^{n_k} [\Phi(\phi^{CSD} - \phi^{Calc})] \quad ; \quad (21)$$

$$F_{resp} = \frac{1}{n_i \sigma_d^2} \sum_{i=1}^{n_i} (d_i^{CSD} - d_i^{Calc})^2 + \frac{1}{n_j \sigma_\theta^2} \sum_{j=1}^{n_j} (\theta_j^{CSD} - \theta_j^{Calc})^2 + \frac{1}{n_k \sigma_\phi^2} \sum_{k=1}^{n_k} [\Phi(\phi^{CSD} - \phi^{Calc})]^2 \quad ; \quad (22)$$

onde i identifica um determinado tipo de distância d (átomo central – átomo coordenante); n_i é o número total das distâncias sendo consideradas (excluindo as indefinidas); n_j e n_k são os números totais de ângulos θ e φ sendo considerados (excluindo-se os indefinidos); sendo ainda $\Phi(\varphi) = [1 - \cos(\varphi)]$. Temos que σ_d^2 é a variância dos erros absolutos das distâncias d , σ_θ^2 é a variância dos erros absolutos dos ângulos θ e σ_φ^2 é a variância dos erros dos ângulos φ . Os ângulos θ e φ estão representados em radianos.

Para a função resposta G_{resp} usamos os desvios-padrões: σ_d , σ_θ e σ_φ que foram obtidos calculando a raiz quadrada das variâncias: σ_d^2 , σ_θ^2 e σ_φ^2 respectivamente. Essas variâncias são obtidas a partir das estimativas de máxima verossimilhança dos parâmetros das distribuições Gama, Beta e von Mises para o caso do hexaaquamagnésio. Desta maneira, estes valores serviram para calibrar a função resposta planetário.

Na função resposta não podemos usar $\cos(\varphi^{CSD} - \varphi^{Calc})$ porque o cosseno de zero é 1 e assim, o cosseno do erro ideal que é zero daria um valor que não zeraria a função resposta. Assim, para zerar a função resposta temos que usar a expressão $(1 - \cos(\varphi^{CSD} - \varphi^{Calc}))$ porque assim um erro zero dá zero para esta função.

O problema é que a distribuição de von Mises gera a variância de φ , porém o que precisamos colocar na fórmula da função resposta é a variância da função $(1 - \cos(\varphi))$.

A demonstração abaixo mostra como podemos obter a variância de $1 - \cos(\varphi)$ a partir da variância de φ . Temos que a variância de $\cos(\varphi)$ pode ser obtida dos parâmetros da distribuição von Mises, μ e κ . Se X segue a distribuição von Mises (μ, κ) , onde podemos representar por $X \sim VM(\mu, \kappa)$, como estamos interessados em calcular a variância de $1 - \cos(X)$, sabemos que $Var(1 - \cos(X)) = Var(\cos(X))$. Pela definição de variância, $Var(W) = E[W^2] - (E[W])^2$, para uma variável aleatória qualquer W . Continuando a demonstração temos que $Var(\cos(X)) = E[(\cos(X))^2] - (E[\cos(X)])^2$. Daí, usando números complexos, temos:

$$\begin{aligned} (\cos(x) + i \sin(x))^2 = \\ \cos(x)^2 - \sin(x)^2 + 2i\cos(x)\sin(x) \end{aligned}$$

$$(\cos(x) + i \sin(x))^2 + 1 = 2 \cos(x)^2 + 2i \cos(x) \sin(x)$$

$$\frac{(\cos(x) + i \sin(x))^2 + 1}{2} = \cos(x)^2 + i \cos(x) \sin(x) \quad (23)$$

agora chamemos o termo de $Z = (\cos(x) + i \sin(x))^2$, onde ficará escrito como,

$$\frac{Z^2 + 1}{2} = \cos(x)^2 + i \cos(x) \sin(x) \quad (24)$$

Sabemos da função de von Mises que:

$$E(Z^2) = \frac{I_2(\kappa)}{I_0(\kappa)} e^{-i2\mu}$$

$$E(\cos(x)^2) = R \left[\frac{E(Z^2) + 1}{2} \right] = \left[\frac{I_2(\kappa)}{I_0(\kappa)} \cos(2\mu) + 1 \right] \frac{1}{2}$$

$$E(\cos(x)) = R[E(Z)] = \frac{I_2(\kappa)}{I_0(\kappa)} \cos(\mu) \quad (25)$$

E por fim temos o que desejávamos calcular, que é apresentado abaixo

$$Var(\cos(x)) = \frac{1}{2} \left[1 + \frac{I_2(\kappa)}{I_0(\kappa)} \cos(2\mu) \right] - \left[\frac{I_1(\kappa)}{I_0(\kappa)} \right]^2 \cos^2(\mu). \quad (26)$$

Agora apresentamos os valores das três variâncias vindas do estudo do $[\text{Mg}(\text{H}_2\text{O})_6]^{2+}$: σ_d^2 é igual a 0.0001274; σ_θ^2 é igual a 0.0001340; e por último σ_ϕ^2 é igual a 6.96882×10^{-7} . Estes são os valores que serão usados na função resposta. Estas variâncias ainda determinam os erros limites empíricos de um método que busque prever a geometria dos átomos coordenantes ao íon Mg^{2+} dentro do poliedro de coordenação na ausência de contra-íons, outras moléculas e efeitos de rede na célula

unitária e que são os seguintes: $0.01\text{\AA}$ para a coordenada radial; 2.08° para a coordenada latitudinal; e 2.34° para a coordenada longitudinal.

Definida a função resposta, resta definir univocamente o sistema de coordenadas esféricas para o complexo. A implementação do planetário em um sistema de referência padrão é muito importante, pois esta não poderá variar durante a parametrização. Resolvemos este dilema a partir das seguintes convenções: (i) O sistema de coordenadas cartesianas é definido na estrutura experimental do banco de dados; (ii) o eixo de referência será transladado para o metal. Em caso de dímeros o eixo de coordenadas será transladado para um dos metais e o segundo metal será tratado como átomo ligante; (iii) No novo sistema de coordenadas definimos que o sentido positivo do eixo Z deve ser definido e projetado diretamente sob o átomo do poliedro de coordenação que tenha o maior comprimento de ligação em relação ao metal; (iv) O sentido positivo do eixo X estará definido no plano formado entre o vetor Z e o vetor representado na direção do átomo do poliedro de coordenação que tenha o menor valor do produto escalar em relação ao vetor Z. A condição de ortogonalidade deste novo sistema de coordenadas cartesianas será dado a partir do método de Gram-Schmidt; (v) por último, o eixo Y será definido pela regra da mão direita como apresentado na Figura 7.

Uma vez definido este sistema unívoco no complexo experimental, teremos também o sistema definido para o poliedro de coordenação da estrutura calculada. Já que a ordem dos átomos representados na matriz submetida ao MOPAC não sofrerá mudança em sua ordem. Desta maneira, basta marcar os átomos sobre o qual o sistema foi definido e construir novamente o sistema de coordenadas cartesianas.

Uma vez definidos os sistemas cartesianos para o complexo experimental e calculado transformamos estes sistemas cartesiano em sistemas de coordenadas polares através da utilização da Equação 9.

2.7- Analisando Geometrias Geradas Através de Métodos *ab-initio*

Com a finalidade de avaliar a geometria do complexo hexaaquamagnésio obtido a partir de cálculos teóricos sem a presença de contra-íons, escolhemos os métodos HF, B3LYP e MP2(Full) com as bases STO-3G, 3-21G, 6-31G, 6-31G*, 6-31+G e 6-31++G.

Obtivemos como resultado desta combinação de métodos e bases o conjunto de valores apresentados na Tabela 10 para as coordenadas radiais de cada átomo coordenado ao magnésio respeitando o critério convencionado e apresentado na Figura 5.

Tabela 10- Valores das distâncias radiais (Å) de cada átomo ligado ao magnésio, obtidos a partir das otimizações de geometria usando diferentes modelos *ab initio*. Cada coluna representa a posição do átomo no poliedro de coordenação segundo a Figura 5.

Método/ Base	Átomos coordenantes do poliedro de coordenação					
	1	2	3	4	5	6
HF/STO-3G	1.93054	1.93054	1.93053	1.93053	1.93053	1.93053
HF/3-21G	2.04361	2.04361	2.04352	2.04352	2.04357	2.04358
HF/6-31G	2.08343	2.08343	2.08306	2.08306	2.08279	2.08279
HF/6-31G*	2.10531	2.10530	2.10530	2.10530	2.10539	2.10540
HF/6-31+G	2.08833	2.08833	2.08795	2.08795	2.08778	2.08778
HF/6-31++G	2.08756	2.08757	2.08719	2.08719	2.08703	2.08703
B3LYP/STO-3G	1.94020	1.94020	1.93965	1.93964	1.93951	1.93950
B3LYP/3-21G	2.04235	2.04234	2.04184	2.04181	2.04124	2.04123
B3LYP/6-31G	2.07827	2.07826	2.07777	2.07777	2.07750	2.07748
B3LYP/6-31G*	2.10124	2.10123	2.09833	2.09832	2.09842	2.09842
B3LYP/6-31+G	2.09188	2.09187	2.09186	2.09186	2.08984	2.08983
B3LYP/6-31++G	2.09184	2.09182	2.09104	2.09103	2.08839	2.08838
MP2/STO-3G	1.92430	1.92429	1.92427	1.92427	1.92426	1.92426
MP2(Full)/3-21G	2.05126	2.05125	2.05122	2.05122	2.05120	2.05120
MP2(Full)/6-31G	2.09866	2.09866	2.09841	2.09840	2.09791	2.09790
MP2(Full)/6-31G*	2.09649	2.09650	2.09669	2.09669	2.09652	2.09653
MP2(Full)/6-31+G	2.10253	2.10253	2.10231	2.10230	2.10195	2.10195
MP2/(Full)6-31++G	2.10119	2.10119	2.10095	2.10094	2.10056	2.10055

Notamos na tabela acima que a base STO-3G quando combinada aos demais métodos resultou nos menores valores para os comprimentos das ligações Mg-O dentro do poliedro de coordenação. Destas, os menores valores encontrados para as distâncias foram obtidos a partir do modelo MP2(Full)/STO-3G. Enquanto os maiores comprimentos de ligação obtidos nos métodos HF e B3LYP estavam associados à base 6-31G* e no método MP2(Full) à base 6-31+G. Destas, o modelo que teve o conjunto de maiores medidas foi o HF/6-31G*.

O segundo conjunto de valores que apresentamos na próxima tabela são os das coordenadas latitudinal dos oxigênios dos poliedros de coordenação dos 15 hexaaquamagnésio.

Tabela 11- Valores dos ângulos latitudinais em graus para cada átomo ligado ao magnésio, obtidos a partir das otimizações de geometria usando diferentes modelos *ab-initio*. Cada coluna representa a posição do átomo no poliedro de coordenação segundo a Figura 5.

Método/Base	Átomos coordenantes do poliedro de coordenação				
	2	3	4	5	6
HF/STO-3G	180.000	90.0023	89.9977	90.0009	89.9985
HF/3-21G	179.999	90.0770	89.9231	89.8701	90.1307
HF/6-31G	180.000	89.9819	90.0180	90.0189	89.9812
HF/6-31G*	180.000	90.0260	89.9739	89.9686	90.0313
HF/6-31+G	180.000	90.0149	89.9851	89.9838	90.0163
HF/6-31++G	180.000	90.0139	89.9860	89.9848	90.0153
B3LYP/STO-3G	179.997	90.9969	89.0062	88.7334	91.2667
B3LYP/3-21G	179.997	90.9589	89.0376	89.9724	90.0268
B3LYP/6-31G	179.998	89.9101	90.0912	89.8728	90.1262
B3LYP/6-31G*	179.996	90.0352	89.9625	89.9638	90.0341
B3LYP/6-31+G	179.996	89.6734	90.3245	89.5969	90.4046
B3LYP/6-31++G	179.997	89.7724	90.2247	89.7344	90.2649
MP2/STO-3G	179.998	91.5512	88.4493	91.5534	88.4456
MP2(Full)/3-21G	179.999	90.0161	89.9835	90.0145	89.9851
MP2(Full)/6-31G	179.999	90.0248	89.9749	89.9745	90.0255
MP2(Full)/6-31G*	179.999	90.0574	89.9424	89.9698	90.0303
MP2(Full)/6-31+G	180.000	90.0175	89.9823	89.9693	90.0306
MP2(Full)/6-31++G	179.999	90.0193	89.9805	89.9660	90.0338

Observamos na tabela acima que a coluna número dois que representa o átomo oposto ao átomo um num poliedro de coordenação perfeito, segundo a Figura 5, sofreu variação angular somente na sexta casa decimal, o que significa que em todos os métodos este átomo não sofreu variação. Em geral, notamos que para as demais colunas os valores dos ângulos não sofrem grande variação. De todas as colunas, a que sofreu uma maior oscilação nos valores foi a coluna seis. A diferença entre o maior e o menor valor desta coluna foi igual a 2.8211 em graus. Sendo assim, estes dados mostram que a variação dos ângulos latitudinais nos cálculos *ab initio* é muito pequena.

Para podermos observar as medidas obtidas para as coordenadas longitudinais apresentamos a Tabela 12:

Tabela 12- Valores dos ângulos longitudinais em graus para cada átomo ligado ao magnésio, obtidos a partir das otimizações de geometria usando diferentes modelos *ab-initio*. Cada coluna representa a posição do átomo no poliedro de coordenação segundo a Figura 5.

Método/Base	Átomos coordenantes do poliedro de coordenação				
	2	3	4	5	6
HF/STO-3G	198.009	90.021	270.020	180.000	0
HF/3-21G	7.599	89.976	269.977	180.000	0
HF/6-31G	48.826	90.019	270.019	180.000	0
HF/6-31G*	334.697	89.996	269.996	180.000	0
HF/6-31+G	295.588	90.013	270.013	180.000	0
HF/6-31++G	298.831	90.012	270.012	180.000	0
B3LYP/STO-3G	162.870	88.695	268.694	180.000	0
B3LYP/3-21G	112.149	89.850	269.858	179.999	0
B3LYP/6-31G	223.637	89.949	269.949	180.000	0
B3LYP/6-31G*	14.275	90.015	270.014	180.002	0
B3LYP/6-31+G	243.822	89.442	269.440	179.999	0
B3LYP/6-31++G	234.817	89.625	269.623	179.999	0
MP2/STO-3G	109.964	91.516	271.518	180.001	0
MP2(Full)/3-21G	233.353	89.988	269.988	180.000	0
MP2(Full)/6-31G	215.031	90.022	270.021	180.000	0
MP2(Full)/6-31G*	350.918	90.006	270.006	180.001	0
MP2(Full)/6-31+G	224.539	90.004	270.004	180.000	0
MP2(Full)/6-31++G	214.877	90.002	270.002	180.000	0

Assim a coluna 1 sofre indeterminação, devido ao fato de o átomo 1 estar perpendicularmente projetado ao eixo X e Y. O átomo 2 também sofre de instabilidade nesta coordenada, pois qualquer pequena variação espacial deste segundo átomo poderá indicar uma grande mudança do ângulo φ , como observamos na respectiva coluna da Tabela 12. A sexta coluna designa o átomo sobre o qual o eixo X é definido, logo suas medidas necessariamente são iguais a zero, por definição. Já a coluna 5 representa justamente o átomo que está oposto a ele, tendo, em geral, os valores de 180^0 . Logo, as colunas que servirão para observarmos mais sensivelmente as variações em torno dos ângulos φ para os vários modelos são as colunas três e quatro. Os maiores valores obtidos nestas colunas para os métodos HF e MP2(Full), foram resultado da combinação com a base STO-3G, diferentemente do método B3LYP, cujo maior valor foi com a base 6-31G*. Destas medidas, os maiores valores para cada coluna foram obtidos para o método MP2(Full)/STO-3G.

Para avaliar estas geometrias otimizadas utilizando diferentes combinações método/base sem a presença dos contra-íons, calculamos a mediana para cada coluna das Tabelas 4, 5 e 6. Pois, a mediana para cada uma destas colunas representa o melhor

valor para tais coordenadas, já que num conjunto de valores a mediana é uma medida de localização do centro da distribuição dos dados, que corresponde ao valor que divide a amostra ao meio. Isto é, a metade dos elementos do conjunto de dados são menores ou iguais à mediana, enquanto que os restantes são superiores ou iguais.

Com a finalidade de avaliar estes valores comparativamente e também testarmos o equilíbrio da função resposta, calculamos cada parcela da função resposta observando equilíbrio na avaliação dos valores. Para a primeira análise apresentamos os valores obtidos para a parcela radial da função resposta:

Tabela 13- Valores calculados a partir da parcela radial da função planetário.

Método /Base	Parcela radial da função planetário
HF/STO-3G	11.8569
HF/3-21G	1.8626
HF/6-31G	1.8517
HF/6-31G*	3.6279
HF/6-31+G	2.0943
HF/6-31++G	2.0339
B3LYP/STO-3G	11.0374
B3LYP/3-21G	2.0001
B3LYP/6-31G	1.6915
B3LYP/6-31G*	3.0956
B3LYP/6-31+G	2.3750
B3LYP/6-31++G	2.3065
MP2/STO-3G	12.4114
MP2(Full)/3-21G	1.6040
MP2(Full)/6-31G	3.0070
MP2(Full)/6-31G*	2.8516
MP2(Full)/6-31+G	3.3559
MP2(Full)/6-31++G	3.2349

O valores do erro médio absoluto calculado para cada combinação método/base, mostrou que a base STO-3G apresentou maior discrepância em termos das medianas calculadas para cada átomo do poliedro de coordenação do conjunto das 45 estruturas experimentais. Podemos observar estes erros através da Figura 15.

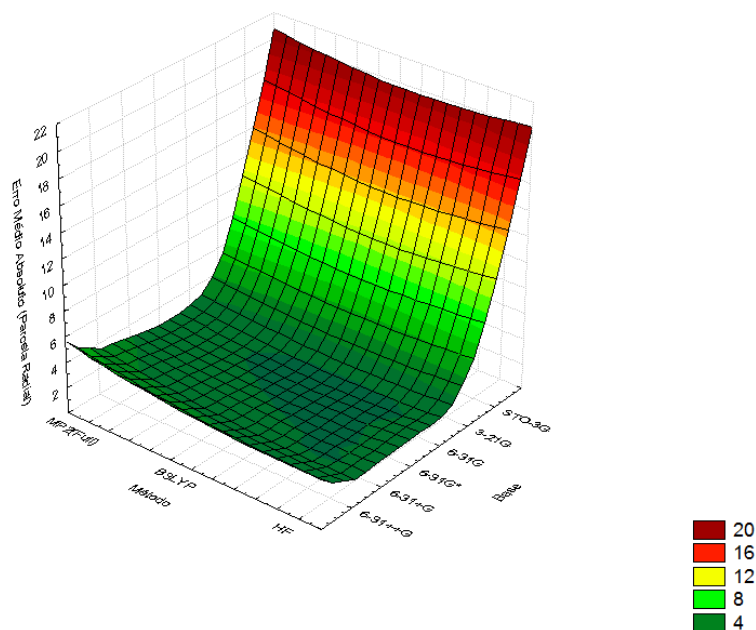


Figura 15 - Gráfico de superfície dos erros médios absoluto dos comprimentos de ligações calculados a partir da parcela radial **d** da função planetário para as diferentes combinações método/base.

Ainda com o objetivo de avaliar os valores e testar o equilíbrio da função resposta planetário, calculamos o erro médio absoluto da variável latitudinal em torno das medianas da mesma maneira como calculamos para a variável radial, obtendo assim os seguintes valores:

Tabela 14- Valores calculados a partir da parcela latitudinal do planetário.

Método/Base	Parcela latitudinal do planetário
HF/STO-3G	0.15952
HF/3-21G	0.11991
HF/6-31G	0.16683
HF/6-31G*	0.14875
HF/6-31+G	0.15376
HF/6-31++G	0.15412
B3LYP/STO-3G	0.27808
B3LYP/3-21G	0.27952
B3LYP/6-31G	0.15265
B3LYP/6-31G*	0.14571
B3LYP/6-31+G	0.14785
B3LYP/6-31++G	0.15196
MP2/STO-3G	0.69652
MP2(Full)/3-21G	0.15941
MP2(Full)/6-31G	0.15002
MP2(Full)/6-31G*	0.14281
MP2(Full)/6-31+G	0.15046
MP2/(Full)6-31++G	0.14949

Em geral, as variações das posições - em termos de θ - dos átomos de oxigênio contidos no poliedro de coordenação do hexaaquamagnésio não variaram muito quando comparados ao complexo cujos valores de d , θ e φ para cada átomo de oxigênio são iguais às medianas de cada coluna das Tabelas 5, 6 e 7. Pois, como podemos ver na Tabela 14, a maior variação observada na parcela latitudinal foi para a combinação MP2/STO-3G com o valor de 0.69652. A combinação HF/3-21G, foi a que apresentou o valor mais exato em termo da variável latitudinal ao apresentar o seguinte valor 0.11991. Abaixo podemos visualizar o gráfico de superfície para estas medidas.

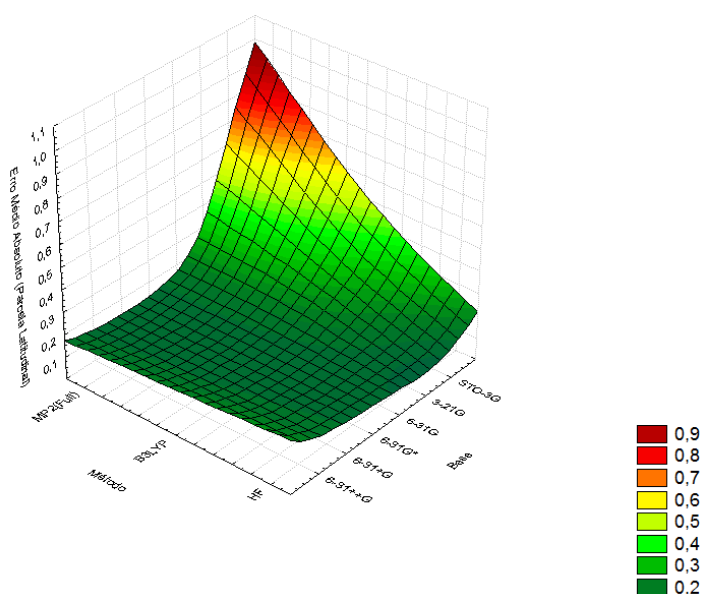


Figura 16 - Gráfico de superfície dos erros médios absoluto dos ângulos latitudinais dos oxigênios do poliedro de coordenação do hexaaquamagnésio calculados a partir da parcela latitudinal da função planetário para as diferentes combinações método/base.

A última parcela da função resposta planetário a ser analisada é referente às coordenadas dos ângulos longitudinais dos átomos de oxigênio de todos complexos (segundo a Figura 5) obtidos a partir da otimização do complexo hexaaquamagnésio nas diferentes combinações método/base. Apresentamos os erros referentes a esta parcela na tabela abaixo.

Tabela 15 - Valores calculados a partir da parcela longitudinal do planetário.

Método/Base	Parcela longitudinal da função resposta
HF/STO-3G	0.01149
HF/3-21G	0.00853
HF/6-31G	0.01136
HF/6-31G*	0.00978
HF/6-31+G	0.01093
HF/6-31++G	0.01085
B3LYP/STO-3G	0.13020
B3LYP/3-21G	0.00282
B3LYP/6-31G	0.00695
B3LYP/6-31G*	0.01105
B3LYP/6-31+G	0.01078
B3LYP/6-31++G	0.00218
MP2/STO-3G	0.39107
MP2(Full)/3-21G	0.00927
MP2(Full)/6-31G	0.01157
MP2(Full)/6-31G*	0.01045
MP2(Full)/6-31+G	0.01035
MP2(Full)/6-31++G	0.01019

Observamos nestas medidas que a parcela longitudinal em geral não sofre grandes variações. Porém, ao comparar os valores entre si observamos uma variação de duas ordens de grandeza o que em termos de métodos é bem significativo, embora em termos absolutos não seja tão significativo. Ao observamos a Tabela 12 percebemos que estes modelos pouco variaram entre si. A maior variação observada foi uma diferença angular longitudinal um pouco maior do que um grau e mesmo assim foi percebida pela parcela longitudinal da função resposta. A maior variação foi novamente dada para a combinação MP2/STO-3G como podemos observar na Figura 17.

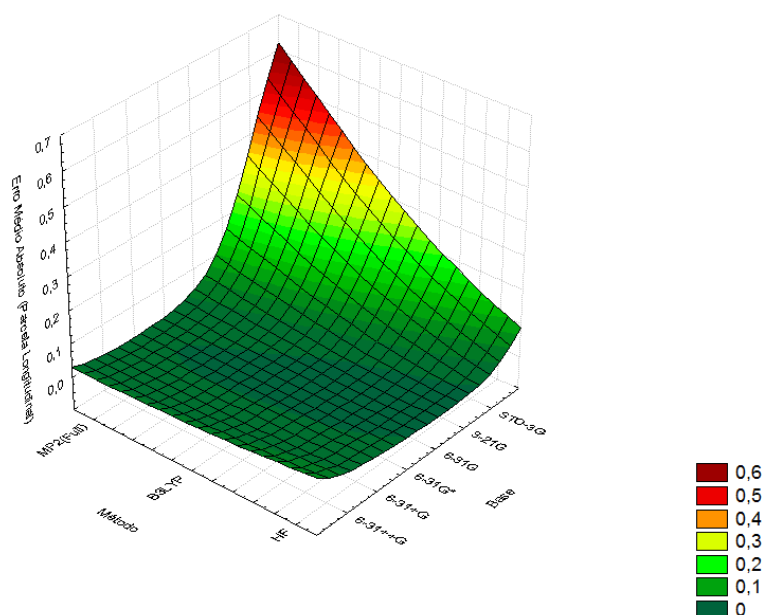


Figura 17 - Gráfico de superfície dos erros médios absoluto dos ângulos longitudinais dos oxigênios do poliedro de coordenação do hexaaquamagnésio calculados a partir da parcela longitudinal da função planetário para as diferentes combinações método/base.

Em especial observamos que a combinação B3LYP 6-31++G gerou o menor erro em relação a estes ângulos longitudinais.

Em geral, observamos que as medidas obtidas em todas parcelas da função resposta planetário se mostraram bem equilibradas. Destas parcelas, a que mais pesou na avaliação das geometrias do hexaaquamagnésio otimizadas a partir dos modelos *ab-initio* foi a parcela radial como podemos observar na Tabela 16.

Tabela 16- Valores do erro médio total calculados pelo planetário.

Método/Base	Erro médio total calculado pelo planetário
HF/STO-3G	12.0279
HF/3-21G	1.9910
HF/6-31G	2.0299
HF/6-31G*	3.7864
HF/6-31+G	2.2590
HF/6-31++G	2.1989
B3LYP/STO-3G	11.4457
B3LYP/3-21G	2.2825
B3LYP/6-31G	1.8511
B3LYP/6-31G*	3.2524
B3LYP/6-31+G	2.5336
B3LYP/6-31++G	2.4607

Tabela 16- Continuação

MP2(Full)/STO-3G	13.4989
MP2(Full)/3-21G	1.7726
MP2(Full)/6-31G	3.1686
MP2(Full)/6-31G*	3.0049
MP2(Full)/6-31+G	3.5167
MP2/(Full)6-31++G	3.3946

Neste estudo, concluímos que a combinação MP2/STO-3G, em todas análises mostrou ser a pior em comparação com os demais modelos aqui apresentados. Pois, gerou o complexo que tem a geometria mais diferente do poliedro de coordenação cujas medidas para cada átomo de oxigênio são iguais às respectivas medianas de cada uma das colunas apresentadas nas Tabelas 5, 6 e 7. Destes modelos, o que apresentou uma geometria mais próxima do complexo desejado foi o MP2/3-21G.

Podemos avaliar o comportamento geral destes métodos através da Figura 18.

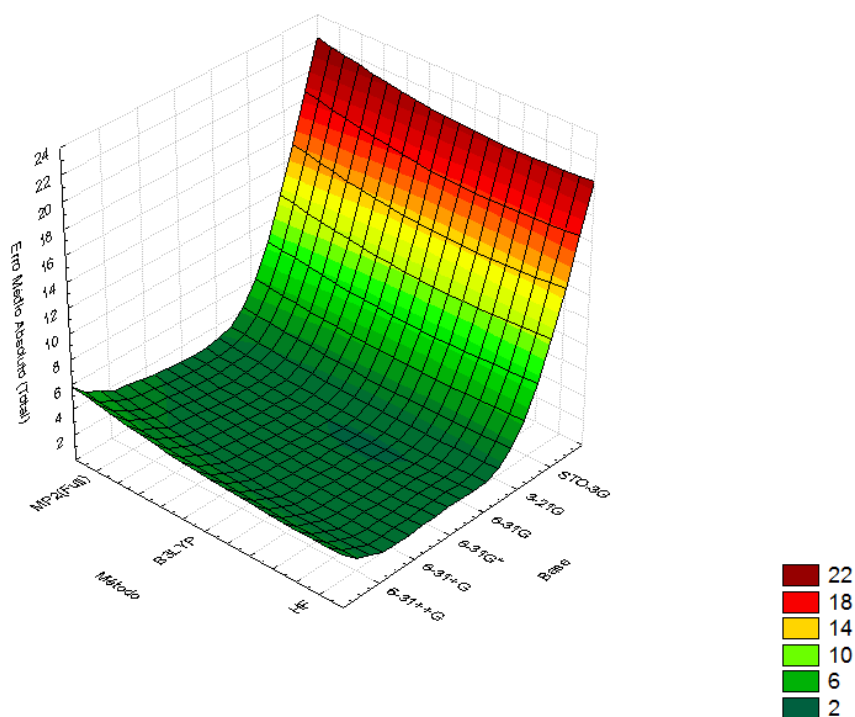


Figura 18 - Gráfico de superfície dos erros médios absoluto total dos oxigênios do poliedro de coordenação do hexaaquamagnésio calculados a partir da função planetário para as diferentes combinações método/base.

3. Construção do Banco de Parametrização

O banco de Parametrização deve ser um conjunto de complexos eleitos que tenham como propriedade representar um determinado universo de complexos. Este conjunto servirá de referência durante todo o processo de parametrização. Logo, um banco de dados bem escolhido aumentará a chance de se encontrar bons parâmetros. Justamente pela importância desta etapa é que discorreremos detalhadamente as etapas necessárias para a preparação não somente do banco de dados como um todo, mas também da construção do banco de parametrização.

Optamos ainda neste capítulo por apresentar de maneira inédita como faremos a escolha do banco de parametrização do modelo Sparkle que representará o íon magnésio utilizando o método RM1³⁰.

3.1- A Escolha dos Complexos a partir do CSD¹⁵⁻¹⁷

Os complexos cristalinos aqui considerados são aqueles cujas estruturas encontram-se disponíveis no CSD¹⁵⁻¹⁷. Destas, selecionamos aquelas que formarão o conjunto universo, o qual tem por objetivo avaliar os parâmetros encontrados pelo programa de parametrização. A partir do conjunto universo, selecionamos um subgrupo que chamaremos conjunto de parametrização. Este deve possuir a menor quantidade possível de estruturas capazes de representar o conjunto universo.

Para definir o conjunto universo, devemos, a partir do CSD, preparar três conjuntos: (i) conjunto “original” que contém estruturas contendo os átomos originalmente parametrizados pelo modelo RM1 (Recife Model 1). que até o momento são os seguintes: C, N, O, F, P, S, Cl, Br, I e H. Este conjunto também poderá conter átomos de outros elementos químicos; (ii) conjunto “Metal” que contém estruturas com átomos do elemento metálico escolhido para ser parametrizado, neste caso o Mg (cada complexo deve conter uma ou mais unidades do cátion metálico); (iii) conjunto “Demais elementos” que contém as estruturas com o restante dos átomos da tabela periódica, podendo ainda haver neste, estruturas já contidas nos conjuntos acima.

O universo será formado por todos os complexos que tenham como átomo central - no poliedro de coordenação - o elemento a ser parametrizado. Assim, se na

célula cristalográfica o cátion magnésio unicamente estiver isolado como contra-íon do complexo, este não participará do conjunto. O metal deverá possuir, como ligantes unicamente os átomos já parametrizados pelo modelo RM1. Desta maneira, para definir este universo executamos a seguinte operação nos conjuntos: o conjunto “original” fará intersecção com o conjunto “Mg” (B) e o complementar do terceiro (C), como observamos na Figura 19.

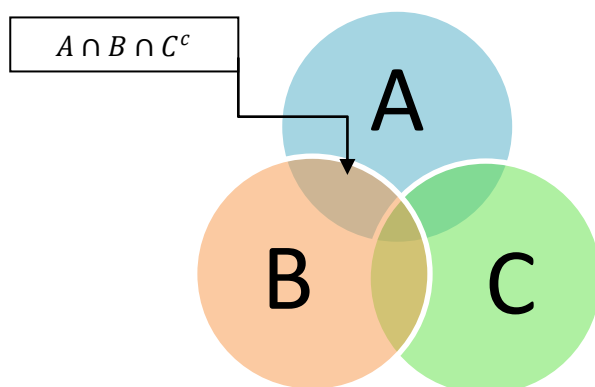


Figura 19 - Representação de Venn para a obtenção do conjunto universo de estruturas cristalográficas utilizadas no processo de parametrização. Onde; A = conjunto “original”; B = conjunto “Mg”; C = conjunto “Demais elementos”. A operação. $A \cap B \cap C^c$, resulta o conjunto universo de estruturas em estudo, que aqui chamamos grupo teste.

Ainda uma filtragem de estruturas desse universo deverá ser feita, antes que seja iniciada a parametrização. Devemos verificar se o fator residual (fator-R) é o menor possível, já que este relaciona, por exemplo, a estrutura calculada a partir do difratograma simulado com a experimental, como observamos na Equação 26:

$$R = \frac{\sum |F^{obs}| - |F^{calc}|}{\sum |F^{obs}|} \quad (27)$$

onde; R é o fator-R; F^{obs} é o fator da estrutura observada experimentalmente que pode ser calculado a partir de $\sqrt{I_{corr}}$, onde I_{corr} é a intensidade experimental do feixe difratado de raio-x corrigido e F^{calc} é o fator de estrutura simulado que pode ser calculado para qualquer reflexão num sistema de coordenadas h,k,l desde que as coordenadas dos átomos na célula unitária sejam previamente conhecidas⁷⁵. O fator-r representa um importante auxílio na resolução de estruturas cristalográficas

desconhecidas, pois geralmente quando este valor está entre 0.1 a 0.2 a estrutura está essencialmente correta. Porém, uma estrutura que foi totalmente resolvida usando dados de intensidade de boa qualidade, apresenta tipicamente um R no intervalo de 0.02 a 0.06⁷⁶. Desta maneira, optamos por utilizar estruturas que tenham um R igual ou menor que 0.05.

Ainda optamos por selecionar no programa a opção *not disorder*, que está diretamente relacionada à disposição dos átomos dentro da célula unitária do cristal. Em outras palavras, nos nossos bancos de dados só deverá haver cristais que tenham uma boa uniformidade em termos de configuração em suas células unitárias. Acima do zero absoluto todos os cristais possuem vacâncias em sua rede. Estas lacunas introduzem desordem no cristal quebrando assim a sua periodicidade perfeita, podendo aumentar de acordo com o tipo de cristal. Ligas costumam apresentar grande desordem, a tal ponto que um átomo poderá ocupar uma determinada posição em uma célula unitária e em outra célula unitária do mesmo cristal o átomo equivalente estará ocupando a posição espacial de outro átomo dentro do mesmo cristal⁷⁶. Desta maneira, cada célula unitária do cristal poderá apresentar diferentes configurações. Para obter estruturas com um erro menor desta propriedade inerente à própria natureza dos cristais é que escolhemos esta opção, “*not disorder*”.

Não optamos pelo item *no ions*, pois não há problema se o complexo estiver carregado, já que neste estudo o que realmente importa é a geometria da estrutura.

A próxima etapa se realiza artesanalmente, pois as estruturas advindas do banco de dados devem ser revisadas e avaliadas individualmente. Para isso devemos salvar as estruturas em um formato que seja escrito em coordenadas cartesianas, para que possam ser visualizadas no *hyperchem* que é o *GUI (Graphical User Interface* ou interface gráfica do usuário) escolhido pelo grupo para observar individualmente os complexos cristalinos. A extensão de arquivo convencional para ser utilizada nesta etapa foi o .pdb (*protein data bank*). A seguir apresentamos o formato da matriz na Figura 20.

HETATM	1	Mg		1	-0.266	0.058	0.081
HETATM	2	O		2	-0.484	2.057	-0.164
HETATM	3	O		3	1.745	0.206	-0.115
HETATM	4	O		4	-0.058	0.280	2.084
HETATM	5	O		5	-0.100	-1.946	0.327
HETATM	6	O		6	-2.277	-0.018	0.311
HETATM	7	O		7	-0.479	-0.231	-1.913
HETATM	8	H		8	-1.354	2.134	-0.518
HETATM	9	H		9	-0.582	2.367	0.720
HETATM	10	H		10	-2.399	0.292	1.192
HETATM	11	H		11	-2.446	-0.943	0.379
HETATM	12	H		12	1.893	-0.060	-1.007
HETATM	13	H		13	1.888	1.137	-0.134
HETATM	14	H		14	0.824	-2.065	0.480
HETATM	15	H		15	-0.220	-2.276	-0.547
HETATM	16	H		16	-0.324	0.627	-2.271
HETATM	17	H	A	17	-1.409	-0.348	-2.012
HETATM	18	H		18	0.845	0.531	2.177
HETATM	19	H		19	-0.070	-0.610	2.395
CONNECT	1	2	3	4	5	6	7
CONNECT	2	1	8	9			
CONNECT	3	1	12	13			
CONNECT	4	1	18	19			
CONNECT	5	1	14	15			
CONNECT	6	1	10	11			
CONNECT	7	1	16	17			
CONNECT	8	2				C	
CONNECT	9	2					D
CONNECT	10	6					E
CONNECT	11	6					
CONNECT	12	3			B		
CONNECT	13	3					
CONNECT	14	5					
CONNECT	15	5					
CONNECT	16	7					
CONNECT	17	7					
CONNECT	18	4					
CONNECT	19	4					
END							

Figura 20 - Matriz representativa em coordenadas cartesianas (x,y,z) no formato .pdb, para o complexo cristalino hexaaquamagnésio, onde cada linha HETATM designa um átomo do complexo. A coluna A diz qual é o elemento, B determina um número que identifica o átomo que será utilizado nas linhas CONECT para identificar como os átomos estão conectados C, D e E indicam respectivamente as coordenadas X, Y e Z de cada elemento.

No formato .pdb, além das coordenadas cartesianas, as conectividades são informadas e podemos verificá-las nas linhas “CONNECT”, onde cada número simboliza um átomo. No exemplo acima podemos ver que na primeira linha CONNECT o átomo 1, que representa o elemento magnésio, está seguido respectivamente pelos números 2,3,4,5,6 e 7 que representam necessariamente o elemento oxigênio como indica a coluna A em conformidade com a B. Desta maneira, podemos inclusive representar o complexo desta matriz no modelo de bolas e bastões como vemos na Figura 21.

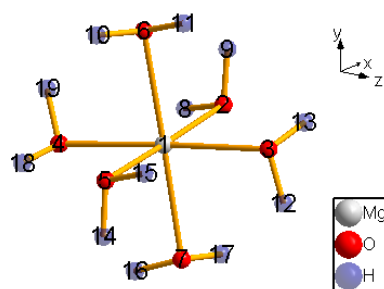


Figura 21 - Representação esquemática de bolas e bastões para o complexo cristalino hexaaquamagnésio em conformidade com matriz da Figura 20.

O *hyperchem* é muito importante nesta etapa do processo, pois a difração de raios-x como ferramenta analítica não permite ao cristalógrafo visualizar os hidrogênios da estrutura. Desta maneira, estes devem ser adicionados manualmente. Justamente por essa razão, é parte fundamental do estudo observar individualmente cada estrutura avaliando se há incoerência no complexo e se os hidrogênios foram adicionados coerentemente. Em caso negativo, adicionamos os mesmos manualmente. Se alguma estrutura defeituosa escapar à filtragem, todo o processo de parametrização poderá ser comprometido.

Outra tarefa importante é retirar todos os polímeros e íons em que o metal não esteja complexado.

De uma maneira geral, os hidrogênios presentes nos complexos foram adicionados depois da análise de raios-X. Desta maneira, todos os hidrogênios das estruturas deverão ser otimizados pelo programa MOPAC 2007. Para a otimização dos hidrogênios, optamos pelo modelo PM6. Porém, antes de otimizar as estruturas deveremos construir os inputs utilizando as seguintes normas:

- 1) Quando o complexo for um monômero, então o metal vem na primeira posição e os átomos diretamente coordenados a ele vêm logo abaixo dele.
- 2) Quando o complexo for um dímero, então temos que ver se existem átomos que estão coordenados aos dois íons metálicos. Se não houver, então a estrutura deve ser organizada da seguinte forma: primeiro o íon metálico, seguido pelos átomos que estão coordenados a ele, depois o segundo íon metálico seguido dos átomos que estão coordenados a ele, etc.

3) Se houver átomos que estão coordenados aos dois íons metálicos, então esses átomos devem ser identificados (só usaremos estruturas com dois átomos coordenados aos dois íons metálicos, pois é a grande maioria dos casos) e a organização dos poliedros fica da seguinte forma: o primeiro íon metálico seguido dos átomos que só estão coordenados a ele, depois o segundo íon metálico seguido dos dois átomos que estão coordenados aos dois íons metálicos seguidos dos átomos que estão coordenados apenas ao segundo íon metálico.

Neste momento, podemos dizer que temos o universo de estruturas em estudo que origina o grupo de parametrização. O critério para distinguir estes dois é simples: os que não participarem do grupo de parametrização farão parte do grupo teste, que ajudará a validar os parâmetros encontrados pela comparação direta.

3.2- Definição do Grupo de Parametrização

Para definir o grupo de parametrização, deve-se realizar primeiro a análise de cluster como apresentado na Fig.22.

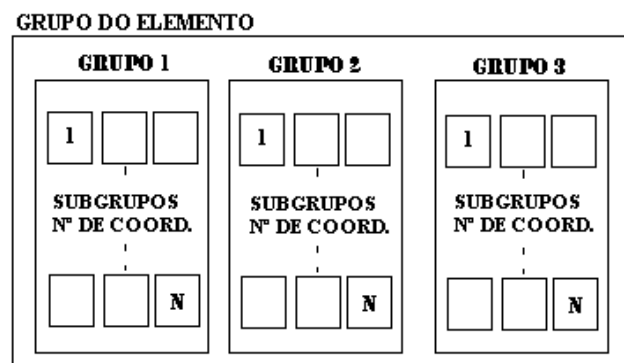


Figura 22- Representação esquemática da análise de cluster, para escolha do grupo de parametrização.

No processo devemos dividir todas as estruturas em três grandes grupos. O primeiro grupo deve conter somente complexos que tenham átomos de carbono no poliedro de coordenação, enquanto no segundo poderá haver carbono mas também deverá possuir oxigênio e/ou nitrogênio. Se a estrutura não se encaixar em nenhum desses grupos, deverá participar do grupo 3.

Dentro destes grupos existem subgrupos. Cada um destes deve conter estruturas com um único número de coordenação. Para poder separar estes grupos utilizamos uma planilha. Na primeira coluna, *Refcode*, se escreve o código de referência da estrutura no CSD¹⁵⁻¹⁷; na segunda coluna o número de coordenação; na terceira, a carga do complexo; na quarta, a quantidade do elemento a ser parametrizado presente na célula do cristal; na quinta, o número de *ligantes livre*, que representa os ligantes que não se associam a nenhum elemento que seja ou faça parte dos outros ligantes do poliedro de coordenação; a sexta coluna, “heter.” diz quais são os diferentes elementos que compõem o poliedro de coordenação; a sétima diz o número de átomos que fazem parte da estrutura (esta informação ajuda o pesquisador a ter uma idéia sobre o tamanho do cristal); como observamos no anexo 1.

3.2.1- Análise de Cluster DIANA

A partir dos subgrupos descritos na seção anterior. serão definidas as estruturas que participarão da parametrização mediante uma segunda análise de cluster feita a partir do método de desagrupamento *DIANA*⁷⁷ (*Divisive Analyses Cluster*). Esta análise foi utilizada pela primeira vez em nosso laboratório por Caroline A. C. de Lima na obtenção do conjunto de parametrização para o átomo de selênio⁷⁸, se mostrando muito eficaz em reunir um conjunto amostral bem representativo do universo de seu banco de dados. O algoritmo do método DIANA consiste na divisão de um cluster em dois até que na última divisão cada membro do conjunto seja, individualmente, um cluster. Desta maneira, a hierarquia do cluster é construída em $N - 1$ passos, onde N é o número de objetos como observamos na Figura 23.

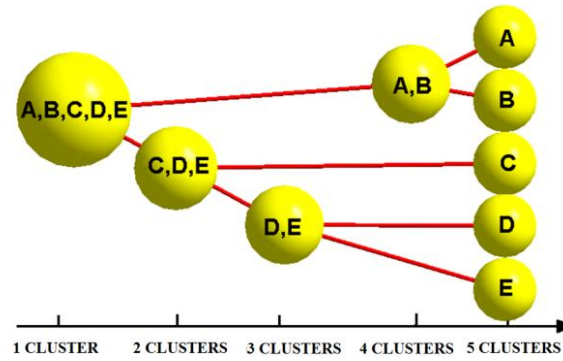


Figura 23 - Representação esquemática do método de análise de desagrupamento DIANA, onde o cluster inicial contém cinco elementos (A, B, C, D e E). O primeiro passo consiste em dividi-lo em dois clusters ((C, D, E) e (A, B)) e assim por diante. Ao total, esta análise hierárquica é feita em quatro passos a partir do cluster inicial.

O critério utilizado para separar os clusters é a dissimilaridade entre os objetos, que é dada por:

$$\delta_i = \frac{1}{|A| - 1} \sum_{j \in A, j \neq i} \delta_{ij} \quad (28)$$

onde, δ_i é a dissimilaridade média entre o objeto i e todos os outros objetos no cluster A ; $|A|$ indica o número de objetos. Isto acontece em cada etapa do processo. A métrica utilizada para aferir a dissimilaridade que indicamos é a euclideana. O algoritmo do DIANA inicialmente encontra o objeto mais dissimilar do grupo. Este será o primeiro elemento do próximo cluster. Em seguida, o algoritmo reagrupa os demais objetos que estejam mais próximos do novo grupo quando comparado ao grupo original. A consequência deste processo é a divisão em dois grupos até que no último passo cada objeto será um cluster.

Observamos através da figura abaixo esta análise hierárquica divisiva, através de seus níveis obtidos a partir do critério de dissimilaridade:

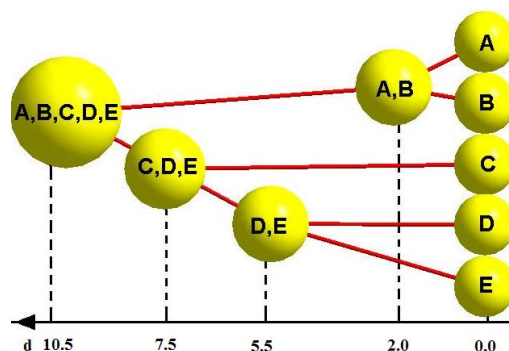


Figura 24 - Gráfico apresentado na Figura 23, apresentando o nível de cada divisão

Abaixo simulamos um dendrograma genérico para exemplificar como seria a escolha do conjunto de complexos de magnésio obtidos a partir do erro calculado pela Equação 4.

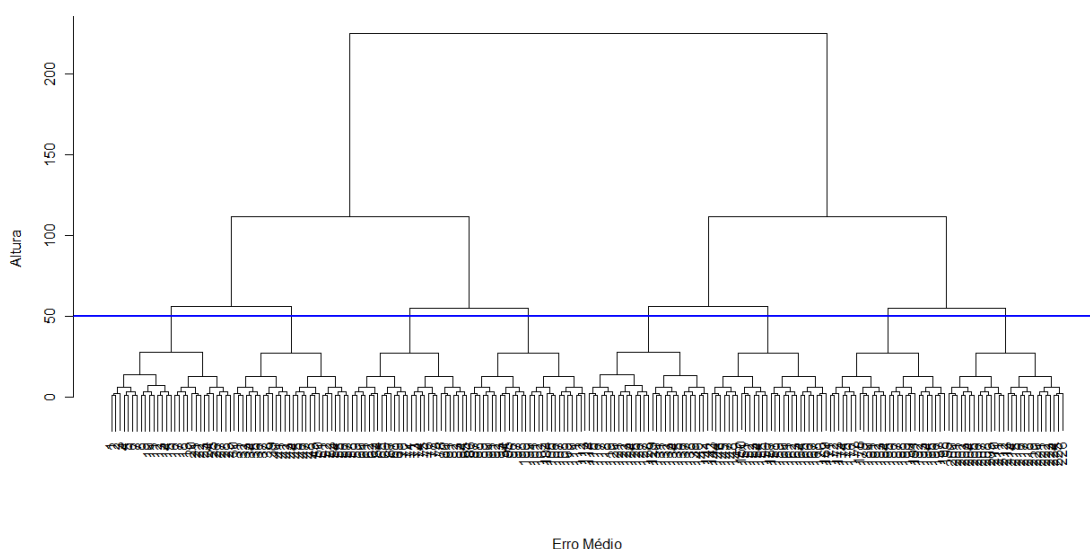


Figura 25 - Dendrograma simulado para um conjunto de 226 complexos considerando o erro médio das geometrias dos complexos. A altura do corte para a separação de novos clusters é feita na altura 50.

Neste exemplo escolhemos cortar o dendrograma na altura 50, pois assim poderemos avaliar o grupo em 8 clusters como mostramos abaixo.

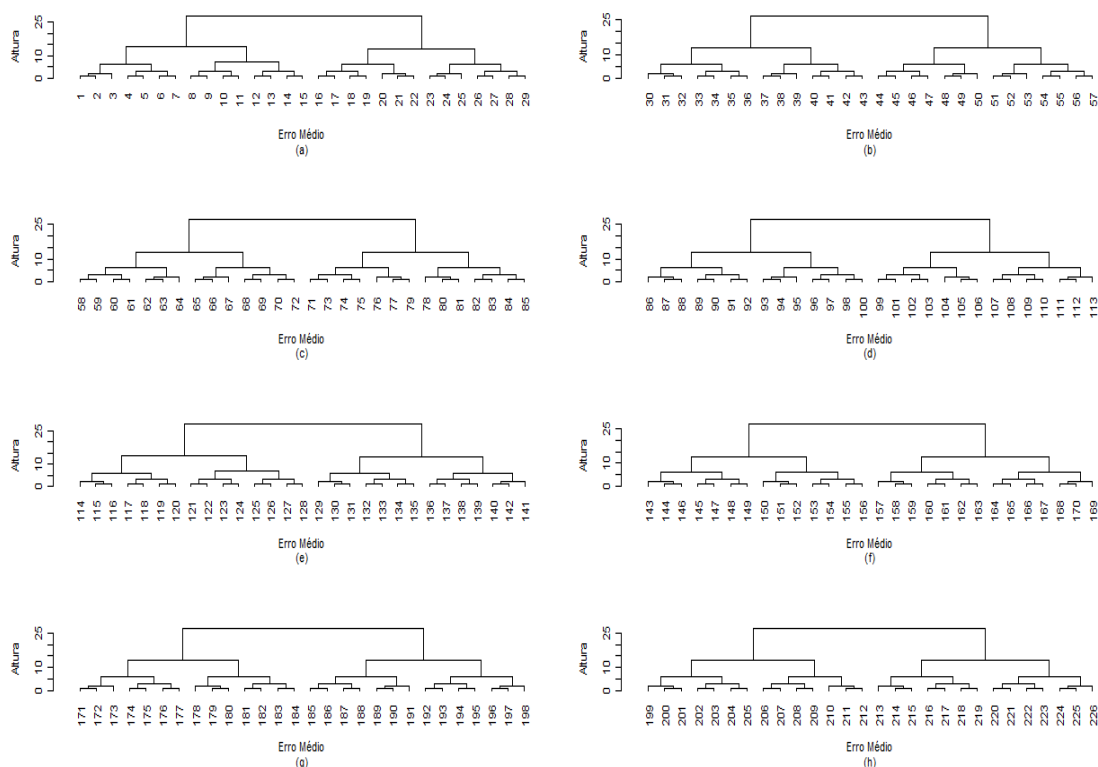


Figura 26 - Clusters obtidos a partir do corte na altura 50 do dendrograma da Figura 25.

Chamaremos estes clusters de a, b, c, d, e, f, g, h. Neste caso os 8 clusters contêm ao total duzentos e vinte e seis diferentes valores para o erro médio. A quantidade de objetos contidos em cada cluster poderá variar de acordo com os erros gerados, neste caso os clusters tiveram quase que o mesmo número de objetos. Em geral, a quantidade de estruturas químicas que participarão do grupo de parametrização poderá ser dada em termos de uma porcentagem ou de uma quantidade fixa de estruturas a ser definida pelo pesquisador. A próxima equação poderá ser usada para os dois critérios que irão definir o conjunto de parametrização:

$$N_p = \sum_{i=1}^w \Psi M_i . \quad (29)$$

Onde, N_p é o número total de complexos que participarão do conjunto de parametrização; Ψ é a porcentagem de estruturas em relação ao grupo teste que participarão do grupo de parametrização; M_i é a quantidade de complexos contidos no cluster i e w é a quantidade total de clusters.

Como o processo de encontrar parâmetros envolve o cálculo repetitivo das moléculas do conjunto de parametrização, estes complexos, não deverão ser grandes a ponto de atrasarem o processo de parametrização. Definida a quantidade de estruturas que serão retiradas de cada cluster, o critério de escolha será dado agora a partir de duas considerações: o tempo de cálculo e ambiente químico do poliedro de coordenação. Ou seja, as estruturas deverão ter o menor tempo de cálculo possível e estar no ambiente químico mais diferente possível. Para esta segunda análise utilizamos a tabela preparada na primeira análise de cluster (anexo 1). A finalidade destas análises é obter um conjunto de parametrização que represente bem o conjunto teste.

Tendo selecionado o conjunto de parametrização, procederemos a parametrização dos átomos escolhidos com programas de parametrização já existentes no grupo. Esta etapa será essencialmente um processo artesanal. A partir deste momento restará apenas acompanhar o processo para detectar possíveis problemas que possam acontecer durante este. A escolha do melhor conjunto de parâmetros de partida será realizada com base no novo critério que será estabelecido na próxima seção.

4. O PROCESSO DE PARAMETRIZAÇÃO

4.1- A Parametrização

Concluída a escolha do grupo de parametrização, resta executar o processo em si da busca dos parâmetros. Durante anos nosso grupo vem associando e desenvolvendo um arsenal de programas que auxiliam a execução desta tarefa.

A busca e validação dos parâmetros não se resumem à utilização da função resposta, o químico precisa certificar que os valores obtidos a partir das parcelas da função resposta estejam equilibrados. Por exemplo, o programa de parametrização poderá retornar um conjunto de parâmetros que quando utilizados, na otimização de geometria do conjunto de parametrização, resulte em estruturas químicas com distâncias muito boas e ângulos horríveis. Desta maneira, o resultado da função resposta para esse grupo poderá ter um valor pequeno, mas quando analisarmos as suas respectivas parcelas observaremos que a variável radial terá um pequeno valor enquanto que a parte angular terá um valor maior. Logo, o melhor valor que a função resposta retornar deverá estar equilibrada em suas parcelas.

Justamente para poder avaliar os complexos cristalinos em suas coordenadas radiais, latitudinais e longitudinais é que a função resposta planetário foi desenvolvida. Este desenvolvimento deve permitir ao químico avaliar as variáveis do planetário ponderadamente, já que os termos das equações estarão equilibrados.

O processo de parametrização em si acontece em algumas etapas: (i) escolhe-se um algoritmo de minimização não linear multivariada para procurar os parâmetros desejados que descreveremos melhor neste capítulo; (ii) estabelece-se os parâmetros para o *sparkle* que sejam capazes de reproduzirem a melhor geometria possível para os complexos do grupo de parametrização; (iii) cada novo conjunto de parâmetros encontrado pelo programa de parametrização é submetido ao software MOPAC para que a geometria dos complexos possa ser calculada. Terminada a otimização da geometria dos complexos contidos no banco de parametrização resta avaliar se os novos parâmetros designam boas geometrias; (iv) para a avaliação da geometria do poliedro de coordenação do complexo é que se aplica a função resposta.

Como observamos o procedimento de buscar e avaliar parâmetros é um processo iterativo, ou seja, pode se repetir quantas vezes for necessário, podendo demorar tanto uma semana quanto anos. Justamente para tentar minimizar este tempo é que estamos desenvolvendo novas técnicas de parametrização, desde a escolha dos complexos a partir do banco de dados até o programa que será utilizado no processo de parametrização.

Neste capítulo abordaremos algumas destas etapas e inclusive sugeriremos melhorias na etapa da escolha do parâmetro inicial para o modelo Sparkle. Descreveremos então para começar, de forma sucinta na próxima seção como os algoritmos de minimização funcionam na busca de novos parâmetros.

4.2- Técnicas de minimização não-linear multivariada

- **Simplex**

O simplex é um algoritmo que encontra o mínimo para uma função de n variáveis. O método é iterativo, pois depende de uma sequência finita de operações, onde cada uma é regida pelo resultado obtido na etapa anterior.

Um simplex é uma figura geométrica de $n + 1$ pontos que é construído no espaço dimensional das variáveis (por exemplo: em duas dimensões o simplex é um triângulo). O ponto inicial é fornecido pelo pesquisador, este passa a ser o primeiro vértice do simplex, os demais n vértices são gerados a partir da rotina de Parkinson e Hutchinson⁷⁹. O vértice do simplex com o maior valor da função é refletido no centro de gravidade dos vértices restantes e o valor da função para este novo ponto é comparado com os valores dos outros pontos. Dependendo deste teste, o resultado será aceito ou rejeitado, podendo assim haver uma expansão ou contração do simplex^{8,79}. Quando não há avanço, o simplex é contraído e o método é repetido.

Este método tende a ser demorado, mas é robusto e portanto muito utilizado para funções que são sujeitas a imprecisões. Em nossa pesquisa utilizamos o software parametric desenvolvido no grupo, onde já foi implementado a rotina do simplex E04CCF da NAG (*Numerical Algorithms Group*), que minimiza uma função qualquer $\Lambda(x)$ de n variáveis independentes $x = (x_1, x_2, \dots, x_n)^T$. As derivadas da função não precisam ser fornecidas.

- **GSA – Generalized Simulated Annealing**

O GSA é um método de otimização global, baseado no SA (*simulated annealing*) que também é conhecido como máquina de Boltzman, pois está baseado na estatística de Boltzman-Gibbs, proposto por Kirkpatrick e outros⁸⁰. Este modelo foi inspirado no processo de recozimento de sistemas físico, técnica muito utilizada na metalurgia, ou seja, um metal é aquecido e resfriado lentamente até que atinja a conformação mais estável. Foi a partir da estatística de Tsallis^{81,82} que se desenvolveu o algoritmo do GSA⁸⁰, também conhecido por máquina de Tsallis. A família dos SA em geral segue o seguinte algoritmo: (i) dada a função custo, escolhemos aleatoriamente ou não, um conjunto de parâmetros que represente uma solução inicial de “energia”, E_{ref} e também um valor para o parâmetro Θ (equivalente à temperatura na termodinâmica estatística); (ii) perturbamos os parâmetros da função custo utilizando a função visita, e calculamos a “nova energia” do sistema, E_N ; (iii) calculamos a variação da energia como segue: $\Delta E = E_N - E_{ref}$; se $\Delta E \leq 0$, então E_N passa a ser o novo E_{ref} ; (iv) se $\Delta E > 0$, então, o ponto de referência é melhor que o novo ponto, porém pode ser rejeitado dependendo da função de probabilidade e de um valor aleatório, como definido nos critérios de Metrópolis⁸³; (v) a temperatura Θ é reduzida de acordo com a função de decaimento da temperatura; (vi) e na última etapa a segunda e quinta etapas são repetidas até que Θ seja tão pequeno que nenhum movimento seja aceito, ou até que alguma das condições de paradas seja satisfeita.

4.3- A Escolha do Parâmetro Inicial para a Parametrização

Até hoje, o critério utilizado para escolher os parâmetros iniciais do *sparkle* que iriam ser utilizados na parametrização, está baseado em pegar os parâmetros do *sparkle* que representam o átomo já parametrizado que estivesse mais próximo dele na tabela periódica. Porém estes parâmetros geralmente não designam boas geometrias para o conjunto de parametrização. Sabendo que muitas vezes o gasto computacional é muito alto para encontrar um parâmetro que dê uma qualidade mínima a geometria dos poliedros de coordenação, optamos por desenvolver uma técnica baseada no modelo Sparkle para encontrar um conjunto de parâmetros que possam representar melhores geometrias.

O novo critério de análise dos parâmetros iniciais está diretamente baseado na simulação e observação do gráfico da energia de repulsão caroço-caroço em função dos comprimentos de ligação metal-átomo coordenante. A equação da energia de repulsão caroço-caroço é composta por duas gaussianas, apresentando assim alguns máximos e mínimos. Sabemos que um ponto de mínimo pode representar uma ligação química, logo, ao conhecer os valores destes mínimos, poderemos saber qual região de distâncias de ligação estes parâmetros representarão bem.

De posse destas informações basta ajustar os parâmetros das gaussianas de maneira que seus mínimos contemplem exatamente as distâncias das ligações químicas.

Já mostramos no capítulo 1 que os parâmetros a_{k_A} e a_{k_B} são os parâmetros de intensidade da gaussiana, enquanto que os parâmetros $b_{k_{AB}}$ e b_{k_B} estão relacionados à largura da gaussiana. Desta maneira, sugerimos a análise direta da função pela variação destes parâmetros.

Os parâmetros deverão ser ajustados de acordo com as distâncias entre o metal e os átomos que compõem o poliedro de coordenação. Para mostrar como estas distâncias podem variar apresentamos na Figura 27 um histograma dos comprimentos de ligação Mg – L.

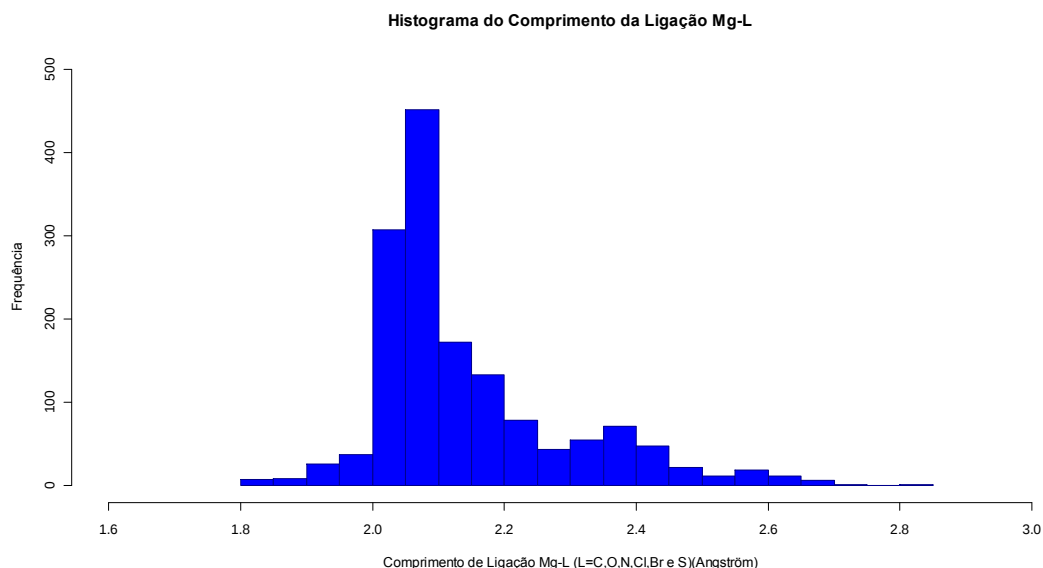


Figura 27 - Histograma das distâncias Mg-L, onde L poderá ser qualquer um dos átomos: C, O, N, Cl, Br e S. A maior quantidade de medidas é observada no intervalo de 2.0 Å a 2.2 Å.

Podemos observar que o histograma acima apresenta 1508 medidas, que é justamente o número de medidas para os diferentes comprimentos de ligações Mg-L,

onde L representa os átomos O, N, C, S, Br e Cl. dos poliedros de coordenação contidos em todo o nosso banco de dados para o Mg. Observamos ainda que estes comprimentos de ligação Mg-L, estão distribuídos numa faixa aproximada de 1.8 Å a um pouco mais de 2.8 Å, o que significa uma extensa faixa de comprimentos de ligação. Esta observação mostra que não é possível ajustar as gaussianas para cada uma destas distâncias simultaneamente, o que implica em um ajuste do modelo a uma determinada quantidade de comprimentos de ligação. Em outras palavras isso quer dizer que teremos que escolher quais átomos ligantes do poliedro de coordenação são os mais importantes para serem representados. Com essa finalidade apresentamos abaixo o histograma de cada comprimento de ligação Mg-L a começar pelo carbono e enxofre.

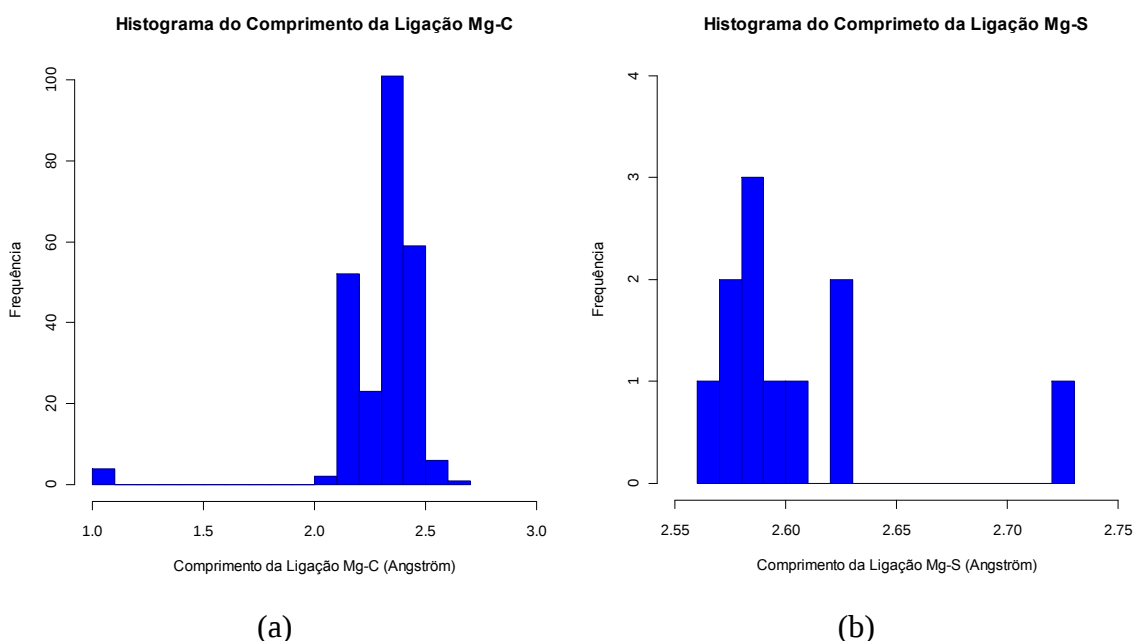


Figura 28 – (a) Histograma para o comprimento de ligação Mg-C, possuindo a maioria de suas medidas no intervalo de 2.3 Å a 2.4 Å. (b) Histograma para o comprimento de ligação Mg-S, que tem as suas medidas a partir de 2.55 Å.

Podemos observar através dos histogramas que a maioria das medidas para a distância Mg-C estão acima de 2 Å, apresentando apenas 4 medidas em torno de 1Å, enquanto que a maioria dos comprimentos da ligação Mg-S, estão além de 2.55 Å, Logo, percebemos que, mesmo tendo poucas medidas para analisar estes dois tipos de ligação, ambas ocupam regiões de distâncias diferentes.

A seguir apresentamos os histogramas referentes aos comprimentos de ligação Mg-Br e Mg-Cl respectivamente.

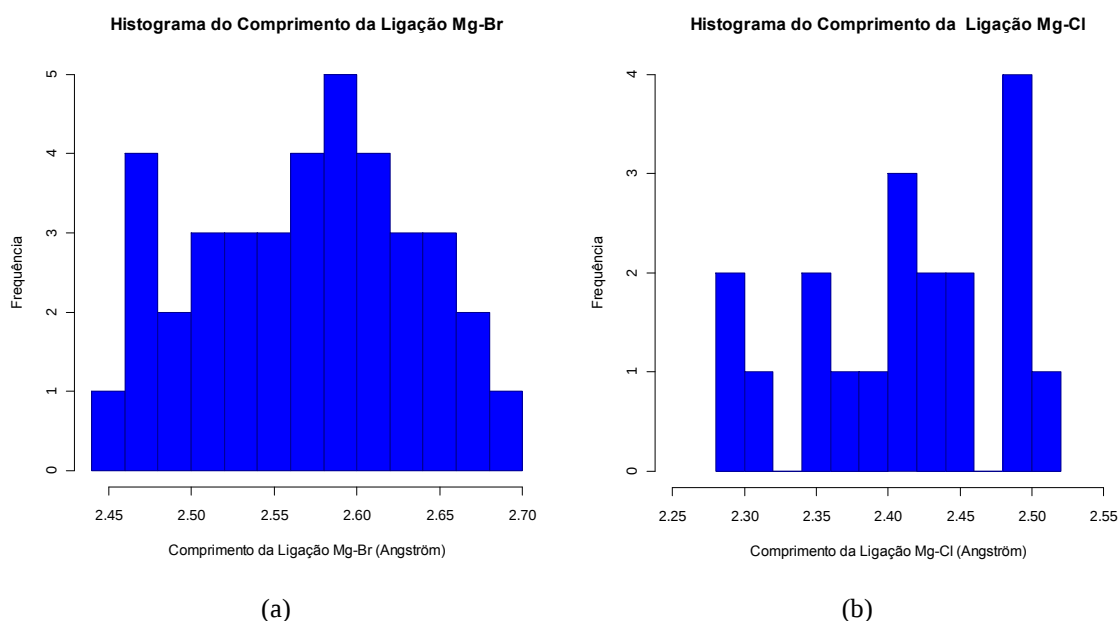


Figura 29 - (a) Histograma para o comprimento de ligação Mg-Br, cujos valores se distribuem no intervalo de 2.45 Å a 2.70 Å. (b) Histograma para o comprimento de ligação Mg-Cl, que tem as suas medidas distribuídas no intervalo de aproximadamente 2.28 Å a 2.52 Å.

Nestes histogramas observamos que o comprimento da ligação Mg-Br está distribuído na faixa de 2.45 Å a 2.70 Å aproximadamente, tendo a maior concentração de medidas próxima a 2.59Å, Apesar de haver poucas medidas para o estudo dessa ligação, observamos que é nesta região que provavelmente observaremos as medidas para o comprimento Mg-Br.

Para o Histograma Mg-Cl, observamos que a distribuição de suas medidas não são contínuas tendo determinadas regiões vacantes. Desta maneira, se esta análise representar todo o universo de comprimentos da ligação Mg-Cl, futuras medidas Mg-L de novos poliedros deverão pertencer a esta mesma região observada no histograma. Também vale a pena observar que a maioria de suas medidas são encontradas fora da região dos comprimentos da ligação Mg-Br.

Por fim apresentaremos os dois últimos histogramas correspondentes aos comprimentos das ligações Mg-O e Mg-N.

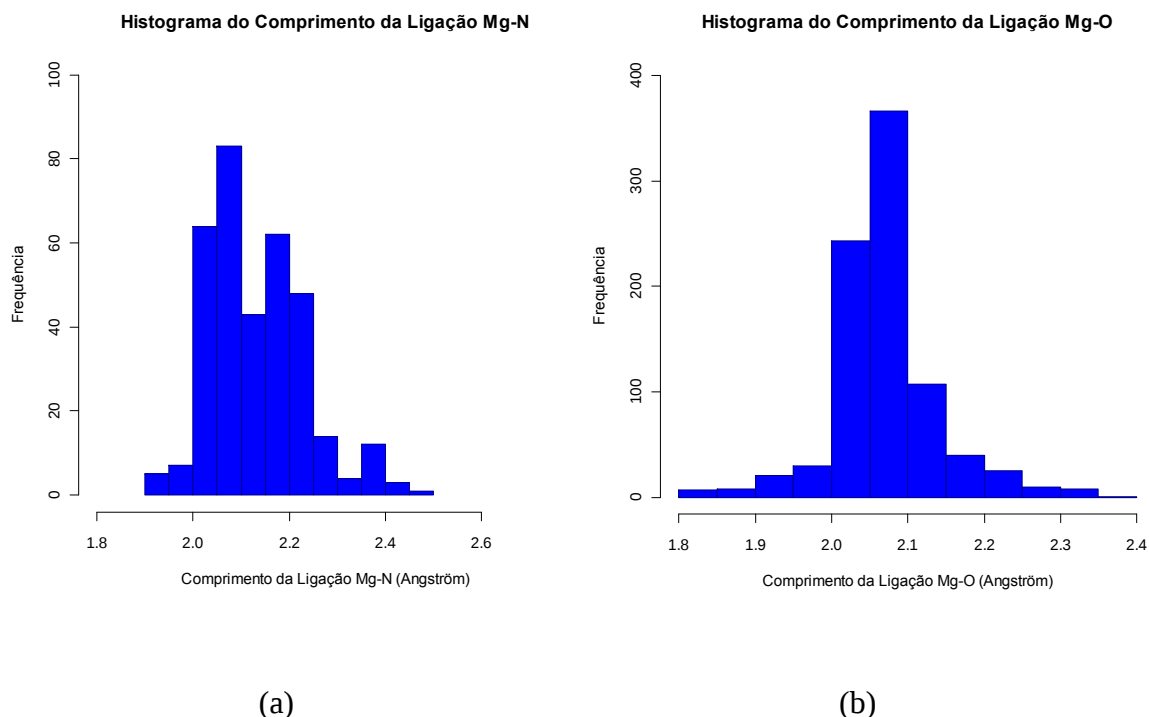


Figura 30 - (a) Histograma para o comprimento de ligação Mg-N, onde a maioria das medidas se encontram no intervalo de 2.0 Å a 2.2 Å. (b) Histograma para o comprimento de ligação Mg-O, cujo a maioria das medidas coincidem com o intervalo onde estão a maioria das medidas da ligação Mg-N variando de 2.0 Å a 2.2 Å.

Esta última análise para os histogramas nos permite perceber que o maior conjunto de medidas do banco de dados são exatamente para as ligações Mg-O e Mg-N, contidas no poliedro de coordenação do complexo cristalino. Outra interessante características destes dois histogramas é que as medidas do comprimento destas duas ligações se encontram praticamente na mesma região. Desta maneira, isto mostra que se resolvermos ajustar os parâmetros da gaussiana para o intervalo do oxigênio, também resolveremos a distância do nitrogênio.

Como podemos observar até agora, cada tipo de ligação apresenta diferentes comprimentos podendo até em muitos casos terem medidas iguais. Porém, um ajuste utilizando a Equação 2 para todas essas medidas ao mesmo tempo seria inviável, pois se alargarmos demais as gaussianas, restringiríamos a descrição das ligações a um pequeno grupo, já que a região de mínimo seria menor. Por outro lado se afinarmos muito, perderíamos a descrição de muitas das ligações existentes. Desta maneira, o equilíbrio deve ser utilizado. Para experimentar esta técnica, decidimos moldar empiricamente as gaussianas privilegiando as ligações Mg-O e Mg-N, já que estas duas ligações representam mais de 80% das estruturas do banco de dados.

Os parâmetros das gaussianas da Equação 2 ajustadas considerando as distâncias das ligações Mg-O e Mg-N são apresentados na Tabela 17.

Tabela 17- Valores dos parâmetros de partida ajustados para a equação da energia de repulsão caroço-caroço do *sparkle* que representam o íon Mg(II).

GSS	54.7468
ALP	4.0911
a_1	0.4619
b_2	8.0724
c_3	1.5071
a_1	1.3153
b_2	9.7447
c_3	2.7316

Para uma melhor análise destes parâmetros de partida, calculamos individualmente a energia de repulsão caroço-caroço para as ligações Mg-O e Mg-N em função da distância no intervalo de 0 a 5.0 Å. Este intervalo está muito além das distâncias apresentadas por todos os comprimentos de ligação química do magnésio com os demais átomos coordenantes, contidos em nosso banco como podemos observar na Figura 27. Ainda, calculamos a contribuição energética do átomo de magnésio para estes gráficos de potenciais, como podemos observar na Figura 32.

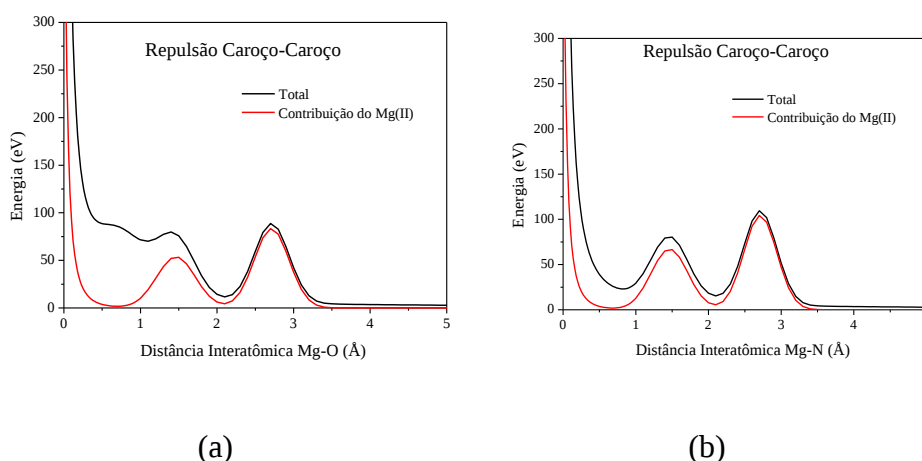


Figura 31 – (a) Gráfico da energia de repulsão caroço-caroço em função das distâncias interatômica Mg-O para o modelo RM1. (b) Gráfico da energia de repulsão caroço-caroço em função das distâncias interatômica Mg-N para o modelo RM1.

Podemos observar nos gráficos da energia total de repulsão caroço-caroço para as ligações Mg-O e Mg-N, que estes se mostraram bem diferentes no intervalo de 0 a 1.5 e bem semelhantes a partir deste intervalo. Na ligação Mg-O observamos que a energia de repulsão total caroço-caroço na região entre 0.3 e 1.0 Å apresenta uma energia superior quando comparada à mesma região no gráfico da ligação Mg-N. Em geral, o potencial da ligação Mg-N difere muito pouco do potencial do Mg isolado, ao contrário do potencial da ligação Mg-O, como podemos observar nos gráficos da Figura 32.

A partir destes gráficos podemos ainda observar que a maioria das distâncias nas ligações Mg-O e Mg-N observadas nos poliedros de coordenação do banco de dados do CSD estão justamente na região de 2 Å, que é onde colocamos o poço do potencial. Sendo assim, estes parâmetros devem nessa região de mínima energia, descrever de maneira equilibrada em geral as distâncias de ligações que estiverem contidas neste poço.

Com objetivo de avaliar os parâmetros de partida obtidos para o íon magnésio, otimizamos as geometrias dos complexos do grupo teste no MOPAC usando o algoritmo de otimização BFGS⁸⁴⁻⁸⁷ com um GNORM = 0.250 (gradiente de normalização). Os complexos foram ainda otimizados usando os modelos AM1, PM3 e PM6. Em seguida, calculamos o erro médio absoluto dos comprimentos de todas as ligações do magnésio com os átomos coordenantes para cada modelo. Os valores estão apresentados na Tabela 18.

Tabela 18 - Média para os erros médios absolutos dos comprimentos de ligações Mg- L, onde L representa os átomos:O; N; C; S; Br; e Cl.

Métodos Semi-empíricos					
Distâncias	N ^o de medidas	RM1	PM3	PM6	AM1
Sp - Sp	35	0.3471	0.4065	0.2176	0.3726
Sp - O	801	0.0753	0.2144	0.0547	0.0552
Sp - N	338	0.0724	0.2303	0.0534	0.0536
Sp - C	222	0.2414	0.1164	0.3160	0.1052
Sp - S	11	0.5752	0.1561	0.2097	0.0283
Sp - P	0	0.0000	0.0000	0.0000	0.0000
Sp - F	0	0.0000	0.0000	0.0000	0.0000
Sp - Cl	24	0.3851	0.0897	0.0529	0.0500
Sp - Br	37	0.5623	0.1760	0.0584	0.0736
Sp - I	1	0.5300	0.3787	0.2047	0.2119
Sp - L	1475	0.1288	0.2056	0.1001	0.0708
L - L	3473	0.2261	0.2272	0.2426	0.2021

Esta tabela mostrou que o ajuste dos parâmetros das gaussianas em relação à distância observadas experimentalmente entre o *sparkle* (Sp), que neste caso está representando o íon de magnésio, e os átomos coordenantes, se mostrou eficaz na predição teórica da geometria de poliedros de coordenação dos complexos, cuja maioria das distâncias magnésio-átomos coordenantes se encontraram justamente no poço da energia de repulsão caroço-caroço.

Estes parâmetros do *sparkle* para o íon magnésio no método RM1, quando utilizados para a otimização das geometrias dos complexos, no geral levaram a melhores distâncias das ligações Sp-O e Sp-N dentro dos poliedros de coordenação, quando comparados ao método PM3, e muito próximos aos métodos AM1 e PM6 para estas ligações. Desta maneira, acreditamos ser esta técnica uma nova ferramenta que ajudará no processo de parametrização.

Outra importante informação que podemos retirar destes gráficos de energia de repulsão caroço-caroço e dos histogramas apresentados acima para distintas ligações magnésio-átomo coordenante, é que este tipo de equação possui uma limitação em representar todas as ligações químicas ao mesmo tempo. Podemos observar também que além de todos os métodos semi-empíricos aqui apresentados, assim como os parâmetros de partida escolhidos para o modelo Sparkle, se apresentaram muito bem na predição teórica de alguns tipos de ligações química e de forma menos satisfatória em outras, não se adequando tão bem ao mesmo tempo a todas as predições do comprimento das

ligações químicas. Este comportamento indica uma tendência geral dos modelos semi-empíricos na predição do comprimento de ligação Sp-L. Outra importante observação a ser feita é que todos estes métodos que foram submetidos ao teste, tem orbitais em seus modelos, enquanto que o modelo Sparkle na sua concepção atual não possui orbitais no metal. Desta maneira, mais uma alternativa para a busca de parâmetros que possam calcular complexos com uma geometria próxima ao limite teórico previsto neste estudo, poderá ser a inclusão de orbitais s e p no processo de parametrização do *sparkle*.

Conclusões

Observamos neste trabalho que é válido utilizar o modelo Sparkle para o cálculo de sistemas químicos que possuam contra-íons e/ou outras moléculas em sua célula unitária. Para chegar a esta conclusão avaliamos 45 compostos diferentes contendo o complexo hexaaquamagnésio, $[\text{Mg}(\text{H}_2\text{O})_6]^{+2}$. Observamos o quanto os contra-íons e/ou outras moléculas influenciam este complexo, exprimindo suas geometrias em coordenadas polares com a origem no próprio íon de magnésio.

Utilizamos métodos estatísticos para justificar a validade do modelo Sparkle, assumindo em todos os testes estatísticos aplicados neste estudo um nível de significância para o p-valor igual a 0.05. Utilizamos o teste estatístico de Bartels, para verificar aleatoriedade nas variáveis radial, latitudinal e longitudinal. Os valores encontrados para estas parcelas foram, respectivamente, 0.2147, 0.7070 e 0.2243. Este resultado mostrou que os contra-íons e/ou outras moléculas produzem perturbações na geometria do complexo que, no seu conjunto, se comportam de forma pseudo-aleatória. Uma vez que os dados se mostraram aleatórios sugerimos a distribuição gama, para a coordenada radial, a distribuição beta, para a coordenada latitudinal e por último a distribuição von Mises, para a coordenada longitudinal. Mostramos através do teste estatístico de Kolmogorov Smirnov que as coordenadas radiais e latitudinais seguiram as respectivas distribuições sugeridas, enquanto o teste estatístico de Watson mostrou que a coordenada longitudinal segue a distribuição von Mises.

O método de máxima verossimilhança permitiu encontrar os parâmetros de cada distribuição e através destes parâmetros chegamos aos valores das variâncias para as coordenadas radial (0.0001274), latitudinal (0.0001340) e longitudinal (6.96882×10^{-7}).

Os desvios médios absolutos das distâncias na geometria do complexo, provocados pelos contra-íons e/ou outras moléculas (0.01Å), foram calculados utilizando a variância obtida a partir do método de máxima verossimilhança e ficaram uma ordem de grandeza abaixo do erro médio normalmente observado para o modelo Sparkle/AM1 para os lantanídeos (0.2Å). Ainda calculamos para as coordenadas latitudinal e longitudinal, respectivamente, os valores de 2.08° e 2.34°. Estes resultados validam a realização de cálculos de otimização da geometria de complexos metálicos sem a inclusão explícita de contra-íons e/ou moléculas, ou até mesmo da rede cristalina,

tanto para o modelo Sparkle, quanto para cálculos *ab initio*. Ainda mostramos que o modelo Sparkle pode ser melhorado para que sejam obtidas geometrias de poliedros de coordenação mais exatas em, ao menos, uma ordem de grandeza, quando comparadas às geometrias obtidas no modelo atual, antes que sejam incluídos de maneira explícita nos cálculos os contra-íons e/ou outras moléculas.

Ainda mostramos que cálculos *ab initio* de complexos isolados, por mais sofisticados que sejam, não poderão apresentar erros médios absolutos para as distâncias inferiores a 0.01Å, sem que haja inclusão explícita de contra-íons e/ou outras moléculas, ou até mesmo do efeito do cristal como um todo.

Com o objetivo de atingir o erro teórico previsto pelo modelo, sugerimos duas novas ferramentas que deverão auxiliar na obtenção de melhores parâmetros que gerem complexos que representem esse limite teórico. A primeira é uma nova função resposta que chamamos planetário, ajustada pela primeira vez nas variáveis radial, latitudinal e longitudinal e calibrada utilizando-se os desvios padrão obtidos dos parâmetros resultantes do estudo das geometrias dos 45 complexos hexaaquamagnésio. A segunda ferramenta é um novo critério para a obtenção dos parâmetros de partida utilizados no processo de parametrização baseado na análise direta do gráfico da energia do potencial caroço-caroço em função da distância metal-átomo coordenante dos poliedros de coordenação experimentais.

Neste trabalho, otimizamos ainda a geometria do complexo hexaaquamagnésio isolado utilizando os métodos HF, B3LYP e MP2(Full) e as bases STO-3G, 3-21G, 6-31G, 6-31G*, 6-31+G e 6-31++G. Com a finalidade de avaliar a predição destes métodos, utilizamos a função resposta planetário. Esta se mostrou equilibrada, indicando inclusive que a variável radial prevista nestes modelos foi a que sofreu maior variação, constituindo-se assim na mais difícil de ser predita. Dos modelos *ab initio*, o que apresentou melhor resultado foi o MP2/3-21G.

Não somente testamos o equilíbrio da função resposta, mas também o critério de obtenção dos parâmetros de partida. O parâmetro obtido sem a realização de parametrização para o modelo Sparkle no método RM1 obteve erros menores que o método PM3 nas ligações Mg-O e Mg-N, erros que foram ainda muito semelhantes aos dos métodos PM3 e PM6. Este resultado foi muito importante, pois no CSD mais de 80% dos átomos coordenados ao Mg são O e N.

Desta maneira, ambas as ferramentas surgem como novas tecnologias que poderão ser aplicadas, não somente ao modelo Sparkle, mas também em outros modelos semi empíricos que nosso grupo porventura venha a desenvolver.

Perspectivas:

Começaremos o Sparkle/RM1 com os metais alcalinos e alcalino-terrosos, mais especificamente com o Mg^{2+} , como teste para a nova metodologia. Tudo funcionando a contento, a generalização inicialmente para Ca^{2+} e em seguida para Na^+ e K^+ será relativamente simples. Esses quatro cátions serão os primeiros a serem parametrizados devido à sua importância biológica.

Em seguida, parametrizaremos com esta nova metodologia os íons de todos os restantes metais alcalinos e alcalino-terrosos e inclusive faremos a inclusão de orbitais s e p no modelo Sparkle/RM1 com o intuito de tentar minimizar seu erro. Esperamos que agora os ângulos fiquem muito mais bem descritos, o que fará o Sparkle/RM1 ser o método mais exato em termos de predição da geometria destes complexos – e altamente competitivo com qualquer técnica *ab initio*.

Finalmente, começaremos uma tentativa de descrever as ligações presentes nos complexos dos metais de transição através, ou do modelo Sparkle, ou de uma combinação do modelo Sparkle com orbitais d explicitamente considerados. Não nos comprometemos a resolver o problema dos complexos dos metais de transição, mas apenas a iniciar estudos que apontem a melhor direção de tratamento deste importante problema no contexto dos métodos semi-empíricos desenvolvidos por nosso grupo.

Referências

1. Pascal, B. *Pensées*; L' Institute de France: Paris, 1670.pp. 38.
2. Stewart, J. J. P. Application of the PM6 method to modeling proteins. *Journal of Molecular Modeling* **2009**, 15 (7), 765-805.
3. De Sa, G. F.; Malta, O. L.; Donega, C. D.; Simas, A. M.; Longo, R. L.; Santa-Cruz, P. A.; da Silva, E. F. Spectroscopic properties and design of highly luminescent lanthanide coordination complexes. *Coordination Chemistry Reviews* **2000**, 196, 165-195.
4. Deandrade, A. V. M.; Dacosta, N. B.; Simas, A. M.; Desa, G. F. Sparkle Model for the Quantum-Chemical AM1 Calculation of Europium Complexes. *Chemical Physics Letters* **1994**, 227 (3), 349-353.
5. James J.P.Stewart . MOPAC Manual, [HTTP://openmopac.net/manual/index.html](http://openmopac.net/manual/index.html). 2007. 3-3-2009.
Ref Type: Online Source
6. Dewar, M. J. S.; Zebisch, E. G.; Healy, E. F.; Stewart, J. J. P. The Development and Use of Quantum-Mechanical Molecular-Models .76. Am1 - A New General-Purpose Quantum-Mechanical Molecular-Model. *Journal of the American Chemical Society* **1985**, 107 (13), 3902-3909.
7. Dean, J. A.; Lange, N. A. *Lange's Handbook of Chemistry*; 11 ed.; McGraw-Hill, New York: 1985.
8. Nelder, J. A.; Mead, R. A Simplex-Method for Function Minimization. *Computer Journal* **1965**, 7 (4), 308-313.
9. Watson, W. H.; Williams, R. J.; Stemple, N. R. The crystal structure of tris(acetylacetonato)(1,10-phenanthroline) europium(III). *Journal of Inorganic and Nuclear Chemistry* **1972**, 34 (2), 501-508.
10. De andrade, A. V. M.; Dacosta, N. B.; Simas, A. M.; Desa, G. F. Sparkle Model for the Quantum-Chemical AM1 Calculation of Europium Complexes of Coordination-Number-9. *Journal of Alloys and Compounds* **1995**, 225 (1-2), 55-59.
11. Rocha, G. B. SMLC II. Nova Versão do Modelo Sparkles para o Cálculo de Complexos de Lantanídeos: Parâmetros para AM1, PM3 e MNDO e Testes, dissertação de mestrado em química, Universidade Federal de Pernambuco. 1998.
12. Rocha, G. B.; Freire, R. O.; da Costa, N. B.; De Sa, G. F.; Simas, A. M. Sparkle model for AM1 calculation of lanthanide complexes: Improved parameters for europium. *Inorganic Chemistry* **2004**, 43 (7), 2346-2354.
13. *MOPAC 93 R2*, Fujitsu Limited, Tokyo (1994)., 2009
14. Rocha, G. B. Desenvolvimento e Parametrização de Métodos Semi- empíricos para Calculo de Molecula Orgânicas e Complexos de Íons Lantanídeos e Actinídeos, tese de doutorado em química, Universidade Federal de Pernambuco. 2002.

15. Allen, F. The Cambridge Structural Database: a quarter of a million crystal structures and rising. *Acta Crystallographica Section B* **2002**, 58 (3), 380-388.
16. Bruno, I. J.; Cole, J. C.; Edgington, P. R.; Kessler, M.; Macrae, C. F.; McCabe, P.; Pearson, J.; Taylor, R. New software for searching the Cambridge Structural Database and visualizing crystal structures. *Acta Crystallographica Section B* **2002**, 58 (3), 389-397.
17. Allen, F. H.; Motherwell, W. D. S. Applications of the Cambridge Structural Database in organic chemistry and crystal chemistry. *Acta Crystallographica Section B* **2002**, 58 (3), 407-422.
18. Bastos, C. C.; Freire, R. O.; Rocha, G. B.; Simas, A. M. Sparkle model for AM1 calculation of neodymium(III) coordination compounds. *Journal of Photochemistry and Photobiology A: Chemistry* **2006**, 177 (2-3), 225-237.
19. da Costa, J.; Freire, R. O.; Rocha, G. B.; Simas, A. M. Sparkle/AM1 modeling of holmium (III) complexes. *Polyhedron* **2005**, 24 (18), 3046-3051.
20. da, C., Jr.; Freire, R. O.; Rocha, G. B.; Simas, A. M. Sparkle model for the AM1 calculation of dysprosium (III) complexes. *Inorganic Chemistry Communications* **2005**, 8 (9), 831-835.
21. Freire, R. O.; Rocha, G. B.; Simas, A. M. Sparkle model for the calculation of lanthanide complexes: AM1 parameters for Eu(III), Gd(III), and Tb(III). *Inorganic Chemistry* **2005**, 44 (9), 3299-3310.
22. Freire, R. O.; da Costa, N. B.; Rocha, G. B.; Simas, A. M. Modeling lanthanide coordination compounds: Sparkle/AM1 parameters for praseodymium (III). *Journal of Organometallic Chemistry* **2005**, 690 (18), 4099-4102.
23. Freire, R. O.; Rocha, G. B.; Simas, A. M. Modeling lanthanide complexes: Sparkle/AM1 parameters for ytterbium (III). *Journal of Computational Chemistry* **2005**, 26 (14), 1524-1528.
24. Freire, R. O.; Rocha, G. B.; Simas, A. M. Modeling lanthanide complexes: Sparkle/AM1 parameters for ytterbium (III). *Journal of Computational Chemistry* **2005**, 26 (14), 1524-1528.
25. Freire, R. O.; do Monte, E. V.; Rocha, G. B.; Simas, A. M. AM1 sparkle modeling of Er(III) and Ce(III) coordination compounds. *Journal of Organometallic Chemistry* **2006**, 691 (11), 2584-2588.
26. Freire, R. O.; Rocha, G. B.; Simas, A. M. Modeling rare earth complexes: Sparkle/AM1 parameters for thulium (III). *Chemical Physics Letters* **2005**, 411 (1-3), 61-65.
27. Freire, R. O.; da Costa, N. B.; Rocha, G. B.; Simas, A. M. Sparkle/AM1 Structure Modeling of Lanthanum (III) and Lutetium (III) Complexes. *The Journal of Physical Chemistry A* **2006**, 110 (17), 5897-5900.
28. Freire, R. O.; da Costa, N. B.; Rocha, G. B.; Simas, A. M. Sparkle/AM1 Parameters for the Modeling of Samarium(III) and Promethium(III) Complexes. *Journal of Chemical Theory and Computation* **2006**, 2 (1), 64-74.

29. Freire, R. O. Modelo Sparkle: Novos Avanços para o Estudo Químico-Quântico de Complexos de Lântanídeos, tese de doutorado em química, Universidade Federal de Pernambuco. 2007.
30. Rocha, G. B.; Freire, R. O.; Simas, A. M.; Stewart, J. J. P. RM1: A reparameterization of AM1 for H, C, N, O, P, S, F, Cl, Br, and I. *Journal of Computational Chemistry* **2006**, 27 (10), 1101-1111.
31. SPARTAN, Wavefunction, Inc. 18401 Von Karman, Suite 370 Irvine, CA 92612, [HTTP://wavefun.com](http://wavefun.com), 2008
32. Hyperchem 8.0, Hypercube, Inc. All Rights Reserved, (c) 1985-2007, [HTTP://www.hyper.com](http://www.hyper.com), 2009
33. Dewar, M. J. S.; Thiel, W. Ground-States of Molecules .38. Mndo Method - Approximations and Parameters. *Journal of the American Chemical Society* **1977**, 99 (15), 4899-4907.
34. Dolg, M.; Stoll, H.; Savin, A.; Preuss, H. Energy-Adjusted Pseudopotentials for the Rare-Earth Elements. *Theoretica Chimica Acta* **1989**, 75 (3), 173-194.
35. GAUSSIAN 98, Gaussian, Inc., Pittsburgh, PA, 1998
36. Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods .1. Method. *Journal of Computational Chemistry* **1989**, 10 (2), 209-220.
37. Stewart, J. J. P. Optimization of Parameters for Semiempirical Methods .2. Applications. *Journal of Computational Chemistry* **1989**, 10 (2), 221-264.
38. da, C.; Freire, R. O.; Simas, A. M.; Rocha, G. B. Structure Modeling of Trivalent Lanthanum and Lutetium Complexes: Sparkle/PM3. *The Journal of Physical Chemistry A* **2007**, 111 (23), 5015-5018.
39. Freire, R. O.; Rocha, G. B.; Simas, A. M. Sparkle/PM3 parameters for praseodymium(III) and ytterbium(III). *Chemical Physics Letters* **2007**, 441 (4-6), 354-357.
40. Freire, R. O.; da, C.; Rocha, G. B.; Simas, A. M. Sparkle/PM3 Parameters for the Modeling of Neodymium(III), Promethium(III), and Samarium(III) Complexes. *Journal of Chemical Theory and Computation* **2007**, 3 (4), 1588-1596.
41. Freire, R. O.; Rocha, G. B.; Simas, A. M. Modeling rare earth complexes: Sparkle/PM3 parameters for thulium(III). *Chemical Physics Letters* **2006**, 425 (1-3), 138-141.
42. Simas, A. M.; Freire, R. O.; Rocha, G. B. Lanthanide coordination compounds modeling: Sparkle/PM3 parameters for dysprosium (III), holmium (III) and erbium (III). *Journal of Organometallic Chemistry* **2008**, 693 (10), 1952-1956.
43. Stewart, J. J. P. Optimization of parameters for semiempirical methods V: Modification of NDDO approximations and application to 70 elements. *Journal of Molecular Modeling* **2007**, 13 (12), 1173-1213.
44. Freire, R. O.; Rocha, G. B.; Albuquerque, R. Q.; Simas, A. M. Efficacy of the semiempirical sparkle model as compared to ECP ab-initio calculations for the prediction of ligand field parameters of europium(III) complexes. *Journal of Luminescence* **2005**, 111 (1-2), 81-87.

45. Cao, X. Y.; Dolg, M. Segmented contraction scheme for small-core lanthanide pseudopotential basis sets. *Journal of Molecular Structure-Theochem* **2002**, *581*, 139-147.
46. Cundari, T. R.; Stevens, W. J. Effective Core Potential Methods for the Lanthanides. *Journal of Chemical Physics* **1993**, *98* (7), 5555-5565.
47. Dolg *Modern Methods and Algorithms of Quantum Chemistry*, Jülich, 2000; Vol. 1; pp. 479.
48. Dolg, M.; Stoll, H.; Preuss, H. Energy-Adjusted Abinitio Pseudopotentials for the Rare-Earth Elements. *Journal of Chemical Physics* **1989**, *90* (3), 1730-1734.
49. Ross, R. B.; Gayen, S.; Ermler, W. C. Ab-Initio Relativistic Effective Potentials with Spin-Orbit Operators .V. Ce Through Lu. *Journal of Chemical Physics* **1994**, *100* (11), 8145-8155.
50. Tsuchiya, T.; Nakajima, T.; Hirao, K.; Seijo, L. A third-order Douglas-Kroll ab initio model potential for the lanthanides. *Chemical Physics Letters* **2002**, *361* (3-4), 334-340.
51. Wang, D. Q.; Zhao, C. Y.; Phillips, D. L. A theoretical study of divalent lanthanide (Sm and Yb) complexes with a triazacyclononane-functionalized tetramethylcyclopentadienyl ligand. *Organometallics* **2004**, *23* (8), 1953-1960.
52. Hay, P. J. Theoretical studies of organometallic complexes of uranium involving nitrogen ligands using density functional approaches. *Faraday Discussions* **2003**, *124*, 69-83.
53. Vazquez, J.; Bo, C.; Poblet, J. M.; de Pablo, J.; Bruno, J. DFT studies of uranyl acetate, carbonate, and malonate, complexes in solution. *Inorganic Chemistry* **2003**, *42* (19), 6136-6141.
54. Joubert, L.; Maldivi, P. A structural and vibrational study of uranium(III) molecules by density functional methods. *Journal of Physical Chemistry A* **2001**, *105* (39), 9068-9076.
55. Bolvin, H.; Wahlgren, U.; Moll, H.; Reich, T.; Geipel, G.; Fanghanel, T.; Grenthe, I. On the structure of Np(VI) and Np(VII) species in alkaline solution studied by EXAFS and quantum chemical methods. *Journal of Physical Chemistry A* **2001**, *105* (51), 11441-11445.
56. Maron, L.; Eisenstein, O. Do f electrons play a role in the lanthanide-ligand bonds? A DFT study of Ln(NR₂)(3); R = H, SiH₃. *Journal of Physical Chemistry A* **2000**, *104* (30), 7140-7143.
57. Hay, P. J.; Martin, R. L. Theoretical studies of the structures and vibrational frequencies of actinide compounds using relativistic effective core potentials with Hartree-Fock and density functional methods: UF₆, NpF₆, and PuF₆. *Journal of Chemical Physics* **1998**, *109* (10), 3875-3881.
58. Freire, R. O.; Rocha, G. B.; Simas, A. M. Lanthanide complex coordination polyhedron geometry prediction accuracies of ab initio effective core potential calculations. *Journal of Molecular Modeling* **2006**, *12* (4), 373-389.

59. Cosentino, U.; Moro, G.; Pitea, D.; Calabi, L.; Maiocchi, A. Ab initio effective core potential calculations on lanthanide complexes: Basis sets and electron correlation effects in the study of $[\text{Gd}-(\text{H}_2\text{O})_9](3+)$. *Theochem-Journal of Molecular Structure* **1997**, 392, 75-85.
60. Zerner, M. C. *ZINDO manual, QTP, University of Florida, Gainesville, FL*. 32611; 1990.
61. Zerner, M. C.; Loew, G. H.; Kirchner, R. F.; Muellerwesterhoff, U. T. Intermediate Neglect of Differential-Overlap Technique for Spectroscopy of Transition-Metal Complexes - Ferrocene. *Journal of the American Chemical Society* **1980**, 102 (2), 589-599.
62. Ridley, J. E.; Zerner, M. C. Triplet-States Via Intermediate Neglect of Differential Overlap - Benzene, Pyridine and Diazines. *Theoretica Chimica Acta* **1976**, 42 (3), 223-236.
63. Malta, O. L. Theoretical Crystal-Field Parameters for the Yoc1 - Eu-3+ System - A Simple Overlap Model. *Chemical Physics Letters* **1982**, 88 (3), 353-356.
64. Malta, O. L. A Simple Overlap Model in Lanthanide Crystal-Field Theory. *Chemical Physics Letters* **1982**, 87 (1), 27-29.
65. R Development Core Team (2005). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, 2009
66. Araujo, A.; Gine, E. *The Central Limit Theorem for Real and Banach Valued Random Variables (Probability & Mathematical Statistics)* ; John Wiley & Sons, New York: 1980.
67. Bartels, R. The Rank Version of Vonneumann Ratio Test for Randomness. *Journal of the American Statistical Association* **1982**, 77 (377), 40-46.
68. Durbin, J. Weak Convergence of Sample Distribution Function When Parameters Are Estimated. *Annals of Statistics* **1973**, 1 (2), 279-290.
69. Mardia, K. V. Multi-Sample Uniform Scores Test on A Circle and Its Parametric Competitor. *Journal of the Royal Statistical Society Series B-Statistical Methodology* **1972**, 34 (1), 102-&.
70. THOM, H. C. S. A NOTE ON THE GAMMA DISTRIBUTION. *Monthly Weather Review* **1958**, 86 (4), 117-122.
71. Fisher, R. A. On the Mathematical Foundations of Theoretical Statistics. *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* **1922**, 222 (594-604), 309-368.
72. Harvey, A. C. *The Econometric Analysis of Time Series*; 2 ed.; Philip Allan, Londres: 1990.
73. Cordeiro, G. M. *Introdução a Teoria de Verossimilhança*. 1992.
74. S.Kay Estimation Theory. In *Fundamentals of Statistical Signal Processing*, New Jersey: Prentice Hall: 1993.

75. Gale Rhodes Obtaining Phases. In *Crystallography Made Crystal Clear : A Guide for Users of Macromolecular Models*, Academic Press: 1993; pp 125-129.
76. Anthony R. West *Basic Solid State Chemistry*; 2 ed.; 1999.
77. Kaufman, L.; Rousseeuw, P. *Finding Groups in Data: An Introduction to Cluster Analysis (Wiley Series in Probability and Statistics)*; Wiley-Interscience: 2005.
78. Lima, C. A. C. RM1:Parametrização para o Átomo de Selênio, monografia de graduação em química, Universidade Federal de Pernambuco. 2008.
79. Parkinson, J. M.; Hutchinson, D. An investigation into the efficiency of variants of the simplex method. In *Numerical Methods for Nonlinear Optimization*, F.A.Lootsma, Ed.; Academic Press: 1972.
80. Kirkpatrick, S.; Gelatt, C. D.; Vecchi, M. P. Optimization by Simulated Annealing. *Science* **1983**, 220 (4598), 671-680.
81. Curado, E. M. F.; Tsallis, C. Generalized Statistical-Mechanics - Connection with Thermodynamics. *Journal of Physics A-Mathematical and General* **1991**, 24 (2), L69-L72.
82. Tsallis, C. Possible Generalization of Boltzmann-Gibbs Statistics. *Journal of Statistical Physics* **1988**, 52 (1-2), 479-487.
83. Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. Equation of State Calculations by Fast Computing Machines. *Journal of Chemical Physics* **1953**, 21 (6), 1087-1092.
84. Fletcher, R. A New Approach to Variable Metric Algorithms. *Computer Journal* **1970**, 13 (3), 317-&.
85. Goldfarb, D. A Family of Variable-Metric Methods Derived by Variational Means. *Mathematics of Computation* /1, 24 (109), 23-26.
86. Shanno, D. F. Conditioning of Quasi-Newton Methods for Function Minimization. *Mathematics of Computation* /7, 24 (111), 647-656.
87. BROYDEN, C. G. The Convergence of a Class of Double-rank Minimization Algorithms 1. General Considerations. *IMA J Appl Math* **1970**, 6 (1), 76-90.

ANEXO

Anexo 1 – Tabela de Referência para Análise de Cluster do Íon Magnésio.

<i>Código de referência -CSD</i>	nº Coord.	Carga	nº de Mg	lig. livres	heter.	nº de atom.
ABOCOK	6	-2	1	2	O.N	35
AECLPA01	5	0	1	1	O.N	88
AHOTEW	3	0	2	1	C	54
ALEBEY	7	0	1	1	C.N	41
ALEBIC	11	0	1	1	C.N	47
ALOHIS	6	1+	1	2	C.O	50
ALOHY	6	0	1	2	Br	48
ALOHUE	6	1+	1	2	C.O	50
AQFVMG	6	0	1	4	O	44
AQMEDA	6	2+	1	6	O	19
BAYJIU	6	2+	1	6	O.N	35
BESXON	6	0	1	2	O	35
BIHLIO	6	2+	1	6	O	19
BIHNEM	6	2+	1	6	O	19
BIKPUG	6	2+	1	6	O	19
CIRVAA	6	2+	1	6	O	19
CIRVAA01	6	2+	1	6	O	19
CULXEM	6	0	1	2	N	105
DAMZAT	6	2+	1	6	O	19
DEMCII	4	0	1	0	N	101
DEZNIG	6	0	1	2	O.N	29
DIRBUB	6	0	1	4	O	25
DOLGAM	4	0	2	1	C.N	52
DOLGEQ	4	0	2	1	C.N	78
EBITIT	8	0	2	1	Br	80
ECERUA	4	0	1	2	Cl	105
ECESIO	4	0	2	2	O	140
ECESOU	4	0	2	2	O.N	136
EGIJUZ	3	0	2	1	C.O	118
EHUFOC	6	0	1	2	O.N.S	55
EMIFAH	4	0	1	0	N	109
EROQAD	6	2+	1	6	O	61
ESONUV	4	0	1	2	C.N	65
ESOPFH	4	0	1	2	C.N	53
ESOPIL	4	0	1	2	C.O.N	100
ESOPUX	4	0	2	0	O.N	176
EXUBOO	6	2+	1	6	O	19
FAMWIA	10	0	1	0	C	65
FATQEX	6	2+	1	4	O.N	35
FEDQUA	6	2+	1	6	O	19
FEDQUA01	6	2+	1	6	O	19
FEDZIY	6	2+	1	2	O.N	51
FELGUY01	6	0	1	6	O	27
FETHAO	6	2+	1	6	O	19
FIYXUG	7	0	1	2	Cl	45
FOMPIG	4	0	1	2	C.O	60
FUYCAD	6	0	1	6	O.N	19

GAJPAI	9	0	1	1	N	72
GANQIV	6	0	1	6	S.N	69
GATLUI	6	2+	1	6	O	19
GELKEO	6	1+	1	6	O	33
GEPBIM	6	0	1	2	O.N	89
GEQBIO	6	4-	1	2	O	27
GIKCIM	5	0	2	0	O.N	70
GIKCOS	5	0	1	1	O.N	62
GOLPIG	6	2+	1	6	O	19
GUHHOG	6	2+	1	6	O	19
GUHQAB	6	0	1	0	O.N	101
GUPJAC	7	0	1	2	Br. O	40
GUQTAN	4	0	1	2	O.N	100
GUZZAC	3	0	2	1	O.N	116
GUZZEG	4	0	1	4	O	99
GUZZIK	4	0	1	2	O.N	97
HABGIB	6	0	1	2	O	31
HALZOJ	6	0	1	2	O.N	111
HAMFAC01	6	0	1	2	N	99
HEDMEJ	5	0	1	1	O.N	50
HETZUB	6	2+	1	6	O	19
HUXRUN	6	2+	1	6	O	61
HUXSAU	8	0	1	3	C.O	47
IHEKOV	5	0	1	1	O.N	60
IHEKOV02	5	0	1	1	O.N	60
ILUXAO	6	2+	1	6	O	19
ISUBUT	6	0	2	3	O.N	60
IWONOX	6	0	1	6	O	47
IWOYEEY	6	2+	1	6	O	19
IZEYAN	6	2+	1	0	N	50
JAMQAQ	6	2+	1	6	O	19
JAMQAQ01	6	-2	1	0	O.N	36
JAWQIH01	7	2+	1	2	O	42
JAWQIH02	7	2+	1	2	O	42
JAWQON	6	1+	1	6	O	20
JAYPEE	6	0	1	6	O	31
JEYCEW	6	2+	1	6	O	19
JIVSAI	4	0	1	0	N	77
KABRAG01	6	0	1	6	O	43
KAFGUU	3	0	1	1	C.N	84
KAFHOP	3	0	1	1	N	94
KALJIR	4	0	1	0	O.N	63
KAWRUW	4	0	2	1	C.N	70
KEHJIR	6	0	1	2	O	79
KIDLUF	6	2+	1	6	O	19
KIMNID	6	2+	1	6	O	19
KINMAV	6	0	1	6	O	43
KIWSEO	6	0	1	2	O.S	31
KOLBES	6	0	2	2	Br.O	104
KUVJIU	6	0	1	6	O.N	71
LAKLIS	8	0	2	1	O.Cl	84
LEMJOC	3	0	2	1	N	118
LEMJOC01	3	0	2	1	N	118
LERVAF	4	0	2	2	Cl	92

LIBVUN	5	0	1	2	C.O	84
LIGPAS	6	0	1	6	N	51
LIWJUW	6	0	1	2	C.O.N	35
LUSJUE	4	0	2	1	C.N	102
LUWLOE	7	0	2	0	C.N	68
LUWLUK	7	0	2	0	C.N	70
LUWMAR	7	0	2	0	C.N	84
LUWMEV	7	0	2	0	C.N	72
LUWMOF	8	0	1	1	C.O.N	57
MAVDOD	6	2+	1	6	O	19
MEYGEC	6	0	1	2	O	35
MGEDTA01	6	2+	1	6	O	19
MGEDTA02	7	-2	1	1	O.N	36
MGHMAL10	6	0	1	4	O	27
MGMLTE	6	0	1	4	O	26
MGNTSP	6	2+	1	6	O	19
MGSACA10	6	1+	1	6	O	32
MINKOJ	6	2+	1	6	O	19
MINNAY	6	0	1	2	N	79
MIRMEF	4	0	1	4	N	131
MISXER	6	2+	1	6	O	19
MOYTAV	6	2+	1	6	O	19
MUHXES	5	0	1	1	N	96
NACMIO	4	0	1	0	N	69
NEBZOJ	4	0	2	2	O.N	144
NECBEC	4	0	1	2	N	83
NERRUX	4	0	2	0	N	110
NILSEG	6	0	1	6	O	41
NIVNUB	4	0	1	4	Cl	29
NOGLAW	6	2+	1	6	O	19
NUMYID	6	2+	1	6	O	49
NUZXEL	10	0	1	0	C	55
OEPORM	6	0	1	2	N	107
OFOZUE	5	0	1	4	C.O.N	60
OFUCAT	4	0	1	4	O.N	105
OKURUH	6	0	1	2	Br	44
OKUSAO	6	0	1	2	Br	44
OKUSES	6	0	1	2	Br	42
OKUSUI	6	0	1	2	Br	33
OKUTAP	6	2+	2	0	Br	64
OKUTAP1	4	-2	2	2	Br	14
OMARID	6	2+	1	6	O	19
PAPBUD	4	0	1	4	C.O	67
PAPCAK	4	0	1	4	C.O	103
PAPCAK01	4	0	1	4	C.O	103
PEPTMG	6	0	1	6	O.N	83
PESWOA	6	0	1	2	O	97
PEYKAG	5	0	1	1	O.N	60
PIJMAW01	5	0	1	2	C.N	44
PINFEX	6	2+	1	6	O	37
PIPCUM	6	0	1	0	O.N	81
POHGIC	6	2+	1	6	O	19
PUFGEC	4	0	2	2	Cl.O.N	132
QALKOD	4	0	1	0	P	111

QERJIH	6	2+	1	6	O	19
QEYRAN	10	0	1	0	C	45
QIBCEJ	5	0	1	5	O.N	70
QIBCUZ	5	0	1	5	O	64
RADXEZ	4	0	1	4	C.O	65
RAKYOR	6	2+	1	6	O	19
RAMXOS	6	2+	1	6	O	19
RANBEO	3	0	2	1	C.N	86
RANBIS	3	0	2	1	C.N	70
RANRII	4	0	1	1	C.N	88
RATRIN	6	2+	1	6	O	19
REFTEC	6	2+	1	6	O	19
REHHES	6	2+	1	6	O	73
REHHES1	4	-2	1	4	Cl	5
RETNAG	4	0	1	0	N	157
REVNIQ	8	0	2	1	Br	77
REVNOW	8	0	2	1	Br	46
REYYEA	4	0	1	0	O.N	87
RIVCEE	6	2+	1	6	O	45
RIZJIT	5	0	1	1	O.N	102
SAVQAH	6	2+	1	6	O	49
SEBMIV	6	0	1	4	O.N	30
SENJOL	6	0	1	2	S	55
SIMZUJ	6	2+	1	6	O	19
SOWWOQ	6	0	1	2	O	43
TAGKAO	5	0	2	2	Br	80
TAGKES	5	0	2	2	Br	88
TAPJEZ	6	2+	1	6	O.N	47
TAPLAX	6	2+	1	0	N	41
TAPLIF	6	2+	1	0	N	41
TAQMON	6	2+	1	0	N	37
TAYROA	6	0	1	6	O	43
TAYSIV	5	-1	1	1	Cl	60
TEKFIY	6	0	1	3	O.N	26
THFMGB01	6	2+	1	6	O	59
TONSUK	6	0	1	3	O.N	32
TOWKUL	6	0	1	2	Br	53
TOXDMG	6	2+	1	6	O	19
TULLER	4	2+	1	4	N	69
TUXYOA	6	2+	1	6	O	19
UDARAS	6	1+	1	6	I	67
UFIZUE	4	0	2	0	O.N	116
UGEQIG	10	0	1	0	C	93
VACSIB	4	0	1	1	C.O	49
VAQBOF	4	0	1	2	O.N	105
VAQLII	4	0	1	2	O.N	89
VAXHUX01	6	0	1	6	O	43
VAXJEJ01	6	0	1	3	O.N	25
VOPCEI	6	2+	1	6	O	19
WABHAJ	6	-1	2	0	O	65
WADLOD	4	0	1	4	Cl	73
WAMXIS	6	0	1	3	S	23
WAWJOU	5	0	1	1	O.N	63
WEDCAJ	6	0	1	2	Cl	35

WEGQEE	4	0	1	0	N	69
WEZQUO	6	0	1	6	O.N	57
WIKXAP	6	2+	1	6	O	19
WIMFED	6	0	1	3	Br	39
XAGREC	6	0	1	2	Br	42
XAZKOZ	6	0	1	2	O.N	31
XECJIY	6	2+	1	6	O	19
XEXCUZ	6	0	1	6	O.N	23
XEZCIP	6	2+	1	6	O	19
XITDUZ	6	2+	1	6	O	19
XIWKET	6	1+	1	2	O.N	36
XUXLUX	6	2+	1	6	O	19
YASKEI	6	2+	1	6	O	97
YIQGOU	6	0	1	3	O.N	30
YIQHEL	6	0	1	2	O.N	54
YUHJIU	6	0	1	6	O	51
ZARMEK	6	2+	1	6	O	19