UNIVERSIDADE DE SÃO PAULO
INSTITUTO DE QUÍMICA

ESTUDO DO MÉTODO DE EQUALIZAÇÃO DA ELETRONEGATIVIDADE NO
CÁLCULO DE ENERGIAS LIVRES DE SOLVATAÇÃO – GBEEM-ELR

Karina Shimizu
Tese de Doutorado

Orientador: Prof. Dr. João Pedro Simon Farah

SÃO PAULO
15 de Dezembro de 2005

Ficha Catalográfica
Elaborada pela Divisão de Biblioteca e
Documenação do Conjunto das Químicas da USP.

S556e Shimizu, Karina
Estudo de método de equalização da eletronegatividade
no cálculo de energias livres de solvatação – GBEEM-ELR /
Karina Shimizu -- São Paulo, 2005.
124p.

Tese (doutorado) - Instituto de Química da Universidade
de São Paulo. Departamento de Química Fundamental.
Orientador: Farah, João Pedro Simon

1. Transferência de elétrons : Química 2. Densidade
funcional : Química Teórica I. T. II. Farah, João Pedro
Simon, orientador.

541.28 CDD

***“Estudo do método de equalização da
eletronegatividade no cálculo de energias
livres de solvatação – GBEEM-ELR”***

KARINA SHIMIZU

Tese de Doutorado submetida ao Instituto de Química da Universidade de São Paulo como parte dos requisitos necessários à obtenção do grau de Doutor em Química – Área: Físico-Química.

Aprovado(a) por:

Prof. Dr. JOÃO PEDRO SIMON FARAH
(Orientador e Presidente)

Prof. Dr. FRANK HERBERT QUINA
IQ – USP

Prof. Dr. MAURO CARLOS COSTA RIBEIRO
IQ – USP

Prof. Dr. SÉRGIO EMANUEL GALEMBECK
FFCL – RP – USP

Prof. Dr. LUÍS GUSTAVO DIAS
Inmetro

SÃO PAULO
15 DE DEZEMBRO 2005

Agradecimentos

Ao Prof. Dr. João Pedro Simon Farah, pela orientação e principalmente pela amizade durante esses anos.

Ao Adilson, por estar sempre ao meu lado.

Ao Luís Gustavo Dias, pela “co-orientação” e sobretudo pela amizade.

Ao Prof. Dr. Frank Herbert Quina, pela colaboração com livros e pela amizade.

Aos colegas de laboratório.

À FAPESP, Fundação de Amparo à Pesquisa do Estado de São Paulo, pela bolsa concedida (projeto no. 01/05852-8).

Índice

Índice de Figuras.....	iv
Índice de Tabelas	vii
Símbolos e Abreviaturas.....	ix
Resumo	xi
Abstract.....	xii
1 – Introdução	1
1.1 – Solvatação	1
1.2 – Distribuição de Cargas do Solute	8
2 – Objetivos.....	9
3 – Formalismo Teórico.....	10
3.1 – Método de Equalização da Eletronegatividade (EEM, Electronegativity Equalization Method)	10
3.2 – Modelo de Cargas 1 (CM1, Charge Model 1)	17
3.3 – Cálculo da redistribuição de cargas do soluto em um modelo de solvente implícito	18
3.4 – Cálculo da energia livre de solvatação em modelo de solvente contínuo	24
4 – Resultados e Discussão.....	28
4.1 – Cálculo da redistribuição de cargas do soluto em um modelo de solvente implícito	28
4.2 – Calibração do Método de Equalização da Eletronegatividade Baseada no Charge Model 1	39
4.3 – O método de equalização e o cálculo de momentos de dipolo	52
4.4 – Cálculo da energia livre de solvatação	70
5 – Conclusões.....	93
6 – Apêndice: Código fonte do programa GBEEM.....	95
7 – Bibliografia	105

Índice de Figuras

Figura 1: Representação esquemática de um soluto rodeado por moléculas de água.....	3
Figura 2: Representação esquemática do soluto em um ambiente onde o solvente é considerado implicitamente.	3
Figura 3: Definições das superfícies moleculares (linhas cheias): (a) superfície de van der Waals (vdWS); (b) superfície excluída do solvente (SES) e c) superfície acessível ao solvente (SAS).	4
Figura 4: Partição em <i>tesseraes</i> da superfície de Van der Waals da molécula de água usando: a) áreas de $0.2 \text{ \AA}^2$ e b) áreas de $0.4 \text{ \AA}^2$	23
Figura 5: Gráfico das cargas calculadas por EEM-PD versus cargas calculadas por PM3-CM1. A linha tracejada representa a identidade e a linha contínua representa a correlação linear entre os dados ($y = ax$, onde $a = 1,100 \pm 0,070$, $R = 0,940$, $SD = 0,155$, $F = 234$ e $N = 32$).	32
Figura 6: Gráfico das cargas calculadas por GEEM-PD vs PM3-CM1 (constante dielétrica = 78,5 e geometrias otimizadas na fase condensada). A linha tracejada representa a identidade e a linha contínua representa a correlação linear entre os dados ($y = ax$, onde $a = 1,062 \pm 0,071$, $R = 0,941$, $SD = 0,165$, $F = 218$ e $N = 29$).	35
Figura 7: Gráfico das diferenças relativas dos dipolos calculados pelo GBEEM-PD vs as diferenças relativas dos dipolos calculados pelo PM3-CM1 ($y = ax$, onde $a = 1,068 \pm 0,072$, $R = 0,663$, $SD = 12$, $F = 8$ e $N = 14$).	38
Figura 8: Gráficos dos desvios calculados para eletronegatividades (●) e durezas (■) dos átomos de carbono e hidrogênio.....	43
Figura 9: Cargas EEM vs cargas CM1. As linhas tracejada e sólida são ajustes lineares às cargas calculadas pelo EEM-PD ($y = ax + b$, $a = 1,30$, $b = -6,1 \times 10^{-7}$, $R = 0,931$, $SD = 0,079$, $F = 31627$, $N = 4903$) e EEM-CM1 ($y = ax + b$, $a = 0,94$, $b = -5,0 \times 10^{-7}$, $R = 0,977$, $SD = 0,032$, $F = 102977$, $N = 4903$). A linha pontilhada é a linha identidade.	46
Figura 10: Resíduo vs cargas CM1. As funções orgânicas são apresentadas com diferentes cores/símbolos. Os resíduos foram calculados utilizando-se a regressão robusta LTS (Least Trimmed Squares). Nesta regressão o ajuste é	

feito minimizando-se a soma dos quadrados do conjunto formado pelos menores resíduos (cerca de 80% do total).	49
Figura 11: Cargas EEM-CM1 para o açúcar e o tetrapeptídeo. As linhas sólidas são ajustes lineares para cargas EEM-CM1 do açúcar ($y=ax+b$, $a=1,03$, $b=3,3 \times 10^{-17}$, $R=0,992$, $SD=0,038$, $N=66$) e o tetrapeptídeo ($y=ax+b$, $a=0,96$, $b=-2,8 \times 10^{-6}$, $R=0,991$, $SD=0,042$, $N=41$) em função das cargas PM3-CM1. A linha pontilhada é a linha identidade.	50
Figura 12: Cargas EEM versus cargas CM1, sendo que as cargas do método de equalização foram calculadas com os parâmetros EEM-CM1.	51
Figura 13: Momento de dipolo EEM-CM1 vs momento de dipolo PM3-CM1. As funções orgânicas estão apresentadas com diferentes cores e símbolos. A linha contínua é a identidade.	54
Figura 14: Efeito da aproximação de cargas predefinidas nas séries de aldeídos, ácidos carboxílicos e nitro compostos. A linha contínua é a linha identidade.	55
Figura 15: Histograma dos dipolos calculados no vácuo para cada aminoácido pelos diferentes métodos.	57
Figura 16: Histograma dos dipolos calculados na fase condensada para cada aminoácido pelos diferentes métodos.	57
Figura 17: Representações em estéreo das estruturas do polipeptídeo (GAS) ₅ , sendo a) na conformação α -hélice e b) uma conformação torcida.	63
Figura 18: Representações em estéreo das estruturas do polipeptídeo (Ala) ₁₅ , sendo a) na conformação α -hélice e b) uma conformação torcida.	64
Figura 19: Momentos de dipolo calculados pelo EEM-CM1 versus momentos de dipolo calculados em nível B3LYP/6-311G(d,p) no vácuo. A linha segmentada é a linha identidade.	71
Figura 20: Momentos de dipolo calculados pelo GBEEM-CM1 versus momentos de dipolo calculados em nível IEFPCM-B3LYP/6-311G(d,p) na presença de solvente, usando constante dielétrica de 78,5 e raio atômico de Bondi. A linha segmentada é a linha identidade.	72
Figura 21: Energias livres de solvatação calculadas pelos modelos GBEEM/NP1 e GBEEM/NP2 versus energias de solvatação experimentais. A linha pontilhada é a identidade.	78

Figura 22: Energias livres de solvatação calculadas pelos modelos GBEEM/NP3 e GBEEM/NP4 versus energias de solvatação experimentais. A linha pontilhada é a identidade.	79
Figura 23: Energias livres de solvatação calculadas pelos modelos GBEEM/NP5 e GBEEM/NP6 versus energias de solvatação experimentais. A linha pontilhada é a identidade.	80
Figura 24: Energias livres de solvatação calculadas pelo modelo SM5.42R versus energias de solvatação experimentais. A linha pontilhada é a identidade.	84
Figura 25: Energias livres de solvatação calculadas pelo modelo IEFPCM versus energias de solvatação experimentais. A linha pontilhada é a identidade.	85
Figura 26: Energias livres de solvatação calculadas pelo modelo GBEEM-ELR versus energias de solvatação experimentais. A linha pontilhada é a identidade.	86

Índice de Tabelas

Tabela 1: Momentos de dipolo no vácuo (μ_g) calculados pelos métodos QEq-PD, EEM-PD e PM3-CM1 (valores em Debye).....	30
Tabela 2: Valências atômicas para moléculas isoladas. Os índices I-J representam o primeiro e segundo átomos na coluna q_I e q_J , respectivamente. Para as moléculas que apresentam mais de um átomo representado por I ou J (Metano, Etano, Metanol, Formaldeído, Metilamina, Água e Nitrometano), os valores das valências são médias aritméticas.....	31
Tabela 3: Cargas atômicas na água (valores em unidade atômica, $\varepsilon \cong 78,5$). Os termos I-J representam o primeiro e segundo átomos na coluna q_I e q_J . As cargas atômicas são médias aritméticas para as moléculas que apresentam mais de um átomo do tipo I ou J.....	34
Tabela 4: Momentos de dipolo no vácuo e na água calculados pelos métodos GBEEEM-PD e PM3-CM1 (valores em Debye).	37
Tabela 5: Eletronegatividades atômicas usadas em diferentes calibrações do método de equalização.....	41
Tabela 6: Durezas atômicas usadas em diferentes calibrações do EEM.	41
Tabela 7: Eletronegatividades (χ) e durezas (η) atômicas com os desvios relativos (σ_{rel}) estimados para cada parâmetro.	45
Tabela 8: Vetor momento de dipolo para o polipeptídeo (Ala) ₇ na conformação α -hélice no vácuo.	58
Tabela 9: Vetor momento de dipolo para o polipeptídeo (GAS) ₅ na conformação α -hélice no vácuo.	59
Tabela 10: Vetor momento de dipolo para o polipeptídeo (GAS) ₅ no vácuo.....	60
Tabela 11: Vetor momento de dipolo para o polipeptídeo (Ala) ₁₅ na conformação α -hélice no vácuo.	61
Tabela 12: Vetor momento de dipolo para o polipeptídeo (Ala) ₁₅ no vácuo.....	62
Tabela 13: Vetor momento de dipolo para o polipeptídeo (Ala) ₇ na conformação α -hélice em dielétrico contínuo ($\varepsilon=78,5$).	66
Tabela 14: Vetor momento de dipolo para o polipeptídeo (GAS) ₅ na conformação α -hélice em dielétrico contínuo ($\varepsilon=78,5$).	67

Tabela 15: Vetor momento de dipolo para o polipeptídeo (GAS) ₅ em dielétrico contínuo ($\epsilon=78,5$).....	68
Tabela 16: Vetor momento de dipolo para o polipeptídeo (Ala) ₁₅ na conformação α -hélice em dielétrico contínuo ($\epsilon=78,5$).	68
Tabela 17: Vetor momento de dipolo para o polipeptídeo (Ala) ₁₅ em dielétrico contínuo ($\epsilon=78,5$).....	69
Tabela 18: Parâmetros otimizados para o cálculo da energia livre de solvatação em GBEEM/NP1. As unidades dos parametros C_{iw} , a e b são kcal mol ⁻¹ Å ⁶ , kcal mol ⁻¹ e kcal mol ⁻¹ Å ⁻² , respectivamente. O parametro c é adimensional.	75
Tabela 19: Coeficientes lineares obtidos pela regressão linear múltipla do modelo GBEEM-ELR.....	82
Tabela 20: Coeficientes lineares obtidos pela análise de regressão passo a passo do modelo GBEEM-ELR.....	82
Tabela 21: Os desvios médios absolutos, MAD, das energias livres de hidratação de 66 moléculas orgânicas neutras calculadas pelos modelos SM5.42R, IEFPCM-HF(UAHF)/6-31G e GBEEM-ELR. Assim como os tempos de cálculo dos três modelos.	83
Tabela 22: Energias livres de solvatação experimentais (<i>Chambers et al., 1996</i>) e calculadas pelos métodos SM5.42R, IEFPCM e GBEEM-ELR (em kcal.mol ⁻¹).	88
Tabela 23: Energias livres de solvatação experimentais (<i>Chambers et al., 1996</i>) e calculadas pelos métodos GBEEM/NP, PCM e AMSOL (em kcal mol ⁻¹).	90

Símbolos e Abreviaturas

EEM	Electronegativity Equalization Method (Método de Equalização da Eletronegatividade)
GB	Generalized Born (Born Generalizado)
GBEEM	Generalized Born Electronegativity Equalization Method
QEq-PD	Calibração do EEM empregando cargas derivadas do potencial eletrostático no nível RHF/6-31G(d) como alvo
EEM-CM1	Calibração do EEM empregando cargas CM1 como alvo
PM3-CM1	Modelo de carga CM1 em nível semi-empírico PM3
EEM-PD	Calibração do EEM empregando os parâmetros da calibração QEq-PD com uma correção de 6% na dureza
PM3-SM5.4	Modelo de carga CM1 em nível semi-empírico PM3 no solvente
GBEEM-PD	GBEEM empregando os parâmetros da calibração EEM-PD
EEM-M	Calibração do EEM empregando cargas Mulliken no nível RHF/STO-3G como alvo
QEq-M	Calibração do EEM empregando cargas Mulliken no nível RHF/6-31G(d) como alvo
GBEEM-CM1	GBEEM empregando cargas EEM-CM1
GBEEM-ELR	Modelo de solvatação que combina o GBEEM com o modelo de Floris-Tomasi via resposta linear estendida
SCRf	Self-Consistent Reaction Field (Campo de Reação Auto-consistente)
PCM	Polarized Continuum Model (Modelo de Contínuo Polarizado)
UAHF	United Atom Model for Hartree-Fock (Modelo de Átomo Unificado para Hartree-Fock)
PCM(UAHF)	Modelo de Contínuo Polarizado empregando o raio UAHF
SM5.4	Solvation Model 5.4
CM1	Charge Model 1
ΔG_{solv}	Energia Livre de Gibbs de Solvatação
ΔG_{np}	Componente Não Polar de ΔG_{solv}
ΔG_{el}	Componente eletrostático de ΔG_{solv}
ΔE_{EN}	Diferença de energia nuclear e eletrônica entre o vácuo e a fase condensada

U_{disp}	Energia média de interação de dispersão entre o soluto e o solvente
C_{iw}	Parâmetro que descreve a interação de dispersão entre o átomo “i” da molécula e o solvente
SES	Solvent-Excluding Surface (Superfície Excluída do Solvente)
SAS	Solvent Accessible Surface (Superfície Acessível ao Solvente)
ϵ	Constante dielétrica
DFT	Density Functional Theory (Teoria do Funcional da Densidade)
N	Número total de elétrons
v	Potencial externo
μ	Momento de dipolo
η	Dureza química
η^*	Dureza química parametrizável
η^0	Dureza química do átomo isolado
$f(\mathbf{r})$	Função de Fukui
E	Energia eletrônica molecular
I	Potencial de ionização
A	Afinidade eletrônica
χ	Eletronegatividade
χ^*	Eletronegatividade parametrizável
χ^0	Eletronegatividade do átomo isolado
ρ	Densidade eletrônica
q	Carga atômica
a	Raio atômico
R_{ij}	Distância entre os átomos i e j
ϕ	Potencial eletrostático
f_{GB}	Função de Born generalizado
α_i^B	Raio efetivo de Born
GA	Genetic Algorithm (Algoritmo Genético)
CCA	Constrained Charge Approximation (Aproximação de Cargas Predefinidas)
MAD	Mean Absolute Deviation (Desvio Absoluto Médio)

Resumo

O método de equalização da eletronegatividade (Electronegativity Equalization Method, EEM), fundamentado em teoria do funcional da densidade eletrônica, foi combinado à aproximação de Born generalizada para moléculas (Generalized Born, GB), e denominado GBEEM (*Dias et al., 2002*). Os momentos de dipolo permanente no vácuo e em meio condensado ($\epsilon \sim 80$), e distribuições de cargas atômicas, mostraram boa concordância com modelo SM5.4 baseado em cargas CM1 em nível PM3 (12 moléculas, correspondendo a 29 cargas atômicas). Este resultado é interessante devido à simplicidade inerente do GBEEM e seu baixo custo computacional.

Uma nova parametrização das durezas e eletronegatividades foi feita com o objetivo de melhorar a distribuição de cargas atômicas em moléculas isoladas em relação ao modelo CM1. Um conjunto de 250 estruturas/cargas PM3/CM1 de moléculas neutras pertencentes a 13 funções orgânicas foi utilizado como alvo na parametrização, utilizando uma metodologia Algoritmo Genético/Simplex de pesquisa de mínimos (*Menegon et al., 2002*). Boa concordância entre os modelos foi obtida. A validação da parametrização e do EEM foi efetuada usando moléculas bifuncionais (tetrapeptídeo e trisacarídeo) mostrando também boa concordância e robustez.

Entretanto, a análise do momento de dipolo permanente das 250 moléculas mostrou uma séria limitação do EEM, e portanto do GBEEM, apesar da boa concordância entre as cargas EEM e CM1. O EEM superestimou os momentos de dipolo. Tal fato pode decorrer de vários fatores, dentre os quais, o truncamento da expansão nas cargas atômicas e ausência de tratamento explícito de interação de troca (exchange). Foi sugerida uma aproximação que restringe a transferência de carga entre grupos na molécula que contornou a limitação do método na predição de momentos de dipolo no vácuo e meio condensado (*Shimizu et al., 2004*).

Com base nos recentes resultados, foi desenvolvido um modelo de solvatação baseado no GBEEM e no modelo de Floris-Tomasi. A calibração foi feita com um conjunto de 62 moléculas neutras (13 grupos funcionais) tendo como alvo as energias livres de hidratação experimentais. Os resultados apresentaram um desvio médio absoluto de 0,71 kcal/mol em relação aos valores experimentais.

Abstract

The electronegativity equalization method (EEM), founded on density functional theory (DFT), has been combined to the generalized Born approximation (GB) for molecules, and called GBEEM (*Dias et al., 2002*). The permanent dipole moment in vacuum and condensed phase ($\epsilon \sim 80$), and atomic charges distributions, have shown good agreement with SM5.4 solvation model based on CM1 charges at PM3 level (12 molecules, corresponding to 29 atomic charges). This result is interesting due the simplicity of GBEEM and its low computational cost.

A new parameterization of the hardness and electronegativities was done with the aim to improve the atomic charges distribution on isolated molecules in comparison to CM1 model. The training set with 250 PM3/CM1 structures/charges of neutral molecules in 13 different organic functions was employed as target in the parameterization. A new optimization approach composed of Genetic and Simplex algorithms was used to fit parameters (*Menegon et al., 2002*). Good agreement between the models was found. The validation of parameterization and EEM was done using bifunctional molecules (tri-glucose and tetra-peptide) showing good agreement and robustness.

However, analysis of permanent dipole moments of 250 molecules shown a serious caveat of EEM and GBEEM, beside the good agreement between EEM and CM1 charges. EEM has overestimated the dipole moments. Such result may be due to the truncated expansion in atomic charges and lacking of explicit treatment of exchange interaction. A new approximation was proposed constraining the charge transfer between groups within the molecule. This approximation corrected the caveat of EEM in the prediction of dipole moments in vacuum and condensed phase (*Shimizu et al., 2004*).

Based on these results, a new solvation model was developed founded in GBEEM and Floris-Tomasi model. The parameterization was done with a training set of 62 neutral molecules (13 functional groups) and experimental hydration free energies as target. This new solvation model has produced a mean absolute deviation, MAD, of 0.71 kcal/mol comparing to experimental data.

1 – Introdução

1.1 – Solvatação

A análise dos efeitos do solvente sobre as estruturas químicas é fundamental para a descrição dos processos químicos que ocorrem na fase condensada, sendo que mesmo informações extraídas de modelos teóricos simplificados são muito úteis na busca da compreensão do processo de solvatação (*Grant & Richards, 1995*).

Nos modelos discretos, a molécula do soluto está rodeada por um grande número de moléculas do solvente. As interações entre o soluto-solvente e solvente-solvente são descritas por um campo de força empírico (*Cornell et al., 1995; MacKerell et al., 1998; Jorgensen et al., 1996*) ou *ab initio* (*Tomasi & Persico, 1994; Orozco & Luque, 2000; Miertus et al., 1981; Klamt, 1995*). Propriedades estruturais e termodinâmicas são obtidas por técnicas de simulação computacional, como dinâmica molecular, Monte Carlo (*Allen & Tildesley, 1987*) ou solvatação dinâmica por DFT (*Gao & Liu, 2005*). Para a simulação ser “realista”, um número considerável de moléculas de solvente deve ser empregado, significando em termos práticos que a maior parte do tempo é usada na amostragem de configurações do solvente, mesmo que o soluto seja o foco do estudo. Na Figura 1 é apresentada uma representação esquemática de um soluto interagindo com moléculas de solvente.

Uma abordagem mais geral para os modelos de solvatação discretos é a combinação do tratamento explícito para um número pequeno de moléculas do solvente com o modelo de contínuo para a vizinhança (*Aleman & Galembeck, 1998; Aleman et al., 1998*). Modelos pertencentes a esta abordagem podem produzir bons resultados mesmo nos sistemas onde as moléculas do solvente estão fortemente ligadas ao soluto, além de fornecer possíveis informações do processo de solvatação (*Bandyopadhyay et al., 2002*). O maior problema associado com estes modelos de solvatação discreto/contínuo é o custo e a qualidade da parte discreta do sistema. Outro modelo desenvolvido que procura fornecer bons resultados sem utilizar uma quantidade elevada de moléculas do solvente é o EFP (Effective Fragment Potential). O modelo EFP representa as interações intermoleculares fracas entre as moléculas do soluto, descritas por mecânica quântica, e as moléculas do solvente representadas pelos fragmentos do potencial efetivo (*Bandyopadhyay et al., 2002*).

Um modelo alternativo considera o solvente de forma implícita (Figura 2). Este modelo não considera o solvente composto por moléculas individuais, mas sim como um efeito de campo médio. Uma vantagem dos modelos de solvente implícito é o baixo custo computacional, muito inferior quando comparado com modelos explícitos. Adicionalmente, os modelos implícitos são apropriados para a estimativa rápida de energias livres de solvatação (Roux, 2001).

A energia livre de solvatação em modelos de solvente implícito pode ser convenientemente dividida em dois termos. Um *termo não-polar* onde as interações de curto alcance entre o soluto e o solvente são consideradas (trabalho de cavitação e interação de dispersão soluto-solvente). Um *termo polar-eletrostático* onde as interações eletrostáticas são tratadas. Conseqüentemente, a energia livre de solvatação pode ser escrita como a soma das energias livres não polar (ΔG_{np}) e eletrostática (ΔG_{el}):

$$\Delta G_{solv} = \Delta G_{np} + \Delta G_{el} \quad (1)$$

Em modelos de solvente implícito, a forma e tamanho da cavidade que acomoda o soluto são elementos críticos do método (Tomasi & Persico, 1994). A forma mais simples de se definir a cavidade é através de uma esfera ou de um elipsóide contendo o soluto. Outro modo de se obter a cavidade seria uma sobreposição de esferas localizadas em cada núcleo do soluto, com o uso de raios atômicos de van der Waals (Figura 3a). Se as superfícies moleculares forem muito irregulares, poderão existir pequenas regiões inacessíveis ao solvente. Foi proposta então a criação de uma superfície excluída do solvente (Solvent-Excluding Surface, SES), que é definida pela rolagem de uma esfera de diâmetro adequado na superfície de van der Waals (Figura 3b). Adicionalmente, uma nova superfície foi definida e denominada de superfície acessível ao solvente (Solvent Accessible Surface, SAS), correspondente à superfície traçada pelo raio da esfera rolada na superfície de van der Waals (Figura 3c).

Figura 1: Representação esquemática de um soluto rodeado por moléculas de água.

Figura 2: Representação esquemática do soluto em um ambiente onde o solvente é considerado implicitamente.

Figura 3: Definições das superfícies moleculares (linhas cheias): (a) superfície de van der Waals (vdWS); (b) superfície excluída do solvente (SES) e c) superfície acessível ao solvente (SAS).

Uma boa correlação entre a energia livre de solvatação de alcanos lineares com a SAS sugeriu que a contribuição não-polar poderia ser descrita simplesmente como um processo diretamente proporcional à área da superfície molecular do soluto. Este comportamento não foi verificado para outros sistemas, indicando que a área não fornece uma representação completa do termo não polar, sendo necessária uma descrição específica para a energia de dispersão (*Su & Gallicchio, 2004*).

As interações eletrostáticas entre o soluto e o solvente são comumente modeladas com base na abordagem eletrostática do contínuo. Uma ferramenta útil para este tipo de aproximação é o conceito de campo de reação, introduzido por Onsager (1936). O campo de reação é o campo elétrico exercido pelo meio polarizado na distribuição de cargas do soluto. O solvente é tratado como um dielétrico-contínuo, homogêneo e isotrópico, caracterizado geralmente por apenas um único valor, a constante dielétrica ϵ (*Cramer & Truhlar, 1999*).

Born desenvolveu em 1920 um modelo para calcular a energia livre de íons esféricos no qual o solvente polar é representado como um meio dielétrico contínuo sem estrutura definida. Este modelo quantifica a diferença de energia eletrostática da transferência de uma carga do vácuo para um meio com uma constante dielétrica ϵ . Posteriormente, Kirkwood e Onsager estenderam o modelo para o tratamento de distribuições arbitrárias de carga dentro de uma cavidade esférica. A esfera pode ser considerada uma cavidade apropriada para um número limitado de moléculas. Esferóides e elipsóides mantêm as vantagens de serem matematicamente simples e são apropriados para um número maior de moléculas. Atualmente, os recursos computacionais permitem o tratamento do problema eletrostático com as mais diversas formas moleculares (*Tomasi & Persico, 1994; Orozco & Luque, 2000; Roux, 2001*).

Nos modelos mais simples de contínuo dielétrico, o soluto é descrito por uma distribuição rígida de carga atômica e a parte eletrostática é calculada do trabalho eletrostático de Born. Adicionalmente, um termo não-polar, calculado como uma somatória de tensões superficiais atômicas multiplicadas pela área acessível atômica ao solvente, é empregado para completar o cálculo da energia livre de solvatação. A desvantagem destes modelos é a imprecisão oriunda do cálculo da energia do soluto pelos campos de força empíricos.

Por outro lado, modelos de solvatação onde o soluto é polarizado têm se mostrados práticos e precisos, principalmente quando cavidades com formatos moleculares são usadas para acomodar o soluto. O mais popular desse formalismo é o

modelo de contínuo polarizado (Polarized Continuum Model, PCM) elaborado inicialmente por Tomasi e colaboradores (1981). O PCM trata o soluto via mecânica quântica, calculando a sua energia e função de onda no vácuo e, posteriormente, incluindo a resposta do campo dielétrico-contínuo do solvente no Hamiltoniano do soluto. O Hamiltoniano modificado é empregado em um novo cálculo de energia do soluto e o ciclo se repete até que um resultado consistente seja obtido (*Cossi et al., 1996; Cramer & Truhlar, 1999*). Para os cálculos de PCM existem três diferentes aproximações: o método original, chamado de PCM dielétrico (D-PCM, *Miertus et al., 1981; Cammi & Tomasi, 1995*), um modelo alternativo, onde o meio circundante é modelado como um condutor ao invés de um dielétrico (C-PCM, *Barone & Cossi, 1998*) e uma implementação onde as equações do PCM são recolocadas em um formalismo de equações integrais (IEF-PCM, *Cancès et al., 1997; Mennucci et al., 1997*).

No cálculo do potencial eletrostático, os modelos PCM geralmente utilizam uma cavidade do soluto definida como a união de esferas sobrepostas tendo os raios determinados empiricamente. A escolha mais comum para muitos átomos é o raio de van der Waals multiplicado por um fator de 1,2. Barone e colaboradores (*Barone et al., 1997*) propuseram uma abordagem para raios atômicos considerando que os átomos de H não estão incluídos em esferas individuais, mas sim incorporados aos átomos de outros elementos aos quais eles estão ligados. Estes raios atômicos são denominados de UAHF (United Atom Model for Hartree-Fock) e foram desenvolvidos com o objetivo de calcular energias de solvatação mais precisas.

Klamt e colaboradores (*Klamt & Schüürmann, 1993*), desenvolveram um modelo muito similar ao PCM, chamado de COSMO (conductor-like screening model). Este modelo calcula cargas e energias empregando uma superfície de van der Waals e assume que o meio é modelado como um condutor, o que simplifica o cálculo eletrostático. O modelo COSMO foi elaborado independentemente e precede uma variação ainda mais similar ao PCM, o C-PCM. Klamt propôs um esquema estatístico para incorporar o termo não-polar e chamou de COSMO-RS, o COSMO para solventes reais, que calcula a energia livre de solvatação para solutos neutros (*Klamt, 1995*).

As formulações que derivam do modelo PCM original, como o IEFPCM, são mais utilizadas do que o modelo modificado COSMO. Os modelos de PCM são bastante precisos, porém demandam um custo computacional razoável, já que o cálculo das interações eletrostáticas entre o soluto e o solvente é feito auto-consistentemente e a

partir da resolução da equação de Poisson. Uma desvantagem adicional do PCM(UAHF) é o emprego de raios atômicos dependentes da hibridização de cada átomo da molécula, porque resulta em um conjunto grande de parâmetros, dificultando a compreensão do ponto de vista químico.

Um modelo que se destaca pelo compromisso entre rapidez e precisão é a aproximação generalizada de Born (Generalized Born, GB). Neste modelo, a distribuição rígida de cargas atômicas é novamente empregada no cálculo da energia eletrostática mas, ao invés de resolver formalmente a equação de Poisson, o GB usa o algoritmo de Still e colaboradores (*Still et al., 1990; Cramer & Truhlar, 1999; Feig & Brooks, 2004*). A primeira formulação do GB, sugerida por Still e colaboradores, era baseada na aproximação de campo central coulômbico para o cálculo do raio generalizado de Born. Na aproximação de campo central, as linhas do vetor deslocamento elétrico são radiais, originárias de posições atômicas e ignoram a mudança de dielétrico (interior do soluto, com $\epsilon = 1$ ou 2; exterior do soluto, com $\epsilon \gg 1$, no caso de solvente polar). Novos algoritmos corrigem a aproximação de campo Coulômbico e levam os novos GBs à quase identidade com soluções numéricas da equação de Poisson (*Feig & Brooks, 2004*). O GB é usado em conjunto com campos de força clássicos, combinando seu termo eletrostático com o tratamento da energia do soluto por mecânica molecular. As implementações do GB estão disponíveis nos programas CHARMM, Amber, MacroModel, Tinker e NAB e GROMOS96 (*Bashford & Case, 2000; Fan et al., 2005*). Tensões superficiais atômicas são geralmente usadas no tratamento do termo não-polar.

Os modelos de solvatação SMx de Cramer e colaboradores utilizam a energia eletrostática calculada pelo GB com o tratamento quântico do soluto (*Cramer & Truhlar, 1999; Cramer & Truhlar, 1991; Giesen et al., 1995; Chambers et al., 1996*). SMx também utiliza tensões superficiais atômicas para o termo não-polar (*Chambers et al, 1996; Cramer & Truhlar, 1999; Orozco et al., 2000*). Nos modelos SMx (x variando de 1 a 5.43), diferentes descrições de cargas atômicas e raios, do algoritmo e das formas funcionais da tensão superficial atômica são trabalhadas. Uma característica dos modelos SMx é que os valores dos parâmetros da tensão superficial são otimizados tendo dados experimentais de solvatação como referência. Na geração SM5.x (SM5.0 à SM5.43), as tensões superficiais são funções complicadas da geometria do soluto.

1.2 – Distribuição de Cargas do Soluta

Cargas atômicas são comumente empregadas para caracterizar as propriedades eletrostáticas do soluto. Tanto em modelos de solvente explícito ou implícito, o cálculo da distribuição de carga em uma molécula tem um papel fundamental na estimativa do componente eletrostático da energia livre de solvatação.

Entretanto, carga atômica em uma molécula não pode ser determinada experimentalmente, e conseqüentemente, não há como afirmar qual das definições existentes de “carga atômica em molécula” é a correta. Portanto, o modelo de “carga atômica em molécula” é julgado posteriormente pela qualidade da informação obtida (Storer *et al.*, 1995).

Na análise de população de Mulliken, um dos modelos mais usados nos cálculos de cargas atômicas em moléculas, a carga no orbital molecular está particionada em contribuições associadas com funções de base atômicas centradas no núcleo (Mulliken, 1955). Um segundo modelo similar e bastante utilizado é a análise de população de Löwdin, sendo que neste formalismo a partição das cargas em funções base monoeletrônicas é obtida por um procedimento de ortogonalização simétrica (Lowdin, 1950).

A grande desvantagem destes dois formalismos (Mulliken e Löwdin) é que ambos dependem da função base e do nível de teoria empregados na descrição da correlação eletrônica. Novos modelos de cargas foram propostos com o objetivo de aliviar este problema (Li *et al.*, 1998). Nestes modelos, as cargas são determinadas pela análise de um observável previsto pela função de onda, por exemplo, ajustando cargas ao potencial eletrostático (Chirlian & Francl, 1987; Breneman & Wiberg, 1989) ou ao momento de dipolo permanente da molécula (Hall & Williams, 1975).

Storer e colaboradores (1995) elaboraram um novo grupo de cargas, denominados de modelos de cargas da classe IV. Estes modelos são definidos por procedimentos de parametrização que utilizam cargas obtidas por análise de população, mapeando-as para reproduzir cargas que dependem de observáveis obtidos experimentalmente (ou obtidos por cálculo mecânico quântico de moléculas pequenas). Este procedimento corrige os erros sistemáticos que podem ocorrer na análise de população da função de onda. O modelo de carga 1 (Charge Model 1, CM1) é um exemplo de modelo que pertence à classe IV em nível semi-empírico. As cargas CM1

são calculadas aplicando-se um fator de correção nas cargas Mulliken para que estas reproduzam momentos de dipolo experimentais.

Uma outra forma de se calcular cargas atômicas parciais é pelo método da equalização da eletronegatividade (Electronegativity Equalization Method, EEM). O EEM é baseado na teoria do funcional da densidade (Density Functional Theory, DFT) fazendo aproximações no tratamento das interações intra-atômicas e interatômicas (Mortier *et al.*, 1986; Parr & Yang, 1989; Baekelandt *et al.*, 1993). O EEM usa eletronegatividades e durezas intrínsecas como parâmetros no cálculo formal das cargas atômicas e quantifica o *princípio de Sanderson* (Sanderson, 1951; Baekelandt *et al.*, 1993), estabelecendo que quando átomos se combinam para formar uma molécula, suas eletronegatividades são igualadas (Mortier *et al.*, 1986; Baekelandt *et al.*, 1993; Bultinck *et al.*, 2002). O EEM vem sendo utilizado no cálculo de cargas atômicas em moléculas de interesse biológico (Stern *et al.*, 1999; Cho *et al.*, 2001), análise de transferência de cargas durante a formação de ligações químicas (Yang & Wang, 1997) e simulação de sistemas polarizáveis (York & Yang, 1996; Liu *et al.*, 1998; Banks *et al.*, 1999).

2 – Objetivos

O presente estudo tem por objetivo desenvolver uma metodologia para o cálculo de energias livres de solvatação baseada no método de equalização da eletronegatividade e da aproximação de Born generalizada. O método de equalização da eletronegatividade foi usado no cálculo de cargas atômicas polarizáveis no vácuo e a sua combinação com a aproximação de Born generalizada possibilita o cálculo da redistribuição de cargas atômicas do soluto no solvente.

A aproximação de Born generalizada foi utilizada também no cálculo da energia eletrostática, um dos principais componentes da energia livre de solvatação. A contribuição não polar para a energia total, que consiste nas energias de cavitação e dispersão, foi calculada baseando-se no modelo proposto por Floris-Tomasi.

3 – Formalismo Teórico

3.1 – Método de Equalização da Eletronegatividade (EEM, Electronegativity Equalization Method)

A descrição precisa das interações atômicas e moleculares é fundamental na compreensão das reações químicas. Métodos *ab initio* e DFT (Mennucci *et al.*, 1998) geralmente produzem bons resultados, porém esses métodos são computacionalmente muito custosos, o que limita a aplicabilidade dessas técnicas em muitos sistemas. Por outro lado, modelos empíricos não são suficientemente precisos para várias importantes aplicações. Os campos de força de mecânica molecular convencionais representam a distribuição de cargas como cargas e/ou dipolos pontuais estáticas negligenciando a transferência de carga e a polarizabilidade molecular. Progressos foram feitos introduzindo-se n-polos atômicos polarizáveis (York & Yang, 1996).

O comportamento da densidade eletrônica está intrinsecamente ligado à natureza das interações moleculares, sendo importante saber como a densidade eletrônica é influenciada pela mudança da configuração molecular e o meio ambiente químico. A Teoria do Funcional da Densidade (DFT, density functional theory) fornece uma boa abordagem para resolver este problema, uma vez que a teoria do funcional da densidade é uma das metodologias que trata diretamente a densidade eletrônica como variável básica. Esta metodologia exige computação intensiva para moléculas grandes e então, métodos aproximados menos custosos computacionalmente são necessários (York & Yang, 1996).

Uma classe particular de métodos que procura resolver este problema é baseada no conceito da equalização da eletronegatividade. O método de equalização da eletronegatividade pode ser descrito como uma versão semi-empírica da teoria do funcional da densidade eletrônica (Parr & Yang, 1989; Baekelandt *et al.*, 1993). O cálculo direto da primeira e segunda derivada da energia com respeito ao número de elétrons (N) e ao potencial externo (v) permite que propriedades como eletronegatividade, dureza, funções de Fukui, entre outras, possam ser obtidas.

Segundo o teorema de Hohenberg e Kohn, a função de distribuição da densidade eletrônica contém toda a informação sobre um sistema de elétrons e núcleos no estado fundamental, ou seja, conhecendo-se o número total de elétrons (N) e o potencial externo em cada ponto, $v(\mathbf{r})$, o sistema é completamente caracterizado. O interessante

neste formalismo é a possibilidade de poder estudar a reatividade química. Uma forma de estudá-la é perturbar o sistema, isto pode ser feito de três formas, mudando o número de elétrons, N , ou modificando o potencial externo, $v(\mathbf{r})$, ou ainda mudando os dois ao mesmo tempo (*Baekelandt et al., 1993*).

A sensibilidade definida por Baekelandt e colaboradores é a variação de uma determinada propriedade X sob influência de uma perturbação, ou seja, a derivada de X em função de N ou $v(\mathbf{r})$. Neste trabalho, estudou-se apenas a primeira forma de perturbar o sistema. Segunda essa definição, propriedades fisicamente importantes do sistema podem ser determinadas para pequenas perturbações em termos de uma derivada independente da função da energia original $E[N, v]$, o Esquema 1 apresenta as propriedades que são determinadas pela primeira e segunda derivadas da energia E , onde μ é o potencial químico, η é a dureza e $f(\mathbf{r})$ é a função de Fukui.

Esquema 1: Primeira e Segunda derivada da energia, $E[N, v]$ e as propriedades por elas determinadas.

Parr e colaboradores (*Parr et al., 1978*) mostraram que o potencial químico é o negativo da eletronegatividade para qualquer espécie química. Uma vez que a aproximação do potencial químico $(\partial E / \partial N)$ pela diferença finita é

$$\mu \approx -\frac{I + A}{2} \quad (2)$$

sendo I o potencial de ionização e A a afinidade eletrônica. A definição da eletronegatividade de Mulliken é

$$\chi_M = \frac{I + A}{2} \quad (3)$$

portanto, $\mu \approx -\chi_M$.

Até o momento foram apresentados alguns conceitos de DFT, porém não como eles são utilizados no método de equalização da eletronegatividade. A expressão da energia eletrônica molecular descrita pelo funcional da densidade é

$$E[\rho] = F[\rho] + \int \rho(\mathbf{r})v(\mathbf{r})d\mathbf{r} \quad (4)$$

onde $\rho(\mathbf{r})$ é a função de distribuição da densidade eletrônica. O termo $F[\rho]$ do funcional da densidade é descrito pela soma das energias cinética eletrônica (T) e energia de repulsão elétron-elétron (V_{ee}), $F[\rho] = T[\rho] + V_{ee}[\rho]$. A expressão de energia, $E[\rho]$, é introduzida no formalismo do EEM como a soma das contribuições atômicas, como mostra a equação 5. A integração do funcional da densidade é então definida localmente por quantidades atômicas sobre a região atômica.

$$E[q] = F^*[q] + \sum_i \frac{1}{2} q_i v_i \quad (5)$$

onde v_i é o potencial eletrostático no átomo i e pode ser descrito da seguinte forma:

$$v_i = \sum_{j \neq i} \frac{q_j}{R_{ij}} \quad (6)$$

R_{ij} é a distância internuclear entre os átomos i e j .

A expressão de $F^*[q]$ no EEM, será introduzida como uma expansão de segunda ordem da energia em cargas atômicas (*Baekelandt et al., 1993*):

$$F^*[q] = \sum_i \left\{ E_i^* + \left(\frac{dE_i}{dq_i} \right) q_i + \frac{1}{2} \left(\frac{d^2 E_i}{dq_i^2} \right) q_i^2 \right\} \quad (7)$$

Sabendo-se que $\chi_i = \left(\frac{dE_i}{dq_i} \right)$ e $\eta_i = \left(\frac{d^2 E_i}{dq_i^2} \right)$ a equação 7 pode ser reescrita como:

$$F^*[q] = \sum_i \left\{ E_i^* + \chi_i^* q_i + \frac{\eta_i^*}{2} q_i^2 \right\} \quad (8)$$

Substituindo-se a equação 8 na equação 5, temos a expressão da energia eletrônica molecular pelo formalismo do EEM:

$$E[q] = \sum_i \left\{ E_i^* + \chi_i^* q_i + \frac{\eta_i^*}{2} q_i^2 + \frac{1}{2} q_i v_i \right\} \quad (9)$$

O potencial eletrostático, v_i , pode ser substituído por $\eta_{ij}(r_{ij})$, uma forma modificada do operador de Coulomb. Este termo, que é expresso na aproximação de Klopman-Ohno-Mataga-Nishimoto, considera os sítios carregados como nuvens de elétrons distribuídas entre os núcleos atômicos e não apenas simples cargas puntiformes (*Baekelandt et al., 1993; Bakowies & Thiel, 1996*):

$$\eta_{ij}(r_{ij}) = \frac{1}{\sqrt{r_{ij}^2 + \left(\frac{2}{\eta_i^* + \eta_j^*} \right)^2}} \quad (10)$$

apresentando os seguintes limites:

$$\begin{cases} r_{ij} \rightarrow 0^+ & \eta_{ij}(r_{ij}) \rightarrow \frac{(\eta_i^* + \eta_j^*)}{2} \\ r_{ij} \rightarrow +\infty & \eta_{ij}(r_{ij}) \rightarrow \frac{1}{r_{ij}} \end{cases}$$

Portanto, a expressão final da energia de uma molécula com N sítios atômicos pelo formalismo do EEM é:

$$E[q] = \sum_{i=1}^N \left(E_i^* + \chi_i^* q_i + \frac{\eta_i^*}{2} q_i^2 \right) + \frac{1}{2} \sum_{i=1}^N \sum_{\substack{j=1 \\ j \neq i}}^N \eta_{ij} (r_{ij}) q_i q_j \quad (11)$$

Para o cálculo das cargas nos sítios atômicos, o EEM quantifica o *princípio de Sanderson* que afirma que quando dois ou mais átomos se combinam para formar uma molécula suas eletronegatividades são equalizadas (Parr & Yang, 1989; Baekelandt et al., 1993).

$$\chi_1 \left(\equiv \frac{\partial E}{\partial q_1} \right) = \chi_2 \left(\equiv \frac{\partial E}{\partial q_2} \right) = \chi_3 \left(\equiv \frac{\partial E}{\partial q_3} \right) = \dots = \chi_N \left(\equiv \frac{\partial E}{\partial q_N} \right) = \chi_{eq} \quad (12)$$

Considerando como exemplo uma molécula diatômica AB, a sua expressão para a energia molecular será:

$$E = \left(E_A^* + \chi_A^* q_A + \frac{\eta_A^*}{2} q_A^2 \right) + \left(E_B^* + \chi_B^* q_B + \frac{\eta_B^*}{2} q_B^2 \right) + \eta_{AB} (r_{AB}) q_A q_B \quad (13)$$

Usando-se a definição da eletronegatividade, que é derivada parcial de E em função da carga, e o princípio da equalização da eletronegatividade, que no equilíbrio todos os átomos possuem a mesma eletronegatividade:

$$\begin{aligned} \chi_A \left(\equiv \frac{\partial E}{\partial q_A} \right) &= \chi_B \left(\equiv \frac{\partial E}{\partial q_B} \right) = \text{constante} \\ \chi_A^* + \eta_A^* q_A + \eta_{AB} q_B &= \chi_B^* + \eta_B^* q_B + \eta_{AB} q_A \\ \left\{ \begin{array}{l} \chi_B^* - \chi_A^* = (\eta_A^* - \eta_{AB}) q_A + (\eta_{AB} - \eta_B^*) q_B \\ q_A + q_B = Q \end{array} \right. \end{aligned} \quad (14)$$

Os termos χ_A^* , χ_B^* , η_A^* e η_B^* são as eletronegatividades e durezas parametrizáveis dos átomos A e B e podem ser obtidos por diferentes formas. As formas mais usadas como alvo na calibração do EEM são as cargas atômicas, energia de interação de pequenos “cluster” e os dados experimentais de potências de ionização e afinidades eletrônicas (Menegon *et al.*, 2002; Mortier *et al.*, 1986; Yang & Wang, 1997; Liu *et al.*, 1998; Mortier *et al.*, 1985). É importante ressaltar que χ_A^* é diferente de χ_A e diferente da eletronegatividade do átomo isolado, χ_A^0 , sendo esta afirmação válida para qualquer átomo da molécula. O termo χ_A^* é a eletronegatividade parametrizável, ou seja, o valor de χ_A^* é otimizado para reproduzir determinada propriedade usada como alvo na calibração do EEM, enquanto que χ_A é a eletronegatividade do átomo A dentro da molécula no equilíbrio e, por fim, χ_A^0 é a eletronegatividade do átomo A isolado, sendo a escala de eletronegatividade de Pauling (1932; 1932b) uma das mais conhecidas. A partir das duas equações acima pode-se calcular as cargas atômicas da molécula AB, uma vez que tem-se duas incógnitas e duas equações. Para generalizar o cálculo de cargas atômicas pelo formalismo do EEM para moléculas de qualquer tamanho, reescreveu-se a equação 14 para um *sistema linear* de N linhas com N incógnitas (as cargas atômicas):

$$\begin{cases} Q = \sum_{j=1}^N q_j \\ \chi_1^* - \chi_2^* = \eta_2^* q_2 - \eta_1^* q_1 + \sum_{\substack{j=1 \\ (j \neq 2)}}^N \eta_{2j}(r_{2j}) q_j - \sum_{\substack{k=1 \\ (k \neq 1)}}^N \eta_{1k}(r_{1k}) q_k \\ \vdots \quad \quad \quad \vdots \\ \chi_1^* - \chi_N^* = \eta_N^* q_N - \eta_1^* q_1 + \sum_{\substack{j=1 \\ (j \neq N)}}^N \eta_{Nj}(r_{Nj}) q_j - \sum_{\substack{k=1 \\ (k \neq 1)}}^N \eta_{1k}(r_{1k}) q_k \end{cases} \quad (15)$$

Portanto, para se calcular as cargas atômicas, os termos χ_i^* e η_i^* devem ser conhecidos, assim como a geometria da molécula que é usada para calcular $\eta_{ij}(r_{ij})$. Deve-se notar que a grande vantagem de se utilizar o EEM no cálculo de distribuição de carga é a sua simplicidade e rapidez em ser implementado computacionalmente (Liu *et al.*, 1998; Ribeiro & Almeida, 2000). Muitos métodos de equalização da eletronegatividade foram formulados.

Mortier e colaboradores (*Mortier et al., 1986*) desenvolveram um método de equalização da eletronegatividade para o cálculo de cargas atômicas e energia eletrônica da molécula. A expressão da eletronegatividade efetiva do átomo na molécula, que é igual a eletronegatividade molecular, foi derivada a partir da teoria do funcional da densidade. Os parâmetros, eletronegatividade e dureza, foram obtidos pela calibração de cargas atômicas de moléculas pequenas, tendo cargas Mulliken STO-3G como cargas alvo. Os valores encontrados foram considerados pelos autores transferíveis e consistentes para o cálculo de cargas atômicas de qualquer molécula.

No e colaboradores estabeleceram um método modificado para a determinação de cargas atômicas, utilizando-se um esquema já existente de equalização parcial da eletronegatividade orbital (*No et al., 1990*). Esta metodologia introduz dipolos experimentais e quando possível, momentos de quadrupolo de uma lista de pequenas moléculas escolhidas tendo diferentes “ambientes” atômicos encontrados em proteínas. Conseqüentemente, as cargas atômicas obtidas reproduzem os valores experimentais de momentos de dipolo e quadrupolo. As cargas encontradas foram o dobro em magnitude das cargas atômicas calculadas em nível *ab initio*. Outro ponto questionável desta metodologia é a transferabilidade e aplicação do esquema para moléculas grandes como proteínas, por exemplo.

Rappé e Goddard definiram um potencial químico usando potenciais de ionizações atômicos, afinidades eletrônicas e raios atômicos experimentais junto com as interações eletrostáticas entre todas as cargas (*Rappé & Goddard III, 1991*). A equalização dos potenciais químicos produz um equilíbrio das cargas, que depende da geometria. Esta aproximação de cargas equilibradas (QEq) fornece cargas que reproduzem os momentos de dipolos experimentais e as cargas atômicas obtidas a partir de potenciais eletrostáticos calculados em nível *ab initio*, sendo uma metodologia que pode ser usada para prever cargas atômicas de polímeros, cerâmicas, semicondutores ou sistemas biológicos.

Cioslowski e Stefanov mostraram que o cálculo de cargas fixadas permite uma rigorosa análise do fluxo de elétrons e da equalização da eletronegatividade no processo de formação da ligação química, fornecendo evidências e explicações para a descontinuidade da derivada da energia que são observadas nos átomos isolados e nas moléculas (*Cioslowski & Stefanov, 1993*).

De Profít e colaboradores formularam um esquema de equalização da eletronegatividade não empírico, onde a expressão para a eletronegatividade de um

átomo na molécula foi derivada baseando-se na mudança do número de elétrons e do potencial externo deste átomo, sobre a formulação da molécula, em relação ao átomo isolado (*De Proft et al., 1995*).

York e Yang apresentaram uma formulação do princípio de equalização do potencial químico através da perspectiva da teoria do funcional da densidade (*York & Yang, 1996*). As equações utilizadas para descrever a redistribuição de cargas provocada pela perturbação do sistema pela aplicação de um campo foram obtidas por uma aproximação de reposta linear. Neste trabalho também foi apresentada uma formulação que fornece uma base para um modelo semi-empírico promissor de polarização e transferência de carga em simulações moleculares.

Yang e Wang desenvolveram um método de equalização da eletronegatividade de átomo-ligação para o cálculo da distribuição de cargas, que depende da geometria e da conectividade, baseando-se no princípio da equalização da eletronegatividade e na teoria do funcional da densidade (*Yang & Wang, 1997*). A equação da eletronegatividade efetiva de um átomo e uma ligação na molécula é descrita por termos como eletronegatividades e durezas e contribuições oriundas de outras cargas atômicas e de ligações.

3.2 – Modelo de Cargas 1 (CM1, Charge Model 1)

O modelo de carga 1 implementou a idéia dos modelos de cargas da classe IV (procedimento que corrige cargas obtidas por análise de população para que estas reproduzam observáveis obtidas experimentalmente) em nível semi-empírico, utilizando os Hamiltonianos AM1 e PM3. A idéia básica do CM1 é simples e intuitiva, baseando-se no fato de que os momentos de dipolo obtidos a partir de cargas de análise de população possam ser imprecisos, mas que a contribuição para o erro de cada tipo de ligação é sistemática. Então foram propostas correções sistemáticas para os momentos de dipolo de ligação, uma correção para cada tipo de ligação de forma semi-empírica (*Storer et al., 1995*). A parametrização do modelo CM1 é feita a partir de algumas escolhas:

- 1) Hamiltoniano usado na obtenção da função de onda,
- 2) a análise de população de função de onda,

- 3) a forma da funcional de mapeamento,
- 4) a lista de moléculas usadas na parametrização.

O mapeamento foi feito em nível semi-empírico, empregando-se os Hamiltonianos citados acima, com cargas Mulliken em um conjunto de 186 moléculas neutras. A funcional usada no mapeamento adiciona a transferência de carga de um átomo para outro em uma dada ligação, $T_{kk'}$, no valor da carga Mulliken, q_k^0 .

$$q_k = q_k^0 + \sum_{k' \neq k} T_{kk'}(B_{kk'}) \quad (16)$$

A transferência de carga depende de parâmetros empíricos, $D_{kk'}$ e $C_{kk'}$, e da ordem de ligação ($B_{kk'}$).

$$T_{kk'} = B_{kk'}(D_{kk'} + C_{kk'} B_{kk'}) \quad (17)$$

3.3 – Cálculo da redistribuição de cargas do soluto em um modelo de solvente implícito

A maioria dos processos químicos ocorre na presença de um solvente e portanto, seus efeitos nas energias e estruturas moleculares devem ser considerados (*Cossi et al., 2002*). A redistribuição de carga ocorre quando o campo de reação eletrostático do solvente acopla-se à estrutura eletrônica do soluto (*Cramer & Truhlar, 1999; Tomasi & Persico, 1994*).

O primeiro ingrediente do modelo proposto para incluir esse efeito em solvatação implícita foi escolher um método de estrutura eletrônica de grande versatilidade e simplicidade: o *método de equalização da eletronegatividade* (EEM, Electronegativity Equalization Method).

O segundo ingrediente foi definir o ciclo termodinâmico usado no cálculo da energia de solvatação (ΔG_{solv}) e acoplamento eletrostático soluto-solvente.

As contribuições para ΔG_{solv} estão relacionadas pelo seguinte ciclo (*Cramer & Truhlar, 1999; Nina et al., 1997; Tomasi & Persico, 1994*): (1) desconsiderar as interações internas do soluto no vácuo; (2) formação de uma cavidade adequada no

solvente considerando as interações de dispersão soluto-solvente; (3) considerar as interações internas do soluto no solvente com os termos de polarização soluto-solvente.

$$\Delta G_{solv} = \Delta E_{EN} + \Delta G_{np} + \Delta G_{el} \quad (18)$$

O termo ΔE_{EN} ($\Delta E_{EN} = E_{EN,S} - E_{EN,vac}$) é a diferença de energia nuclear e eletrônica entre o vácuo e a fase condensada. ΔG_{np} é a contribuição não polar para a energia livre de solvatação e ΔG_{el} é o trabalho eletrostático reversível.

Usando o método de equalização da eletronegatividade para expressar ΔE_{EN} em termos da eletronegatividade atômica, χ_i^* , e dureza atômica, η_i^* , obtêm-se a seguinte expressão:

$$\begin{aligned} \Delta G_{solv} = & \sum_{i=1}^N \left(E_i^* + \chi_i^* q_i + \frac{\eta_i^*}{2} q_i^2 \right) + \frac{1}{2} \sum_{i=1}^N \sum_{\substack{j=1 \\ (j \neq i)}}^N \eta_{ij}(r_{ij}) q_i q_j + \Delta G_{np} + \Delta G_{el} \\ & - \sum_{i=1}^N \left(E_i^* + \chi_i^* q_{i,vac} + \frac{\eta_i^*}{2} q_{i,vac}^2 \right) - \frac{1}{2} \sum_{i=1}^N \sum_{\substack{j=1 \\ (j \neq i)}}^N \eta_{ij}(r_{ij}) q_{i,vac} q_{j,vac} \end{aligned} \quad (19)$$

Os $\eta_{ij}(r_{ij})$ são os termos de interação de distribuição das cargas e q_i e $q_{i,vac}$ são as cargas atômicas do soluto no solvente e no vácuo, respectivamente.

O componente ΔG_{np} não depende da distribuição de carga e pode ser calculado usando a área superficial de acesso (Bashford & Case, 2000; Cramer & Truhlar, 1999; Tomasi & Persico, 1994). Este termo empírico inclui cavitação, interações de dispersão

e termos adicionais, tais como contribuições de ligações de hidrogênio para a energia livre de interação soluto-solvente.

O terceiro ingrediente foi escolher um modelo de trabalho eletrostático expresso como um polinômio de segunda ordem nas cargas atômicas (como o EEM).

O cálculo do potencial eletrostático, ϕ , por este modelo, envolve a resolução exata da equação de Poisson,

$$\nabla[\epsilon(\mathbf{r})\nabla\phi(\mathbf{r})] = -4\pi\rho(\mathbf{r}) \quad (20)$$

onde ρ é a distribuição de cargas, e a constante dielétrica, ϵ , possui o valor de ϵ_{in} no interior do soluto e ϵ_{ex} em qualquer lugar do dielétrico. Nas condições da fase gasosa, $\epsilon_{ex} = 1$, enquanto que nas condições do solvente, $\epsilon_{ex} = \epsilon_w$, a constante dielétrica do solvente. Resolvendo a equação 20 nas duas condições descritas, obtêm-se os potenciais eletrostáticos na fase gasosa e no solvente, ϕ_{vac} e ϕ_{sol} , respectivamente. A diferença entre esses potenciais é o campo de reação, $\phi_{reac} = \phi_{sol} - \phi_{vac}$, e o componente eletrostático da energia livre de solvatação é

$$\Delta G_{el} = \frac{1}{2} \int \phi_{reac}(\mathbf{r}) \rho(\mathbf{r}) dV \quad (21)$$

ou, se a distribuição da carga molecular for aproximada para cargas atômicas parciais q_i :

$$\Delta G_{el} = \frac{1}{2} \sum_i q_i \phi_{reac}(\mathbf{r}_i) \quad (22)$$

No caso de um íon simples de raio a e carga q , os potenciais podem ser encontrados analiticamente e o resultado é a fórmula de Born:

$$\Delta G_{Born} = -\frac{q^2}{2a} \left(1 - \frac{1}{\epsilon_w} \right) \quad (23)$$

Considere uma molécula de N átomos com cargas $q_1, q_2, \dots, q_N$ inseridas em esferas de raios $a_1, a_2, \dots, a_N$, e que a separação r_{ij} entre quaisquer duas esferas seja suficientemente

grande em comparação com o raio, então a energia eletrostática pode ser dada pela soma individual dos termos de Born e dos termos de Coulomb:

$$\Delta G_{el} = \sum_i^N \frac{q_i^2}{2a_i} \left(\frac{1}{\epsilon_w} - 1 \right) + \frac{1}{2} \sum_i^N \sum_{j \neq i}^N \frac{q_i q_j}{r_{ij}} \left(\frac{1}{\epsilon_w} - 1 \right) \quad (24)$$

onde o fator $\left(\frac{1}{\epsilon_w} - 1 \right)$ aparece no termo de Coulomb, porque a interação coulômbica foi reescalada para a mudança de constante dielétrica do vácuo para o solvente. O solvente foi tratado como um meio isotrópico e homogêneo (*Bashford & Case, 2000*).

A aproximação de Born generalizado representa o trabalho eletrostático como uma expansão quadrática em relação às cargas (*Bashford & Case, 2000; Cramer & Truhlar, 1999; Still et al., 1990*) semelhante a equação 24. Troca-se o r_{ij} na equação 24 por uma função que tem como limites r_{ij} e $\sqrt{\alpha_i^B \alpha_j^B}$ para longas e curtas distâncias, respectivamente. Essa função de Born generalizado, f_{GB} , é:

$$f_{GB} = \sqrt{r_{ij}^2 + \alpha_i^B \alpha_j^B} e^{-r_{ij}^2 / 4\alpha_i^B \alpha_j^B} \quad (25)$$

sendo que, o α_i^B é o raio efetivo de Born (*Bashford & Case, 2000; Still et al., 1990*), que não depende somente do raio atômico intrínseco, mas também da posição relativa de todos os outros átomos. Na aproximação de campo de Coulomb o vetor deslocamento elétrico é centro simétrico e não muda de direção quando encontra uma interface que separa meios de constante dielétrica diferentes. Usando esta aproximação, a expressão do raio efetivo de Born pode ser descrita como:

$$\frac{1}{\alpha_i^B} = \frac{1}{4\pi} \int_{ex} \frac{1}{r^4} dV \quad (26)$$

A expressão do raio efetivo de Born pode ser reescrita considerando que a integração é feita sobre a região ocupada pela molécula, excluindo-se a esfera de raio a_i , e com o átomo i localizado na origem da coordenada do sistema:

$$\frac{1}{\alpha_i^B} = \frac{1}{a_i} - \frac{1}{4\pi} \int_{in, r>a_i} \frac{1}{r^4} dV \quad (27)$$

O α_i^B foi calculado usando a aproximação de campo de Coulomb e a abordagem de integral de superfície (*Bashford & Case, 2000*). A superfície molecular de van der Waals é a superfície gerada por intersecções de esferas centradas em átomos. Usando o programa GEPOL93 (*Pascual-Ahuir et al., 1993*), que particiona a superfície molecular em pequenos domínios chamados de *tesseraes* (Figura 4), as coordenadas dos centros de cada *téssera*, $\mathbf{r}_j$, o vetor normal perpendicular a superfície deste ponto, $\mathbf{n}(\mathbf{r}_j)$, e a área de cada *téssera*, Δa_j , são utilizados na expressão abaixo para obtenção de α_i^B (letras em negrito correspondem à vetores):

$$\frac{1}{\alpha_i^B} = \frac{1}{4\pi} \int \frac{(\mathbf{r} - \mathbf{R}_i) \cdot \mathbf{n}(\mathbf{r})}{|\mathbf{r} - \mathbf{R}_i|^4} d^2 r = \frac{1}{4\pi} \sum_{j=1}^{N_{tesserae}} \frac{(\mathbf{r}_j - \mathbf{R}_i) \cdot \mathbf{n}(\mathbf{r}_j)}{|\mathbf{r}_j - \mathbf{R}_i|^4} \Delta a_j \quad (28)$$

onde $\mathbf{R}_i$ é o vetor posição de cada átomo i .

A aproximação de Born generalizado tem por objetivo encontrar uma fórmula analítica relativamente simples (equação 29), que descreva o quanto possível, a equação de Poisson (*Bashford & Case, 2000*).

$$\Delta G_{el} = \sum_{i=1}^N \frac{q_i^2}{2\alpha_i^B} \left(\frac{1}{\epsilon} - 1 \right) + \frac{1}{2} \sum_{i=1}^N \sum_{\substack{j=1 \\ (j \neq i)}}^N \frac{q_i q_j}{f_{GB}(r_{ij})} \left(\frac{1}{\epsilon} - 1 \right) \quad (29)$$

Como último ingrediente, usou-se o princípio variacional na energia de solvatação para determinar as cargas atômicas referentes ao acoplamento eletrostático soluto-solvente (similarmente ao EEM). Exemplificando, seja um soluto diatômico (AB) com carga total $q_a + q_b = 0$. Usando as equações 18 e 19 para minimizar a energia livre de solvatação com respeito à carga A, obtêm-se:

$$\frac{d\Delta G_{solv}}{dq_A} = 0 = \chi_A^* - \chi_B^* + (\eta_A^* + \eta_B^* - 2\eta_{AB}(r_{AB}))q_A + \left(\frac{1}{\epsilon} - 1 \right) \left(\frac{1}{\alpha_A^B} + \frac{1}{\alpha_B^B} - \frac{2}{f_{GB}(r_{AB})} \right) q_A \quad (30)$$

Figura 4: Partição em *tesseraes* da superfície de Van der Waals da molécula de água usando: a) áreas de $0.2 \text{ \AA}^2$ e b) áreas de $0.4 \text{ \AA}^2$.

e portanto:

$$q_A = \frac{\chi_B^* - \chi_A^*}{\left(\eta_A^* + \eta_B^* - 2\eta_{AB}(r_{AB})\right) + \left(\frac{1}{\varepsilon} - 1\right) \left(\frac{1}{\alpha_A^B} + \frac{1}{\alpha_B^B} - \frac{2}{f_{GB}(r_{AB})}\right)} \quad (31)$$

A expressão para a carga atômica A num solvente implícito, quando comparada à expressão do EEM para A no vácuo, mostra um termo eletrostático adicional (último fator no denominador da equação 31). De fato, a mesma expressão pode ser obtida definindo as eletronegatividades efetivas dos sítios atômicos do soluto no solvente por:

$$\langle \chi_i \rangle = \frac{\partial \Delta G_{solv}}{\partial q_i} = \frac{\partial (E_{EN,s} + \Delta G_{el})}{\partial q_i} \quad (32)$$

e usando um princípio de equalização da eletronegatividade generalizado:

$$\langle \chi_A \rangle = \langle \chi_B \rangle \quad (33)$$

Este novo modelo que acopla estrutura eletrônica via EEM e trabalho eletrostático via Born generalizado foi denominado de GBEEM (Generalized Born-Electronegativity Equalization Method) (*Dias et al., 2002*).

Como o GBEEM é baseado em um tipo de modelo de Born, este pode ser classificado como um método de campo de reação auto-consistente na aproximação de resposta linear (*Cramer & Truhlar, 1999; Chambers et al., 1996; Tapia et al., 1982*).

3.4 – Cálculo da energia livre de solvatação em modelo de solvente contínuo

Energias livres de solvatação podem ser calculadas por modelos de contínuo e muitas propriedades eletrônicas importantes podem ser obtidas pela integração de diferentes características do solvente. As propriedades são calculadas usando-se uma combinação da primeira camada de solvatação, com moléculas explícitas do solvente, com o modelo de contínuo para representar o solvente. Resultados ainda melhores são

obtidos quando a primeira camada de solvatação é determinada a partir de cálculos de dinâmica molecular (*Cossi et al., 2002*).

Modelos de solvente implícito para cálculo de energias livres de solvatação estão sendo desenvolvidos devido a sua eficiência e rapidez (*Roux, 2001; Tomasi & Persico, 1994*). Nas últimas décadas aumentou o interesse em se calcular energias livres de solvatação por metodologias de campo de reação auto-consistente (Self-Consistent Reaction Field, SCRF), envolvendo o cálculo da estrutura eletrônica com métodos de dielétrico contínuo (*Marten et al., 1996*).

No cálculo da energia livre de solvatação utilizou-se o mesmo ciclo termodinâmico descrito no item anterior (cálculo de cargas no solvente) para definir as contribuições de ΔG_{solv} (*Cramer & Truhlar, 1999; Nina et al., 1997; Tomasi & Persico, 1994*): (1) desligar as interações internas do soluto no vácuo; (2) formação de uma cavidade adequada no solvente considerando as interações de dispersão soluto-solvente; (3) ligar as interações internas do soluto no solvente com os termos de polarização soluto-solvente.

$$\Delta G_{\text{solv}} = \Delta E_{\text{EN}} + \Delta G_{\text{np}} + \Delta G_{\text{el}} \quad (18)$$

A equação 18 mostra que serão necessários calcular ΔE_{EN} , ΔG_{np} e ΔG_{el} para se obter a energia livre de solvatação, o interessante que deve ser ressaltado aqui, é que o ΔE_{EN} e o ΔG_{el} serão calculados a partir das metodologias desenvolvidas anteriormente neste trabalho.

O termo ΔG_{np} não depende da distribuição de carga e é geralmente estimado usando um único parâmetro de tensão superficial microscópica a partir da área superficial de acesso (*Bashford & Case, 2000; Cramer & Truhlar, 1999; Tomasi & Persico, 1994*). O uso de um único parâmetro de tensão superficial está baseado no fato que para alcanos lineares a energia livre de solvatação, a energia livre padrão de transferência gás $\rightarrow$ água, aumenta quase que linearmente com a área superficial total acessível ao solvente (*Zacharias, 2003*):

$$\Delta G_{\text{solv}} = \beta + \gamma \text{SAS} \quad (34)$$

onde SAS (Solvent Accessible Surface) é a área superficial total acessível ao solvente.

SAS é gerada pelo centro de uma esfera rolando sobre a superfície molecular de van der Waals do soluto, onde o raio da esfera está relacionado com algo como uma “espessura média” da primeira camada de solvatação e portanto, com o “tamanho” da molécula do solvente (*Cramer e Truhlar, 1999*), Figura 3c. Para o cálculo da SAS, raios atômicos são necessários como também a geometria da molécula. Os raios atômicos definidos por Bondi (*Bondi, 1964*) junto com o programa GEPOL93 (*Pascual-Ahuir et al., 1993*) foram usados para computar a SAS.

A parte não-polar da energia livre de solvatação pode ser calculado por outros métodos, além do apresentado na equação 34, que é um dos mais simples. Note que para alcanos, no modelo descrito anteriormente, é assumido que as contribuições ΔE_{EN} e ΔG_{el} para a energia de solvatação são desprezíveis e apenas o termo ΔG_{np} contribui para o processo. Entretanto, este comportamento não é observado para alcanos cíclicos. Zacharias sugere que uma melhor estimativa do termo ΔG_{np} seria obtido dividindo-o em duas partes: a-) cavitação, que inclui contribuições entrópicas e entálpicas devidas a reordenação e reorientação das moléculas de água ao redor do soluto; b-) interações de van der Waals soluto-solvente.

Assim:

$$\Delta G_{solv} = a + bSAS + cU_{disp} \quad (35)$$

U_{disp} é a energia média de interação de dispersão entre soluto e solvente (*Zacharias, 2003; Barone et al., 1997; Tomasi & Persico, 1994*). A expressão para U_{disp} é obtida considerando que a interação de dispersão é aditiva para os átomos que constituem o soluto:

$$U_{disp} = \sum_i^{N_s} \rho_w C_{iw} \int_{V_w} dr_{iw} g(r_{iw}) r_{iw}^{-6} \quad (36)$$

onde: ρ_w é a densidade do solvente (água), C_{iw} é o parâmetro que descreve a interação de dispersão entre o átomo “i” da molécula e o solvente (no caso, a água), r_{iw} é a distância entre a água e os átomos do soluto e $g(r_{iw})$ é a função de distribuição radial entre o solvente e os sítios atômicos do soluto.

A soma é feita para todos os átomos do soluto (N_s) e a integração é feita a partir do volume excludente do soluto (V_w). Tomasi e Persico (Tomasi & Persico, 1994) mostraram que esta integral de volume pode ser transformada em uma integral de superfície do volume excludente do soluto. Na aproximação uniforme a função de distribuição radial solvente-sítio atômico do soluto, $g(r_{iw})$, é zero dentro do soluto e 1 fora, ou seja que não existe solvente entrando na cavidade do soluto e que fora da cavidade o solvente é distribuído como um gás ideal:

$$U_{disp} = \sum_i^{N_s} \rho_w C_{iw} \int_{S_s} d\sigma_s \mathbf{A}(\mathbf{r}_{iw}) \mathbf{n}_s \quad (37)$$

sendo $\mathbf{A}(\mathbf{r}_{iw}) = \mathbf{r}_{iw}/(3r_{iw}^6)$. A interação de dispersão pode ser estimada a partir da soma dos elementos da superfície, ΔS_k , ao redor da molécula:

$$U_{disp} = \sum_i^{N_s} \rho_w C_{iw} \sum_k \Delta S_k A(r_{ik}) \mathbf{n}_{S_k} \quad (38)$$

onde r_{ik} é a distância de pontos do átomo i para o elemento de superfície k e $\mathbf{n}_{S_k}$ é o vetor normal associado com o elemento de superfície ΔS_k .

O volume excludente do soluto e sua superfície são determinados por GEPOL93 considerando os mesmos raios atômicos do soluto e solvente utilizados no cálculo da SAS. A soma da equação 38 foi dividida e reagrupado por tipos atômicos: H, C, O e N em $U_{disp}(H)$, $U_{disp}(C)$, $U_{disp}(O)$ e $U_{disp}(N)$, sendo que o U_{disp} total é a soma dos componentes por átomos. A equação 34 pode ser reescrita:

$$\Delta G_{solv} = a + bSAS + c_H U_{disp}(H) + c_C U_{disp}(C) + c_O U_{disp}(O) + c_N U_{disp}(N) \quad (39)$$

Neste estudo combinamos a parte eletrostática, calculada com o GBEEM, com a parte não polar, calculada do modelo de Floris-Tomasi:

$$\Delta G_{solv} = \Delta E_{EN} + \Delta G_{el} + a + bSAS + c_H U_{disp}(H) + c_C U_{disp}(C) + c_O U_{disp}(O) + c_N U_{disp}(N) \quad (40)$$

Alguns parâmetros utilizados no cálculo da energia de solvatação foram fixados: eletronegatividades e durezas para o EEM-CM1 (parâmetros obtidos também neste estudo), raio do solvente (0,17nm), raios atômicos extraídos da compilação de Bondi (*Bondi, 1964*), constante dielétrica do solvente ($\epsilon = 78,5$) e densidade do solvente ($\rho_w = 0,03346$ moléculas/Å³).

4 – Resultados e Discussão

4.1 – Cálculo da redistribuição de cargas do soluto em um modelo de solvente implícito

O cálculo da redistribuição de carga de solutos neutros imersos em um solvente implícito (tipo *continuum* dielétrico) foi efetuado pela combinação do método de equalização da eletronegatividade com o modelo de Born generalizado, utilizando-se os parâmetros de eletronegatividades e durezas químicas da calibração QEq-PD (*Bakowies & Thiel, 1996*), que usa cargas atômicas derivadas de potencial eletrostático em nível *ab initio* RHF/6-31G como alvo. Inicialmente foram calculados no vácuo (constante dielétrica 1) os dipolos (Tabela 1) e cargas (Tabela 2), de um conjunto de 12 moléculas neutras.

O primeiro ponto interessante é que os momentos de dipolo calculados com esta calibração (QEq-PD), são maiores do que os calculados por cargas atômicas CM1 em nível semi-empírico PM3 (denominado de PM3-CM1). A escolha deste modelo para comparação foi feita pelo fato das cargas CM1 produzirem momentos de dipolo similares aos calculados em nível *ab initio*, porém com as vantagens dos modelos semi-empíricos, em termos de recursos computacionais. Foi verificado empiricamente que um aumento em cerca de 6% nas durezas atômicas do QEq-PD resulta em momentos de dipolo (Tabela 1) e cargas (Tabela 2) comparáveis aos valores das cargas e momentos de dipolo calculados por PM3-CM1. Estas novas durezas atômicas corrigidas em conjunto com as eletronegatividades originais foram denominadas EEM-PD.

A análise dos resultados organizados na Tabela 2 mostra que os valores de cargas atômicas calculadas no vácuo por QEq-PD apresentam inversões de sinal para metanol e metilamina, em comparação com os resultados calculados por PM3-CM1. Com a correção de 6% nas durezas atômicas do método EEM-PD, os valores de cargas atômicas calculadas estão mais próximos dos valores PM3-CM1. Apenas o átomo de nitrogênio do composto nitrometano apresenta inversão de sinal.

O gráfico das cargas atômicas EEM-PD versus PM3-CM1 no vácuo (Figura 5) mostra que existe uma boa correlação entre as cargas calculadas por ambos os modelos (coeficiente de correlação, R , de 0,94 e desvio padrão, SD , de 0,15). Este resultado é comparável à correlação obtida entre as cargas QEq-PD e as cargas derivadas de potencial eletrostático calculadas em nível RHF, cujo coeficiente de correlação e desvio padrão são 0,96 e 0,10, respectivamente (*Bakowies & Thiel, 1996*).

Tabela 1: Momentos de dipolo no vácuo (μ_g) calculados pelas métodos QEq-PD, EEM-PD e PM3-CM1 (valores em Debye).

Molécula	QEq-PD	EEM-PD	PM3-CM1
	μ_g	μ_g	μ_g
Dimetil éter	1,58	1,45	1,22
Formaldeído	2,59	2,36	2,41
Acetona	3,01	2,77	2,92
Metanoato de metila	1,75	1,61	1,79
Acetonitrila	3,40	3,03	3,89
Metilamina	1,43	1,24	1,21
Nitrometano	4,54	4,23	3,70
Formamida	3,19	2,83	3,56
Água	2,13	1,87	1,92
Metanol	1,83	1,68	1,59
Ciclohexilamina	1,09	0,96	1,03
Metil n-propil éter	1,26	1,18	1,14
Etanal	3,03	2,77	2,70
Butanona	3,06	2,82	2,88

Tabela 2: Valências atômicas para moléculas isoladas. Os índices I-J representam o primeiro e segundo átomos na coluna q_I e q_J , respectivamente. Para as moléculas que apresentam mais de um átomo representado por I ou J (Metano, Etano, Metanol, Formaldeído, Metilamina, Água e Nitrometano), os valores das valências são médias aritméticas.

I-J	Molécula	QEq-PD ^a		EEM-PD ^b		PM3-CM1	
		q_I	q_J	q_I	q_J	q_I	q_J
H-C	Metano	0,123	-0,490	0,133	-0,532	0,028	-0,110
	Etano	-	-	0,090	-0,135	0,086	-0,129
	Metanol	-0,005	0,276	0,062	0,008	0,021	0,110
	Metilamina	-0,035	0,379	0,051	-0,012	0,035	-0,026
	Formaldeído	0,009	0,441	0,048	0,266	0,014	0,376
H-N	Metilamina	0,379	-1,030	0,317	-0,774	0,292	-0,665
	Formamida	0,431	-0,994	0,374	-0,904	0,323	-0,671
H-O	Água	0,409	-0,817	0,340	-0,680	0,357	-0,714
	Metanol	0,427	-0,688	0,362	-0,532	0,355	-0,528
C-N	Metilamina	0,379	-1,030	-0,012	-0,774	-0,026	-0,665
	Formamida	-	-0,994	0,653	-0,904	0,386	-0,671
C $\equiv$ N	Acetonitrila	0,468	-0,515	0,482	-0,437	0,107	-0,341
C=O	Formaldeído	0,441	-0,459	0,266	-0,362	0,376	-0,403
	Metanoato de metila	-	-	0,678	-0,470	0,496	-0,464
	Acetona	-	-	0,439	-0,410	0,362	-0,407
C-O	Metanol	0,276	-0,688	0,008	-0,536	0,110	-0,528
	Dimetil éter	-	-	0,013	-0,395	0,095	-0,356
N-O	Nitrometano	0,870	-0,485	-0,082	-0,140	1,017	-0,520

a-) geometrias otimizadas no nível RHF/6-31G*, as cargas foram calculadas pelo método de equalização da eletronegatividade, QEq-PD;

b-) geometrias otimizadas no nível semi-empírico PM3, as cargas foram calculadas pelo método de equalização da eletronegatividade, EEM-PD.

Figura 5: Gráfico das cargas calculadas por EEM-PD versus cargas calculadas por PM3-CM1. A linha tracejada representa a identidade e a linha contínua representa a correlação linear entre os dados ($y = ax$, onde $a = 1,100 \pm 0,070$, $R = 0,940$, $SD = 0,155$, $F = 234$ e $N = 32$).

Utilizando uma constante dielétrica igual a 78,5 e os mesmos parâmetros corrigidos em 6%, foram calculadas as redistribuições de cargas atômicas em água (GBEEM/PD) e, para comparação, foram obtidas cargas PM3-CM1 e momento de dipolo, usando o modelo de solvatação SM5.4 (*Chambers et al., 1996*), com a geometria otimizada em fase condensada PM3-SM5.4, como mostra a Tabela 3 para cargas e Tabela 4 para dipolos. Os valores das cargas atômicas GBEEM-PD estão em boa concordância com os valores obtidos por PM3-CM1. A principal diferença ocorre no nitrogênio de nitrometano, com inversão de sinal, como observado nos cálculos no vácuo e nos átomos de carbono que estão ligados aos átomos de nitrogênio, principalmente na metilamina. Uma explicação plausível para o desvio do grupo NO₂ reside no fato de que todas as possíveis ligações químicas formadas pelo átomo de nitrogênio estão sendo descritas por um único par eletronegatividade e dureza química. Este par se mostrou eficiente no cálculo de cargas em outros grupos funcionais envolvendo o átomo de nitrogênio. Entretanto, no caso do grupo NO₂ talvez fosse necessária a inclusão de um conjunto de parâmetros específico. No caso dos átomos de carbono ligados aos átomos de nitrogênios o desvio pode ocorrer pelo mesmo motivo do grupo NO₂ ou ainda como uma forma de compensar os desvios associados às cargas do nitrogênio. É importante ressaltar que a discussão acerca deste tipo de inversão no sinal das cargas atômicas e também dos desvios deve ser feita com cautela, pois os cálculos de referência em nível semi-empírico podem sofrer dos mesmos problemas apresentados para o EEM.

A regressão linear entre as cargas atômicas dos solutos em água, calculadas usando GBEEM-PD e PM3-CM1 (Figura 6) exibe valores estatísticos muito próximos aos apresentados pelo EEM-PD para os mesmos solutos no vácuo, sendo $R = 0,94$ e $SD = 0,16$ para o GBEEM-PD e $R = 0,94$ e $SD = 0,15$ para o EEM-PD. Um ponto que merece destaque é que a redistribuição de carga no solvente calculada pelo GBEEM-PD foi efetuada sem a necessidade de uma reparametrização das eletronegatividades e durezas no solvente, ou seja, o GBEEM-PD emprega os mesmos parâmetros obtidos para o cálculo no vácuo. Adicionalmente, os resultados obtidos no vácuo e no solvente sugerem que os modelos propostos neste trabalho podem ser generalizados e aplicados para outras moléculas que não pertencem ao conjunto de calibração.

Tabela 3: Cargas atômicas na água (valores em unidade atômica, $\epsilon \cong 78,5$). Os termos I-J representam o primeiro e segundo átomos na coluna q_I e q_J . As cargas atômicas são médias aritméticas para as moléculas que apresentam mais de um átomo do tipo I ou J.

I-J	Molécula	GBEEM-PD ^{a,b}		PM3-CM1	
		q_I	q_J	q_I	q_J
H-C	Metanol	0,084 (35)	0,000 (100)	0,024 (14)	0,115 (5)
	Metilamina	0,156 (206)	-0,288 (2300)	0,034 (3)	-0,032 (23)
	Formaldeído	0,096 (100)	0,271 (2)	0,052 (271)	0,439 (17)
H-N	Metilamina	0,339 (7)	-0,858 (11)	0,314 (8)	-0,698 (5)
	Formamida	0,445 (19)	-0,975 (8)	0,353 (9)	-0,585 (13)
H-O	Água	0,417 (23)	-0,834 (23)	0,402 (13)	-0,803 (12)
	Metanol	0,476 (31)	-0,644 (21)	0,403 (14)	-0,591 (12)
C-N	Metilamina	-0,288 (2300)	-0,858 (11)	-0,032 (23)	-0,698 (5)
	Formamida	0,668 (2)	-0,975 (8)	0,389 (1)	-0,585 (13)
C $\equiv$ N	Acetonitrila	0,493 (2)	-0,572 (31)	0,199 (86)	-0,523 (53)
C=O	Formaldeído	0,271 (2)	-0,462 (28)	0,439 (17)	-0,543 (35)
	Metanoato de metila	0,661 (3)	-0,540 (15)	0,532 (7)	-0,566 (22)
	Acetona	0,420 (4)	-0,510 (24)	0,386 (7)	-0,533 (31)
C-O	Metanol	0,000 (100)	-0,644 (20)	0,115 (5)	-0,591 (12)
	Dimetil éter	-0,009 (169)	-0,438 (11)	0,089 (6)	-0,404 (13)
N-O	Nitrometano	-1,598 (1849)	-0,208 (49)	1,046 (3)	-0,590 (13)

a-) geometrias otimizadas na fase condensada (PM3-SM5.4), as cargas foram calculadas pelo método de equalização da eletronegatividade.

b-) os valores entre parêntesis são as diferenças relativas entre as cargas calculadas no vácuo e no solvente: $\% \Delta q_I = 100(q_{Iw} - q_{Ig}) / q_{Ig}$ e $\% \Delta q_J = 100(q_{Jw} - q_{Jg}) / q_{Jg}$

Figura 6: Gráfico das cargas calculadas por GEEM-PD vs PM3-CM1 (constante dielétrica = 78,5 e geometrias otimizadas na fase condensada). A linha tracejada representa a identidade e a linha contínua representa a correlação linear entre os dados ($y = ax$, onde $a = 1,062 \pm 0,071$, $R = 0,941$, $SD = 0,165$, $F = 218$ e $N = 29$).

A distribuição de carga em uma molécula permite o cálculo dos componentes do momento de dipolo, bem como, obviamente, seu momento de dipolo total (módulo do vetor dipolo). Os resultados obtidos para momentos de dipolo calculados de EEM-PD e GBEEM-PD (Tabela 4) são comparáveis aos de PM3-CM1, um dos modelos mais utilizados no cálculo de dipolos, tanto no vácuo quanto em solvente. Uma forma bastante conveniente de se comparar os momentos de dipolo é através das diferenças relativas entre os dipolos calculados no vácuo e no solvente. Observando-se as diferenças relativas de ambos os modelos verifica-se que existe uma boa concordância entre os resultados (Tabela 4). Era esperado que o modelo GBEEM-PD fosse comparável ao modelo SM5.4 que usa cargas PM3-CM1 na previsão de momento de dipolo, tanto no vácuo como na água, devido a correlação anteriormente observada para as cargas entre os modelos. O gráfico entre as diferenças relativas dos dipolos GBEEM-PD versus PM3-CM1 (Figura 7) mostra que os maiores desvios pertencem as moléculas que possuem o átomo de nitrogênio na sua composição (ciclohexilamina, metil n-propil éter e nitrometano), sugerindo que as cargas do nitrogênio não estão sendo reproduzidas fielmente. Este é outro indicio de que a calibração QEq-PD, assim como a EEM-PD, com um único conjunto de parâmetros não descreve corretamente todos os tipos de ligações que o nitrogênio pode fazer. Portanto, uma nova calibração mais refinada talvez pudesse corrigir esta imperfeição, pois o conjunto de moléculas usadas na calibração do QEq-PD foi bastante restrito (cerca de 20 moléculas) e sem muitos representantes das diversas hibridizações do átomo de nitrogênio.

Tabela 4: Momentos de dipolo no vácuo e na água calculados pelos métodos GBEEM-PD e PM3-CM1 (valores em Debye).

Molécula	GBEEM-PD			PM3-CM1		
	μ_g	μ_w	$\sim \% \Delta \mu^a$	μ_g	μ_w	$\sim \% \Delta \mu^a$
Dimetil éter	1,45	1,81	25	1,22	1,61	32
Formaldeído	2,36	3,20	36	2,41	3,45	43
Acetona	2,77	4,19	51	2,92	4,50	54
Metanoato de metila	1,61	2,32	44	1,79	2,63	47
Acetonitrila	3,03	4,88	61	3,89	5,74	48
Metilamina	1,24	1,61	30	1,21	1,50	24
Nitrometano	4,23	6,76	60	3,70	5,19	40
Formamida	2,83	4,23	49	3,56	5,77	62
Água	1,87	2,29	22	1,92	2,27	18
Metanol	1,68	2,11	26	1,59	2,08	31
Ciclohexilamina	0,96	1,55	61	1,03	1,40	36
Metil n-propil éter	1,18	1,93	64	1,14	1,65	45
Etanal	2,77	4,21	52	2,70	4,01	49
Butanona	2,82	4,43	57	2,88	4,48	56

a-) $\% \Delta \mu = 100(\mu_w - \mu_g) / \mu_g$

Figura 7: Gráfico das diferenças relativas dos dipolos calculados pelo GBEEM-PD vs as diferenças relativas dos dipolos calculados pelo PM3-CM1 ($y = ax$, onde $a = 1,068 \pm 0,072$, $R = 0,663$, $SD = 12$, $F = 8$ e $N = 14$).

4.2 – Calibração do Método de Equalização da Eletronegatividade Baseada no Charge Model 1

Uma nova parametrização das durezas e eletronegatividades foi feita com o objetivo de melhorar a distribuição de cargas atômicas em moléculas isoladas em relação ao modelo CM1. Uma vez que as cargas CM1 calculadas em nível semi-empírico reproduzem adequadamente os momentos de dipolo *ab initio*, era esperado que o método de equalização reproduzisse corretamente as cargas CM1 e, conseqüentemente, os momentos de dipolo calculados pelo método de equalização também reproduziriam corretamente o momento de dipolo CM1 para qualquer molécula, independente do seu tamanho. Em outras palavras, a idéia era que a falha na previsão dos momentos de dipolo estaria relacionada com a calibração EEM-PD e que uma calibração mais detalhada corrigiria essa falha. Outra vantagem das cargas CM1 é a possibilidade de se utilizar um conjunto maior de moléculas na parametrização, pois o cálculo semi-empírico tem um reduzido custo computacional.

Os parâmetros atômicos eletronegatividade e dureza foram então otimizados utilizando-se um conjunto de 250 moléculas contendo somente átomos de C, H, O e N em diferentes funções orgânicas: alcanos, alcenos, alcinos, éteres, aldeídos, cetonas, ésteres, nitrilas, aminas, amidas, ácidos carboxílicos e álcoois. Essas moléculas foram otimizadas no vácuo em nível PM3 no programa AMSOL7.0 (Hawkins *et al.*, 2003), de onde foram obtidas as cargas CM1.

A função objetiva para a otimização dos parâmetros atômicos empregados no EEM foi definida como sendo a soma dos quadrados das diferenças entre cargas atômicas PM3-CM1 e cargas obtidas pelo método de equalização da eletronegatividade para um determinado conjunto de eletronegatividades e durezas ponderadas por elemento químico.

Os algoritmos Genético (GA) e Simplex foram combinados para pesquisar a hipersuperfície da função objetiva, pois o algoritmo genético não converge tão rápido quanto o Simplex para pontos estacionários locais, enquanto que o Simplex não é tão eficiente quanto o algoritmo genético na pesquisa de mínimos ou máximos globais.

O algoritmo genético procura o mínimo global em hipersuperfícies usando uma forma de algoritmo estocástico que imita o processo evolucionário, com a sobrevivência do indivíduo mais apto, troca de material genético, mutações entre outros (Goldberg, 1989), e trabalha sem precisar de derivadas da função a ser otimizada. Para a operação

do algoritmo genético é necessário estabelecer o número de indivíduos de uma população e número de gerações que esta população pode evoluir. Cada indivíduo na população é definido como um conjunto de parâmetros de eletronegatividades e durezas. Empiricamente, foi definido que uma população de 32 indivíduos era adequada para esta parametrização. O critério de convergência adotado na execução do algoritmo genético foi a minimização da função objetiva, ou seja, era considerada finalizada a otimização quando o erro não diminuía após um mínimo de 100 gerações.

A última geração do GA foi empregada como parâmetros iniciais na execução do programa Simplex. O Simplex (*Press et al., 1992*) é um algoritmo que explora as propriedades topológicas de espaços convexos e também não precisa de informação sobre derivadas da função. Assim como no GA é definido o número de gerações, no Simplex também é definido o número máximo de interações possíveis. O critério de convergência utilizado neste caso foi a minimização do erro, ou seja, novas interações foram feitas até que não houvesse diminuição do erro por pelo menos 50 interações. O conjunto de parâmetros finais obtidos da execução do Simplex era então transformado em um indivíduo e inserido em uma nova execução do GA, com o objetivo de verificar a auto-consistência do resultado.

Na Tabela 5 e na Tabela 6 são apresentados para efeito de comparação os conjuntos de parâmetros obtidos para este trabalho os usados em outras calibrações do EEM. Em geral, quase todas as calibrações de eletronegatividades apresentam o mesmo comportamento, exceto o EEM-M, uma vez que as eletronegatividades dos átomos N e O estão invertidas em relação às outras escalas. Esta inversão pode ser explicada observando-se a equação de χ_{eq} (Equação 14), pois esta equação mostra que deve-se empregar mais de um tipo de hibridização por elemento químico no conjunto de calibração para evitar a sobre-polarização, que acarretaria em um parâmetro sem significância real para o elemento químico em questão. No caso do EEM-M, o conjunto de calibração possuía apenas moléculas com um tipo de hibridização para o nitrogênio (*Bultinick et al., 2004*).

Tabela 5: Eletronegatividades atômicas usadas em diferentes calibrações do método de equalização.

Átomo	Eletronegatividade / kcal mol ⁻¹ e ⁻¹				
	EEM-M ^a	QEq-PD ^b	QEq-M ^b	EEM-PD ^c	EEM-CM1 ^d
H	101,67	101,98	89,33	101,98	89,58
C	131,00	116,99	119,38	116,99	107,36
N	244,43	178,42	174,05	178,42	156,45
O	196,02	190,92	205,55	190,92	231,93

a-) EEM-M foi calibrado usando cargas Mulliken no nível RHF/STO-3G (*Baekelandt et al., 1993*);

b-) QEq-PD e QEq-M são as versões EEM de Bakowies e Thiel calibradas usando modelos de cargas derivadas do potencial eletrostático e Mulliken no nível RHF/6-31G*, respectivamente (*Bakowies & Thiel, 1996*);

c-) EEM-PD tem o mesmo conjunto de eletronegatividades que QEq-PD (*Dias et al., 2002*);

d-) esta calibração.

Tabela 6: Durezas atômicas usadas em diferentes calibrações do EEM.

Átomo	Dureza / kcal mol ⁻¹ e ⁻²				
	EEM-M ^a	QEq-PD ^b	QEq-M ^b	EEM-PD ^c	EEM-CM1 ^d
H	317,62	319,17	319,17	338,32	424,40
C	208,71	232,09	258,89	246,02	328,93
N	304,09	299,08	325,87	317,02	309,59
O	255,58	344,35	371,15	365,01	477,13

a-) EEM-M foi calibrado usando cargas Mulliken no nível RHF/STO-3G (*Baekelandt et al., 1993*);

b-) QEq-PD e QEq-M são as versões EEM de Bakowies e Thiel calibradas usando modelos de cargas derivadas do potencial eletrostático e Mulliken no nível RHF/6-31G*, respectivamente (*Bakowies & Thiel, 1996*);

c-) Parâmetros iniciais (*Dias et al., 2002*). O EEM-PD tem durezas 6% maiores do que QEq-PD;

d-) esta calibração.

Para a escala EEM-M foram observadas inversões nas durezas dos átomos de N e O, assim como as eletronegatividades. A alta eletronegatividade do EEM-M para N induziu uma dureza com um valor elevado, pois altas durezas são necessárias para compensar altas eletronegatividades (*Bultinick et al., 2004*). Apesar da escala obtida neste estudo, EEM-CM1, ter sido feita utilizando-se um método diferente dos empregados em outras calibrações de eletronegatividades e durezas (EEM-M, QEq-PD, QEq-M e EEM-PD usaram RHF, enquanto que o EEM-CM1 usou o PM3), dos diferentes modelos de cargas empregados como alvo (EEM-M e QEq-M baseiam-se em cargas Mulliken, QEq-PD e EEM-PD baseiam-se em cargas derivadas de potencial eletrostático, enquanto que EEM-CM1 baseia-se em cargas CM1) e a ausência de carbono aromático no conjunto de calibração do EEM-CM1, que são elementos mais macios que carbonos não aromáticos, pois os sistemas π conjugados são facilmente polarizáveis, só foram observadas inversões nas durezas dos átomos de C e N em relação às demais escalas.

Um fator que foi sistematicamente desprezado em outros trabalhos de parametrização do EEM é a estimativa do desvio do valor médio para cada parâmetro atômico de eletronegatividade e dureza, usado no cálculo de cargas. O método utilizado na estimativa do desvio dos parâmetros obtidos neste trabalho foi variar individualmente o valor de cada parâmetro em cerca de 10% em torno do valor mínimo, mantendo os demais fixados. Após cada variação, as cargas foram recalculadas e o desvio para este novo conjunto de cargas foi determinado (Figura 8). A curva obtida pela variação de cada parâmetro foi ajustada por uma função do tipo $a(x-b)^2 + c$, onde b é o valor minimizado do parâmetro, c é o desvio associado com os parâmetros obtidos na calibração do EEM e a é o termo que determinará o desvio do parâmetro. A partir do termo a tem-se que:

$$desvio = \frac{1}{\sqrt{2a}} \quad (41)$$

Figura 8: Gráficos dos desvios calculados para eletronegatividades (●) e durezas (■) dos átomos de carbono e hidrogênio.

Na Tabela 7 são mostrados os parâmetros obtidos pela calibração EEM-CM1 e os respectivos desvios. Nesta tabela também são listados os parâmetros da calibração QEq-M e os desvios associados, calculados pela metodologia descrita acima. Pode-se notar que os desvios calculados por ambas as parametrizações são semelhantes, sendo que os desvios relacionados às durezas são maiores do que os desvios das eletronegatividades. Este fato sugere que as cargas atômicas são mais sensíveis às variações na eletronegatividade.

O cálculo de cargas parece não ser influenciado pela dureza do átomo de hidrogênio, pois na calibração QEq-M o parâmetro obtido não está no mínimo e o valor estimado como mínimo gera um desvio na mesma ordem de grandeza do parâmetro. No caso do EEM-CM1 além da dureza do hidrogênio, a dureza do carbono também apresenta um desvio na mesma ordem de grandeza do parâmetro otimizado. Este fato não é inesperado, uma vez que as cargas dos átomos hidrogênio e carbono são muito pequenas (próximas de zero) e, conseqüentemente, os desvios associados a estes átomos serão grandes. O valor do desvio relativo da dureza do carbono por QEq-PD é pequeno, o que é surpreendente, e possivelmente este resultado foi obtido para um grupo específico de moléculas. Como há uma grande incerteza associada às durezas (e portanto uma insensibilidade no cálculo das cargas), não é possível uma discussão precisa sobre as divergências nas escalas de durezas obtidas pelas várias calibrações presentes na literatura.

A análise das correlações entre as cargas EEM calculadas com a nova parametrização EEM-CM1 e as cargas calculadas pelo EEM-PD (Figura 9) mostra que, quando o conjunto utilizado na parametrização inicial era formado por 12 moléculas, a correção das durezas EEM-PD pareceu suficiente e as cargas calculadas correlacionaram bem com as cargas PM3-CM1, pois este conjunto de moléculas era relativamente restrito. Entretanto, com o emprego de um conjunto de 250 moléculas na parametrização EEM-CM1, fica evidente no gráfico que a nova parametrização reproduz melhor as cargas PM3-CM1, tanto que a regressão linear entre as cargas EEM-CM1 e PM3-CM1 e a linha identidade ($f(x) = x$) são muito próximas e há pouca dispersão dos pontos em torno da linha identidade.

Tabela 7: Eletronegatividades (χ) e durezas (η) atômicas com os desvios relativos (σ_{rel}) estimados para cada parâmetro.

Átomo	χ / kcal.mol ⁻¹ .e ⁻¹	% σ_{rel} ^a	η / kcal.mol ⁻¹ .e ⁻²	% σ_{rel} ^b
EEM-CM1				
H	90	41	424	108
C	107	39	329	110
N	156	38	310	42
O	232	48	477	53
QEq-M				
H	89	24	319	*
C	119	21	259	24
N	174	39	326	48
O	206	36	371	43

a) % $\sigma_{\text{rel}} = 100(\sigma/\chi)$

b) % $\sigma_{\text{rel}} = 100(\sigma/\eta)$

* Este parâmetro não está no mínimo, por isso não foi possível estimar o desvio. O mínimo foi estimado em 274 com um desvio relativo estimado em 90%.

Figura 9: Cargas EEM vs cargas CM1. As linhas tracejada e sólida são ajustes lineares às cargas calculadas pelo EEM-PD ($y = ax + b$, $a = 1,30$, $b = -6,1 \times 10^{-7}$, $R = 0,931$, $SD = 0,079$, $F = 31627$, $N = 4903$) e EEM-CM1 ($y = ax + b$, $a = 0,94$, $b = -5,0 \times 10^{-7}$, $R = 0,977$, $SD = 0,032$, $F = 102977$, $N = 4903$). A linha pontilhada é a linha identidade.

Os gráficos dos resíduos, que são a diferença entre as cargas EEM-CM1 e PM3-CM1, são apresentados na Figura 10 com diferentes símbolos para ajudar na identificação dos grupos funcionais. No geral, os resíduos estão distribuídos de maneira aleatória, mas para os alcenos e alcinos é possível observar que existe uma tendência de inversão no sinal das cargas dos átomos de carbono. Resíduos maiores do que 0,15 (valor absoluto) aparecem nas seguintes funções orgânicas: ésteres, aldeídos e nitrilas. Os maiores resíduos em ésteres devem-se ao átomo de oxigênio pois, assim como no caso do nitrometano, apenas um valor de eletronegatividade e um valor de dureza química estão sendo usados para descrever tanto o oxigênio acila quanto o alquila. Já nos aldeídos são os átomos de carbonos carbonílicos que apresentam os maiores resíduos, enquanto que nas nitrilas são os C_{α} e $C(sp)$, respectivamente, que apresentam os maiores resíduos.

Como forma de se avaliar a robustez da parametrização EEM-CM1, as cargas atômicas parciais no vácuo de duas moléculas relevantes biologicamente, um tetrapeptídeo (Glicina-Alanina-Serina-Alanina) e um açúcar, foram calculadas e os resultados obtidos foram comparados com PM3-CM1. Essas moléculas possuem mais de um grupo funcional (amida, álcool e ácido carboxílico no caso do peptídeo; álcool e acetal no caso do açúcar) e não foram incluídas no grupo de moléculas usadas na otimização, assim como também não foram usadas moléculas com mais de um grupo funcional. As geometrias moleculares foram otimizadas no vácuo no nível PM3 antes de calcular as cargas atômicas. O resultado das regressões entre as cargas EEM-CM1 e as cargas PM3-CM1 para as duas moléculas em questão (Figura 11) mostra uma boa concordância, indicando que os parâmetros obtidos em nosso trabalho são válidos e transferíveis para moléculas que não participaram da otimização (*Menegon et al., 2002*).

Na parametrização EEM-PD foi verificado que as eletronegatividades e durezas atômicas otimizadas para o cálculo de cargas parciais no vácuo também eram aplicáveis ao cálculo da redistribuição de cargas em fase condensada. Para verificar se esta hipótese também é válida com relação a parametrização EEM-CM1, as cargas atômicas de um polipeptídeo formado 15 unidades do aminoácido alanina foram calculadas no vácuo e em água. Os resultados mostram que existe uma boa correspondência entre as cargas PM3-CM1 e as cargas calculadas pelo EEM-CM1, assim como entre as cargas GBEEM-CM1 e as cargas PM3-CM1 calculadas com efeito de solvente (Figura 12). Isto indica que os parâmetros eletronegatividade e dureza atômicas obtidos pela

parametrização EEM-CM1 são transferíveis para o cálculo de cargas na presença de solvente com o modelo GBEEM-CM1.

Apesar da melhoria na distribuição de cargas com esta nova parametrização, em relação a EEM-PD, o cálculo dos momentos de dipolo para este conjunto de 250 moléculas mostrou que a nova parametrização feita apresenta um desvio elevado dos dipolos no vácuo. Com relação ao solvente, os momentos de dipolo calculados com estes parâmetros também apresentam dipolos maiores quando comparados com dipolos calculados por outros modelos de cargas ou ainda com o valor esperado do operador momento de dipolo. Os resultados obtidos até aqui para os dipolos sugerem que mais estudos deveriam ser feitos, pois independente da parametrização empregada, do modelo de carga e da fase (gasosa ou condensada), o momento de dipolo sempre aumentava desproporcionalmente com o incremento da cadeia carbônica, sendo, portanto, um indício de que poderia existir um problema intrínseco no EEM.

Figura 10: Resíduo vs cargas CM1. As funções orgânicas são apresentadas com diferentes cores/símbolos. Os resíduos foram calculados utilizando-se a regressão robusta LTS (Least Trimmed Squares). Nesta regressão o ajuste é feito minimizando-se a soma dos quadrados do conjunto formado pelos menores resíduos (cerca de 80% do total).

Figura 11: Cargas EEM-CM1 para o açúcar e o tetrapeptídeo. As linhas sólidas são ajustes lineares para cargas EEM-CM1 do açúcar ($y=ax+b$, $a=1,03$, $b=3,3 \times 10^{-17}$, $R=0,992$, $SD=0,038$, $N=66$) e o tetrapeptídeo ($y=ax+b$, $a=0,96$, $b=-2,8 \times 10^{-6}$, $R=0,991$, $SD=0,042$, $N=41$) em função das cargas PM3-CM1. A linha pontilhada é a linha identidade.

4.3 – O método de equalização e o cálculo de momentos de dipolo

Um problema intrínseco ao método de equalização da eletronegatividade ressurgiu no estudo de momentos de dipolo em moléculas pertencentes a diferentes funções orgânicas. Na Figura 13 estão apresentados os momentos de dipolo calculados a partir de cargas PM3-CM1 e momentos de dipolo calculados a partir de cargas EEM-CM1. Esta figura é muito interessante, pois se verifica como o momento de dipolo aumenta com o aumento da cadeia carbônica. Outro fato que pode ser observado a partir dessa figura é que alguns grupos funcionais são mais afetados do que outros, por exemplo, os grupos aldeídos, ácidos carboxílicos e nitro são os que apresentam maiores desvios em relação ao momento de dipolo PM3-CM1. É conhecido que o EEM sobreestima a polarizabilidade com o crescimento da cadeia molecular devido a uma exagerada delocalização das cargas (como se as moléculas fossem “condutores metálicos”) (*Chelli et al., 1999*).

Uma simples aproximação, baseada no trabalho de Chelli e colaboradores (*Chelli et al., 1999*), foi testada no cálculo dos momentos de dipolo: a soma parcial das cargas pertencentes a grupos atômicos predefinidos, por exemplo, grupos funcionais, é mantida constante e não é permitida a transferência de cargas entre os grupos.

As séries dos aldeídos, ácidos carboxílicos e nitro foram usadas para testar a aproximação de cargas predefinidas (constrained charge approximation, CCA). Cada molécula foi dividida em duas partes: o grupo funcional (-COH, -COOH e -NO₂) e o grupo cadeia (cadeia alquila), e para ambos os grupos a soma das cargas foi mantida constante em zero (note que nenhuma nova parametrização foi feita). Os momentos de dipolo calculados com essa aproximação estão apresentados na Figura 14, onde é evidente a redução da sobrepolarização do EEM (*Shimizu et al., 2004*).

Então, a aproximação foi testada em moléculas maiores (e.g., polipeptídeos). Estes testes serviram também para verificar a independência da aproximação de cargas predefinidas ao conjunto de parâmetros, utilizando durezas e eletronegatividades atômicas de outras calibrações encontradas na literatura: os QEq-M e QEq-PD, parametrizações feitas por Bakowies e Thiel (*Bakowies & Thiel, 1996*), baseadas em cargas Mulliken e cargas derivadas de potencial eletrostático, respectivamente, calculadas em nível RHF/6-31G* e o EEM-M, parametrização feita por Mortier e colaboradores (*Mortier et al., 1986*), baseada em cargas Mulliken calculadas em nível RHF/STO-3G. Os valores dos parâmetros estão apresentados nas Tabela 5 e Tabela 6.

Os momentos de dipolo calculados usando a aproximação de cargas predefinidas estão representados por um “*” nas siglas de cada parametrização EEM, por exemplo, EEM*-CM1.

Os polipeptídeos utilizados nos cálculos foram um heptapeptídeo composto por 7 unidades do aminoácido alanina, (Ala)₇, e dois pentadecapeptídeos, o primeiro formado por 15 unidades de alanina, (Ala)₁₅, e o segundo formado por 5 repetições de uma sequência de glicina, alanina e serina, (GAS)₅. Na aproximação de cargas predefinidas, a soma das cargas atômicas em cada aminoácido é igual a zero.

Cálculos semi-empíricos em nível PM3 foram feitos para os três polipeptídeos com AMSOL7.0 (*Hawkins et al., 2003*) e em níveis *ab initio* B3LYP/6-31G(d) e B3LYP/6-311G(d,p) com GAUSSIAN03 (*Frisch et al., 2003*)

Figura 13: Momento de dipolo EEM-CM1 vs momento de dipolo PM3-CM1. As funções orgânicas estão apresentadas com diferentes cores e símbolos. A linha contínua é a identidade.

Figura 14: Efeito da aproximação de cargas predefinidas nas séries de aldeídos, ácidos carboxílicos e nitro compostos. A linha contínua é a linha identidade.

Na Tabela 8 estão apresentados os valores dos momentos de dipolos calculados para o heptapeptídeo no vácuo, com as diferentes parametrizações EEM, com e sem a aproximação das cargas predefinidas e tendo valores *ab initio* como referência para comparar a precisão dos métodos. Em geral, todos os modelos EEM* apresentam momentos de dipolo mais “precisos” do que os seus correspondentes modelos EEM, sendo que as melhores previsões entre os métodos semi-empíricos foram conseguidas com os EEM*-M, QEq*-M e PM3-CM1, nesta ordem. Apesar do EEM*-M apresentar a melhor concordância com o valor *ab initio*, os momentos de dipolo calculados para cada aminoácido (~6D) é maior do que os valores das outras parametrizações EEM (Figura 15), do OPLS-AA e do PM3-CM1, e são sobreestimados em solvente (Figura 16). Por isso, a parametrização EEM-M foi excluída dos próximos testes.

As Tabela 9, Tabela 10, Tabela 11 e Tabela 12 apresentam os momentos de dipolo calculados para os pentadecapeptídeos, (Ala)₁₅ e (GAS)₅ (Figura 17 e Figura 18), em duas conformações diferentes para cada polipeptídeo.

Assim como para o heptapeptídeo, todos os cálculos feitos com os modelos EEM* usando a aproximação de cargas predefinidas são mais precisos do que os modelos EEM correspondentes sem a aproximação, tendo os valores *ab initio* como referência. Para os pentadecapeptídeos, os modelos semi-empíricos que tiveram maior concordância com os valores *ab initio* foram o PM3-CM1 e QEq*-M, produzindo inclusive valores similares.

Com os bons resultados obtidos no vácuo, os momentos de dipolo no solvente foram calculados utilizando a combinação do EEM com o modelo de Born generalizado (Dias *et al.*, 2002).

Cálculos *single-point* em dielétrico contínuo foram feitos com AMSOL7.0, em nível PM3, usando o modelo de solvatação SM5.4P (Chambers *et al.*, 1996) e, em nível *ab initio* B3LYP/6-31G*, usando o *Integral Equation Formalism for the Polarizable Continuum Model* (Mennucci *et al.*, 1998), IEFPCM, implementado em GAUSSIAN03.

Figura 15: Histograma dos dipolos calculados no vácuo para cada aminoácido pelos diferentes métodos.

Figura 16: Histograma dos dipolos calculados na fase condensada para cada aminoácido pelos diferentes métodos.

Tabela 8: Módulo e componentes do vetor momento de dipolo para o polipeptídeo (Ala)₇ na conformação α -hélice no vácuo.

Modelo	Momento de Dipolo (valores em Debye)				$\sim\% \Delta \mu^a$
	μ	μ_x	μ_y	μ_z	
EEM-CM1	11,45	-11,11	0,45	2,77	55
EEM-M	19,51	-18,96	0,53	4,59	23
QEq-M	13,02	-12,65	0,01	3,05	49
QEq-PD	10,19	-9,93	0,01	2,27	60
EEM*-CM1	17,92	-17,34	0,72	4,45	30
EEM*-M	26,50	-24,74	0,37	9,48	4
QEq*-M	23,67	-23,08	0,30	5,25	7
QEq*-PD	14,84	-14,05	0,19	4,77	42
OPLS-AA	22,22	-21,76	0,54	4,50	13
PM3-CM1	27,64	-26,89	0,59	6,36	9
B3LYP/6-31G(d)	25,35	-24,87	0,36	4,86	0,5
B3LYP/6-311G(d,p)	25,47	-24,99	0,33	4,90	0

* utilização da aproximação de cargas predefinidas nos métodos de equalização da eletronegatividade.

a-) $\% \Delta \mu = 100 |\mu_{B3LYP/6-311G(d,p)} - \mu_{model}| / \mu_{B3LYP/6-311G(d,p)}$

Tabela 9: Módulo e componentes do vetor momento de dipolo para o polipeptídeo (GAS)₅ na conformação α -hélice no vácuo.

Modelo	Momento de Dipolo (valores em Debye)				
	μ	μ_x	μ_y	μ_z	$\sim\% \Delta\mu^a$
EEM-CM1	20,81	20,69	-1,56	-1,64	72
QEq-M	23,61	23,50	-1,64	-1,56	68
QEq-PD	18,83	18,75	-1,32	-1,14	75
EEM*-CM1	49,78	49,72	-0,92	-2,15	33
QEq*-M	69,64	69,59	-0,72	-2,51	6
QEq*-PD	45,52	45,46	-0,38	-2,28	38
OPLS-AA	61,47	61,40	-2,15	-2,04	17
PM3-CM1	74,84	74,77	-1,79	-2,77	1
B3LYP/6-31G(d)	72,72	72,68	-2,05	-1,64	1,5
B3LYP/6-311G(d,p)	73,85	73,81	-2,09	-1,72	0

* utilização da aproximação de cargas predefinidas nos métodos de equalização da eletronegatividade.

a-) $\% \Delta\mu = 100|\mu_{B3LYP/6-311G(d,p)} - \mu_{model}| / \mu_{B3LYP/6-311G(d,p)}$

Tabela 10: Módulo e componentes do vetor momento de dipolo para o polipeptídeo (GAS)₅ no vácuo.

Modelo	Momento de Dipolo (valores em Debye)				
	μ	μ_x	μ_y	μ_z	$\sim\% \Delta \mu^a$
EEM-CM1	15,68	15,51	1,20	-2,00	66
QEq-M	16,40	16,16	1,64	-2,27	64
QEq-PD	12,80	12,58	1,46	-1,85	72
EEM*-CM1	31,67	31,62	1,14	-1,48	31
QEq*-M	42,10	42,04	1,71	-1,59	8
QEq*-PD	25,43	25,39	0,99	-1,04	44
OPLS-AA	38,72	38,54	2,31	-2,91	15
PM3-CM1	46,63	46,45	2,57	-3,22	2
B3LYP/6-31G(d)	45,06	44,86	1,90	-3,80	1
B3LYP/6-311G(d,p)	45,60	45,40	1,98	-3,85	0

* utilização da aproximação de cargas predefinidas nos métodos de equalização da eletronegatividade.

a-) $\% \Delta \mu = 100 |\mu_{B3LYP/6-311G(d,p)} - \mu_{model}| / \mu_{B3LYP/6-311G(d,p)}$

Tabela 11: Módulo e componentes do vetor momento de dipolo para o polipeptídeo (Ala)₁₅ na conformação α -hélice no vácuo.

Modelo	Momento de Dipolo (valores em Debye)				
	μ	μ_x	μ_y	μ_z	$\sim\% \Delta\mu^a$
EEM-CM1	17,87	17,84	0,13	1,02	73
QEq-M	19,87	19,84	0,11	1,02	70
QEq-PD	15,43	15,41	0,04	0,68	77
EEM*-CM1	46,14	46,11	0,00	1,73	30
QEq*-M	65,23	65,20	-0,14	1,90	2
QEq*-PD	41,84	41,80	-0,01	1,91	37
OPLS-AA	49,58	49,57	-0,49	1,01	25
PM3-CM1	67,51	67,47	-0,65	2,41	2
B3LYP/6-31G(d)	65,63	65,62	-0,34	1,16	1
B3LYP/6-311G(d,p)	66,23	66,22	-0,35	1,22	0

* utilização da aproximação de cargas predefinidas nos métodos de equalização da eletronegatividade.

a-) $\% \Delta\mu = 100|\mu_{B3LYP/6-311G(d,p)} - \mu_{model}| / \mu_{B3LYP/6-311G(d,p)}$.

Tabela 12: Módulo e componentes do vetor momento de dipolo para o polipeptídeo (Ala)₁₅ no vácuo.

Modelo	Momento de Dipolo (valores em Debye)				
	μ	μ_x	μ_y	μ_z	$\sim\% \Delta\mu^a$
EEM-CM1	9,72	8,65	0,18	4,43	69
QEq-M	10,69	9,49	0,59	4,89	65
QEq-PD	8,24	7,22	0,57	3,92	73
EEM*-CM1	20,66	19,72	1,62	5,96	33
QEq*-M	29,00	28,07	2,30	6,95	6
QEq*-PD	16,88	15,83	1,42	5,70	45
OPLS-AA	22,96	21,57	2,14	7,58	26
PM3-CM1	31,85	30,28	1,33	9,77	3
B3LYP/6-31G(d)	30,55	29,43	1,70	8,02	1
B3LYP/6-311G(d,p)	30,89	29,78	1,83	8,01	0

* utilização da aproximação de cargas predefinidas nos métodos de equalização da eletronegatividade.

a-) $\% \Delta\mu = 100|\mu_{B3LYP/6-311G(d,p)} - \mu_{model}| / \mu_{B3LYP/6-311G(d,p)}$.

a-)

b-)

Figura 17: Representações em estéreo das estruturas do polipeptídeo (GAS)₅, sendo a) na conformação α -hélice e b) uma conformação torcida.

a-)

b-)

Figura 18: Representações em estéreo das estruturas do polipeptídeo (Ala)₁₅, sendo a) na conformação α -hélice e b) uma conformação torcida.

Na Tabela 13 estão apresentados os momentos de dipolo obtidos para o heptapeptídeo no solvente, com as diferentes parametrizações EEM, com e sem a aproximação de cargas predefinidas. Assim como no vácuo, os modelos EEM* com a aproximação reduziram a sobrepolarização do EEM e conseqüentemente o dipolo. O modelo QEq*-M segue de perto o modelo PM3-CM1 de Cramer e Truhlar (*Cramer & Truhlar, 1999*), nas previsões dos vetores de momento de dipolo.

As tabelas abaixo (Tabela 14, Tabela 15, Tabela 16 e Tabela 17) apresentam os momentos de dipolo no solvente para os pentadecapeptídeos em duas conformações diferentes. O modelo PM3-CM1 possui a melhor concordância com o *ab initio* IEFPCM em nível B3LYP/6-31G(d). O modelo QEq*-M apresenta uma polarização maior quando comparado com os modelos PM3-CM1 e IEFPCM, porém não foi feita nenhuma nova parametrização ou ajuste do tamanho da cavidade molecular que afeta sensivelmente o campo de reação do solvente sobre o soluto (*Curutchet et al., 2003; Barone et al., 1997; Nina et al., 1997; Tomasi & Persico, 1994*). Interessantemente, o modelo QEq-PD produziu bons resultados para os dipolos no solvente, fato não era esperado, uma vez que no vácuo este modelo não produz bons resultados. Este fato pode ter ocorrido, dentre outros motivos, pelo número limitado de moléculas empregadas, portanto um estudo mais detalhado seria necessário para se compreender os bons resultados obtidos pelo QEq-PD, porém este estudo foge ao objetivo deste trabalho.

Tabela 13: Módulo e componentes do vetor momento de dipolo para o polipeptídeo (Ala)₇ na conformação α -hélice em dielétrico contínuo ($\epsilon=78,5$).

Modelo	Momento de Dipolo (valores em Debye)				
	μ	μ_x	μ_y	μ_z	$\sim\% \Delta \mu^a$
EEM-CM1	29,97	-29,38	1,11	5,80	13
EEM-M	92,62	-91,29	2,40	15,43	169
QEq-M	45,40	-44,66	0,17	8,16	32
QEq-PD	36,88	-36,30	0,20	6,54	7
EEM*-CM1	26,94	-25,76	1,38	7,75	22
EEM*-M	60,30	-55,56	2,25	23,33	75
QEq*-M	40,79	-39,18	1,10	11,28	18
QEq*-PD	27,72	-25,81	0,98	10,06	20
PM3-CM1	38,66	-37,14	1,15	10,69	12
B3LYP/6-31G(d)	34,44	-33,44	0,96	8,16	0

* utilização da aproximação de cargas predefinidas nos métodos de equalização da eletronegatividade.

a-) $\% \Delta \mu = 100 |\mu_{B3LYP/6-31G(d)} - \mu_{model}| / \mu_{B3LYP/6-31G(d)}$.

Tabela 14: Módulo e componentes do vetor momento de dipolo para o polipeptídeo (GAS)₅ na conformação α -hélice em dielétrico contínuo ($\epsilon=78,5$).

Modelo	Momento de Dipolo (valores em Debye)				
	μ	μ_x	μ_y	μ_z	$\sim\% \Delta \mu^a$
EEM-CM1	66,04	65,85	-3,80	-3,26	27
QEq-M	101,21	100,97	-5,67	-4,03	13
QEq-PD	84,30	84,11	-4,42	-3,54	6
EEM*-CM1	63,75	63,60	-2,15	-3,79	29
QEq*-M	97,50	97,34	-2,29	-5,19	8
QEq*-PD	66,95	66,78	-1,51	-4,59	26
PM3-CM1	90,91	90,77	-2,21	-4,62	1
B3LYP/6-31G(d)	89,93	89,81	-2,54	-3,75	0

* utilização da aproximação de cargas predefinidas nos métodos de equalização da eletronegatividade.

a-) $\% \Delta \mu = 100 |\mu_{B3LYP/6-31G(d)} - \mu_{model}| / \mu_{B3LYP/6-31G(d)}$.

Tabela 15: Módulo e componentes do vetor momento de dipolo para o polipeptídeo (GAS)₅ em dielétrico contínuo ($\epsilon=78,5$).

Modelo	Momento de Dipolo (valores em Debye)				
	μ	μ_x	μ_y	μ_z	$\sim\% \Delta \mu^a$
EEM-CM1	48,87	48,55	2,72	-4,91	14
QEq-M	70,21	69,63	4,81	-7,67	24
QEq-PD	57,57	57,11	3,83	-6,15	2
EEM*-CM1	42,66	42,53	2,18	-2,62	25
QEq*-M	61,52	61,35	3,39	-3,06	9
QEq*-PD	40,71	40,61	2,14	-2,05	28
PM3-CM1	57,26	56,98	4,38	-3,58	1
B3LYP/6-31G(d)	56,69	56,45	3,35	-4,03	0

* utilização da aproximação de cargas predefinidas nos métodos de equalização da eletronegatividade.

a-) $\% \Delta \mu = 100 |\mu_{B3LYP/6-31G(d)} - \mu_{model}| / \mu_{B3LYP/6-31G(d)}$.

Tabela 16: Módulo e componentes do vetor momento de dipolo para o polipeptídeo (Ala)₁₅ na conformação α -hélice em dielétrico contínuo ($\epsilon=78,5$).

Modelo	Momento de Dipolo (valores em Debye)				
	μ	μ_x	μ_y	μ_z	$\sim\% \Delta \mu^a$
EEM-CM1	61,87	61,84	-0,18	2,18	22
QEq-M	94,38	94,33	-0,28	3,00	19
QEq-PD	76,83	76,79	-0,44	2,33	3
EEM*-CM1	59,67	59,57	0,10	3,44	25
QEq*-M	91,61	91,47	-0,61	5,06	15
QEq*-PD	61,66	61,48	-0,60	4,64	22
PM3-CM1	81,91	81,77	-1,35	4,55	3
B3LYP/6-31G(d)	79,44	79,39	-0,70	2,80	0

* utilização da aproximação de cargas predefinidas nos métodos de equalização da eletronegatividade.

a-) $\% \Delta \mu = 100 |\mu_{B3LYP/6-31G(d)} - \mu_{model}| / \mu_{B3LYP/6-31G(d)}$.

Tabela 17: Módulo e componentes do vetor momento de dipolo para o polipeptídeo (Ala)₁₅ em dielétrico contínuo ($\epsilon=78,5$).

Modelo	Momento de Dipolo (valores em Debye)				
	μ	μ_x	μ_y	μ_z	$\sim\% \Delta \mu^a$
EEM-CM1	32,70	31,18	1,64	9,71	15
QEq-M	50,39	48,30	3,45	13,96	31
QEq-PD	41,42	39,64	2,76	11,68	8
EEM*-CM1	28,52	26,49	2,12	10,35	26
QEq*-M	43,80	41,05	2,73	15,04	14
QEq*-PD	28,96	26,02	1,47	12,62	25
PM3-CM1	40,77	37,12	2,15	16,73	6
B3LYP/6-31G(d)	38,36	35,78	2,45	13,61	0

* utilização da aproximação de cargas predefinidas nos métodos de equalização da eletronegatividade.

a-) $\% \Delta \mu = 100 |\mu_{B3LYP/6-31G(d)} - \mu_{model}| / \mu_{B3LYP/6-31G(d)}$.

4.4 – Cálculo da energia livre de solvatação

Na calibração da energia livre de solvatação foram usadas 62 moléculas contendo átomos de C, H, O e N em diferentes funções orgânicas: alcanos, alcenos, alcinos, éteres, aldeídos, cetonas, ésteres, aminas, ácidos carboxílicos e álcoois. Essas moléculas foram otimizadas no vácuo em nível HF/6-31G(d) com GAUSSIAN03. Posteriormente, cálculos *single-point* na presença de dielétrico contínuo foram realizados para essas moléculas, em nível HF/6-31G(d) usando o *Integral Equation Formalism for the Polarizable Continuum Model* (Mennucci et al., 1998), IEFPCM com raio UAHF e, em nível PM3 para o modelo SM5.42R (Chambers et al., 1996).

O momento de dipolo para este conjunto de moléculas foi calculado no vácuo e em solvente contínuo, usando raio de Bondi para manter a relação com o GBEEM, em nível B3LYP/6-311G(d,p). A forma como o EEM foi implementado neste trabalho ainda não permite a otimização de geometria uma vez que não existe um campo de força associado ao modelo de carga flutuante. Portanto, foram escolhidas moléculas com pouca liberdade conformacional e para o cálculo das propriedades moleculares foram utilizadas geometrias no vácuo.

As Figura 19 e Figura 20 apresentam os momentos de dipolo no vácuo e em dielétrico contínuo calculados pelo EEM/GBEEM usando os parâmetros eletronegatividade e dureza atômicas da calibração EEM-CM1, tendo os momentos de dipolo calculados em nível B3LYP/6-311G(d,p) no vácuo e IEFPCM-B3LYP/6-311G(d,p) em dielétrico. As regressões lineares $\mu_{\text{EEM/GBEEM}} = a\mu_{\text{B3LYP/6-311G(d,p)}}$ calculadas para ambas as fases, vácuo e solvente, atestam a precisão do modelo EEM/GBEEM. No vácuo, os valores estatísticos da regressão linear obtidos são $a = 0,96 \pm 0,03$, $r = 0,936$, $\text{rmsd} = 0,30$ e $F = 427$, já no solvente os valores estatísticos são $a = 1,03 \pm 0,03$, $r = 0,951$, $\text{rmsd} = 0,43$ e $F = 581$.

Figura 19: Momentos de dipolo calculados pelo EEM-CM1 versus momentos de dipolo calculados em nível B3LYP/6-311G(d,p) no vácuo. A linha segmentada é a linha identidade.

Figura 20: Momentos de dipolo calculados pelo GBEEM-CM1 versus momentos de dipolo calculados em nível IEFPCM-B3LYP/6-311G(d,p) na presença de solvente, usando constante dielétrica de 78,5 e raio atômico de Bondi. A linha segmentada é a linha identidade.

As piores previsões dos momentos de dipolo usando o EEM e o GBEEM foram corrigidos usando a aproximação de cargas predefinidas (constrained charge approximation, CCA) descrita anteriormente (Shimizu *et al.*, 2004). No caso dos momentos de dipolo calculados pelo EEM as moléculas corrigidas são: m-cresol ($\mu_{ab\ initio} = 1,04$ D, $\mu_{EEM} = 1,77$ D e $\mu_{EEM-CCA} = 1,29$ D), 2,6-dimetilpiridina ($\mu_{ab\ initio} = 1,43$ D, $\mu_{EEM} = 0,58$ D e $\mu_{EEM-CCA} = 1,36$ D), 2,5-dimetilpiridina ($\mu_{ab\ initio} = 2,00$ D, $\mu_{EEM} = 1,28$ D e $\mu_{EEM-CCA} = 1,53$ D), etanoato de metila ($\mu_{ab\ initio} = 2,06$ D, $\mu_{EEM} = 3,04$ D e $\mu_{EEM-CCA} = 1,16$ D) e 2-propen-1-ol ($\mu_{ab\ initio} = 1,84$ D, $\mu_{EEM} = 1,06$ D e $\mu_{EEM-CCA} = 1,83$ D).

As moléculas corrigidas com o CCA no cálculo dos momentos de dipolo pelo GBEEM são: anilina ($\mu_{ab\ initio} = 2,24$ D, $\mu_{GEEM} = 1,16$ D e $\mu_{GEEM-CCA} = 2,52$ D), 2,6-dimetilpiridina ($\mu_{ab\ initio} = 2,41$ D, $\mu_{GEEM} = 0,89$ D e $\mu_{GEEM-CCA} = 2,86$ D), 2,5-dimetilpiridina ($\mu_{ab\ initio} = 3,18$ D, $\mu_{GEEM} = 2,49$ D e $\mu_{GEEM-CCA} = 2,80$ D), etanoato de metila ($\mu_{ab\ initio} = 2,62$ D, $\mu_{GEEM} = 6,11$ D e $\mu_{GEEM-CCA} = 1,97$ D) e 2-propen-1-ol ($\mu_{ab\ initio} = 2,69$ D, $\mu_{GEEM} = 1,35$ D e $\mu_{GEEM-CCA} = 2,92$ D).

A previsão correta do momento de dipolo é importante, pois a contribuição polar da energia livre de solvatação depende do momento de n-polo no solvente. Portanto, espera-se que os primeiros momentos de n-polo no solvente sejam preponderantes na contribuição polar, e geralmente, para moléculas polares neutras, a maior contribuição é devida ao momento de dipolo permanente. Uma vez que os momentos de dipolo, no vácuo e no solvente, calculados pelo EEM/GBEEM estão próximos aos momentos de dipolo *ab initio*, prosseguiu-se com o cálculo da energia livre de solvatação.

Assumindo a energia livre de solvatação como uma soma das contribuições eletrostáticas, de cavitação e de dispersão e que existem várias possibilidades de se calcular a energia de cavitação e de dispersão, alguns modelos foram testados antes de se concluir que o modelo que melhor descreve ΔG_{solv} é o representado pela combinação da parte eletrostática de GB com a parte não polar envolvendo a SAS e U_{disp} particionado por elemento químico.

$$\Delta G_{solv} = c_0 + c_1 \Delta E_{EN} + c_2 \Delta G_{el} + c_3 SAS + c_H U_{disp}(H) + c_C U_{disp}(C) + c_O U_{disp}(O) + c_N U_{disp}(N) \quad (42)$$

É importante ressaltar que em todos os modelos propostos variou-se apenas a parte não polar da energia de solvatação, ou seja, a parte eletrostática calculada pelo

GBEEM foi sempre mantida constante. A primeira abordagem, GBEEM/NP1, consiste em utilizar U_{disp} como um único valor e a SAS.

$$\Delta G_{solv} = \Delta E_{EN} + \Delta G_{el} + a + bSAS + c U_{disp} \quad (43)$$

Os parâmetros C_{iw} utilizados no cálculo da energia de dispersão, assim como os parâmetros a , b e c foram otimizados usando uma combinação de algoritmo genético (Goldberg, 1989) e Simplex (Press et al., 1992). Energias livres de solvatação experimentais foram usadas como alvo e uma função objetiva definida como a soma dos quadrados das diferenças entre os valores experimentais e preditos foi minimizada. Os parâmetros obtidos são apresentados na Tabela 18.

Um ponto negativo na abordagem GBEEM/NP1 é o uso de C_{iw} dependente da hibridização dos átomos, deixando o modelo muito específico e gerando um número elevado de parâmetros. Portanto, uma nova metodologia, denominada de GBEEM/NP2, foi proposta mantendo-se a estrutura do GBEEM/NP1, porém utilizando-se C_{iw} retirados do OPLS-AA, sendo apenas um valor para cada tipo atômico. O resultado obtido foi muito pouco diferente do anterior com muitos valores para C_{iw} , Figura 21, mostrando que a simplificação é razoável.

Como os resultados foram similares, Optou-se por modificar, nas próximas abordagens, apenas a energia de cavitação, mantendo-se a energia de dispersão calculada com C_{iw} retirado do OPLS-AA. O modelo GBEEM/NP3 emprega a energia de cavitação proposta por Tuñón, Silla e Pascual-Ahuir:

$$\Delta G_{cav} = \gamma S_M - RT \ln(1 - V_S n_S) \quad (44)$$

onde S_M é a área da superfície da cavidade, V_S é o volume da molécula de solvente e n_S é a densidade numérica do solvente.

Tabela 18: Parâmetros otimizados para o cálculo da energia livre de solvatação em GBEEM/NP1. As unidades dos parâmetros C_{iw} , a e b são $\text{kcal mol}^{-1} \text{\AA}^6$, kcal mol^{-1} e $\text{kcal mol}^{-1} \text{\AA}^{-2}$, respectivamente. O parâmetro c é adimensional.

Parâmetro*	Valor
C_{iw} H-O	64,55
C_{iw} H-N	2955,87
C_{iw} H-C	734,58
C_{iw} H-O-H	2085,19
C_{iw} C=C e Caromático	1901,81
C_{iw} C=O	368,92
C_{iw} C≡C	201,33
C_{iw} C-C	703,50
C_{iw} N-H	1905,61
C_{iw} N-N	1982,65
C_{iw} Naromático	1076,81
C_{iw} O-H e O-C	557,93
C_{iw} O=C	104,92
a	4,296
b	0,002
c	0,110

* Para os parâmetros C_{iw} foram usados diferentes tipos atômicos. Na tabela os valores de C_{iw} correspondem aos átomos em negrito pertencentes aos grupos funcionais descritos.

Outra possibilidade para se calcular energia livre de cavitação foi sugerida por Pierotti-Claverie, a implementação dessa energia no cálculo da energia livre de solvatação foi chamada de GBEEM/NP4.

$$\Delta G_{cav} = \sum_i^{esferas} \frac{A_i}{4\pi R_i^2} \Delta G_{cav}(R_i) \quad (45)$$

sendo R_i é o raio do átomo i e A_i é a área superficial acessível ao solvente do átomo i .

No modelo de solvatação implícita de Cramer e Trular (*Chambers et al., 1996*), o termo não polar, chamado de termo da primeira camada de solvatação (G_{CDS}^0), tem a seguinte forma:

$$G_{CDS}^0 = \sum_i \sigma_i A_i(R_i) \quad (46)$$

onde $A_i(R_i)$ é a área superficial acessível ao solvente do átomo i e σ_i é a tensão superficial microscópica do átomo.

O parâmetro σ_i foi também parametrizado com a combinação de algoritmo genético (*Goldberg, 1989*) e Simplex (*Press et al., 1992*), tendo como alvo as energias livres de solvatação experimentais. A combinação desse modelo proposto por Cramer e Truhlar com a parte eletrostática do GBEEM foi denominada de GBEEM/NP5.

Outra abordagem testada, GBEEM/NP6, foi a utilização dos termos do ciclo termodinâmico do ΔG_{solv} como descritores em uma regressão linear múltipla, onde a parte não polar foi calculada usando as áreas superficiais dos átomos C, O e N, acessíveis ao solvente de cada molécula, A_C , A_O e A_N , respectivamente.

$$\Delta G_{solv} = a + b \Delta E_{EN} + c \Delta G_{el} + d A_C + e A_O + f A_N \quad (47)$$

sendo $a = 2,833$, $b = 0,930$, $c = 0,488$, $d = -0,033$, $e = -0,154$ e $f = -0,237$.

A análise dos gráficos de energias de solvatação calculadas por GBEEM/NPx (onde x varia de 1 a 6) versus energias de solvatação experimentais (Figura 21, Figura 22 e Figura 23) mostra que os piores modelos são o GBEEM/NP3, GBEEM/NP4 e GBEEM/NP5, pois estes são os modelos que apresentam os maiores desvios. O modelo GBEEM/NP5 não apresenta sensibilidade no cálculo de energias livres de solvatação

maiores do que zero, pois os pontos se alinham sobre o eixo x. Os modelos GBEEM/NP1 e GBEEM/NP2, apresentando resultados parecidos entre si, são melhores do que os descritos anteriormente, pois as energias de solvatação calculadas por estes modelos são mais concordantes com as energias de solvatação experimentais. O modelo GBEEM/NP2 mostra-se bastante sensível para as energias livres de solvatação maiores do que zero, uma vez que os pontos concentram-se sobre a linha identidade, inclusive este é o modelo que melhor prevê a energia de solvatação nesta faixa. O modelo GBEEM/NP6, dentre os modelos vistos até agora, é o que melhor reproduz as energias livres de solvatação, uma vez que os pontos estão bem alinhados com a linha identidade e apenas dois pontos com valores inversos.

O bom resultado apresentado pelo modelo GBEEM/NP6 sugeriu que a utilização das energias de dispersão particionadas por elemento químico pudesse produzir resultados ainda melhores.

Figura 21: Energias livres de solvatação calculadas pelos modelos GBEEM/NP1 e GBEEM/NP2 versus energias de solvatação experimentais. A linha pontilhada é a identidade.

Figura 22: Energias livres de solvatação calculadas pelos modelos GBEEM/NP3 e GBEEM/NP4 versus energias de solvatação experimentais. A linha pontilhada é a identidade.

Figura 23: Energias livres de solvatação calculadas pelos modelos GBEEM/NP5 e GBEEM/NP6 versus energias de solvatação experimentais. A linha pontilhada é a identidade.

Os termos ΔG_{el} , SAS, U_{disp} particionado por elemento químico e ΔE_{EN} foram calculados e usados em uma regressão linear múltipla de acordo com a equação abaixo, sendo que as energias livres de solvatação aquosas experimentais são as variáveis independentes. Os coeficientes lineares obtidos na regressão linear múltipla definem o modelo GBEEM-ELR. A Tabela 19 mostra o coeficiente do SAS (c_3) fora da faixa de valores geralmente obtidos em SAS nos modelos de hidratação (Zacharias, 2003), porém está próximo ao valor da tensão superficial da água ($0,104 \text{ kcal.mol}^{-1}.\text{\AA}^{-2}$). Nesta tabela observa-se também que existe um pequeno mas importante termo de dispersão, originado no átomo H (valor c_H).

$$\Delta G_{solv} = c_0 + c_1 \Delta E_{EN} + c_2 \Delta G_{el} + c_3 SAS + c_H U_{disp}(H) + c_C U_{disp}(C) + c_O U_{disp}(O) + c_N U_{disp}(N) \quad (42)$$

Os desvios dos coeficientes c_3 e c_H são muito próximos aos valores dos próprios coeficientes, portanto foi feita uma análise de regressão passo a passo (“stepwise”) (Wittink, 1988), para se obter significância estatística dos descritores. Esta análise adiciona termo por termo enquanto o Fparcial de Fisher correspondente for significativo (>4). O resultado obtido está apresentado na Tabela 20. É interessante ressaltar que foram descartados os coeficientes c_3 e c_H que estão relacionados com o termo dependente da SAS, e a interação de dispersão entre os átomos de hidrogênio e a água, respectivamente. O coeficiente c_0 é positivo, sugerindo algum trabalho de cavitação médio suficiente para este pequeno conjunto de moléculas, ou seja, pode ser que este coeficiente já possua a contribuição da SAS. A contribuição de dispersão entre os átomos de hidrogênio e a água pode estar sendo compensada pelos outros tipos de átomos (C, O e N).

Tabela 19: Coeficientes lineares obtidos pela regressão linear múltipla do modelo GBEEM-ELR.

$$\Delta G_{solv} = c_0 + c_1 \Delta E_{EN} + c_2 \Delta G_{el} + c_3 SAS + c_H U_{disp}(H) + c_C U_{disp}(C) + c_O U_{disp}(O) + c_N U_{disp}(N)$$

	c_0^a	c_1^b	c_2^b	c_3^c	c_H^d	c_C^d	c_O^d	c_N^d
Coef	1,51	1,02	2,46	0,09	-10,06	-89,92	-99,09	-209,39
SD	0,67	0,15	0,49	0,08	11,71	56,24	45,47	45,01
Fparcial	-	26	46	1	1	3	5	22

- a) a unidade da intersecção é kcal.mol⁻¹;
b) estes coeficientes lineares são adimensionais;
c) a unidade é kcal.mol⁻¹.Å⁻²;
d) a unidade é kcal.mol⁻¹.Å⁻⁶.

Tabela 20: Coeficientes lineares obtidos pela análise de regressão passo a passo do modelo GBEEM-ELR.

$$\Delta G_{solv} = c_0 + c_1 \Delta E_{EN} + c_2 \Delta G_{el} + c_3 SAS + c_H U_{disp}(H) + c_C U_{disp}(C) + c_O U_{disp}(O) + c_N U_{disp}(N)$$

	c_0^a	c_1^b	c_2^b	c_3^c	c_H^d	c_C^d	c_O^d	c_N^d
Coef	2,45	1,06	2,52	-	-	-27,36	-53,41	-163,72
SD	0,40	0,14	0,45	-	-	4,11	14,37	16,34
Fparcial	-	31	55	-	-	44	14	100

- a) a unidade da intersecção é kcal.mol⁻¹;
b) estes coeficientes lineares são adimensionais;
c) a unidade é kcal.mol⁻¹.Å⁻²;
d) a unidade é kcal.mol⁻¹.Å⁻⁶.

A performance do GBEEM-ELR foi comparada com os modelos SM5.42R e IEFPCM através do cálculo do desvio médio absoluto (mean absolute deviation, MAD) entre os valores previstos e os valores experimentais das energias livres de solvatação. A Tabela 21 mostra que os modelos SM5.42R e IEFPCM são evidentemente melhores do que o modelo GBEEM-ELR. A observação dos gráficos de energias livres de solvatação calculadas versus energias livres de solvatação experimentais (Figura 24, Figura 25 e Figura 26) sugere algumas tendências. Na faixa experimental menor do que $-6,0 \text{ kcal.mol}^{-1}$ ($\Delta G_{\text{solv}} < -6,0$), o modelo IEFPCM apresentou valores de energias livres mais negativos do que o modelo SM5.42R. Os MADs parciais para esta faixa são $0,45 \text{ kcal.mol}^{-1}$ (SM5.42R), $0,39 \text{ kcal.mol}^{-1}$ (IEFPCM) e $1,07 \text{ kcal.mol}^{-1}$ (GBEEM-ELR). Na faixa intermediária, de $-6,0 \text{ kcal.mol}^{-1}$ a $-4,0 \text{ kcal.mol}^{-1}$ ($-6,0 \leq \Delta G_{\text{solv}} < -4,0$), os MADs sugerem que o modelo SM5.42R apresenta os melhores valores de energias livres de solvatação e os valores deles são $0,40 \text{ kcal.mol}^{-1}$ (SM5.42R), $0,69 \text{ kcal.mol}^{-1}$ (IEFPCM) e $0,82 \text{ kcal.mol}^{-1}$ (GBEEM-ELR). Entre $-4,0 \text{ kcal.mol}^{-1}$ e 0 kcal.mol^{-1} ($-4,0 \leq \Delta G_{\text{solv}} < 0$), o modelo IEFPCM apresenta novamente os melhores valores para a energia livre e os MADs são $0,51 \text{ kcal.mol}^{-1}$ (SM5.42R), $0,40 \text{ kcal.mol}^{-1}$ (IEFPCM) e $0,46 \text{ kcal.mol}^{-1}$ (GBEEM-ELR). Na faixa positiva das energias livres de solvatação ($\Delta G_{\text{solv}} \geq 0$), as melhores previsões foram feitas pelo modelo IEFPCM, com os seguintes MADs: $0,38 \text{ kcal.mol}^{-1}$ (SM5.42R), $0,34 \text{ kcal.mol}^{-1}$ (IEFPCM) e $0,72 \text{ kcal.mol}^{-1}$ (GBEEM-ELR).

Tabela 21: Os desvios médios absolutos, MAD, das energias livres de hidratação de 66 moléculas orgânicas neutras calculadas pelos modelos SM5.42R, IEFPCM-HF(UAHF)/6-31G e GBEEM-ELR. Assim como os tempos de cálculo dos três modelos.

Modelo	MAD (kcal.mol ⁻¹)	Tempo (s)
SM5.42R	0,44	4,27
IEFPCM	0,51	4954
GBEEM-ELR	0,71	1,43

Figura 24: Energias livres de solvatação calculadas pelo modelo SM5.42R versus energias de solvatação experimentais. A linha pontilhada é a identidade.

Figura 25: Energias livres de solvatação calculadas pelo modelo IEFPCM versus energias de solvatação experimentais. A linha pontilhada é a identidade.

Figura 26: Energias livres de solvatação calculadas pelo modelo GBEEM-ELR versus energias de solvatação experimentais. A linha pontilhada é a identidade.

Na Tabela 21 estão apresentados também os tempos de cálculo das energias livres de solvatação do conjunto de 62 moléculas neutras, para os três modelos de solvatação. O modelo IEFPCM é bem mais lento do que os dois outros modelos, cerca de mil vezes mais lento do que o SM5.42R e cerca de quatro mil vezes mais lento do que o GBEEM-ELR. Porém, quando se compara o tempo de cálculo do SM5.42R e do GBEEM-ELR, a diferença não é significativa, pois em uma relação custo/benefício o SM5.42R seria a melhor opção. Por outro lado, o tempo de 1,43 s do GBEEM-ELR deve-se em maior parte ao cálculo da SAS pelo GEPOL (0,91 s), seguido do cálculo dos raios de Born (0,39 s), sendo que apenas 0,13 s é o tempo de cálculo do EEM. Portanto, o tempo total do GBEEM-ELR poderia ser bastante reduzido com uma otimização do código fonte do GEPOL. Outra vantagem do modelo GBEEM-ELR é a sua base em descritores de reatividade química, como eletronegatividade e dureza química, além da possibilidade de se calcular dureza e maciez para cada átomo na molécula.

Com relação aos desvios entre os valores calculados pelo GBEEM-ELR e os valores experimentais, uma possibilidade de se melhorar os resultados seria através de um aprimoramento do cálculo do termo eletrostático da energia livre de solvatação pois, em alguns casos, aparentemente o termo não-polar está compensando um valor alto do termo eletrostático.

Na Tabela 22 estão apresentadas as energias de solvatação experimentais e as energias de solvatação calculadas com os modelos de solvatação SM5.42R, IEFPCM e GBEEM-ELR. Já na Tabela 23 estão apresentados os valores dos descritores usados na regressão múltipla para calcular as energias livres de solvatação pelo modelo GBEEM-ELR.

Tabela 22: Energias livres de solvatação experimentais (*Chambers et al., 1996*) e calculadas pelos métodos SM5.42R, IEFPCM e GBEEM-ELR (em kcal.mol⁻¹).

Molécula	ΔG Experimental	ΔG SM5.42R	ΔG IEFPCM	ΔG GBEEM-ELR
água	-6,31	-5,78	-6,31	-5,04
anilina	-5,49	-5,27	-4,87	-6,98
antraceno	-4,23	-4,20	-2,46	-3,53
benzeno	-0,87	-1,37	-1,36	-0,52
ciclopentano	1,20	0,66	0,6	0,46
ciclopentanol	-5,49	-5,13	-5,67	-3,71
ciclopropano	0,75	0,25	1,05	0,92
dimetilamina	-4,29	-3,87	-4,35	-3,37
1,2-dimetoxietano	-4,84	-2,47	-4,05	-6,25
2,4-dimetilpiridina	-4,86	-4,17	-3,57	-4,87
2,5-dimetilpiridina	-4,72	-4,13	-3,19	-4,65
2,6-dimetilpiridina	-4,60	-4,34	-2,82	-4,14
1,4-dioxano	-5,05	-5,04	-7,63	-6,74
dipropilamina	-3,66	-2,48	-2,2	-3,23
etanal	-3,50	-4,09	-3,57	-3,37
etano	1,83	1,32	1,96	1,35
ácido etanóico	-6,70	-6,32	-6,97	-7,95
etanol	-5,01	-5,22	-5,35	-3,51
eteno	1,27	1,33	1,47	1,16
etilamina	-4,50	-3,99	-4,65	-4,37
etilbenzeno	-0,80	-0,92	-1,23	-0,89
dietil éter	-1,76	-1,04	-2,11	-2,81
etino	-0,01	0,29	0,16	1,06
metanoato de etila	-2,65	-3,45	-3,34	-3,44
fenol	-6,62	-5,87	-6,88	-4,75
hidrazina	-9,30	-9,21	-9,81	-10,59
m-cresol	-5,49	-5,63	-6,56	-6,44
2,2-dimetilpropano	2,50	2,10	2,75	0,60
trimetilamina	-3,23	-2,90	-2,77	-2,68
etanoato de metila	-3,32	-3,57	-3,95	-3,23
metil isopropil éter	-2,01	-0,80	-1,79	-2,73
metanoato de metila	-2,78	-3,75	-3,69	-3,16
2-metoxietanol	-6,77	-6,61	-7,43	-6,81
2-metilpiridina	-4,63	-4,52	-3,74	-4,22
3-metilpiridina	-4,77	-4,31	-4,09	-4,66
4-metilpiridina	-4,94	-4,34	-4,48	-4,93
metil propil éter	-1,66	-0,82	-1,42	-2,79
2-metilpropano	2,32	1,72	2,36	0,79
2-metil-2-propanol	-4,51	-3,86	-4,69	-3,18
2-metilpropeno	1,16	1,29	0,79	0,48
metano	2,00	1,61	1,13	1,78

Tabela 22: continuação

Molécula	ΔG	ΔG	ΔG	ΔG
	Experimental	SM5.42R	IEFPCM	GBEEM-ELR
metanol	-5,11	-5,22	-5,14	-3,65
metilamina	-4,56	-4,28	-4,76	-4,64
dimetil éter	-1,92	-1,08	-2,11	-3,03
m-xileno	-0,84	-0,88	-0,78	-1,06
naftaleno	-2,39	-2,96	-1,98	-2,03
o-cresol	-5,87	-5,54	-6,04	-6,26
o-xileno	-0,90	-1,02	-0,94	-0,95
piridina	-4,70	-4,70	-4,62	-4,23
pirrolidina	-5,48	-4,81	-6,32	-3,49
propanal	-3,44	-3,69	-3,76	-3,54
propano	1,96	1,40	1,96	1,03
ácido propanóico	-6,47	-5,68	-7,1	-5,76
1-propanol	-4,83	-4,95	-4,67	-3,53
2-propanol	-4,76	-4,62	-5,12	-3,30
propanona	-3,85	-3,89	-3,58	-4,05
propeno	1,27	1,18	0,61	0,81
2-propen-1-ol	-5,08	-5,11	-5,59	-4,81
propilamina	-4,39	-3,77	-3,93	-4,49
p-xileno	-0,81	-0,85	-0,77	-1,09
tetrahidrofurano	-3,47	-2,63	-4,3	-3,09
tolueno	-0,89	-1,12	-1,09	-0,76

Tabela 23: Energias livres de solvatação experimentais (*Chambers et al., 1996*) e calculadas pelos métodos GBEEM/NP, PCM e AMSOL (em kcal mol⁻¹).

Molécula	ΔG Exp.	ΔG GBEEM-ELR	ΔG_{el}	ΔE_{EN}	SAS	U_{disp} (H)	U_{disp} (C)	U_{disp} (N)	U_{dips} (O)
água	-6,31	-5,04	-7,95	1,09	36,11	0,11	0,00	0,00	0,03
anilina	-5,49	-6,98	-6,55	1,42	123,91	0,33	0,09	0,02	0,00
antraceno	-4,23	-3,53	-1,39	0,31	201,97	0,46	0,19	0,00	0,00
benzeno	-0,87	-0,52	-0,60	0,10	108,62	0,29	0,09	0,00	0,00
ciclopentano	1,20	0,46	-0,02	0,00	115,12	0,46	0,07	0,00	0,00
ciclopentanol	-5,49	-3,71	-4,47	0,71	125,85	0,46	0,07	0,00	0,03
ciclopropano	0,75	0,92	-0,12	0,01	79,12	0,29	0,05	0,00	0,00
dimetilamina	-4,29	-3,37	-2,39	0,31	85,75	0,33	0,04	0,02	0,00
1,2-dimetoxietano	-4,84	-6,25	-6,83	1,00	135,32	0,46	0,07	0,00	0,04
2,4-dimetilpiridina	-4,86	-4,87	-4,08	1,20	146,17	0,42	0,10	0,02	0,00
2,5-dimetilpiridina	-4,72	-4,65	-4,21	1,34	146,37	0,42	0,10	0,02	0,00
2,6-dimetilpiridina	-4,60	-4,14	-4,42	1,56	146,28	0,42	0,11	0,02	0,00
1,4-dioxano	-5,05	-6,74	-6,90	0,76	113,50	0,37	0,06	0,00	0,04
dipropilamina	-3,66	-3,23	-1,38	0,36	172,54	0,68	0,09	0,02	0,00
etanal	-3,50	-3,37	-10,90	3,30	72,53	0,19	0,04	0,00	0,03
etano	1,83	1,35	-0,06	0,00	69,93	0,29	0,04	0,00	0,00
ácido etanóico	-6,70	-7,95	-13,44	3,05	82,60	0,19	0,03	0,00	0,05
etanol	-5,01	-3,51	-4,95	0,65	80,68	0,29	0,03	0,00	0,03
eteno	1,27	1,16	-0,21	0,02	61,55	0,20	0,04	0,00	0,00
etilamina	-4,50	-4,37	-3,23	0,48	85,75	0,34	0,03	0,02	0,00
etilbenzeno	-0,80	-0,89	-0,82	0,30	150,20	0,46	0,12	0,00	0,00
dietil éter	-1,76	-2,81	-3,52	0,51	124,28	0,46	0,07	0,00	0,02

Tabela 23: continuação

Molécula	ΔG Exp.	ΔG GBEEM-ELR	ΔG_{el}	ΔE_{EN}	SAS	U_{disp} (H)	U_{disp} (C)	U_{disp} (N)	U_{dips} (O)
etino	-0,01	1,06	-0,35	0,07	54,33	0,10	0,04	0,00	0,00
metanoato de etila	-2,65	-3,44	-2,35	0,20	104,14	0,28	0,05	0,00	0,05
fenol	-6,62	-4,75	-5,42	0,95	118,88	0,28	0,09	0,00	0,03
hidrazina	-9,30	-10,59	-7,17	1,07	58,65	0,20	0,00	0,05	0,00
m-cresol	-5,49	-6,44	-9,81	2,26	140,31	0,37	0,10	0,00	0,03
2,2-dimetilpropano	2,50	0,60	0,07	0,00	132,64	0,54	0,07	0,00	0,00
trimetilamina	-3,23	-2,68	-1,82	0,22	106,64	0,41	0,05	0,01	0,00
etanoato de metila	-3,32	-3,23	-2,62	0,42	104,41	0,28	0,05	0,00	0,05
metil isopropil éter	-2,01	-2,73	-3,83	0,65	123,13	0,46	0,06	0,00	0,02
metanoato de metila	-2,78	-3,16	-2,60	0,32	82,65	0,19	0,04	0,00	0,05
2-metoxietanol	-6,77	-6,81	-7,94	1,18	113,32	0,37	0,05	0,00	0,05
2-metilpiridina	-4,63	-4,22	-4,67	1,58	124,76	0,33	0,09	0,02	0,00
3-metilpiridina	-4,77	-4,66	-4,48	1,39	124,74	0,33	0,09	0,02	0,00
4-metilpiridina	-4,94	-4,93	-4,35	1,23	124,78	0,33	0,09	0,02	0,00
metil propil éter	-1,66	-2,79	-4,25	0,84	124,27	0,46	0,06	0,00	0,02
2-metilpropano	2,32	0,79	0,01	0,01	112,58	0,46	0,06	0,00	0,00
2-metil-2-propanol	-4,51	-3,18	-4,00	0,57	122,41	0,46	0,06	0,00	0,02
2-metilpropeno	1,16	0,48	-0,24	0,04	104,45	0,37	0,07	0,00	0,00
metano	2,00	1,78	-0,05	0,00	47,76	0,20	0,02	0,00	0,00
metanol	-5,11	-3,65	-5,77	0,81	59,01	0,20	0,02	0,00	0,03
metilamina	-4,56	-4,64	-3,89	0,59	64,04	0,25	0,02	0,02	0,00
dimetil éter	-1,92	-3,03	-4,71	0,68	80,99	0,29	0,04	0,00	0,02
m-xileno	-0,84	-1,06	-0,78	0,21	151,51	0,46	0,12	0,00	0,00
naftaleno	-2,39	-2,03	-1,00	0,21	155,30	0,37	0,14	0,00	0,00
o-cresol	-5,87	-6,26	-9,50	2,16	139,42	0,37	0,10	0,00	0,02

Tabela 23: continuação

Molécula	ΔG Exp.	ΔG GBEEM-ELR	ΔG_{el}	ΔE_{EN}	SAS	U_{disp} (H)	U_{disp} (C)	U_{disp} (N)	U_{dips} (O)
o-xileno	-0,90	-0,95	-0,84	0,26	150,28	0,46	0,12	0,00	0,00
piridina	-4,70	-4,23	-4,95	1,63	103,13	0,24	0,08	0,02	0,00
pirrolidina	-5,48	-3,49	-2,33	0,46	109,89	0,42	0,06	0,02	0,00
propanal	-3,44	-3,54	-9,93	2,94	93,66	0,28	0,05	0,00	0,03
propano	1,96	1,03	-0,04	0,00	91,60	0,38	0,05	0,00	0,00
ácido propanóico	-6,47	-5,76	-11,21	3,09	103,74	0,28	0,05	0,00	0,05
1-propanol	-4,83	-3,53	-4,81	0,72	102,24	0,37	0,05	0,00	0,03
2-propanol	-4,76	-3,30	-4,43	0,62	101,92	0,37	0,05	0,00	0,02
propanona	-3,85	-4,05	-10,01	2,78	94,17	0,28	0,05	0,00	0,03
propeno	1,27	0,81	-0,23	0,04	83,43	0,29	0,05	0,00	0,00
2-propen-1-ol	-5,08	-4,81	-10,04	2,43	92,91	0,28	0,05	0,00	0,02
propilamina	-4,39	-4,49	-3,25	0,58	107,36	0,42	0,05	0,02	0,00
p-xileno	-0,81	-1,09	-0,77	0,20	151,76	0,46	0,12	0,00	0,00
tetrahidrofurano	-3,47	-3,09	-4,89	0,97	105,18	0,37	0,06	0,00	0,02
tolueno	-0,89	-0,76	-0,75	0,20	130,20	0,37	0,11	0,00	0,00

5 – Conclusões

Foi proposto um novo método para cálculo de distribuição de cargas atômicas em meio condensado, fundamentado em teoria do funcional da densidade. O método trata de forma clara e computacionalmente eficiente o acoplamento entre a estrutura eletrônica da molécula (soluto) e o solvente implícito. Desenvolvimento de um novo modelo de solvente implícito foi também apresentado e estudado com resultados promissores.

O método de equalização da eletronegatividade não prevê corretamente momentos de dipolo permanentes em moléculas com elevados números de átomos (como por exemplo, nonanal) quando comparados com dipolos calculados por outros modelos de cargas ou ainda com o valor esperado do operador momento de dipolo. O EEM sobreestima a polarizabilidade com o crescimento da cadeia molecular exagerando na delocalização das cargas (como se as moléculas fossem “condutores metálicos”). Inicialmente, essa imperfeição do EEM foi tratada fazendo uma nova parametrização das eletronegatividades e durezas atômicas, tendo cargas CM1 como alvo, porém essa nova parametrização não corrigiu os momentos de dipolo, indicando que esta imperfeição é intrínseca às aproximações usadas na derivação do EEM (*Cong & Yang, 2000; Chelli et al., 1999; Baekelandt et al., 1993; Mortier et al., 1986*).

A aproximação de cargas predefinidas melhora significativamente o EEM na previsão de observáveis moleculares sendo que agora bons resultados para vetores de momento de dipolo são produzidos. O GBEEM foi comparado com outros modelos de dielétrico contínuo, como o semi-empírico SM5.4 implementado em AMSOL7.0 e o *ab initio* IEFPCM implementado em GAUSSIAN03, obtendo bons resultados para dipolos de polipeptídeos quando eletronegatividades e durezas atômicas da parametrização QEq-M são usadas com a aproximação de cargas predefinidas.

Foi proposto um modelo de energia de solvatação e este foi comparado com outros modelos de dielétrico contínuo, assim como foi feito para o momento de dipolo, mostrando ser promissor, uma vez que tanto o GBEEM-ELR, quanto o IEFPCM(UAHF) em nível HF/6-31G(d) e o SM5.42 em nível PM3 apresentaram MADS menores do 1 kcal/mol na previsão de energias livres de hidratação. Apesar dos modelos PCM e SMx serem claramente melhores do que o GBEEM-ELR, os resultados não são desanimadores. Acredita-se que a pesquisa de novos e aprimorados modelos de EEM/campo de reação irão não apenas descrever melhor os processos de solvatação,

mas serão também mais rápidos computacionalmente quando se considerar a polarização do soluto.

Futuros avanços no método EEM e GBEEM podem incluir a criação de descritores de reatividade (teoria do funcional da densidade conceitual), desenvolvimento de campos de força polarizável, tratamento de excitação eletrônica $S_0 \rightarrow S_1$ para grandes moléculas.

6 – Apêndice: Código fonte do programa GBEEM

A seguir é apresentado o código fonte do programa GBEEM, que calcula cargas atômicas no vácuo, em fase condensada e a parte eletrostática da energia livre de solvatação. O programa foi escrito na linguagem Fortran77.

```
C
C
C
      program GBSA plus EEM
*
      parameter(mxatm=200,mxesp=20)
*
      common / matriz / q,qv,c,d,natm,nesp,natt
      common / coord / elq,x,y,z,t,rij
      common / eletroneg / qsi,eta,qt,alfa
      common / gbsa / fq,eps,ae,gel,at,npol,tot,en,en1,en2,en3,en4
*
      real q(mxatm),qv(mxatm),c(mxatm,mxatm+1),d(mxatm,mxatm+1)
      real qsi(mxatm),eta(mxatm,mxatm),qt(mxesp),alfa
      real x(mxatm),y(mxatm),z(mxatm),rij(mxatm,mxatm)
*
      real xqsi(10),xeta(10)
      real fq(mxatm,mxatm),eps,ae(mxatm),gel,at(mxatm),
#      npol,tot,en,en1,en2,en3,en4
*
      integer natm(mxatm),nesp,natt
      integer t(mxatm)
*
      integer i,j,k,l,ii
      real tp1,tp2
*
      character*2 elq(mxatm)
*
*
*
*
      open(4,file='qsieteta')
      do 5 i=1,4
      read(4,*)xqsi(i),xeta(i)
5      continue
      read(4,*) nesp,eps
      read(4,*) natt
      do 7 i=1,nesp
      read(4,*) natm(i)
7      continue
      close(4)
*
      call read
*
      do 10 i=1,natt
***
      if (elq(i).eq.'H') then
*
      qsi(i)=xqsi(1)
      eta(i,i)=xeta(1)
```

```

*
      go to 10
    endif
***
    if (elq(i).eq.'O') then
*
      qsi(i)=xqsi(2)
      eta(i,i)=xeta(2)
*
      go to 10
    endif
***
    if (elq(i).eq.'N') then
*
      qsi(i)=xqsi(3)
      eta(i,i)=xeta(3)
*
      go to 10
    endif
***
    if (elq(i).eq.'C') then
*
      qsi(i)=xqsi(4)
      eta(i,i)=xeta(4)
*
      go to 10
    endif
***
    if (elq(i).ne.'H'.and.
+elq(i).ne.'O'.and.
+elq(i).ne.'N'.and.
+elq(i).ne.'C' ) stop 'Atomo nao definido'
10    continue
*
    call qsieta
*
    call solvat
*
    call sl
*
    call eletro
*
    call qsietav
*
    call slv
*
    call eletronuclear
*
    call write
*
  end
*****
C
C  leitura de arquivo externo
C
      subroutine read
*
      parameter(mxatm=200,mxesp=20)
*
      common / matriz / q,qv,c,d,natm,nesp,natt
      common / coord / elq,x,y,z,t,rij

```

```

common / eletroneg / qsi,eta,qt,alfa
common / gbsa / fq,eps,ae,gel,at,npol,tot,en,en1,en2,en3,en4
*
real q(mxatm),qv(mxatm),c(mxatm,mxatm+1),d(mxatm,mxatm+1)
real qsi(mxatm),eta(mxatm,mxatm),qt(mxesp),alfa
real x(mxatm),y(mxatm),z(mxatm),rij(mxatm,mxatm)
real fq(mxatm,mxatm),eps,ae(mxatm),gel,at(mxatm),
#      npol,tot,en,en1,en2,en3,en4
*
integer natm(mxatm),nesp,natt
integer t(mxatm),atm
*
integer i,j,k,l,m
real tp1,tp2,aa
*
character*2 elq(mxatm),lab(mxatm)
*
*
*
j=1
open(3,file='raio')
read(*,*) n
do 7 k=1,nesp
  do 10 i=j,natm(k)+j-1
    read(*,*) elq(i),x(i),y(i),z(i)
10  continue
  do 30 i=j,natm(k)+j-1
    read(3,*) n,aa,fq(i,i)
    fq(i,i)=fq(i,i)*3.01140889e-3
30  continue
  qt(k)=0
  j=j+natm(k)
7   continue
close(3)
natt=j-1
return
end
*****
C
C  Calcula termos de eletronegatividade, dureza e carga no solvente
C  com parametros retirados de ajuste a cargas
C
subroutine qsieteta
*
parameter(mxatm=200,mxesp=20)
*
common / matriz / q,qv,c,d,natm,nesp,natt
common / coord / elq,x,y,z,t,rij
common / eletroneg / qsi,eta,qt,alfa
common / gbsa / fq,eps,ae,gel,at,npol,tot,en,en1,en2,en3,en4
*
real q(mxatm),qv(mxatm),c(mxatm,mxatm+1),d(mxatm,mxatm+1)
real qsi(mxatm),eta(mxatm,mxatm),qt(mxesp),alfa
real x(mxatm),y(mxatm),z(mxatm),rij(mxatm,mxatm)
real fq(mxatm,mxatm),eps,ae(mxatm),gel,at(mxatm),
#      npol,tot,en,en1,en2,en3,en4
*
integer natm(mxatm),nesp,natt
integer t(mxatm)
*
integer i,j,k,l

```

```

      real aij,aij2,rij2
      real tp,sum
*
      character*2 elq(mxatm)
*
*
*
      do 110 i=1,natt
      do 120 j=i+1,natt
        rij2=(x(j)-x(i))*(x(j)-x(i))+(y(j)-y(i))*(y(j)-y(i))+
#      (z(j)-z(i))*(z(j)-z(i))
        rij(i,j)=sqrt(rij2)
        rij(j,i)=rij(i,j)
*
        aij=2.0d0/(eta(i,i)+eta(j,j))
        aij2=aij*aij
*
        eta(i,j)=1.0d0/sqrt(9.06858352e-6*rij2+aij2)
*MN      eta(i,j)=1.0d0/(3.01140889e-3*rij(i,j)+aij)
        eta(j,i)=eta(i,j)
*
120      continue
110      continue
      k=1
      do 130 l=1,nesp
        do 140 i=k,natm(l)+k-1
          do 150 j=1,natt
*
            c(i,j) = eta(i,j) - eta(k,j)
*
150          continue
            c(i,natt+1)=qsi(k)-qsi(i)
            c(k,i)=1
140          continue
            c(k,natt+1)=qt(l)
            k=k+natm(l)
130          continue
          return
        end
*****
C
C  Resolve sistema linear pelo esquema de Khaletsky
C
      subroutine sl
*
      parameter(mxatm=200,mxesp=20)
*
      common / matriz / q,qv,c,d,natm,nesp,natt
      common / gbsa / fq,eps,ae,gel,at,npol,tot,en,en1,en2,en3,en4
*
      real q(mxatm),qv(mxatm),c(mxatm,mxatm+1),mat(mxatm),
#      d(mxatm,mxatm+1)
      real fq(mxatm,mxatm),eps,ae(mxatm),gel,at(mxatm),
#      npol,tot,en,en1,en2,en3,en4
*
      integer natm(mxatm),nesp,natt
      integer t(mxatm)
      integer i,j,k,l
*
*
*

```



```

DO 200 J=2,natt+1
  c(1,J)=c(1,J)/c(1,1)
200  continue
DO 250 I=2,natt
DO 300 J=I,natt
  SUM1=0
DO 350 K=1,I-1
  SUM1=SUM1+c(J,K)*c(K,I)
350  continue
  c(J,I)=c(J,I)-SUM1
300  continue
DO 400 J=I+1,natt+1
  SUM2=0
DO 450 K=1,I-1
  SUM2=SUM2+c(I,K)*c(K,J)
450  continue
  c(I,J)=(c(I,J)-SUM2)/c(I,I)
400  continue
250  CONTINUE
DO 500 I=natt,1,-1
  SUM3=0
DO 550 J=I+1,natt
  SUM3=SUM3+c(I,J)*q(J)
550  continue
  q(I)=c(I,natt+1)-SUM3
500  continue
RETURN
END
*****
C
C escreve em arquivo externo
C
  subroutine write
*
  parameter(mxatm=200,mxesp=20)
*
  common / matriz / q,qv,c,d,natm,nesp,natt
  common / coord / elq,x,y,z,t,rij
  common / eletroneg / qsi,eta,qt,alfa
  common / gbsa / fq,eps,ae,gel,at,npol,tot,en,en1,en2,en3,en4
*
  real q(mxatm),qv(mxatm),c(mxatm,mxatm+1),d(mxatm,mxatm+1)
  real qsi(mxatm),eta(mxatm,mxatm),qt(mxesp),alfa
  real x(mxatm),y(mxatm),z(mxatm),rij(mxatm,mxatm)
  real fq(mxatm,mxatm),eps,ae(mxatm),gel,at(mxatm),
#    npol,tot,en,en1,en2,en3,en4
*
  integer natm(mxatm),nesp,natt
  integer t(mxatm)
*
  integer i,j,k,l
  real tp1,tp2,tp3
*
  character*2 elq(mxatm)
*
*
*
  k=1
do 160 j=1,nesp
  sum=0.0e0
  write(*,*)natm(j)

```

```

        write(*,*) 'cargas no solvente'
        do 170 i=k,natm(j)+k-1
        write(*,*) elq(i),q(i)
        sum=sum+q(i)
170    continue
        write(*,*) 'cargas no vacuo'
        do 180 i=k,natm(j)+k-1
        write(*,*) elq(i),qv(i)
        sum=sum+qv(i)
180    continue
        k=k+natm(j)
160    continue
        return
        end
*****
        subroutine solvat
*
        parameter(mxatm=200,mxesp=20)
*
        common / matriz / q,qv,c,d,natm,nesp,natt
        common / coord / elq,x,y,z,t,rij
        common / eletroneg / qsi,eta,qt,alfa
        common / gbsa / fq,eps,ae,gel,at,npol,tot,en,en1,en2,en3,en4
*
        real q(mxatm),qv(mxatm),c(mxatm,mxatm+1),d(mxatm,mxatm+1)
        real qsi(mxatm),eta(mxatm,mxatm),qt(mxesp),alfa
        real x(mxatm),y(mxatm),z(mxatm),rij(mxatm,mxatm)
*
        real xqsi(10),xeta(10)
        real fq(mxatm,mxatm),eps,ae(mxatm),gel,at(mxatm),
#         npol,tot,en,en1,en2,en3,en4
*
        integer natm(mxatm),nesp,natt
        integer t(mxatm)
*
        integer i,j,k,l
        real tp1,tp2,tp3,rij2
*
        character*2 elq(mxatm)
*
        do 10 i=1,natt
        do 20 j=i+1,natt
*
            rij2=rij(i,j)*rij(i,j)
*
            fq(i,j)=fq(i,i)*fq(j,j)*exp(-0.25*9.06858352e-6*rij2/
#             (fq(i,i)*fq(j,j)))
            fq(i,j)=sqrt(9.06858352e-6*rij2+fq(i,j))
            fq(j,i)=fq(i,j)
*
20        continue
10    continue
        eps=1.0e0/eps
        k=1
        do 30 l=1,nesp
        do 40 i=k,natm(l)+k-1
        do 50 j=1,natt
            tp1=1.0e0/fq(i,j)
            tp2=1.0e0/fq(k,j)
            c(i,j)=(eps-1)*(tp1-tp2) + c(i,j)

```

```

50      continue
        c(k,i)=1
40      continue
        k=k+natm(1)
30      continue
        return
        end
*****
      subroutine eletro
*
      parameter(mxatm=200,mxesp=20)
*
      common / matriz / q,qv,c,d,natm,nesp,natt
      common / coord / elq,x,y,z,t,rij
      common / eletroneg / qsi,eta,qt,alfa
      common / gbsa / fq,eps,ae,gel,at,npol,tot,en,en1,en2,en3,en4
*
      real q(mxatm),qv(mxatm),c(mxatm,mxatm+1),d(mxatm,mxatm+1)
      real qsi(mxatm),eta(mxatm,mxatm),qt(mxesp),alfa
      real x(mxatm),y(mxatm),z(mxatm),rij(mxatm,mxatm)
*
      real xqsi(10),xeta(10)
      real fq(mxatm,mxatm),eps,ae(mxatm),gel,gell,gel2,
#      npol,tot,at(mxatm),en,en1,en2,en3,en4
*
      integer natm(mxatm),nesp,natt
      integer t(mxatm)
*
      integer i,j,k,l
      real tp1,tp2,tp3,rij2,tp,tp4
*
      character*2 elq(mxatm)
      l=0
      gel=0
      npol=0
      tot=0
      do 30 k=1,nesp
        l=l+natm(k)
30      continue
      gell=0
      gel2=0
      do 10 i=1,l
        do 20 j=i+1,l
          gel2=gel2+(0.5*(q(i)*q(j)/fq(i,j)))*(eps-1)
20      continue
        gell=gell+(q(i)*q(i)/(2*fq(i,i)))*(eps-1)
10      continue
      gel=gel+gell+gel2+gel2
      write(*,*) gel
      return
      end
*****
C
C  Calcula termos de eletronegatividade, dureza e carga no vacuo
C  com parametros retirados de ajuste a cargas
C
      subroutine qsietav
*
      parameter(mxatm=200,mxesp=20)
*
      common / matriz / q,qv,c,d,natm,nesp,natt

```

```

common / coord / elq,x,y,z,t,rij
common / eletroneg / qsi,eta,qt,alfa
common / gbsa / fq,eps,ae,gel,at,npol,tot,en,en1,en2,en3,en4
*
real q(mxatm),qv(mxatm),c(mxatm,mxatm+1),d(mxatm,mxatm+1)
real qsi(mxatm),eta(mxatm,mxatm),qt(mxesp),alfa
real x(mxatm),y(mxatm),z(mxatm),rij(mxatm,mxatm)
real fq(mxatm,mxatm),eps,ae(mxatm),gel,at(mxatm),
#      npol,tot,en,en1,en2,en3,en4
*
integer natm(mxatm),nesp,natt
integer t(mxatm)
*
integer i,j,k,l
real aij,aij2,rij2
real tp,sum
*
character*2 elq(mxatm)
*
*
*
k=1
do 130 l=1,nesp
  do 140 i=k,natm(l)+k-1
    do 150 j=1,natt
*
      d(i,j) = eta(i,j) - eta(k,j)
*
150    continue
      d(i,natt+1)=qsi(k)-qsi(i)
      d(k,i)=1
140    continue
      d(k,natt+1)=qt(l)
      k=k+natm(l)
130    continue
      return
      end
*****
C
C  Resolve sistema linear pelo esquema de Khaletsky
C
      subroutine slv
*
      parameter(mxatm=200,mxesp=20)
*
      common / matriz / q,qv,c,d,natm,nesp,natt
      common / gbsa / fq,eps,ae,gel,at,npol,tot,en,en1,en2,en3,en4
*
      real q(mxatm),qv(mxatm),c(mxatm,mxatm+1),mat(mxatm),
#      d(mxatm,mxatm+1)
      real fq(mxatm,mxatm),eps,ae(mxatm),gel,at(mxatm),
#      npol,tot,en,en1,en2,en3,en4
*
      integer natm(mxatm),nesp,natt
      integer t(mxatm)
      integer i,j,k,l
*
*
*
DO 200 J=2,natt+1
  d(1,J)=d(1,J)/d(1,1)

```

```

200    continue
      DO 250 I=2,natt
      DO 300 J=I,natt
        SUM1=0
      DO 350 K=1,I-1
        SUM1=SUM1+d(J,K)*d(K,I)
350    continue
      d(J,I)=d(J,I)-SUM1
300    continue
      DO 400 J=I+1,natt+1
        SUM2=0
      DO 450 K=1,I-1
        SUM2=SUM2+d(I,K)*d(K,J)
450    continue
      d(I,J)=(d(I,J)-SUM2)/d(I,I)
400    continue
250    CONTINUE
      DO 500 I=natt,1,-1
        SUM3=0
      DO 550 J=I+1,natt
        SUM3=SUM3+d(I,J)*qv(J)
550    continue
      qv(I)=d(I,natt+1)-SUM3
500    continue
      RETURN
      END
*****
      subroutine eletronnuclear
*
      parameter(mxatm=200,mxesp=20)
*
      common / matriz / q,qv,c,d,natm,nesp,natt
      common / coord / elq,x,y,z,t,rij
      common / eletroneg / qsi,eta,qt,alfa
      common / gbsa / fq,eps,ae,gel,at,npol,tot,en,en1,en2,en3,en4
*
      real q(mxatm),qv(mxatm),c(mxatm,mxatm+1),d(mxatm,mxatm+1)
      real qsi(mxatm),eta(mxatm,mxatm),qt(mxesp),alfa
      real x(mxatm),y(mxatm),z(mxatm),rij(mxatm,mxatm)
*
      real xqsi(10),xeta(10)
      real fq(mxatm,mxatm),eps,ae(mxatm),gel,gel1,gel2,
#      npol,tot,at(mxatm),en,en1,en2,en3,en4,eng,ens
*
      integer natm(mxatm),nesp,natt
      integer t(mxatm)
*
      integer i,j,k,l
      real tp1,tp2,tp3,rij2,tp,tp4
*
      character*2 elq(mxatm)
      l=0
      en=0
      ens=0
      eng=0
      do 30 k=1,nesp
      l=l+natm(k)
30    continue
      en1=0
      en2=0
      en3=0

```

```

en4=0
do 10 i=1,1
do 20 j=i+1,1
  en3=en3+eta(i,j)*q(i)*q(j)
  en4=en4+eta(i,j)*qv(i)*qv(j)
20  continue
  en1=en1+(qsi(i)+0.5*eta(i,i)*q(i))*q(i)
  en2=en2+(qsi(i)+0.5*eta(i,i)*qv(i))*qv(i)
10  continue
  en=en+en1-en2+en3-en4
  write(*,*) en
return
end
*****

```

7 – Bibliografia

- 1) Alemán, C.; Galembeck, S. E. *Chem. Phys.* **1998**, 232, 151-159.
- 2) Alemán, C.; Ishiki, H. M.; Armelin, E. A.; Abrahão Junior, O.; Galembeck, S. E. *Chem. Phys.* **1998**, 233, 85-96.
- 3) Allen, M. P.; Tildesley, D. J., *Computer Simulation of Liquids*, Clarendon Press, Oxford, 1987.
- 4) Baekelandt, B. G.; Mortier, W. J.; Schoonheydt, R. A., *The EEM Approach to Chemical Hardness in Molecules and Solids: Fundamentals and Applications, Structure and Bonding*, vol.80, 1993.
- 5) Bakowies, D.; Thiel, W. *J. Comput. Chem.* **1996**, 17, 87-108.
- 6) Bandyopadhyay, P.; Gordon, M. S.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **2002**, 5023-5032.
- 7) Banks, J. L.; Kaminski, G. A.; Zhou, R.; Mainz, D. T.; Berne, B. J.; Friesner, R. A. *J. Chem. Phys.* **1999**, 110, 741-754.
- 8) Barone, V.; Cossi, M. *J. Phys. Chem. A* **1998**, 102, 1995-2001.
- 9) Barone, V.; Cossi, M.; Tomasi, J. *J. Chem. Phys.* **1997**, 107, 3210-3221.
- 10) Bashford, D.; Case, D. A. *Annu. Rev. Phys. Chem.* **2000**, 51, 129-152.
- 11) Bondi, A. *J. Phys. Chem.* **1964**, 68, 441-451.
- 12) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, 11, 361-373.
- 13) Bultinck, P.; Langenaeker, W.; Lahorte, P.; De Proft, F.; Geerlings, P.; Waroquier, M.; Tollenaere, J. P. *J. Phys. Chem. A* **2002**, 106, 7887-7894.
- 14) Bultinck, P.; Vanholme, R.; Popelier, P. L. A.; De Proft, F.; Geerlings, P. *J. Phys. Chem. A* **2004**, 108, 10359-10366.
- 15) Cammi, R.; Tomasi, J. *J. Comput. Chem.* **1995**, 16, 1449-1458.
- 16) Cancès, E.; Mennucci, B.; Tomasi, J. *J. Chem. Phys.* **1997**, 107, 3032-3041.
- 17) Chambers, C. C.; Hawkins, G. D.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem.* **1996**, 100, 16385-16398.
- 18) Chelli, R.; Procacci, P.; Righini, R.; Califano, S. *J. Chem. Phys.* **1999**, 111, 8569-8575.
- 19) Chirlian, L. E.; Franci, M. M. *J. Comput. Chem.* **1987**, 8, 894-905.
- 20) Cho, K. H.; Kang, Y. K.; No, K. T.; Scheraga, H. A. *J. Phys. Chem. B* **2001**, 105, 3624-3634.

- 21) Cioslowski, J.; Stefanov, B. B. *J. Chem. Phys.* **1993**, 99, 5151-5162.
- 22) Cong, Y.; Yang, Z.-Z. *Chem. Phys. Lett.* **2000**, 316, 324-329.
- 23) Cornell, W. D.; Cieplak, P.; Bayly, C. I.; Gould, I. R.; Merz, K. M.; Ferguson, D. M.; Spellmeyer, D. C.; Fox, T.; Caldwell, J. W.; Kollman, P. A. *J. Am. Chem. Soc.* **1995**, 117, 5179-5197.
- 24) Cossi, M.; Scalmani, G.; Rega, N.; Barone, V. *J. Chem. Phys.* **2002**, 117, 43-54.
- 25) Cossi, M.; Barone, V.; Cammi, R.; Tomasi, J. *Chem. Phys. Lett.* **1996**, 255, 327-335.
- 26) Cramer, C. J.; Truhlar, D. G. *Chem. Rev.* **1999**, 99, 2161-2200.
- 27) Cramer, C. J.; Truhlar, D. J. *J. Am. Chem. Soc.* **1991**, 113, 8305.
- 28) Curutchet, C.; Cramer, C. J.; Truhlar, D. G.; Ruiz-Lopéz, M. F.; Rinaldi, D.; Orozco, M.; Luque, F. J. *J. Comp. Chem.* **2003**, 24, 284-297.
- 29) De Proft, F.; Langenaeker, W.; Geerlings, P. *J. Mol. Struct.(Theochem)* **1995**, 339, 45-55.
- 30) Dias, L. G.; Shimizu, K.; Farah, J. P. S.; Chaimovich, H. *Chem. Phys.* **2002**, 282, 237-243.
- 31) Fan, H.; Mark, A. E.; Zhu, J.; Honig, B. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, 102, 6760-6764.
- 32) Feig, M.; Brooks III, C. L. *Current Opinion in Structural Biology* **2004**, 14, 217-224.
- 33) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. A.; Montgomery, J. A.; Vreven, T.; Kudin, K. N.; Burant, J. C.; Millam, J. M.; Iyengar, S. S.; Tomasi, J.; Barone, V.; Mennucci, B.; Cossi, M.; Scalmani, G.; Rega, N.; Peterson, G. A.; Nakatsuji, H.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Klene, M.; Li, X.; Know, J. E.; Hratchian, H. P.; Cross, J. B.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Ayala, P. Y.; Morokuma, K.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Zakrzewski, V. G.; Dapprich, S.; Daniels, A. D.; Strain, M. C.; Farkas, O.; Malick, D. K.; Rabuck, A. D.; Raghavachari, K.; Foresman, J. B.; Ortiz, J. V.; Cui, Q.; Baboul, A. G.; Clifford, S.; Cioslowski, J.; Stefanov, B. B.; Liu, G.; Liashenko, A.; Piskorz, P.; Komaromi, I.; Martin, R. L.; Fox, D. J.; Keith, T.; Al-Laham, M. A.; Peng, C. Y.; Nanayakkara, A.; Challacombe, M.; Gill, P. M. W.; Johnson, B.; Chen, W.;

- Wong, M. W.; Gonzalez, C.; Pople, J. A. *Gaussian 03, Revision B.04*, Gaussian, Inc., Pittsburgh PA, 2003.
- 34) Gao, B.; Liu, Z. *J. Phys. Chem. A* **2005**, 9104-9111.
 - 35) Giesen, D. J.; Storer, J. W.; Cramer, C. J.; Truhlar, D. J., *J. Am. Chem. Soc.* **1995**, 117, 1057-1068.
 - 36) Goldberg, D. E., *Genetic Algorithms in Search, Optimization and Machine Learning*, Addison-Wesley, Reading, 1989.
 - 37) Grant, G. H.; Richards, W. G., *Computational Chemistry*, Oxford University Press, Oxford, 1995.
 - 38) Hall, D.; Williams, D. E. *Acta Crystallogr.* **1975**, A31, 56-58.
 - 39) Hawkins, G. D.; Giesen, D. J.; Lynch, G. C.; Chambers, C. C.; Rossi, I.; Storer, J. W.; Li, J.; Winget, P.; Rinaldi, D.; Liotard, D. A.; Cramer, C. J.; Truhlar, D. G., *AMSOL-version 7.0*, 2003.
 - 40) Jensen, F. *Introduction to Computational Chemistry*, John Wiley & Sons, Chichester, 1999.
 - 41) Jorgensen, W. L.; Maxwell, D. S.; Rives, J. T. *J. Am. Chem. Soc.* **1996**, 118, 11225-11236.
 - 42) Klamt, A. *J. Phys. Chem.* **1995**, 99, 2224-2235.
 - 43) Klamt, A.; Schüürmann, G. *J. Chem. Soc. Perkin Trans. 2* **1993**, 799-805.
 - 44) Li, J.; Zhu, T.; Cramer, C. J.; Truhlar, D. G. *J. Phys. Chem. A* **1998**, 102, 1820-1831.
 - 45) Liu, Y. P.; Kim, K.; Berne, B. J.; Friesner, R. A.; Rick, S. W. *J. Chem. Phys.* **1998**, 108, 4739-4755.
 - 46) Liu, Y. P.; Kim, K.; Berne, B. J.; Friesner, R. A.; Rick, S. W. *J. Chem. Phys.* **1998**, 108, 4739-4755.
 - 47) Löwdin, P. *J. Chem. Phys.* **1950**, 18, 365-375.
 - 48) MacKerell Jr., A. D.; Bashford, D.; Bellot, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher III, W. E.; Roux, B.; Schlenkrich, B.; Smith, J.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Karplus, M. *J. Phys. Chem. B* **1998**, 102, 3586-3616.
 - 49) Marten, B.; Kim, K.; Cortis, C.; Friesner, R. A.; Murphy, R. B.; Ringnalda, M. N.; Sitkoff, D.; Honig, B. *J. Phys. Chem.* **1996**, 100, 11775-11788.

- 50) Menegon, G.; Shimizu, K.; Farah, J. P. S.; Dias, L. G.; Chaimovich, H. *Phys. Chem. Chem. Phys.* **2002**, 4, 5933-5936.
- 51) Mennucci, B.; Cammi, R.; Tomasi, J. *J. Chem. Phys.* **1998**, 109, 2798-2807.
- 52) Mennucci, B.; Cancès, E.; Tomasi, J. *J. Phys. Chem. B* **1997**, 101, 10506-10517.
- 53) Miertus, S.; Scrocco, E.; Tomasi, J. *Chem. Phys.* **1981**, 55, 117-129.
- 54) Mortier, W. J.; Genechten, K. V.; Gasteiger, J. *J. Am. Chem. Soc.* **1985**, 107, 829-835.
- 55) Mortier, W. J.; Ghosh, S. K.; Shankar, S. *J. Am. Chem. Soc.* **1986**, 108, 4315-4320.
- 56) Mulliken, R. S. *J. Chem. Phys.* **1955**, 23, 1833-1840.
- 57) Nina, M.; Beglov, D.; Roux, B. *J. Phys. Chem. B* **1997**, 101, 5239-5248.
- 58) No, K. T.; Grant, A.; Scheraga, H. A. *J. Phys. Chem.* **1990**, 94, 4732-4739.
- 59) Onsager, L. *J. Am. Chem. Soc.* **1936**, 58, 1486.
- 60) Orozco, M.; Luque, F. J. *Chem. Rev.* **2000**, 100, 4187-4225.
- 61) Parr, R. G.; Donnelly, R. A.; Levy, M.; Palke, W. E. *J. Chem. Phys.* **1978**, 68, 3801-3807.
- 62) Parr, R.G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*, Oxford University Press, New York, 1989.
- 63) Pascual-Ahuir, J. L.; Silla, E.; Tunon, I. GEPOL93v8, 1993.
- 64) Pauling, L. *J. Am. Chem. Soc.* **1932b**, 54, 3570-3582.
- 65) Pauling, L. *J. Am. Chem. Soc.* **1932**, 54, 988-1003.
- 66) Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P., *Numerical Recipes in Fortran 77: The Art of Scientific Computing*, second edition, Cambridge University Press, New York, 1992.
- 67) Rappé, A. K.; Goddard III, W. A. *J. Phys. Chem.* **1991**, 95, 3358-3363.
- 68) Ribeiro, M. C. C.; Almeida, L. C. J. *J. Chem. Phys.* **2000**, 113, 4722-4731.
- 69) Roux, B. in *Computational Biochemistry & Biophysics*, Eds. Becker, O.; MacKerell Jr., A. D.; Roux, B. and Watanabe, M., Marcel Dekker, New York, 2001.
- 70) Sanderson, R. T. *Science* **1951**, 114, 670-672.
- 71) Shimizu, K.; Chaimovich, H.; Farah, J. P. S.; Dias, L. G.; Bostick, D. L. *J. Phys. Chem. B* **2004**, 108, 4171-4177.
- 72) Stern, H. A.; Kaminski, G. A.; Banks, J. L.; Zhou, R.; Berne, B. J.; Friesner, R. A. *J. Phys. Chem. B* **1999**, 103, 4730-4737.
- 73) Still, W. C.; Tempczyk, A.; Hawley, R. C.; Hendrickson, T. *J. Am. Chem. Soc.* **1990**, 112, 6127-6129.

- 74) Storer, J. W.; Giesen, D. J.; Cramer, C. J.; Truhlar, D. G. *J. Comput.-Aided Mol. Des.* **1995**, 9, 87-110.
- 75) Su, Y.; Gallicchio, E. *Biophys. Chem.* **2004**, 109, 251-260.
- 76) Tapia, O. in: Ratajczak, H.; Orville-Thomas, W. J. (Eds.), *Molecular Interactions*, vol. 3, Wiley, New York, 1982.
- 77) Tomasi, J.; Persico, M. *Chem. Rev.* **1994**, 94, 2027-2094.
- 78) Wittink, D. R., *The Application of Regression Analysis*, Prentice Hall College Div, 1988.
- 79) Yang, Z. Z.; Wang, C. S. *J. Phys. Chem. A* **1997**, 101, 6315-6321.
- 80) York, D. M.; Yang, W. *J. Chem. Phys.* **1996**, 104, 159-172.
- 81) Zacharias, M. *J. Phys. Chem. A* **2003**, 107, 3000-3004.

CURRICULUM VITAE

Dados Pessoais

Nome: Karina Shimizu

Filiação: Shiro Shimizu e Helena Toyoko Shimizu

Nascimento: 19 de fevereiro de 1975, São Paulo – SP

Educação

Ensino Médio

Escola São Teodoro de Nossa Senhora de Sion, São Paulo, SP, 1992

Ensino Superior

Instituto de Química – Universidade de São Paulo – São Paulo, Brasil

Bacharelado em Química – Opção Industrial, 1998

Instituto de Química – Universidade de São Paulo – São Paulo, Brasil

Mestre em Físico-Química, 2001

Título da Dissertação: Cálculos de Solvatação de Reagentes, Intermediários e Complexos Ativados de Reações de Hidrólise

Instituto de Química – Universidade de São Paulo – São Paulo, Brasil

Doutorado em Físico-Química, 08/2001 – presente

Título do Trabalho: Estudo do Método de Equalização da Eletronegatividade no Cálculo de Energias Livres de Solvatação – GBEEM-ELR

Publicações

- 1) Freitas, A. A.; Shimizu, K.; Dias, L. G.; Quina, F. H. (2005) “A Computational Study of Substituted Flavylum Salts and their Quinonoidal Conjugate-Bases: $S_0 \rightarrow S_1$ Electronic Transition and Absolute pK_a Calculations by DFT and Semiempirical Methods”, *J. Phys. Chem. A*, submetido.

- 2) Paulo, L.; Freitas, A. A.; Shimizu, K.; Quina, F. H.; Maçanita, A. L. (2005) “Novel Ground and Excited State Prototropic Reactivity of a Carboxy-Substituted Hydroxyflavylium Salt”, *J. Phys. Chem. B*, no prelo.
- 3) Shimizu, K.; Freitas, A. A.; Farah, J. P. S.; Dias, L. G. (2005) “Predicting Hydration Free Energies of Neutral Compounds by a New Parameterization of the Polarisable Continuum Model”, *J. Phys. Chem. A*, **109**, 11322-11327.
- 4) Silva, P. F.; Lima, J. C.; Freitas, A. A.; Shimizu, K.; Maçanita, A. L.; Quina, F. H. (2005) “Charge-Transfer Complexation as a General Phenomenon in the Copigmentation of Anthocyanins”, *J. Phys. Chem. A*, **109**, 7329-7338.
- 5) Shimizu, K.; Chaimovich, H.; Farah, J. P. S.; Dias, L. G.; Bostick, D. L. (2004) “Calculation of the Dipole Moment for Polypeptides Using the Generalized Born-Electronegativity Equalization Method: Results in Vacuum and Continuum-Dielectric Solvent”, *J. Phys. Chem. B*, **108**, 4171-4177.
- 6) Menegon, G.; Shimizu, K.; Farah, J. P. S.; Dias, L. G.; Chaimovich, H. (2002) “Parameterization of the Electronegativity Equalization Method Based on the Charge Model 1”, *Phys. Chem. Chem. Phys.*, **4**, 5933-5936.
- 7) Dias, L. G.; Shimizu, K.; Farah, J. P. S.; Chaimovich, H. (2002) “A Simple Method For the Fast Calculation of Charge Redistribution of Solutes in a Implicit Solvent Model”, *Chem. Phys.*, **282**, 237-243.
- 8) Nery, A. L. P.; Quina, F. H.; Moreira Jr., P. F.; Medeiros, C. E. R.; Baader, W. J.; Shimizu, K.; Catalani, L. H.; Bechara, E. J. H. (2001) “Does the Photochemical Conversion of Colchicine Into Lumicolchicines Involve Triplet Transients? A Solvent Dependence Study”, *Photochem. Photobiol.*, **73** (3), 213-218.

Comunicações em Reuniões Científicas

- 1) Paulo, L.; Freitas, A. A.; Shimizu, K.; Quina, F. H.; Maçanita, A. L., *Novel Ground and Excited State Prototropic Reactivity of a Hydroxycarboxyflavylium Salt*, 9th International Conference on Methods and Applications of Fluorescence – Spectroscopy, Imaging and Probes, 2005, Lisboa, Portugal.
- 2) Carpentieri, V. M.; Shimizu, K.; Dourado, A. G.; Sório, C. S. N. M.; Farah, J. P. S., *Estudo da solvatação de solutos neutros em misturas solvente orgânico/água e soluções aquosas de micelas*, 27^a Reunião Anual da Sociedade Brasileira de Química, 2004, Salvador, BA.

- 3) Shimizu, K.; Dias, L. G.; Farah, J. P. S., *Semiempirical pK_a Predictions of *m*-Substituted Benzoic Acids in Acetonitrile-Water Mixtures*, 7^a Conferência Latinoamericana de Físico-Química Orgânica, 21/09/2003-26/09/2003, Florianópolis, SC.
- 4) Dias, L. G.; Shimizu, K.; Farah, J. P. S.; Chaimovich, H., *A Simple Method for the Fast Calculation of Charge Redistribution of Solutes in an Implicit Solvent Model*, 1^o Workshop de Modelagem Molecular de Riberão Preto, 06/09/02, Riberão Preto, SP.
- 5) Menegon, G.; Dias, L. G.; Farah, J. P. S.; Shimizu, K., *Reparameterization Of A Simple And Fast Atomic Charge Model*, Encontro Satélite ao Congresso Internacional de Biofísica promovido pela IUPAB (Workshop On Molecular Modeling In Biophysics – Methods And Aplications), 03/05/02-05/05/02, Rio de Janeiro, RJ.
- 6) Shimizu, K.; Sório, C. S. N. M.; Dourado, A. G.; Carmo, A.; Monteiro, V. P.; Farah, J. P. S., *Análise de Grupos em Estudos de Partição de Solutos Neutros em Sistemas Bifásicos e Soluções Micelares por LSER e Redes Neurais Artificiais*, Simpósio Química e Meio Ambiente – Estamos Fazendo Nossa Parte?, 25/03/02 e 26/03/02, São Paulo, SP.
- 7) Shimizu, K.; Ishii, J. T.; Farah, J. P. S., *Solvation Energy Calculations on Solvent Mixtures*, 11^o Simpósio Brasileiro de Química Teórica, 2001, Caxambu, MG.
- 8) Shimizu, K.; Sório, C. S. N. M.; Dourado, A. G.; Monteiro, V. P.; Farah, J. P. S., *Cluster Analysis of LSER and Neural Net Residuals in Two-Phase Systems Partitions and Micellar Solubilization of Neutral Molecules*, 3rd Workshop on Chemical Structure and Biological Activity: Perspectives in QSAR, 08/11/01-10/11/01, São Paulo, SP.
- 9) Bechara, E. J. H.; Farah, J. P. S.; Quina, F. H.; Catalani, L. H.; Shimizu, K.; Nery, A. L. P., *Conversão de Colchicina em Lumicolchicinas em Plantas: Fotoquímica na Ausência de Luz?*, 23^a Reunião Anual da Sociedade Brasileira de Química, 2000, Poços de Caldas, MG.
- 10) Shimizu, K.; Farah, J. P. S., *Cálculos de Solvatação de Intermediários Reativos e de Complexos Ativados*, 6^a Simpósio de Iniciação Científica da Universidade de São Paulo, 1998, São Carlos, SP.

- 11) Shimizu, K.; Meneghelli, P.; Farah, J. P. S., *Cálculos Semipíricos e Ab Initio de Estados de Transição*, 21^a Reunião Anual da Sociedade Brasileira de Química, 1998, Poços de Caldas, MG.
- 12) Shimizu, K.; Meneghelli, P.; Dourado, A. G.; Dias, L. G.; Farah, J. P. S., *Solvatação e Estrutura Molecular. LSER em Sistemas Organizados e Misturas de Solventes*, Sessão da Academia Brasileira de Ciências – Sessão Regular de Ciências Químicas, 1997, São Paulo, SP.
- 13) Shimizu, K.; Meneghelli, P.; Farah, J. P. S., *Cálculos Semi-empíricos e Reatividade de Hidrólise de Ésteres*, 14^a Semana da Química Prof. Dr. Etelvino J. H. Bechara, 1997, São Paulo, SP.
- 14) Shimizu, K.; Meneghelli, P.; Farah, J. P. S., *Estrutura Molecular e Solubilidade*, 5^a Simpósio de Iniciação Científica da Universidade de São Paulo, 1997, São Paulo SP.
- 15) Shimizu, K.; Farah, J. P. S., *Estrutura Molecular e Solubilidade*, 20^a Reunião Anual da Sociedade Brasileira de Química, 1997, Poços de Caldas, MG.