

• 研究论文 •

分子亲脂-亲水性的量子化学描述 II. 氨基酸侧链的亲水指标和亲脂指标

杜奇石^{*,a,b} 刘朋军^a 孙浩^b 李大鹏^b 谢军民^b

(^a海南师范大学化学系 海口 571158)

(^b天津师范大学生物信息与药物开发研究所 天津 300074)

摘要 在启发式亲脂势 HMLP (heuristic molecular lipophilicity potential)的基础上提出了分子、分子片段和原子的亲水指标和亲脂指标. 计算出了 20 个天然氨基酸侧链的亲水、亲脂指标和亲水、亲脂表面积, 并用线性自由能函数表达氨基酸侧链的溶剂化自由能, $\Delta G_{\text{sol},i}^{\ominus} = b_0 + b_1L_i + b_2H_i + b_3S_i^+ + b_4S_i^-$. 应用线性自由能函数和氨基酸侧链的亲水和亲脂指标, 计算了 20 个氨基酸残基的 3 种相转移自由能(蒸气-水、蒸气-正辛醇、正辛醇-水)和正辛醇-水分配系数 $\log P_{\text{ow}}$, 取得了与实验值高度一致的良好效果. HMLP 的亲水和亲脂指标是 HMLP 的指标化, 扩展了这一方法的使用范围. 氨基酸侧链的亲水、亲脂指标和线性自由能函数有望用于生物大分子受体与配体的结合自由能的估算、蛋白质的结构与功能、蛋白-蛋白相互作用和识别的研究.

关键词 启发式亲脂势(HMLP); 亲脂指标; 蛋白质; 氨基酸; 生物信息学

Quantum Chemical Description for Molecular Lipophilicity and Hydrophilicity: II. Lipophilic Indices and Hydrophilic Indices of Aminoacid Side Chains

DU, Qi-Shi^{*,a,b} LIU, Peng-Jun^a SUN, Hao^b LI, Da-Peng^b XIE, Jun-Min^b

(^a Department of Chemistry, Hainan Normal University, Haikou 571158)

(^b Institute of Bioinformatics and Drug Discovery, Tianjin Normal University, Tianjin 300074)

Abstract Heuristic molecular lipophilicity potential (HMLP) was applied to the study of lipophilicity and hydrophilicity of 20 natural aminoacid side chains. The HMLP parameters, surface area A_i , lipophilic indices L_i and hydrophilic indices H_i of aminoacid side chains were derived from lipophilicity potential $L(\mathbf{r})$. The parameters were correlated with the experimental data of phase transfer free energies of vapor-to-water, vapor-to-octanol, and octanol-to-water through a linear free energy equation $\Delta G_{\text{sol},i}^{\ominus} = b_0 + b_1L_i + b_2H_i + b_3S_i^+ + b_4S_i^-$. For all above 3 types of transfer free energies between different medium phases, the HMLP parameters of 20 aminoacid side chains provided good descriptions. HMLP is an *ab initio* quantum chemical approach and a structure-based technique. Except for atomic van der Waals radii, there was no other empirical parameter used. The computational results of HMLP have clear physical and chemical meaning and could provide useful lipophilic and hydrophilic parameters for the studies of proteins and peptides.

Keywords heuristic molecular lipophilicity potential (HMLP); lipophilicity index; protein; aminoacid; biological informatics

* E-mail: duqishi@yahoo.com

Received May 21, 2005; revised July 15, 2005; accepted September 20, 2005.

国家自然科学基金(No. 20373048)和天津市科委自然科学基金资助项目.

20 种天然氨基酸是蛋白质的基本构件, 氨基酸侧链的疏水性和亲水性在生命科学中占有重要地位, 在蛋白质的空间结构与其所表达的生物学功能的研究中起重要作用, 也是许多药物设计和生物分子模拟计算的重要参数^[1,2]. 氨基酸侧链的疏水与亲水性是蛋白质折叠和三维结构预测的决定性因素^[3,4]. 新近的研究发现, 蛋白质在细胞中的位置是由信号引肽的氨基酸残基的疏水性和亲水性的差异决定的^[5,6]. 细胞膜中的各种水通道和离子通道^[7,8]的发现, 更加激发了对氨基酸侧链的疏水与亲水性的研究. 疏水性和亲水性在生命科学的遗传密码的解析中也受到越来越多的重视^[9-11]. Weber 等^[12]和 Lacey 等^[13]分别报道了氨基酸的疏水性与相应的反密码子的疏水性的直接关系, 核苷酸与氨基酸之间的疏水和亲水部分的相互偶合是引发二者间相互制约的重要因素. 准确计算描述氨基酸侧链的疏水和亲水指标, 在生命科学研究中有重要意义.

杜奇石等^[14-19]提出的分子的启发式亲脂势 HMLP (heuristic molecular lipophilicity potential)用量子化学方法描述分子局部的亲水-亲脂性, 有明确的物理意义和严格的计算方法, 除原子的范德华半径外, 不使用其它经验参数, 与大量使用经验参数和复杂原子类型分类法各种经验方法相比, 有明显的优势. 杜奇石等最近从量子力学和统计力学的基本原理出发, 完成了 HMLP 的屏蔽函数的理论推导^[14], 使 HMLP 建立在坚实的理论基础上. 但是, HMLP 方法依赖于量子化学计算, 要把这一方法应用于真实的生物大分子体系, 重要的一步是推导出描述分子片段, 特别是氨基酸残基的亲水、亲脂指标. 本文的目的是在 HMLP 最新进展的基础上, 计算 20 种氨基酸侧链的亲水和亲脂指标, 并使用这些参数计算 3 种相转移自由能, 与实验值相对比, 以证实氨基酸侧链的亲水、亲脂指标和 HMLP 方法的有效性.

1 理论方法

分子的启发式亲脂势 HMLP 认为分子的亲水性和亲脂性是分子自身的性质, 由分子的电子结构决定, 可以从分子表面的静电势分布来计算. 从量子力学和统计力学的基本原理出发, 推导出了溶质分子在极性溶剂中的统一的亲水-亲脂势能函数^[14],

$$u_{\mu-\mu}(\mathbf{r}) = \frac{\mu^4}{(3kT)^2 d_0^2 \epsilon^3} V_\alpha(\mathbf{r}) \left[\sum_{i \neq \alpha}^M \frac{V_i(\mathbf{r}_i)}{|R_i - \mathbf{r}|^3} \right] \quad (1)$$

式中 \mathbf{r} 是溶质分子的原子 α 的表面 S_α 上的一点, $V_\alpha(\mathbf{r})$ 是该点的静电势, $V_i(\mathbf{r}_i)$ 是溶质分子的其它原子 i 在其表面上点 \mathbf{r}_i 的静电势, R_i 是原子 i 的核坐标, 求和号上的 M 是

溶质分子中的原子数目, 不包括原子 α , d_0 是第一水化层的厚度, ϵ 是介电常数, μ 是溶剂分子的偶极矩. HMLP 采用从头算量子化学法计算静电势,

$$V(\mathbf{r}) = \sum_{\alpha} \frac{Z_{\alpha}}{|R_{\alpha} - \mathbf{r}|} - \int_{\infty}^{\rho(\mathbf{r}')} \frac{\rho(\mathbf{r}')}{|\mathbf{r}' - \mathbf{r}|} d\mathbf{r}' \quad (2)$$

式中 $\rho(\mathbf{r}')$ 是分子的单电子密度函数. 式(1)表达了被溶质分子的原子 α 吸引的溶剂偶极子 μ 与被其它原子 ($i=1, \dots, M, i \neq \alpha$) 吸引的溶剂偶极子 μ 间的静电作用能 $u_{\mu-\mu}(\mathbf{r})$. 如果原子 α 的静电势 $V_\alpha(\mathbf{r})$ 与周围原子的静电势 $V_i(\mathbf{r})$ 的极性相反, 溶剂偶极子间的静电作用能 $u_{\mu-\mu}(\mathbf{r})$ 是负值, 原子 α 表面的点 \mathbf{r} 为亲水性; 反之, 若 $V_\alpha(\mathbf{r})$ 与 $V_i(\mathbf{r})$ 的极性相同, 溶剂分子间的静电作用能 $u_{\mu-\mu}(\mathbf{r})$ 是正值, 原子 α 表面的点 \mathbf{r} 为疏水性. 式(1)从理论上证明了我们早先提出的启发式亲脂势^[15-19],

$$L(\mathbf{r}) = V(\mathbf{r}) \sum_{i \neq \alpha}^M \zeta \frac{b_i}{|R_i - \mathbf{r}|^\gamma} \quad (3)$$

与式(1)对比, 式(3)用原子 i 的 ESP (electrostatic potential) 描述符^[15-19] b_i 代替了静电势 $V_i(\mathbf{r})$,

$$b_i = \int_{S_i} V(\mathbf{r}) dS \approx \sum_{k \in S_i} V(\mathbf{r}_k) \Delta S_k \quad (4)$$

式中 S_i 是原子 i 在分子中暴露的表面积, ΔS_k 是面积元. 原子的 ESP 描述符 b_i 含有原子的静电势和表面积两方面的因素, 可以看作是以原子暴露表面积为权重的原子静电势, 全面表达了分子中的原子局部的静电作用的能力. 式(1)给出了式(3)中的两个参数的数值, $\gamma-3$, ζ 与温度 T 、溶剂的偶极矩 μ 、介电常数 ϵ 和水化层的厚度 d_0 有关,

$$\zeta = \frac{\mu^4}{(3kT)^2 d_0^2 \epsilon^3} \quad (5)$$

亲脂势 $L(\mathbf{r})$ 与温度的平方成反比, 与溶剂偶极矩 μ 的 4 次方成正比. 在以上推导中使用了一些近似条件, 因而是近似的. 在 HMLP 中我们把 ζ 和 γ 作为经验参数, 初步的优化表明 $\gamma=2.3 \sim 2.5$ 的效果较好, ζ 值留待与实验值对比后确定.

为了表达分子中的原子的亲脂性, 我们定义如下原子的亲脂势指标,

$$l_i = \int_{S_i} L(\mathbf{r}) dS \approx \sum_{k \in S_i} L(\mathbf{r}_k) \Delta S_k \quad (6)$$

式中 S_i 是原子 i 的暴露表面积. 原子的亲脂势指标含有

原子亲脂势的强弱和表面积大小两方面的因素,能全面说明分子中的某类原子的亲脂和亲水性.原子的亲脂势指标是亲脂势 HMLP 的指标化,在同类化合物中有加和性和可转移性.还可以定义分子亲脂指标,

$$L_M = \sum_{\substack{i \\ (l_i > 0)}} l_i \quad (7)$$

和亲水指标,

$$H_M = \sum_{\substack{i \\ (l_i < 0)}} l_i \quad (8)$$

分子的亲脂指标 L_M 和亲水指标 H_M 分别是分子中所有亲脂性原子的亲脂势指标 ($l_i > 0$) 和所有亲水性原子的亲脂势指标 ($l_i < 0$) 的和.类似地,我们还可以定义一类分子的某个取代基或分子片段的亲脂和亲水指标.

2 氨基酸侧链的亲水和亲脂指标

20 个氨基酸取中性未电离形式,几何构型用 HF/6-31G*方法优化.分子的范德华表面用程序 MS^[21,22]构造,由正三角形面积元组成,点密度为 2500/nm².氨基酸中的 C, H, O, N 和 S 原子的范德华半径^[20]分别为 0.200, 0.117, 0.139, 0.155 和 0.178 nm,所有的物理量用原子单位.用 Gaussian 03 程序^[23]和 HF/6-31G**方法计算面积元的中心的静电势,计算结果列于表 1.

分子的范德华表面由正三角形面积元组成,用 Gaussian03 程序^[23]和 HF/6-31G**方法计算面积元的中心的静电势.在本文中 S^+ 表示氨基酸侧链的亲脂性表面积, S^- 是亲水性表面积, S 是侧链的总表面积. L 是氨基酸侧链的亲脂指标, H 是亲水指标, W 表示水相, O 表示正辛醇相.表 1 的最后一列是 20 种氨基酸残基的正辛醇-水分配系数 $\log P_{O/W}$ 取自文献^[27].从表 1 看出,有 6 个氨基酸侧链的亲水指标 H 是零 (Leu, Ile, Val, Ala, Gly, Pro), 氨基酸 Phe 的侧链的亲水指标非常小 ($H_{Phe} = -0.1195$), 以上 7 种氨基酸是疏水性氨基酸.大多数氨基酸的侧链既有非零的亲脂指标 L , 又有非零的亲水指标 H . 被认为最亲水的氨基酸 Arg 有最大的亲水指标 ($H_{Arg} = -1.4797$), 同时有很大的亲脂指标 ($L_{Arg} = 1.2424$). 这是因为它的侧链 $(CH_2)_3NH(C=NH)NH_2$ 中既有强大的亲水性基团, 又有疏水性很强的基团 $(CH_2)_3$, 可以认为是两亲性氨基酸.类似的两亲性氨基酸还有 His, Gln, Lys, Asn, Glu 和 Asp.

一般认为溶剂化自由能 ΔG_{sol}^\ominus 由三部分组成: 静电作用、疏水作用和洞穴贡献^[25],

$$\Delta G_{sol}^\ominus = \Delta G_{elec}^\ominus + \Delta G_{hdr}^\ominus + \Delta G_{cav}^\ominus \quad (9)$$

表 1 20 个氨基酸侧链的亲脂表面积、亲水表面积、亲脂指标和亲水指标^a

Table 1 Positive surface area, negative surface area, lipophilic indices and hydrophilic indices of 20 aminoacids side chains^a

AA	S^+/nm^2	S^-/nm^2	S/nm^2	L	H	$\log P_{O/W}^b$
Leu	0.8455	0.0000	0.8455	1.2906	0.0000	2.167
Ile	0.8861	0.0000	0.8861	1.1046	0.0000	2.167
Val	0.7781	0.0000	0.7781	0.5324	0.0000	1.640
Phe	1.0571	0.1125	1.1695	0.4412	-0.1195	2.423
Met	0.7036	0.2323	0.9359	1.0768	-0.3068	1.246
Trp	1.3370	0.1488	1.4858	0.8364	-0.4310	1.878
Ala	0.3478	0.0000	0.3478	0.1744	0.0000	0.702
Cys	0.2356	0.3045	0.5401	0.2479	-0.2402	0.987
Gly	0.0376	0.0000	0.0376	0.0208	0.0000	0.184
Tyr	0.8097	0.4272	1.2368	0.4534	-0.5896	1.887
Thr	0.4673	0.1605	0.6278	1.4265	-0.4369	-0.042
Ser	0.2607	0.1596	0.4203	0.2346	-0.6040	-0.453
His	0.8217	0.1386	0.9603	0.8124	-0.7766	-1.321
Gln	0.7009	0.1787	0.8795	1.0036	-0.7211	-0.936
Lys	0.9771	0.0808	1.0579	1.4600	-0.6229	-0.790
Asn	0.5051	0.1778	0.6829	0.6396	-0.7211	-1.003
Glu	0.5716	0.2557	0.8273	1.0315	-0.9298	-1.868
Asp	0.3742	0.2527	0.6269	0.6058	-0.9298	-1.935
Arg	0.9080	0.3531	1.2611	1.2424	-1.4797	-2.061
Pro	0.6923	0.0000	0.6923	0.3226	0.0000	1.128

^a S^+ : Surface area of side chain with positive HMLP; S^- : Surface area of side chain with negative HMLP; S : Total surface area of side chain; W : Water phase; O : *n*-Octanol phase; ^b Ref [27].

分子的亲水指标 H_i 对静电作用 ΔG_{elec}^\ominus 有最大贡献, 分子的疏水指标 L_i 对疏水作用 ΔG_{hdr}^\ominus 有最大贡献, 洞穴部分的作用能 ΔG_{cav}^\ominus 应与分子的表面积有关. 为此, 我们把溶剂化自由能表达为 HMLP 参数的线性函数,

$$\Delta G_{sol,i}^\ominus = b_0 + b_1 L_i + b_2 H_i + b_3 S_i^+ + b_4 S_i^- \quad (10)$$

在式(10)中我们把表面积分为亲水 S^- 和疏水 S^+ 两部分, 这是因为它们对洞穴作用自由能 ΔG_{cav}^\ominus 可能有不同的贡献. 从热力学角度考虑, 溶剂化自由能可以分解为焓和熵两部分的贡献 $\Delta G_{sol}^\ominus = \Delta H_{sol}^\ominus - T\Delta S_{sol}^\ominus$. HMLP 的亲水指标 H_i 对焓 ΔH_{sol}^\ominus 有最大贡献, 亲脂指标 L_i 和表面积 S_i 共同对熵 ΔS_{sol}^\ominus 做出贡献.

根据表 1 中的数据, 我们用最小二乘广义逆矩阵法计算 20 种氨基酸侧链的三种相转移自由能的线性回归系数 b_0, b_1, b_2, b_3 和 b_4 , 计算结果列于表 2. 表 2 中 R^2 是相关系数的平方, S^2 是标准差的平方. 计算结果表明, 20 种氨基酸侧链的亲水和亲脂指标与三种相转移自由能有很高的相关性.

根据热力学原理, 表 2 中的三种相转移自由能间有

如下关系,

$$\Delta G_{O-W}^{\ominus} = \Delta G_{V-W}^{\ominus} - \Delta G_{V-O}^{\ominus} \quad (11)$$

我们发现, 表 2 中的三种线性相转移自由能计算公式(10)的 4 个系数间也有类似关系, $b_k(O \rightarrow W) \approx b_k(V \rightarrow O) - b_k(V \rightarrow W)$ ($k=0, 1, 2, 3, 4$).

应用线性自由能函数式(10)和表 2 中的系数及表 1 中的 HMLP 参数, 计算 20 种氨基酸残基的三种相转移自由能, 表 3 列出了三种相转移自由能的 HMLP 计算值和实验值的对比. 有几种氨基酸残基没有实验值, 由于公式(1)适用于各种极性溶剂, 三种相转移自由能的计算值与实验值有很好的—致性. 20 种氨基酸的 HMLP 参数还可以用于氨基酸残基的正辛醇-水的分配系数的计算. 假设氨基酸侧链的 HMLP 参数与它们的正辛醇-水分配系数 $\log P_{O/W}$ 有如下线性关系,

$$\log P_{O/W} = \log \frac{C_o}{C_w} = a_0 + a_1 L + a_2 H + a_3 S^+ + a_4 S^- \quad (12)$$

用表 1 的 20 种氨基酸侧链的 HMLP 参数和最小二乘回归计算得到式(12)中的系数 $a_0=0.0301$, $a_1=0.0218$, $a_2=0.0604$, $a_3=-0.1528$ 和 $a_4=0.4414$, 相关系数和标准差分别为 $R^2=0.9519$ 和 $s^2=0.1394$. 图 1 是用式(12)计算的 20 个氨基酸残基的正辛醇-水分配系数的计算值与文献值的对比.

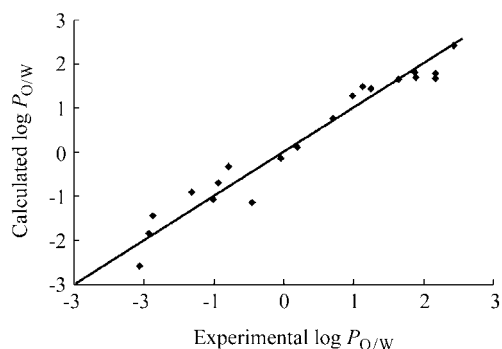


图 1 20 个氨基酸残基的正辛醇-水分配系数 $\log P_{O/W}$ 的 HMLP 计算值与实验值的对比

Figure 1 Comparison of experimental and HMLP calculation partition coefficients $\log P_{O/W}$ of 20 amino acid residues from *n*-octanol to water

表 2 20 种氨基酸残基的三种相转移自由能线性函数的统计计算结果

Table 2 Statistical calculation results for 3 types of phase transfer free energies of 20 amino acid residues

Transfer	b_0	$b_1 (S_i^+)$	$b_2 (S_i^-)$	$b_3 (L_i)$	$b_4 (H_i)$	R^2	S^2
V→W	12.189	-0.07475	0.09904	0.0055	62.963	0.9819	15.32
V→O	12.012	-0.17826	0.09608	1.8466	56.197	0.9784	16.21
O→W	0.0177	0.09364	0.00296	-1.8411	6.766	0.8888	3.017

3 结论

从量子力学和统计力学的基本原理出发推导出的统一的亲水-亲脂势能函数式(1)揭示了分子亲脂势的物理意义, 给出了分子局部的亲水和亲脂性的统一的计算方法. 溶质分子表面交替分布的异号静电势造成了溶质分子与极性溶剂分子及溶剂分子间的吸引相互作用, 表现为亲水性; 溶质分子表面同号分布的静电势造成了极性溶剂分子间的相互排斥作用, 表现为疏水性. 由于在疏水表面的水分子间的排斥作用, 水分子离开疏水表面, 取高度有序的排列方式^[26], 产生负的熵效应. 启发式亲脂势 $L(\mathbf{r})$ 可以看作是这种疏水排列驱动力的度量. 由公式(3)和(4)定义的分子的亲脂势 $L(\mathbf{r})$ 实质上是表达了相对极性的静电势 $V(\mathbf{r})$, 亲脂势 HMLP 保留了静电势描述静电相互作用的能力, 包括氢键、盐键、偶极和 Lewis 酸碱等, 我们把这些性质归于亲水性. 亲脂势 HMLP 用同号的静电势描述疏水性, 把静电势的应用范围扩展到疏水作用. 亲脂势 HMLP 用统一的方法计算分子的亲水性和亲脂性, 概括了分子间的多种作用因素, 与其它方法相比, 在受体与配体结合自由能的估算上有明显的优势.

本文在启发式亲水-亲脂势 HMLP 的基础上提出了分子、分子片段和原子的亲水和亲脂指标, 计算出了 20 个天然氨基酸侧链的亲水、亲脂指标和亲水、亲脂表面积. 与大多数经验的亲脂指标不同, HMLP 分别用亲水指标和亲脂指标描述分子片段的亲水部分和亲脂部分, 有明确的物理意义, 能说明较多的实验现象. 应用亲水和亲脂指标, 提出了计算溶剂化自由能的线性自由能函数 $\Delta G_{sol,i}^{\ominus} = b_0 + b_1 L_i + b_2 H_i + b_3 S_i^+ + b_4 S_i^-$. 把线性自由能函数和 20 种氨基酸的 HMLP 的亲水和亲脂指标用于 3 种相转移自由能(蒸气-水、蒸气-正辛醇、正辛醇-水)的计算, 取得了与实验值高度一致的结果. HMLP 的亲水和亲脂指标是 HMLP 的指标化, 扩展了这一方法的使用范围. 20 种氨基酸是构造蛋白质的基本构件, 氨基酸的侧链的亲水、亲脂指标和线性自由能函数有望用于配体与生物大分子受体的结合自由能的估算、蛋白质的结构与功能、蛋白-蛋白相互识别和作用的研究.

表3 20个氨基酸残基的相转移自由能的HMLP计算值与实验值的对比^{a-c}Table 3 Comparison of phase transfer free energies of 20 amino acid residues between HMLP calculation and experimental data^{a-c}

AA	V→W		V→O		O→W	
	Expt.	Calcd.	Expt.	Calcd.	Expt.	Calcd.
Leu	9.54	5.876	2.18	0.6758	7.37	6.627
Ile	9.00	5.571	0.460	1.743	8.54	7.248
Val	8.33	6.375	3.39	0.8750	4.94	6.980
Phe	-3.18	-2.120	-11.93	-11.651	8.75	9.924
Met	-6.20	-10.081	-11.72	-13.552	5.53	5.38
Trp	-24.61	-23.464	-35.12	-33.067	10.51	9.210
Ala	8.12	9.5903	5.94	6.135	2.18	3.415
Cys	-5.19	-1.678	—	-2.302	—	3.730
Gly	10.00	11.908	10.00	11.380	0.00	0.7102
Tyr	-25.58	-26.754	-32.40	-30.613	6.82	6.750
Thr	-20.43	-17.216	-21.56	-16.694	1.13	0.3909
Ser	-21.18	-26.208	-21.35	-24.611	0.167	-1.779
His	-42.99	-41.473	-46.97	-43.446	3.976	0.6200
Gln	-39.26	-36.678	-38.97	-37.436	-0.293	0.3550
Lys	-39.85	-33.527	-40.18	-36.939	0.335	2.103
Asn	-40.52	-35.225	-40.48	-34.626	-0.042	-1.017
Glu	-42.86	-48.089	-39.56	-46.067	-3.31	-2.090
Asp	-45.84	-48.626	—	-43.363	—	-3.425
Arg	-83.38	-84.262	-77.86	-81.642	-5.53	-3.759
Pro	—	7.0157	—	0.2672	—	6.449

^a Transfer free energies are in kJ/mol; ^b Experimental data are taken from Ref. [24]; ^c The numbers in bold are predicted values.

References

- Adams, R. L.; Knowler, J. T.; Leader, D. P. *The Biochemistry of the Nucleic Acids*, 11th ed., Chapman & Hall, New York, **1992**, pp. 515~585.
- Adessi, C.; Soto, C. *Curr. Med. Chem.* **2002**, *9*, 963.
- Edsall, J. T.; McKenzie, H. A. *Adv. Biophys.* **1983**, *16*, 53.
- Venu, K.; Svensson, L. A.; Halle, B. *Biophys. J.* **1999**, *77*, 1074.
- Nielsen, H.; Brunak, S.; Heijne, G. V. *Protein Eng.* **1999**, *12*, 3.
- Chou, K. C.; Howe, W. J. *Biochem. Biophys. Res. Commun.* **2002**, *292*, 702.
- Agre, P.; King, L. S.; Yasui, M.; Guggino, W. B.; Ottersen, O. L.; Fujiyoshi, Y.; Engel, A.; Nielsen, S. *J. Physiol.* **2002**, *542*, 3.
- Doyle, D. A.; Morais, C. J.; Pfuetzner, R. A.; Kuo, A.; Gulbis, J. M.; Cohen, S. L.; Chait, B. T.; MacKinnon, R. *Science* **1998**, *280*, 69.
- Wang, K. *Chemical Questions in Life Science*, Peking University Press, Beijing, **1990**, p. 56 (in Chinese). (王夔, 生命科学中的化学问题, 北京大学出版社, 北京, **1990**, p. 56.)
- Wang, W.-Q. *Life Science*, Beijing Industry University Press, Beijing, **2001**, p. 43 (in Chinese). (王文清, 生命科学, 北京工业大学出版社, 北京, **2001**, p. 43.)
- Yan, L.-F.; Sun, Z.-R. *Molecular Structure of Protein*, Tsinghua University Press, Beijing, **1999**, p. 18 (in Chinese). (阎隆飞, 孙之荣, 蛋白质的分子结构, 清华大学出版社, 北京, **1999**, p. 18.)
- Weber, A. L.; Lacey, J. C., Jr. *J. Mol. Evol.* **1978**, *11*, 199.
- Lacey, J. C., Jr.; Mullins, D. W., Jr. *Orig. Life* **1983**, *13*, 3.
- Du, Q.; Liu, P.-J.; Mezey, P. G. *J. Chem. Inf. Model.* **2005**, *45*, 347.
- Du, Q.; Mezey, P. G.; Chou, K. C. *J. Comput. Chem.* **2005**, *26*, 461.
- Du, Q.-S.; Wei, D.-Q.; Li, A.-X. *Chem. J. Chin. Univ.* **2004**, *25*, 1706 (in Chinese). (杜奇石, 魏冬青, 李爱秀, 高等学校化学学报, **2004**, *25*, 1706.)
- Du, Q.; Mezey, P. G. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 451.
- Du, Q.; Mezey, P. G.; Arteca, G. A. *J. Comput.-Aided Mol. Des.* **1997**, *11*, 503.
- Du, Q.; Arteca, G. A. *J. Comput.-Aided Mol. Des.* **1996**, *10*, 133.
- Du, Q.; Arteca, G. A. *J. Comput. Chem.* **1996**, *17*, 258.
- Connolly, M. L. *Science* **1983**, *221*, 709.
- Connolly, M. L. *J. Appl. Crystallogr.* **1983**, *16*, 548.
- Foresman, J. B.; Frisch, A. *Exploring Chemistry with Electronic Structure Methods*, 2nd Ed., Gaussian, Pittsburgh, PA, **1995**.
- Radzicka, A.; Wolfenden, R. *Biochemistry* **1988**, *27*, 1664.
- Smith, B. J. *J. Comput. Chem.* **1999**, *20*, 428.
- Head-Gordon, T. *Proc. Natl. Acad. Sci. U. S. A.* **1995**, *92*, 8308.
- Wu, H. *Molecules* **1999**, *4*, 16.