

## 胆红素分子激发态性质的密度泛函理论研究\*

李媛媛 胡竹斌 孙海涛<sup>†</sup> 孙真荣<sup>‡</sup>

(华东师范大学精密光谱科学与技术国家重点实验室, 上海 200241)

(2020年4月8日收到; 2020年5月18日收到修改稿)

胆红素是人体胆汁中的主要色素, 与人体健康密切相关. 结合荧光蛋白的胆红素分子代表一类新型荧光发色团, 在生物成像和生物传感领域有着重要应用. 本文结合隐式溶剂模型和线性响应含时密度泛函理论方法计算了胆红素分子最低单重激发态的垂直激发能、振子强度和垂直发射能. 以实验测量值和高水平 RI-ADC(2) 计算值作为参考, 系统考察了一系列密度泛函方法的预测表现, 结果发现最优化调控区间分离密度泛函方法整体表现最优, 预测的绝对误差和相对误差最小. 这得益于泛函中包含适宜的准确交换项比例能够产生既不离域也不局域的电子结构. 在最优化调控泛函方法计算的波函数基础上, 基于空穴-电子分析和片段间电荷转移方法定量表征了胆红素分子最低单重激发态, 发现其具有杂化局域-电荷转移的激发特征. 相信本工作可以为今后研究胆红素分子的激发态动力学过程和光谱性质提供重要理论依据, 该最优化调控理论模型也可以为接下来其他生物分子体系的激发态性质研究提供可靠、高效的理论工具.

**关键词:** 胆红素, 密度泛函理论, 最优化调控, 激发态特征**PACS:** 31.15.ee, 31.15.E-, 31.15.es**DOI:** 10.7498/aps.69.20200518

## 1 引言

胆红素 (bilirubin) 是人胆汁中的主要色素, 在临床上是判定新生儿黄疸的重要依据, 也是肝功能的重要指标<sup>[1,2]</sup>. 胆红素在人体内浓度过高时有毒性, 会引起对大脑和神经系统不可逆的损害. 如图 1 所示, 胆红素分子是由两个对称的吡咯-吡咯酮单元通过亚甲基连接而成的, 它的构象多变, 主要依赖于分子内部和溶剂环境中形成的氢键, 以及两个羧基基团的电离状态<sup>[3,4]</sup>. 例如, 在非极性溶剂 (如氯仿等) 中, 分子内部形成六个分子内氢键, 将整个分子卷为“脊瓦状” (“ridge-tile”) 的刚性结构<sup>[5,6]</sup>; 在极性溶剂 (如二甲基亚砜等) 中, 部分分子内氢键被打破, 能同时存在分子内氢键和与溶剂分子形成的分子间氢键<sup>[7]</sup>; 而在碱性溶液中, 由于分子处

于完全电离状态而无分子内氢键形成, 呈现出充分伸展的构型<sup>[7,8]</sup>.

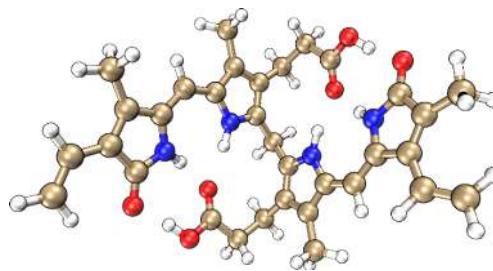


图 1 本文研究的胆红素分子结构

Fig. 1. The molecular structure of bilirubin studied in this work.

由于胆红素独特的分子结构, 针对胆红素分子的电子结构和激发态性质的研究引起了研究者的广泛兴趣. 特别是结合荧光蛋白的胆红素 (Bilin-FPs) 代表了一类新型荧光发色团, 具有高亮度、可

\* 国家自然科学基金 (批准号: 11727810, 61720106009, 21603074) 和上海市科委项目 (批准号: 19JC1412200) 资助的课题.

<sup>†</sup> 通信作者. E-mail: [htsun@phy.ecnu.edu.cn](mailto:htsun@phy.ecnu.edu.cn)<sup>‡</sup> 通信作者. E-mail: [zrsun@phy.ecnu.edu.cn](mailto:zrsun@phy.ecnu.edu.cn)

编码性、光开关效应等优点,有望被广泛应用于生物成像和生物传感等领域<sup>[9]</sup>. Zietz 等<sup>[10]</sup>基于瞬态吸收光谱技术发现胆红素分子最低单重激发态具有亚皮秒量级的衰减时间,而且进一步通过荧光上转换测量发现激发态在氯仿溶剂中的衰减时间比二甲基亚砷中更短<sup>[11]</sup>. Cao 等<sup>[12]</sup>利用皮秒和飞秒瞬态吸收光谱研究了结合 UnaG 蛋白的胆红素分子的激发态动力学行为,发现其激发态的非辐射衰减过程明显受限于激发态质子转移过程. Zietz 和 Blomgren<sup>[13]</sup>基于完全活性空间自洽场方法理论研究了具有一半胆红素分子结构的二吡咯片段的 Z-E 异构化,进一步通过模拟基态-激发态间的圆锥交叉构型揭示了实验观测到的超短衰减时间和低荧光量子产率的来源.

尽管上述研究表明胆红素分子的激发态动力学行为与其激发态构型变化密切相关,但是对于胆红素发色团的荧光发光机理仍然缺乏系统的理论研究,尤其是与胆红素分子荧光性能密切相关的最低单重激发态特征仍缺乏定量精确的表征.然而,首要挑战来自于缺乏一个兼具计算效率和计算精度的电子结构理论方法来描述胆红素分子的电子结构与激发态性质.如图 1 所示,胆红素分子包含 79 个原子,而通常用来处理激发态性质的高水平方法,如 post-Hartree-Fock (PHF) 理论<sup>[14,15]</sup>、完全活性空间二阶微扰理论 (complete active space second-order perturbation theory, CASPT2)<sup>[16]</sup>和基于 GW approximation and Bethe-Salpeter Equation (GW-BSE) 的多体微扰理论<sup>[17,18]</sup>等,由于其计算成本随着体系原子数目的增加而呈指数增加,很难处理具有较大尺寸的分子体系.众所周知,密度泛函理论 (density functional theory, DFT) 和含时密度泛函理论 (time-dependent DFT, TDDFT) 由于兼顾较好的计算精度和较低的计算成本,是当前研究具有较大尺寸分子体系的最流行的理论方法之一<sup>[19,20]</sup>.然而,近期研究表明有机分子的电子激发能<sup>[21,22]</sup>、二阶非线性光学系数<sup>[23]</sup>和激发态特征<sup>[24]</sup>等性质极其依赖于密度泛函中所选取的交换-相关近似项.尤其是有机分子的前线分子轨道能级和电子耦合参数等与密度泛函中包含的准确交换项比例 ( $eX\%$ ) 几乎呈现线性相关<sup>[25]</sup>.因此,针对具有较大分子尺寸的胆红素分子而言,迫切需要寻找一种能够定量准确、定性高效预测胆红素分子激发态性质的电子结构理论方法.

近年来,“最优化调控”区间分离密度泛函 (range-separated density functional) 理论方法被广泛应用于精确预测有机分子的激发态和光学性质计算中<sup>[26-30]</sup>.在区间分离密度泛函理论框架下,泛函中包含的交换项可以用误差函数  $\text{erf}(x)$  和电子间距离  $r_{12}$  的关系表示:

$$\frac{1}{r_{12}} = \frac{1 - \text{erf}(\omega r_{12})}{r_{12}} + \frac{\text{erf}(\omega r_{12})}{r_{12}}, \quad (1)$$

其中泛函的交换项被分为来自短程处 DFT 的交换部分和来自长程处 HF 的交换部分.  $\omega$  是区间分离参数 (单位是 Bohr<sup>-1</sup>), 可以看成是电子密度的函数<sup>[31]</sup>.因此  $\omega$  对体系具有依赖性,即具有不同电子密度分布的分子应对应不同大小的  $\omega$  参数. Baer 和 Kronik 等提出:在精确的 Kohn Sham-DFT 理论框架下,负的最高占据分子轨道 (highest occupied molecular orbital, HOMO) 能量应该无限接近垂直电离能 (ionization potential, IP) 的大小,因而通过构造一个  $J^2(\omega)$  方程来实现对区间分离参数  $\omega$  的最优化调控<sup>[32,33]</sup>:

$$J^2 = \sum_{i=0}^1 [\epsilon_H(N+i) + \text{IP}(N+i)]^2, \quad (2)$$

其中,  $\epsilon_H(N)$  是指包含  $N$  电子的中性体系的 HOMO 能量,  $\text{IP}(N)$  是计算的中性 ( $N$ ) 和阳离子 ( $N-1$ ) 体系的能量差.当忽略轨道弛豫时,阴离子 ( $N+1$ ) 体系的电离能  $\text{IP}(N+1)$  可以近似地对应  $N$  电子体系的电子亲和能  $EA(N)$ .对于任何分子体系,通过满足上述方程 (2) 就可以得到最优化调控的  $\omega$  值.本文基于一系列密度泛函理论方法,包括最优化调控区间分离泛函,系统研究了胆红素分子的激发态性质,包括垂直跃迁能、振子强度和激发态特征等.相信本工作可以为今后基于胆红素分子体系以及其他大尺寸生物分子的激发态性质研究提供可靠高效的计算工具.

## 2 理论计算方法

胆红素分子的基态 ( $S_0$ ) 几何结构是在 B3LYP<sup>[34,35]</sup>-D3(BJ)<sup>[36]</sup>/6-31G(d)<sup>[37-39]</sup> 水平上优化的.最低单重激发态 ( $S_1$ ) 的垂直激发能 ( $E_{V\Lambda}$ ) 和垂直发射能 ( $E_{VE}$ ) 是分别基于基态和最低单重激发态结构利用线性响应 TDDFT 方法计算得到.为了测试各种密度泛函方法在实际预测时的表现,

本文采用了 12 种密度泛函, 包括 1 个纯广义梯度近似泛函 (PBE<sup>[40]</sup>); 1 个杂化泛函 (B3LYP); 3 个明尼苏达系列泛函 (MN15<sup>[41]</sup>、M062X<sup>[42]</sup> 和 M06HF<sup>[43]</sup>); 4 个区间分离密度泛函 (CAM-B3LYP<sup>[44]</sup>、LC- $\omega$ PBE<sup>[45]</sup>、 $\omega$ B97XD<sup>[46]</sup> 和 M11<sup>[47]</sup>); 2 个最优化调控区间分离密度泛函 (LC- $\omega$ PBE\* 和  $\omega$ B97XD\*) 以及 1 个双杂化泛函 (B2GPPLYP<sup>[48]</sup>). 为了区别于默认的区间分离泛函, 用符号\*表示区间分离参数  $\omega$  是经过最优化调控的. 此外, 还采用二阶簇和代数图解构造方法 (algebraic diagrammatic construction method to second order, ADC(2)<sup>[49]</sup>) 的计算结果作为参考, 其中 ADC(2) 方法利用密度拟合 (resolution of the identity, RI) 近似技术来提高计算效率. 在计算垂直发射能 ( $E_{VE}$ ) 时,  $S_1$  的激发态结构是在 TD- $\omega$ B97XD\*/6-31G(d) 水平下优化. 另外, 还考察了基组效应的影响 (见表 1), 从 def-SV(P)<sup>[50]</sup>、def-TZVP<sup>[51]</sup> 到 def2-TZVP<sup>[52]</sup> 基组, 激发能减小约为 0.08 eV, 而对于 RI-ADC(2) 方法, 则减小了约 0.16 eV, 可见常规密度泛函方法 (如 CAM-B3LYP) 对基组的敏感程度要小于 RI-ADC(2) 这类基于波函数的方法. 综合考虑效率与精度, 本文最终采用 def-TZVP 基组进行计算. 胆红素分子的吸收光谱和发射光谱实验数据来自于文献 [12,53], 其中实验值取自氯仿溶剂中吸收光谱和发射光谱的最大吸收波长. 通过与实验结果比较, 进一步统计了各种泛函的绝对误差 (absolute error,  $AE = \text{理论值} - \text{实验值}$ ) 和相对误差 (relative error,  $RE = |\text{理论值} - \text{实验值}| / \text{实验值}$ ). 结合隐式溶剂模型来考虑溶剂氯仿的影响, 常规密度泛函方法采用连续极化介质模型 (PCM)<sup>[54,55]</sup>, 双杂化泛函 B2GPPLYP 采用 SMD<sup>[56]</sup> 模型, RI-ADC(2) 采用 COSMO<sup>[57]</sup> 模型. 所有密度泛函理论计算均使用 Gaussian 09<sup>[58]</sup> 和 Gaussian 16<sup>[59]</sup> 软件完成, 双杂化泛函 B2GPPLYP 和 RI-ADC(2) 方法分别通过 ORCA<sup>[60]</sup> 和 Turbomole<sup>[61]</sup> 软件完成. 空穴-电子分析和片段间电荷转移 (interfragment charge transfer, IFCT) 分析通过 Multiwfn<sup>[62]</sup> 软件完成, 分子结构图和空穴-电子分布图通过 VMD<sup>[63]</sup> 软件完成. 计算软硬件平台由华东师范大学超算中心提供, 其中 CPU 为 64 位 Intel® Xeon 2680 V4, 主频为 2.4 GHz, 内存为 RDIMM DDR4 64 GB, 存储为 300 T 的 GPFS 并行文件系统.

表 1 基组对计算的垂直激发能 ( $E_{VA}$ ) 的影响  
 Table 1. Influence of basis set on the calculated vertical excitation energy ( $E_{VA}$ ).

Basis set	$E_{VA}/\text{eV}$			
	CAM-B3 LYP		RI-ADC(2)	
	PCM	GAS	COSMO	GAS
def-SV(P)	2.98	3.11	2.80	2.97
def-TZVP	2.91	3.05	2.69	2.86
def2-TZVP	2.90	3.04	—	2.81
EXP <sup>a</sup>	2.73	—	—	—

<sup>a</sup> Experimental values are taken from Refs. [12,53].

### 3 结果与讨论

表 2 列出了各种密度泛函计算的胆红素分子的垂直激发能 ( $E_{VA}$ ) 和垂直发射能 ( $E_{VE}$ ) 大小, 以及与实验值相比的绝对误差和相对误差. 首先, 可以看出最优化调控区间分离密度泛函方法 LC- $\omega$ PBE\* 和  $\omega$ B97XD\* 表现优异, 很好地重现了  $E_{VA}$  的实验值, 它们对  $E_{VA}$  的绝对误差分别为 0.03 eV 和 0.12 eV, 而相对误差仅为 1% 和 4%. 相比而言, 未经调控的 LC- $\omega$ PBE 和  $\omega$ B97XD 泛函产生的绝对 (相对) 误差分别为 0.38 eV (14%) 和 0.21 eV (8%). 上述结果说明了对具有默认区间分离参数  $\omega$  的区间分离泛函进行调控的必要性, 其相对误差和绝对误差都明显减小. 相比于未调控 LC- $\omega$ PBE 和  $\omega$ B97XD 泛函, 它们的区间分离参数  $\omega$  明显小于默认值, 即 LC- $\omega$ PBE\*: 0.178 Bohr<sup>-1</sup> vs. LC- $\omega$ PBE: 0.400 Bohr<sup>-1</sup> 和  $\omega$ B97XD\*: 0.137 Bohr<sup>-1</sup> vs.  $\omega$ B97XD: 0.200 Bohr<sup>-1</sup>. 这表明对于胆红素分子而言需要更小的区间分离参数  $\omega$ , 而由方程 (1) 可知更小的  $\omega$  值表示在描述泛函中的交换作用时类似 DFT 交换项将在更远的电子间距离处被类似 HF 交换项取代. 有趣的是, 像 MN15、M062X 和 CAM-B3LYP 泛函的预测表现仍然不错, 绝对 (相对) 误差分别为 0.08 eV (3%)、0.18 eV (7%) 和 0.18 eV (7%). 和预期一致, PBE 和 B3LYP 泛函明显低估了  $E_{VA}$  大小, 绝对误差分别为 -0.90 eV 和 -0.25 eV, 而相对误差达到 33% 和 9%. 而 M06HF 泛函预测时明显高估了  $E_{VA}$  大小, 绝对和相对误差分别为 0.39 eV 和 14%. 对于垂直发射能 ( $E_{VE}$ ) 而言, 除了 PBE (-0.71 eV)、B3LYP (-0.14 eV)、M06HF (0.13 eV) 和未调控的 LC- $\omega$ PBE (0.15 eV) 泛函外, 所有泛函均表现不错, 绝

对(相对)误差均小于 0.1 eV (4%). 这说明胆红素分子的激发态结构对于泛函的依赖程度要小于基态构型. 总的来说, 最优化调控区间分离泛函在预测  $E_{VA}$  和  $E_{VE}$  时的整体误差最小, 很好地重现了吸收和发射实验测量值. 值得注意的是 MN15 泛函整体上也可以给出准确的预测, 对  $E_{VA}$  和  $E_{VE}$  的预测误差分别为 0.08 eV 和 0.01 eV. 对于

$E_{VA}$  的预测, 我们还采用了处在密度泛函“雅各布梯”高阶的双杂化泛函 B2GPPLYP, 发现其对于  $E_{VA}$  的计算预测表现并没有预期的好, 绝对(相对)误差为 0.19 eV(7%). 还考察了高水平二阶簇和代数图解构造计算方法 RI-ADC(2), 其对于  $E_{VA}$  的预测误差仅为 -0.04 eV, 同样很好地预测了  $E_{VA}$  的实验值.

表 2 各种理论方法计算胆红素分子的垂直激发能 ( $E_{VA}$ )、振子强度 ( $f$ ) 和垂直发射能 ( $E_{VE}$ ) 以及与实验值相比的绝对误差和相对误差

Table 2. Vertical absorption energies ( $E_{VA}$ ), oscillator strength ( $f$ ) and vertical emission energies ( $E_{VE}$ ) of bilirubin and the absolute errors and relative errors compared to the available experimental data.

	$\omega$	$E_{VA}/\text{eV}$	$f(S_1)$	$AE/\text{eV}$	$RE/\%$	$E_{VE}/\text{eV}$	$AE/\text{eV}$	$RE/\%$
PBE	—	1.83	0.02	-0.90	33	1.68	-0.71	30
B3LYP	—	2.48	0.12	-0.25	9	2.25	-0.14	6
MN15	—	2.81	1.21	0.08	3	2.40	0.01	0.4
M062X	—	2.91	1.30	0.18	7	2.46	0.07	3
M06HF	—	3.12	1.42	0.39	14	2.52	0.13	6
CAM-B3LYP	0.330	2.91	1.33	0.18	7	2.45	0.06	3
LC- $\omega$ PBE	0.400	3.11	1.43	0.38	14	2.54	0.15	6
$\omega$ B97XD	0.200	2.94	1.36	0.21	8	2.46	0.07	3
M11	0.250	3.01	1.40	0.28	10	2.49	0.10	4
LC- $\omega$ PBE*	0.178	2.76	1.20	0.03	1	2.36	-0.03	1
$\omega$ B97XD*	0.137	2.85	1.28	0.12	4	2.43	0.04	2
B2 GPPLYP	—	2.92	1.13	0.19	7	—	—	—
RI-ADC(2)	—	2.69	1.06	-0.04	1	—	—	—
EXP <sup>a</sup>	—	2.73	—	—	—	2.39	—	—

<sup>a</sup> Experimental values are taken from Refs. [12,53].

此外, 还考察了各种密度泛函方法预测胆红素分子的最低单重激发态振子强度 ( $f(S_1)$ ) 的表现. 振子强度是衡量两个电子态之间跃迁强度的无量纲的量, 通常体系的基态与激发态之间的振子强度越大, 吸收(发射)光谱中相应的吸收(发射)峰也越强. 表 2 列出了各种密度泛函方法计算的胆红素分子的  $f(S_1)$  值, 同时以 RI-ADC(2) 方法计算的振子强度值作为参考. 可以看出, 与 RI-ADC(2) 的计算值相比, PBE 和 B3LYP 泛函产生非常小的  $f$  值, 错误地暗示了胆红素分子的  $S_1$  态具有明显的电荷转移特征. 而除了 PBE 和 B3LYP 泛函外, 其他类型泛函均可以合理描述胆红素分子的最低单重激发态的振子强度大小.

由上述误差统计分析发现, 各种密度泛函方法的预测表现与泛函中包含的准确交换项比例 ( $eX\%$ ) 密切相关. 通常表现为低/高  $eX\%$  的密度泛

函方法倾向于描述具有更离域化/局域化特征的  $\pi$  共轭电子结构, 从而相应低估/高估激发能. 如图 2 所示, 在电子间距离  $r_{12} = 2.75$  Bohr (约 1.455 Å) 处, 即  $\pi$  共轭体系中碳-碳单键和双键的平均距离处, LC- $\omega$ PBE\* ( $\omega = 0.178$  Bohr<sup>-1</sup>) 泛函包含大约 52% 的  $eX\%$ ,  $\omega$ B97XD\* ( $\omega = 0.137$  Bohr<sup>-1</sup>) 包含大约 53% 的  $eX\%$ ; 而未经调控的 LC- $\omega$ PBE, 其  $\omega$  默认为 0.4 Bohr<sup>-1</sup>, 包含近 88% 的  $eX\%$ , 同样未经调控的  $\omega$ B97XD 泛函,  $\omega$  默认为 0.2 Bohr<sup>-1</sup>, 包含 66% 的  $eX\%$ . 其他区间分离泛函如 CAM-B3LYP 泛函包含 56% 的  $eX\%$ , M11 泛函包含了 82% 的  $eX\%$ . 而 MN15、M062X 和 M06HF 泛函在任意  $r_{12}$  距离上都有 44%、54% 和 100% 的  $eX\%$ ; PBE 和 B3LYP 泛函分别包含 0 和 20% 的  $eX\%$ . 有趣的是, 较大的误差均来自于泛函中包含过低的  $eX\%$  (如 PBE 和 B3LYP) 或过高的

$eX\%$ (如未调控 LC- $\omega$ PBE、M06HF 和 M11). 可以看出最优化调控泛函 LC- $\omega$ PBE\*、 $\omega$ B97XD\*、MN15、M062X 和 CAM-B3LYP 等泛函的良好预测表现可以归因于它们所包含的适宜  $eX\%$ , 均在 44%—56% 范围内, 这样一个合理的  $eX\%$  有利于密度泛函方法在描述电子结构时既不过于离域化也不过于局域化. 特别是在对区间分离密度泛函中的区间分离参数进行最优化调控后, 可以使其在离域化作用和局域化作用间达到平衡, 从而可以精确描述胆红素分子的激发态电子结构. 为了证明该最优化调控方法的普适性, 我们从文献中选取了藻青素 (phycocyanobilin) 和四环霉素 (tetracycline) 生物分子 (见补充文件) 进行验证计算, 发现最优化调控方法 LC- $\omega$ PBE\* 整体上仍能保持非常高的预测精度, 且不随生物分子体系的变化而明显变化. 有趣的是, 对于藻青素和四环霉素分子, 它们的最低单重激发态  $f(S_1)$  较大, 属于明显 LE 激发, 因此常用的 B3LYP(甚至 PBE) 泛函表现也还不错, 但是在胆红素分子的预测上这两种泛函是出现定性错误的, 这一点更加突显出最优化调控泛函方法的普适性.

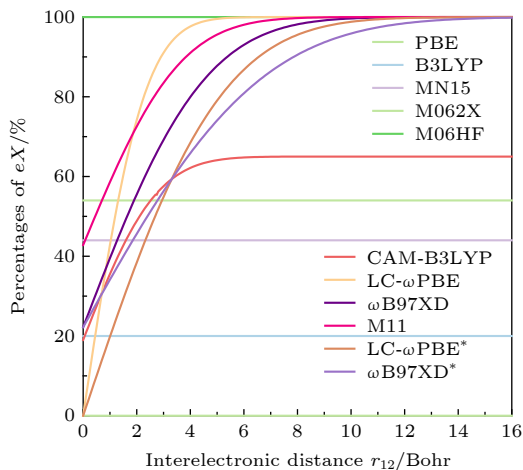


图 2 本文中各种密度泛函中所含准确交换项比例 ( $eX\%$ ) 与电子间距离 ( $r_{12}$ ) 关系示意图

Fig. 2. Percentages of exact-exchange ( $eX\%$ ) included in various density functionals as a function of interelectronic distance ( $r_{12}$ , Bohr).

由上述密度泛函的基准测试发现, 最优化调控区间分离密度泛函 LC- $\omega$ PBE\* 在定量描述胆红素分子激发态的电子跃迁能和振子强度方面都表现最优. 因此, 为进一步描述胆红素分子最低单重激发态 ( $S_1$ ) 的特征, 我们在 LC- $\omega$ PBE\* 泛函计算的

波函数基础上采用空穴-电子分析方法进行分析. 该方法将电子激发过程描述为“空穴 $\rightarrow$ 电子”, 可以通过空穴电子的定义衍生出各种激发态特征指数, 还可以以图形化的方式直观地考察电子的去留, 进而考察其激发态特征. 通常根据空穴和电子的相对位置, 可以将激发态特征分为: 局域激发 (local excitation, LE)、电荷转移激发 (charge transfer, CT) 和杂化局域-电荷转移激发 (hybrid local and charge transfer, HLCT). 首先, 基于 Multiwfn 波函数分析程序定量考察了胆红素分子  $S_1$  态的电子转移距离 ( $D$  指数) 和空穴与电子的分离程度 ( $S_r$  指数) 等指标. 其中,  $D$  指数衡量的是空穴和电子质心之间的距离, 数值越大表示空穴和电子分布范围差距越大 (倾向于 CT 特征), 数值越小表示分布范围越接近 (倾向于 LE 特征).  $S_r$  指数衡量空穴和电子的重叠程度, 数值越大重叠程度越高 (LE), 数值越小说明空穴和电子的分离越显著 (CT), 取值范围是  $0 < S_r < 1$ . 计算得到  $D$  指数为  $1.18 \text{ \AA}$ , 表明空穴与电子主要分布范围很接近;  $S_r$  值为  $0.75 \text{ a.u.}$ , 表明空穴与电子重叠程度较高. 从以上激发态指数可知, 胆红素分子的最低单重激发态应该属于局域激发为主的类型. 另外, 如图 3 所示, 为胆红素分子最低单重激发态的空穴-电子分布, 其中蓝色表示空穴, 绿色表示电子, 空穴和电子几乎对称分布在两个吡咯-吡咯酮片段上, 并且空穴和电子在空间上既有重叠也有分离, 暗示着吡咯-吡咯酮片段内存在电荷转移.

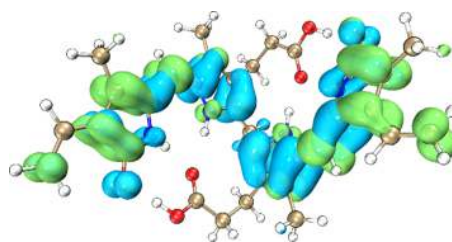


图 3 胆红素分子最低单重激发态的空穴-电子分布示意图  
Fig. 3. Diagram of hole-electron distribution for the lowest singlet excited state (isovalue=0.001).

为了进一步定量探究胆红素分子最低单重激发态的特征, 如图 4 所示, 将胆红素分子划分为 3 个片段, 将亚甲基连接的两个甲基吡咯环 (红色轮廓) 定义为第一片段, 两个乙烯基吡咯酮 (绿色轮廓) 为第二片段, 剩下的羧酸基团 (蓝色轮廓) 为第三片段. 图 4 显示了各片段分别对电子和空穴的

贡献,可以看出,对于空穴而言,片段1贡献有52.9%,而片段2有46.1%的贡献.对于电子而言,片段1有20.6%的贡献,而片段2有79.0%的贡献.片段3对空穴或电子几乎都没有贡献.比较空穴和电子的贡献,激发前后片段1的电子贡献减小说明有失去电子的倾向,而片段2贡献变大则说明具有接受电子的倾向.此外,进一步采用IFCT方法定量考察了电子激发过程中体系片段之间的电子转移量.表3列出了胆红素分子 $S_1$ 态片段间电子转移量以及各片段电子转移的净变化量.片段1失去0.323个电子,片段2得到0.329个电子.如图5所示,各片段电子转移矩阵热图直观展示了片段间发生的电子转移是由片段1转移到片段2,即吡咯环向吡咯酮发生电子转移,其中吡咯-吡咯酮片段内电荷转移成分约为32%.因此可以认为这个激发过程并不是纯局域激发,还包括片段间电荷转移过程,因而胆红素分子最低单重激发态应属于局域-电荷转移杂化激发类型.此外,上述结果也合理解释了为什么PBE和B3LYP这类包含较低 $eX\%$ 的泛函在预测 $S_1$ 态的激发能时产生明显误

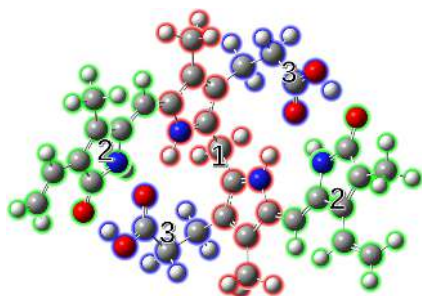


图4 胆红素分子的三个片段划分及各片段对最低单重激发态的空穴和电子的贡献

Fig. 4. The divided three fragments of bilirubin molecule and contribution of each fragment to the hole and electron for the lowest singlet excited state.

表3 胆红素分子最低单重激发态的各片段电子净变化量以及片段间电子转移量

Table 3. Net change of each fragment and electron transfer between fragments for lowest singlet excited state of bilirubin molecule.

	Electron transfer between fragments		
	1	2	3
1	0.109	0.418	0.002
2	0.095	0.364	0.002
3	0.002	0.008	0.000
Net change of each fragment	-0.323	0.329	-0.006

差,并错误预测 $S_1$ 态具有明显的电荷转移特征.

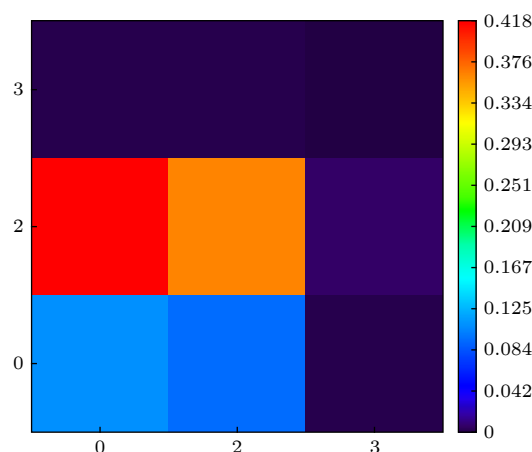


图5 胆红素分子最低单重激发态的各片段电子转移矩阵热图

Fig. 5. Diagram of electron transfer matrix for each fragment of lowest singlet excited state of bilirubin molecule.

## 4 结论

综上所述,本文基于结合隐式溶剂模型的线性响应TDDFT方法计算研究了胆红素分子最低单重激发态的垂直激发能( $E_{VA}$ )、振子强度( $f(S_1)$ )和垂直发射能( $E_{VE}$ ).以实验值和RI-ADC(2)计算值作为参考,系统考察了一系列密度泛函方法在实际预测时的表现,结果发现最优化调控的LC- $\omega$ PBE\*和 $\omega$ B97XD\*泛函对 $E_{VA}$ 的绝对误差分别为0.03 eV和0.12 eV,对 $E_{VE}$ 的绝对误差分别为-0.03 eV和0.04 eV,明显优于未经调控的区间分离密度泛函.进一步分析发现其他泛函的预测误差与泛函中包含的准确交换项比例( $eX\%$ )密切相关.可以看出最优化调控泛函LC- $\omega$ PBE\*、 $\omega$ B97XD\*、MN15、M062X和CAM-B3LYP等泛函的良好预测表现可以归因于它们所包含适宜的 $eX\%$ (44%—56%),这显然有利于在实际计算预测时产生既不过分离域也不过分局域的电子结构.而较大的预测误差均来自于泛函中包含过低的 $eX\%$ (如PBE和B3LYP)或过高的 $eX\%$ (如未调控LC- $\omega$ PBE、M06HF和M11).尤其是,与RI-ADC(2)计算的振子强度相比,PBE和B3LYP泛函预测了非常小的 $f$ 值,显然错误地暗示了胆红素分子的 $S_1$ 态具有明显的电荷转移特征,这一点是格外需要注意的.基于最优化调控方法的足够精度,定量考察了胆红素分子最低单重激发态( $S_1$ )的特征,

采用空穴-电子分析和 IFCT 方法分别计算了电子激发过程中胆红素分子  $S_1$  态的电子转移距离和空穴与电子的分离程度、各片段对空穴-电子的贡献和片段之间的电子转移量等,证实了胆红素分子的最低单重激发态具有局域-电荷转移杂化激发特征的激发态,并且电荷转移成分约为 32%。相信该研究工作可以为今后研究胆红素分子的激发态动力学过程和光谱性质提供重要理论依据,本文提出的基于区间分离密度泛函的最优调控理论方法也可以为接下来其他生物分子体系的激发态性质研究提供更可靠、高效的理论工具。

## 参考文献

- [1] Stocker R, Yamamoto Y, McDonagh A F, Glazer A N, Ames B N 1987 *Science* **235** 1043
- [2] Fevery J 2008 *Liver Int.* **28** 592
- [3] Hissi E G V, Martinez J C G, Zamarbide G N, Estrada M R, Jensen S J K, Tomas-Vert F, Csizmadia I G 2009 *J. Mol. Struct.: THEOCHEM* **911** 24
- [4] Carreira-Blanco C, Singer P, Diller R, Lustres J L P 2016 *Phys. Chem. Chem. Phys.* **18** 7148
- [5] Person R V, Peterson B R, Lightner D A 1994 *J. Am. Chem. Soc.* **116** 42
- [6] Nogales D, Lightner D A 1995 *J. Biol. Chem.* **270** 73
- [7] Boiadjev S E, Watters K, Wolf S, Lai B N, Welch W H, McDonagh A F, Lightner D A 2004 *Biochemistry* **43** 15617
- [8] Lightner D A, Holmes D L, McDonagh A F 1996 *J. Biol. Chem.* **271** 2397
- [9] Braslavsky S E, Holzwarth A R, Schaffner K 1983 *Angew. Chem. Int. Ed.* **22** 656
- [10] Zietz B, Macpherson A N, Gillbro T 2004 *Phys. Chem. Chem. Phys.* **6** 4535
- [11] Zietz B, Gillbro T 2007 *J. Phys. Chem. B* **111** 11997
- [12] Cao X D, Zhang C C, Gao Z H, Liu Y Y, Zhao Y Z, Yang Y, Chen J Q, Jimenez R, Xu J H 2019 *Phys. Chem. Chem. Phys.* **21** 2365
- [13] Zietz B, Blomgren F 2006 *Chem. Phys. Lett.* **420** 556
- [14] Fabiano E, Della Sala F, Cingolani R, Weimer M, Görling A 2005 *J. Phys. Chem. A* **109** 3078
- [15] Hammond J R, Kowalski K 2009 *J. Chem. Phys.* **130** 194108
- [16] Budzák Š, Scalmani G, Jacquemin D 2017 *J. Chem. Theory Comput.* **13** 6237
- [17] Hedin L 1965 *Phys. Rev.* **139** A796
- [18] Hybertsen M S, Louie S G 1986 *Phys. Rev. B* **34** 5390
- [19] Kohn W, Sham L J 1965 *Phys. Rev.* **140** A1133
- [20] Cohen A J, Mori-Sanchez P, Yang W T 2011 *Chem. Rev.* **112** 289
- [21] Jacquemin D, Wathelet V, Perpète E A, Adamo C 2009 *J. Chem. Theory Comput.* **5** 2420
- [22] Tian X H, Sun H T, Zhang Q S, Adachi C 2016 *Chin. Chem. Lett.* **27** 1445
- [23] Sun H T, Autschbach J 2013 *ChemPhysChem* **14** 2450
- [24] Jiang Y R, Hu Z B, Zhou B, Zhong C, Sun Z R, Sun H T 2019 *J. Phys. Chem. C* **123** 5616
- [25] Sutton C, Sears J S, Coropceanu V, Brédas J 2013 *J. Phys. Chem. Lett.* **4** 919
- [26] Stein T, Kronik L, Baer R 2009 *J. Am. Chem. Soc.* **131** 2818
- [27] Sun H T, Zhong C, Brédas J 2015 *J. Chem. Theory Comput.* **11** 3851
- [28] Penfold T J 2015 *J. Phys. Chem. C* **119** 13535
- [29] Kronik L, Stein T, Refaely-Abramson S, Baer R 2012 *J. Chem. Theory Comput.* **8** 1515
- [30] Sun H T, Zhong C, Sun Z R 2016 *Acta Phys.-Chim. Sin.* **32** 2197 (in Chinese) [孙海涛, 钟成, 孙真荣 2016 物理化学学报 **32** 2197]
- [31] Körzdörfer T, Sears J S, Sutton C, Brédas J 2011 *J. Chem. Phys.* **135** 204107
- [32] Baer R, Livshits E, Salzner U 2010 *Annu. Rev. Phys. Chem.* **61** 85
- [33] Stein T, Kronik L, Baer R 2009 *J. Chem. Phys.* **131** 244119
- [34] Becke A D 1993 *J. Chem. Phys.* **98** 5648
- [35] Lee C, Yang W T, Parr R G 1988 *Phys. Rev. B* **37** 785
- [36] Grimme S, Ehrlich S, Goerigk L 2011 *J. Comput. Chem.* **32** 1456
- [37] Ditchfield R, Hehre W J, Pople J A 1971 *J. Chem. Phys.* **54** 724
- [38] Hehre W J, Ditchfield R, Pople J A 1972 *J. Chem. Phys.* **56** 2257
- [39] Hariharan P C, Pople J A 1973 *Theor. Chim. Acta* **28** 213
- [40] Perdew J P, Burke K, Ernzerhof M 1996 *Phys. Rev. Lett.* **77** 3865
- [41] Yu H S, He X, Li S L, Truhlar D G 2016 *Chem. Sci.* **7** 5032
- [42] Zhao Y, Truhlar D G 2008 *Theor. Chem. Acc.* **120** 215
- [43] Zhao Y, Truhlar D G 2006 *J. Phys. Chem. A* **110** 13126
- [44] Yanai T, Tew D P, Handy N C 2004 *Chem. Phys. Lett.* **393** 51
- [45] Vydrov O A, Scuseria G E 2006 *J. Chem. Phys.* **125** 234109
- [46] Chai J D, Head-Gordon M 2008 *Phys. Chem. Chem. Phys.* **10** 6615
- [47] Peverati R, Truhlar D G 2011 *J. Phys. Chem. Lett.* **2** 2810
- [48] Goerigk L, Grimme S 2010 *J. Chem. Phys.* **132** 184103
- [49] Trofimov A B, Schirmer J 1995 *J. Phys. B: At., Mol. Opt. Phys.* **28** 2299
- [50] Schäfer A, Horn H, Ahlrichs R 1992 *J. Chem. Phys.* **97** 2571
- [51] Schäfer A, Huber C, Ahlrichs R 1994 *J. Chem. Phys.* **100** 5829
- [52] Weigend F, Ahlrichs R 2005 *Phys. Chem. Chem. Phys.* **7** 3297
- [53] Kumagai A, Ando R, Miyatake H, Greimel P, Kobayashi T, Hirabayashi Y, Shimogori T, Miyawaki A 2013 *Cell* **153** 1602
- [54] Menucci B 2012 *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2** 386
- [55] Tomasi J, Menucci B, Cammi R 2005 *Chem. Rev.* **105** 2999
- [56] Marenich A V, Cramer C J, Truhlar D G 2009 *J. Phys. Chem. B* **113** 6378
- [57] Klamt A, Schüürmann G 1993 *J. Chem. Soc., Perkin Trans.* **2** 799
- [58] Frisch M J, Trucks G W, Schlegel H B, Scuseria G E, Robb M A, Cheeseman J R, Scalmani G, Barone V, Petersson G A, Nakatsuji H, Li X, Caricato M, Marenich A V, Bloino J, Janesko B G, Gomperts R, Menucci B, Hratchian H P, Ortiz J V, Izmaylov A F, Sonnenberg J L, Williams, Ding F, Lipparini F, Egidi F, Goings J, Peng B, Petrone A, Henderson T, Ranasinghe D, Zakrzewski V G, Gao J, Rega N, Zheng G, Liang W, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Throssell K, Montgomery J J A, Peralta J E, Ogliaro F, Bearpark M J, Heyd J J, Brothers E N, Kudin K N, Staroverov V N, Keith T A, Kobayashi R, Normand J, Raghavachari K, Rendell A P, Burant J C,

- Iyengar S S, Tomasi J, Cossi M, Millam J M, Klene M, Adamo C, Cammi R, Ochterski J W, Martin R L, Morokuma K, Farkas O, Foresman J B, Fox D J 2013 *Gaussian 09 Revision E.01* Wallingford: Gaussian Inc
- [59] Frisch M J, Trucks G W, Schlegel H B, Scuseria G E, Robb M A, Cheeseman J R, Scalmani G, Barone V, Petersson G A, Nakatsuji H, Li X, Caricato M, Marenich A V, Bloino J, Janesko B G, Gomperts R, Mennucci B, Hratchian H P, Ortiz J V, Izmaylov A F, Sonnenberg J L, Williams, Ding F, Lipparini F, Egidi F, Goings J, Peng B, Petrone A, Henderson T, Ranasinghe D, Zakrzewski V G, Gao J, Rega N, Zheng G, Liang W, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Vreven T, Throssell K, Montgomery J J A, Peralta J E, Ogliaro F, Bearpark M J, Heyd J J, Brothers E N, Kudin K N, Staroverov V N, Keith T A, Kobayashi R, Normand J, Raghavachari K, Rendell A P, Burant J C, Iyengar S S, Tomasi J, Cossi M, Millam J M, Klene M, Adamo C, Cammi R, Ochterski J W, Martin R L, Morokuma K, Farkas O, Foresman J B, Fox D J 2016 *Gaussian 16 Revision A.03* Wallingford: Gaussian Inc
- [60] Neese F 2012 *Wiley Interdiscip. Rev.: Comput. Mol. Sci.* **2** 73
- [61] Ahlrichs R, Bär M, Häser M, Horn H, Kölmel C 1989 *Chem. Phys. Lett.* **162** 165
- [62] Lu T, Chen F W 2012 *J. Comput. Chem.* **33** 580
- [63] Humphrey W, Dalke A, Schulten K 1996 *J. Mol. Graphics* **14** 33

## Density functional theory studies on the excited-state properties of Bilirubin molecule\*

Li Yuan-Yuan    Hu Zhu-Bin    Sun Hai-Tao<sup>†</sup>    Sun Zhen-Rong<sup>‡</sup>

(State Key Laboratory of Precision Spectroscopy, East China Normal University, Shanghai 200241, China)

( Received 8 April 2020; revised manuscript received 18 May 2020 )

### Abstract

Bilirubin is the main pigment in human bile, which is closely related to human health. Bilirubin combining with fluorescent protein represents a new type of fluorescent chromophore and has important applications in the field of biological imaging and biosensor. Due to the lack of efficient and accurate electronic structure methods, the electronic structure and excited-state properties of bilirubin molecule are not characterized quantitatively and accurately. Firstly, the vertical absorption energy, oscillator strength and vertical emission energy of the lowest singlet excited state of bilirubin molecule are calculated by combining the implicit solvent model and the linear response time-dependent density functional theory (TDDFT) method. Compared to the experimental data and high-level RI-ADC(2) calculation, the prediction performance of a series of density functional methods is systematically investigated. The results show that the optimally-tuned range separated density functional method has the best overall performance and the minimum absolute and relative errors. This is obviously due to the fact that the suitable proportion of exact exchange included in density functionals can produce neither delocalized nor localized electronic structures. Based on the produced wavefunction by the optimally-tuned method, the excited-state characteristics of the  $S_1$  state of bilirubin molecule indicate a hybrid local and charge transfer excitation, based on the quantitative characterization using hole-electron analysis and interfragment charge transfer method. This work can provide a theoretical basis for the study of excited-state dynamics and spectral properties of bilirubin molecules and the optimally tuned range-separated DFT method also provide a reliable and efficient theoretical tool to study the excited-state properties of other biomolecular systems in the future.

**Keywords:** bilirubin, density functional theory, optimal tuning, excited-state character

**PACS:** 31.15.ee, 31.15.E-, 31.15.es

**DOI:** 10.7498/aps.69.20200518

\* Project supported by National Natural Science Foundation of China (Grant Nos.11727810, 61720106009, 21603074) and Projects from Shanghai Science and Technology Commission (Grant No. 19 JC1412200).

<sup>†</sup> Corresponding author. E-mail: [htsun@phy.ecnu.edu.cn](mailto:htsun@phy.ecnu.edu.cn)

<sup>‡</sup> Corresponding author. E-mail: [zrsun@phy.ecnu.edu.cn](mailto:zrsun@phy.ecnu.edu.cn)