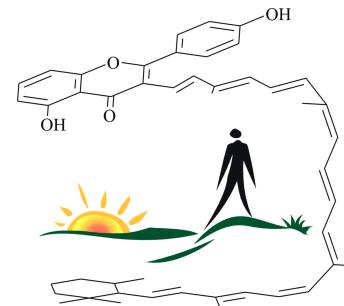


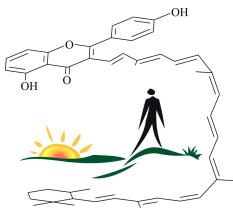
# Coupled-cluster study of model systems: the nucleic acid bases and their dimers

Timo Fleig

Theoretical and Computational Chemistry  
Heinrich Heine University Düsseldorf

April 16, 2007





# Outline

## I. Introduction and Motivation

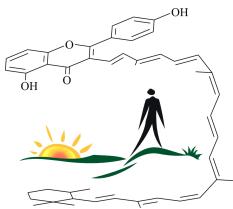
## II. Results for nucleic acid bases

## III. Accuracy of RI-CC2

One-particle basis sets

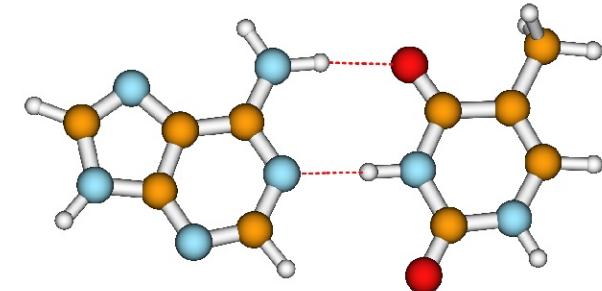
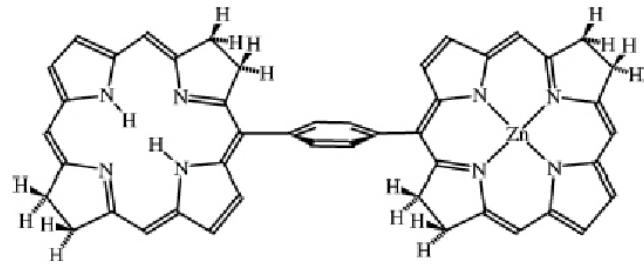
The CC hierarchy: correlation errors

## IV. Results for nucleic acid base dimers



# I. Introduction

## *Ab-initio* Electronic-Structure Studies



C3

Structure optimization

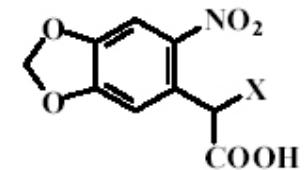
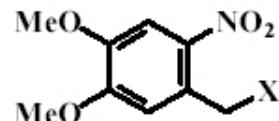
Vertical and adiabatic excitation energies

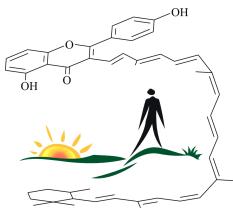
Charge–transfer states and processes

Methods:

Coupled Cluster

Configuration Interaction





# Introduction

## Overview of Current Activities

- **Rigorous study** of monomer units (structures, excited states)

T Fleig, S Knecht, C Hättig, *J. Phys. Chem. A* **2007**, accepted.

- **Dimer studies** at calibrated computational level

Investigation of CT/proton transfer processes

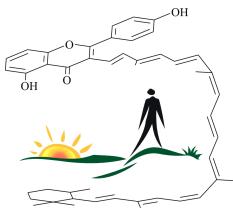
- **1-Methylthymine:** Simulation of solvation effects

M Etinski, T Fleig, with K. Kleinermanns

- 
- **Method development:** Parallel large-scale MRCI

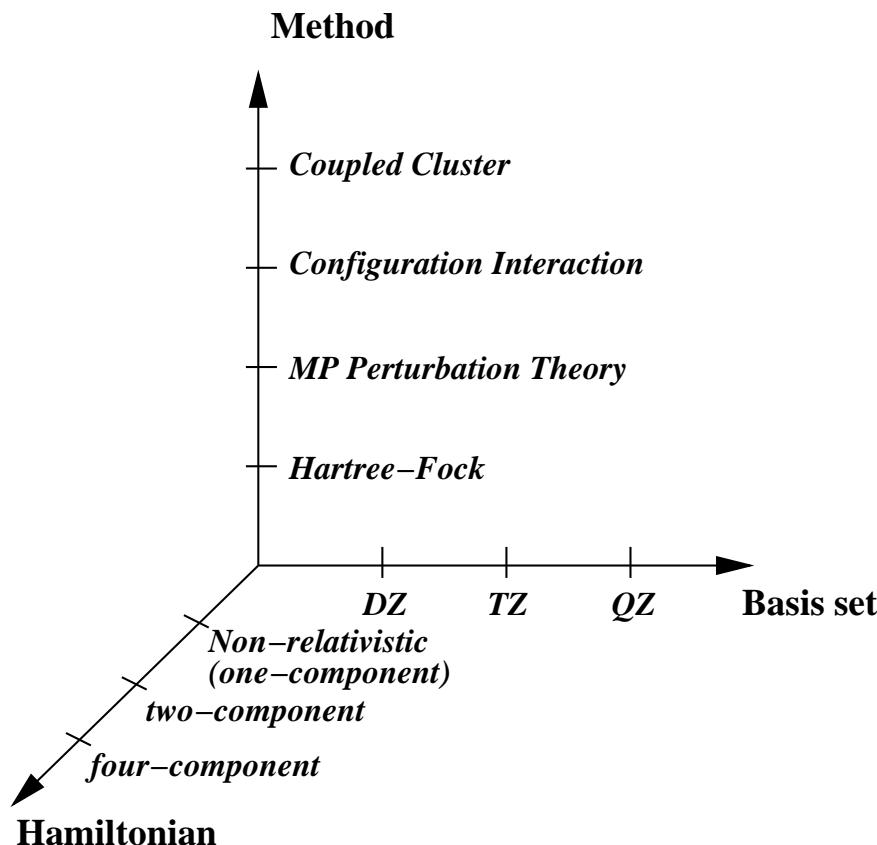
S Knecht, T Fleig, *Chem. Phys. Lett.* **2007**, in preparation.

- Interfacing LUCITA to TURBOMOLE → RI approximation



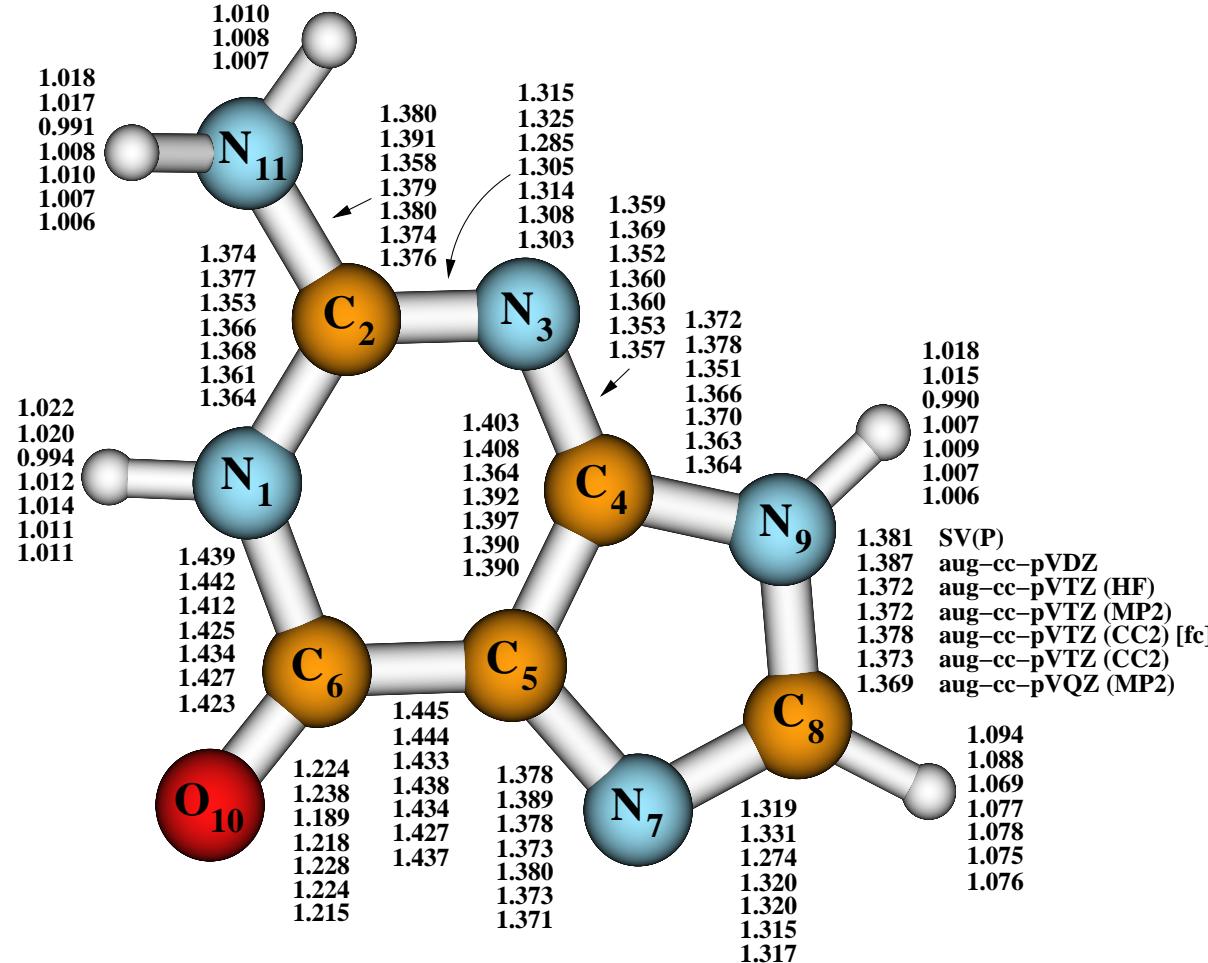
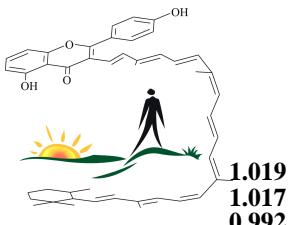
# I. Introduction

## The Accuracy of Quantum-Chemical Models



- **Electron Correlation:**  
*Increasing rigor of treatment*  
Hartree-Fock, MP2, CC2  
*Coupled Cluster* approaches  
Study of model hierarchies
- **Basis set expansion:**  
*Correlation-consistent*  
one-particle basis sets  
SV(P), cc-pVTZ,  
aug-cc-pVXZ (X=D,T,Q)
- **Hamiltonian:**  
*Non-relativistic* Hamiltonian

→ Various approaches in direct comparison



## II. Results

### Monomer structures. (Guanine)

- Basis sets:

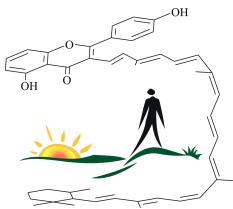
$aDZ \rightarrow aTZ \rightarrow aQZ$   
Bond contraction  $\approx -0.015 \text{ \AA}$

$SV(P)$  error  $\approx 0.02 \text{ \AA}$

$XZ \rightarrow aXZ \approx +0.005 \text{ \AA}$

- Correlation:

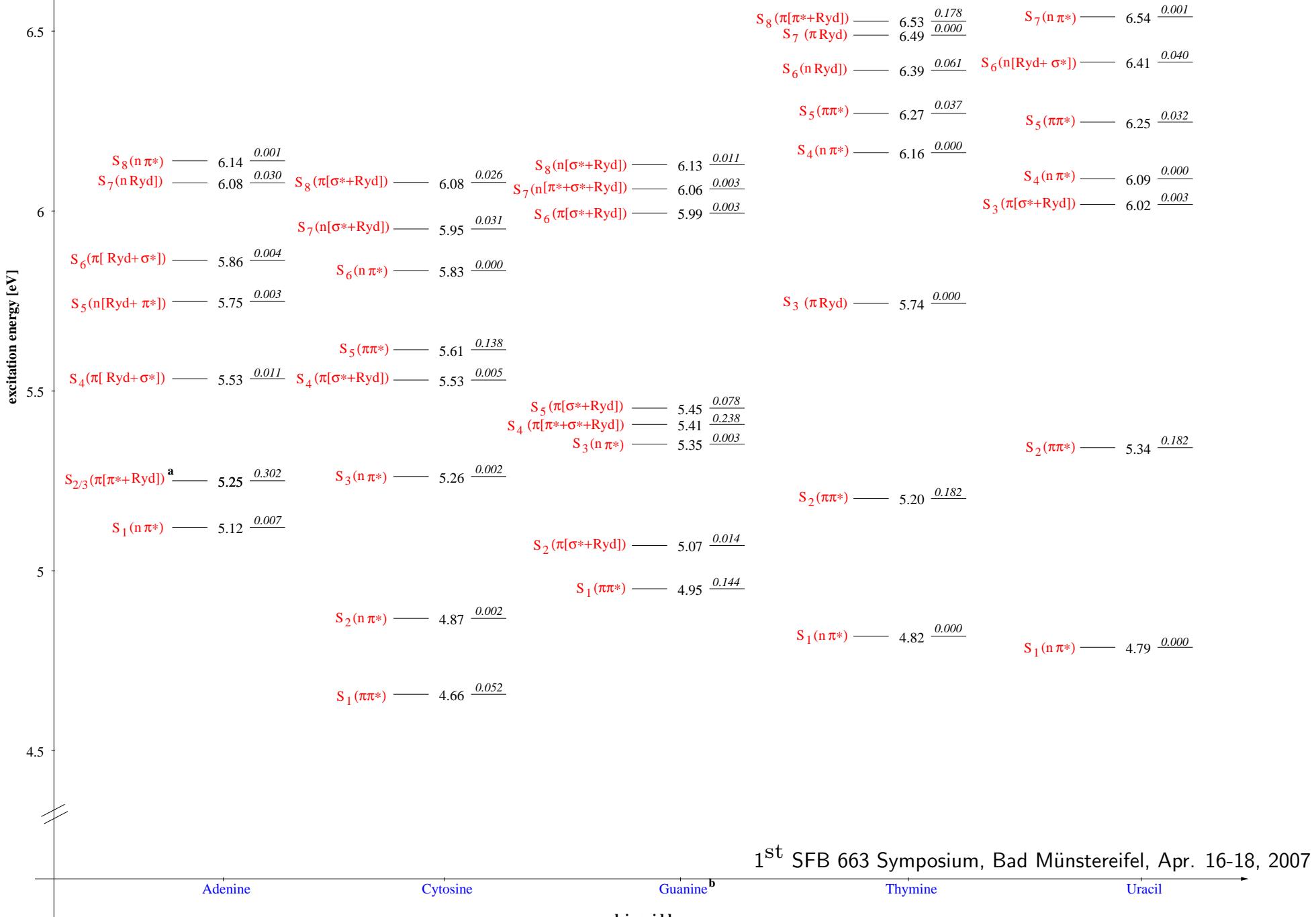
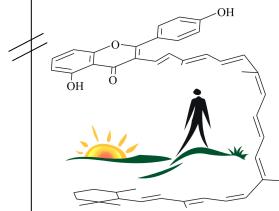
$HF \rightarrow MP2 \approx CC2$   
Bond stretching  $\approx +0.004 \dots + 0.04 \text{ \AA}$

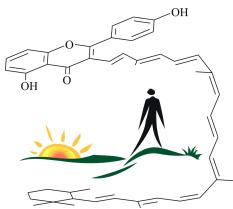


## II. Results

### Nucleic acid base monomers

# Excited states (RI-CC2/aug-cc-pVTZ)





## II. Results

### Other methods and experiment

### 9H-Adenine

Excitation	TD-DFT <sup>1</sup>	CASSCF/CASPT2 <sup>2</sup>	DFT/MRCI <sup>3</sup>	RI-CC2	Exp. <sup>4</sup>	Exp. <sup>E_0-E_5</sup>
$n - \pi^*$	5.04	6.15	5.01	5.12		4.40
$\pi - \pi^*$	5.09	5.13	4.90	5.25	4.98	4.48
$\pi - \pi^*$		5.20	5.04	5.25		

- TD-DFT/B3LYP performs well for monomer units
- CASSCF/CASPT2 suffers from limited active spaces
- Comparable (gas-phase) experiments are rare (vertical excitation)

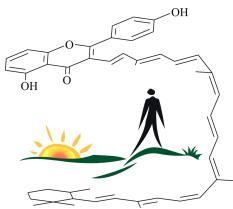
<sup>1</sup>A L Sobolewski, W Domcke, *Eur. Phys. J. D* **2002**, *20*, 369

<sup>2</sup>M P Fülscher, L Serrano-Andrés, B O Roos, *J. Am. Chem. Soc.* **1997**, *119*, 6168

<sup>3</sup>C M Marian, *J. Chem. Phys.* **2005**, *122*, 104314

<sup>4</sup>L B Clark, G G Peschel, J. I. Tinoco, *J. Phys. Chem.* **1965**, *69*, 3615

<sup>5</sup>N J Kim, G Jeong, Y S Kim, J Sung, S K Kim, Y D Park, *J. Chem. Phys.* **2000**, *113*, 10051



## II. Results

### Comparison with Experiment

NA base	State (transition)	RI-CC2 $T_v$ [eV]	Experimental
Adenine	$S_1 (n - \pi^*)$	5.12	$4.40^6$
	$S_2 (\pi - \pi^*)$	5.25	$4.48^6$ $4.98^7$
Thymine	$S_1 (n - \pi^*)$	4.82	
	$S_2 (\pi - \pi^*)$	5.20	$4.8^8$ $4.7^9$
Guanine	$S_1 (\pi - \pi^*)$	4.98	$4.4 - 4.5^{10}$
Cytosine	$S_1 (\pi - \pi^*)$	4.66	$4.28^7$
Uracil	$S_1 (n - \pi^*)$	4.80	
	$S_2 (\pi - \pi^*)$	5.35	$5.08^7$

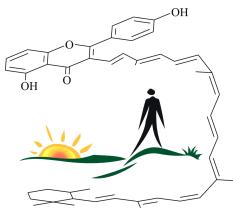
<sup>6</sup> $E_{0-0}$ , vibration corrected; N J Kim, G Jeong, Y S Kim, J Sung, S K Kim, Y D Park, *J. Chem. Phys.* **2000**, 113, 10051

<sup>7</sup>Vapor phase; L B Clark, G G Peschel, J. I. Tinoco, *J. Phys. Chem.* **1965**, 69, 3615

<sup>8</sup>1,3-Dimethyluracil; L B Clark, G G Peschel, J. I. Tinoco, *J. Phys. Chem.* **1965**, 69, 3615

<sup>9</sup>In water and TMP; L B Clark, G G Peschel, J. I. Tinoco, *J. Phys. Chem.* **1965**, 69, 3615

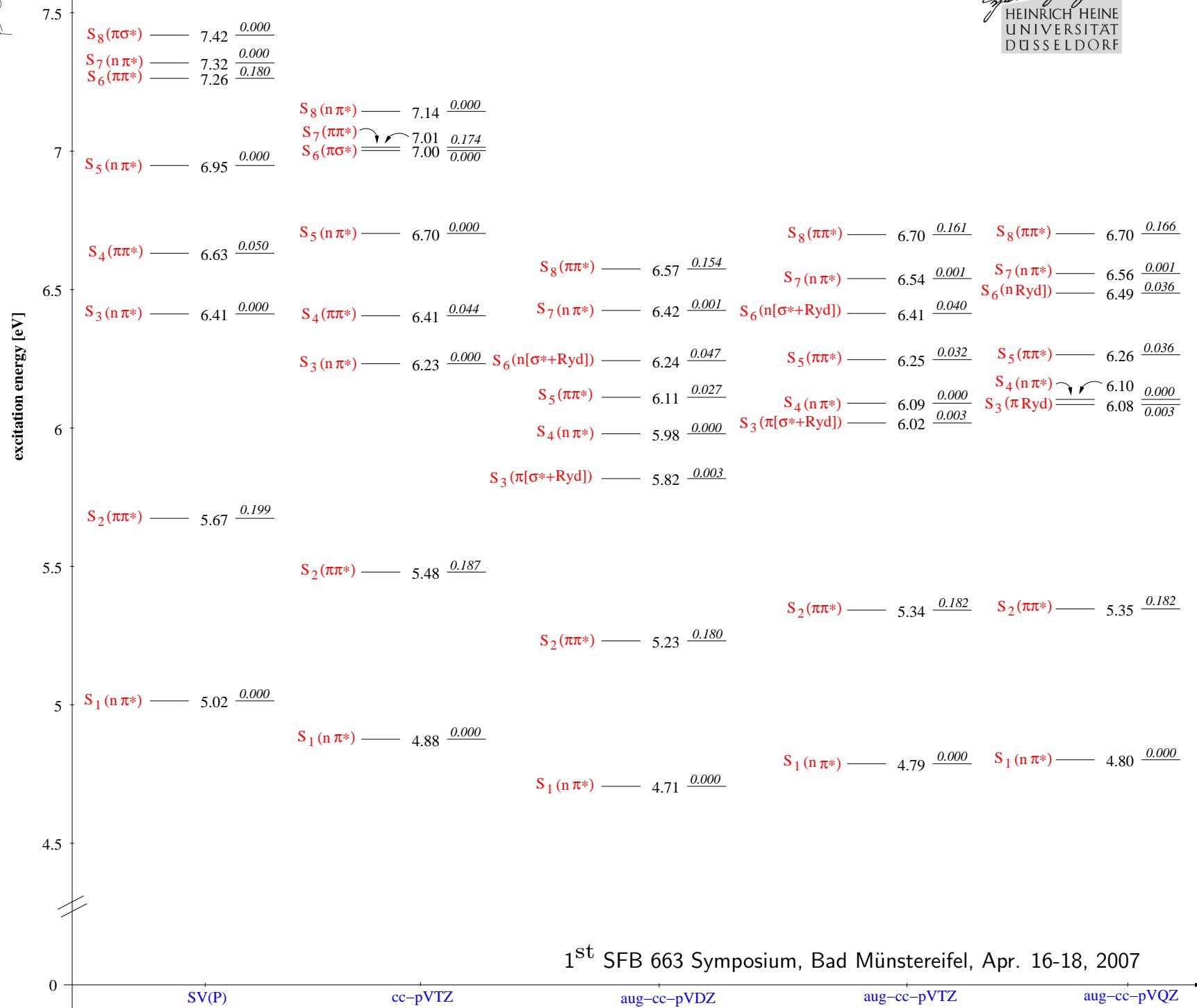
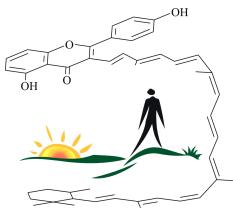
<sup>10</sup>In: M P Fülscher, L Serrano-Andrés, B O Roos, *J. Am. Chem. Soc.* **1997**, 119, 6168

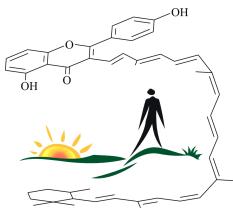


## II. Results

### Nucleic acid base monomers

# Excited states (Uracil, Basis sets)



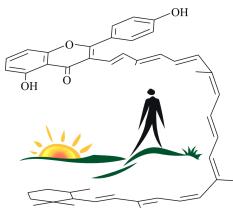


### III. Analysis

#### Correlation-Consistent Basis Sets

##### Some basic features:

- Generalized contraction for occupied orbitals
- Optimized primitives to maximize correlation contribution (CISD)
- cc-pVTZ: 2s 2p 2d 1f primitives
- Construction: Valence-correlated ground states of atoms !
- Excited states: Diffuse electron distributions
- aug-cc-pVTZ: + 1s 1p 1d 1f diffuse primitives
- cc-pVDZ sets do not have sufficient flexibility to describe excited states accurately



### III. Analysis

#### Effect of augmenting functions

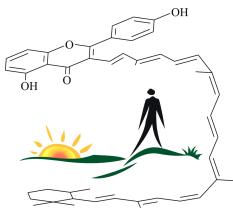
Uracil excited states, vertical RI-CC2/(aug-)cc-pVTZ

$$\Delta E_{T_v} = E_{T_v}^{\text{cc-pVTZ}} - E_{T_v}^{\text{aug-cc-pVTZ}}$$

State	$\Delta E_{T_v}$ [eV]
$n\pi^*$	-0.09
$\pi\pi^*$	-0.14
$n\pi^*$	-0.14
$\pi\pi^*$	-0.16
$n\pi^*$	-0.16
$\pi\pi^*$	<b>-0.31</b>

- Mean deviation of  $-0.17$  eV
- Error increases with excitation energy
- Independent of character of state

- Explanation via orbitals involved in excitations
- Analysis of lowest  $n\pi^*$  and highest  $\pi\pi^*$  states



### III. Analysis

#### Effect of augmenting functions

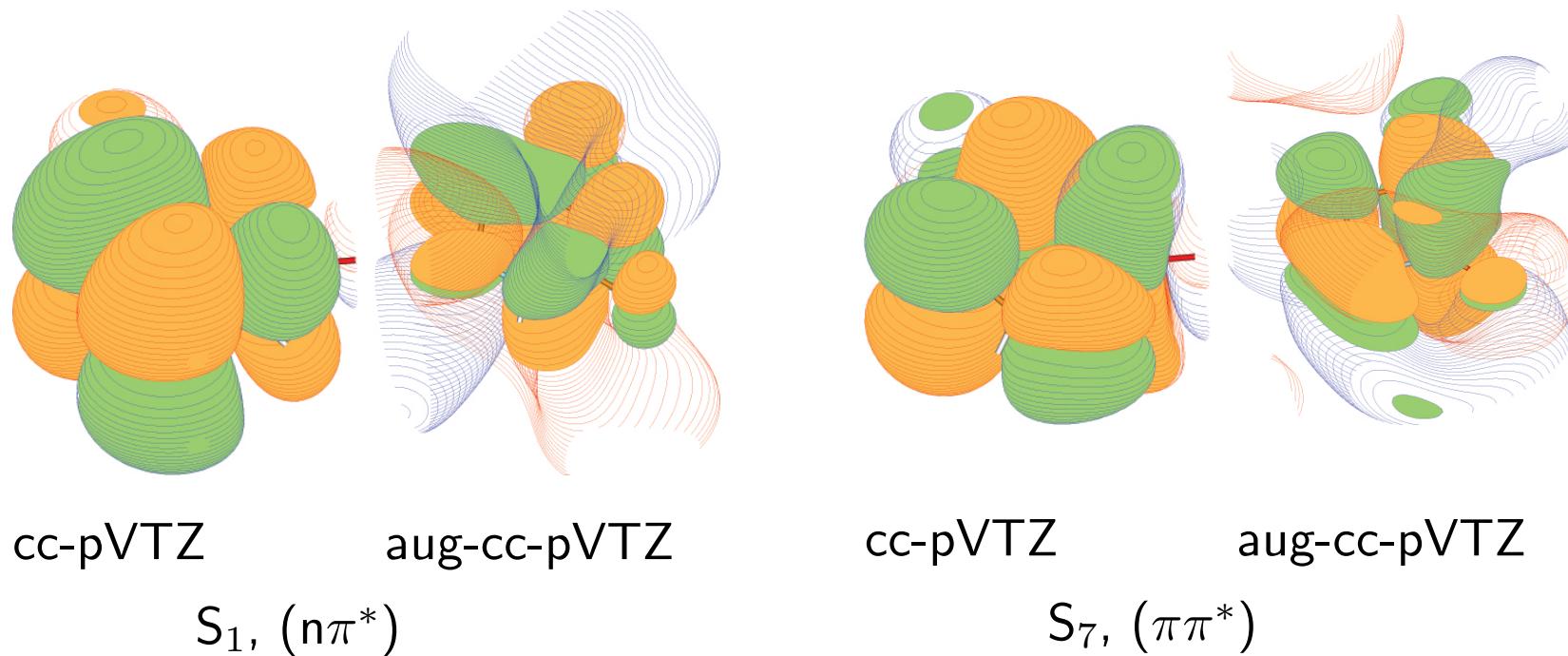
Electron density for involved  $\pi^*$  orbital (contour value 0.01):

$$\langle r^2 \rangle - \langle r \rangle^2 = 12.6$$

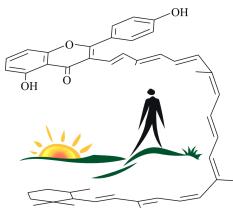
$$35.1$$

$$13.7$$

$$34.8$$

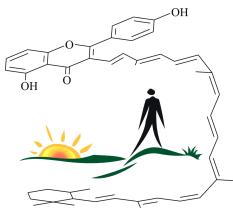


- Involved  $\pi^*$  are most compact virtual orbitals (aug-cc-pVTZ)
- Diffuse functions play significant role for excited states !



## Some Conclusions on Basis Sets

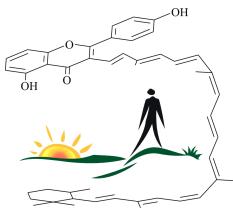
- SV(P) sets **not suited** for the description of excited states
- The cc-pVXZ series is **significantly less accurate** than series aug-cc-pVXZ
- aug-cc-pVXZ sets open for access to Rydberg states
- **Wrong order** of excited states using too small basis sets (guanine)
- Augmentation increases computational effort significantly.



## II. Results

### Nucleic acid base monomers

# Excited states (Uracil, Electron Correlation)



### III. Accuracy: Electron Correlation

#### The CC Hierarchy

O Christiansen, H Koch, P Jørgensen, *Chem. Phys. Lett.* **1995**, *243*, 409-418.

C Hättig, F Weigend, *J. Chem. Phys.* **2000**, *113*, 5154-5161.

$$\text{CC2} \quad \left\langle \mu_2 \left| \hat{H} + [\hat{F}, \hat{T}_2] \right| \text{Ref} \right\rangle = 0$$

$$\text{CCSD} \quad \left\langle \mu_2 \left| \hat{H} + [\hat{H}, \hat{T}_2] + \frac{1}{2} [[\hat{H}, \hat{T}_2], \hat{T}_2] \right| \text{Ref} \right\rangle = 0$$

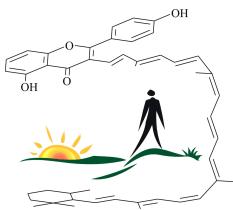
$$\text{CC3} \quad \left\langle \mu_3 \left| [\hat{F}, \hat{T}_3] + [\hat{H}, \hat{T}_2] \right| \text{Ref} \right\rangle = 0$$

CCSDT ...

Scaling of the models  
with system size  $N$ :

CC2	CCSD	CC3	CCSDT
$N^5$	$N^6$	$N^7$	$N^8$

- Fock-operator approximation  
→ reduced cost
- $\hat{F} + \hat{V}^t$   
→ response theory for excited states
- Reduced scaling essential for large-scale applications
- Accuracy?



### III. Accuracy: Electron Correlation

#### The Coupled Cluster Hierarchy

H Koch, O Christiansen, P Jørgensen, J Olsen, *Chem. Phys. Lett.* **1995**, 244, 75-82.

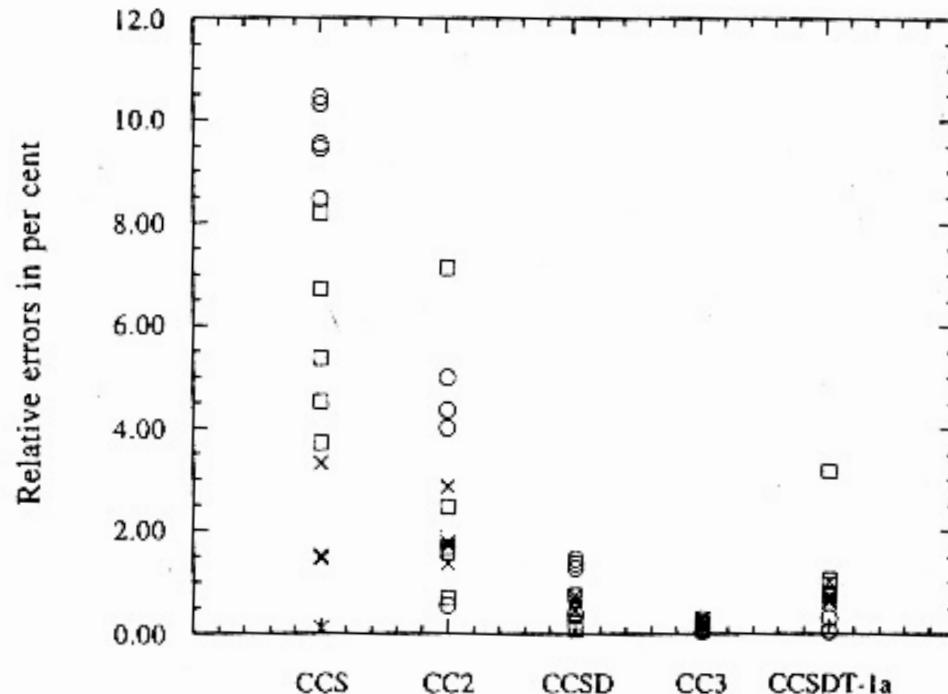
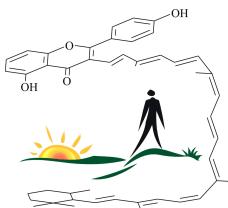


Fig. 1. The percent deviation from FCI excitation energies for the single replacement dominant excitation in Ne (○), BH (×) and CH<sub>2</sub> (□).

- Relative errors in small molecules
- (aug)-cc-pVDZ basis set
- Largest CC2 errors **0.04 eV**
- CCSD consistently better than CC2
- CC3 close to exact values



### III. Accuracy: Electron Correlation

#### The Coupled Cluster Hierarchy

O Christiansen, H Koch, P Jørgensen, J Olsen, *Chem. Phys. Lett.* **1996**, 256, 185-194.

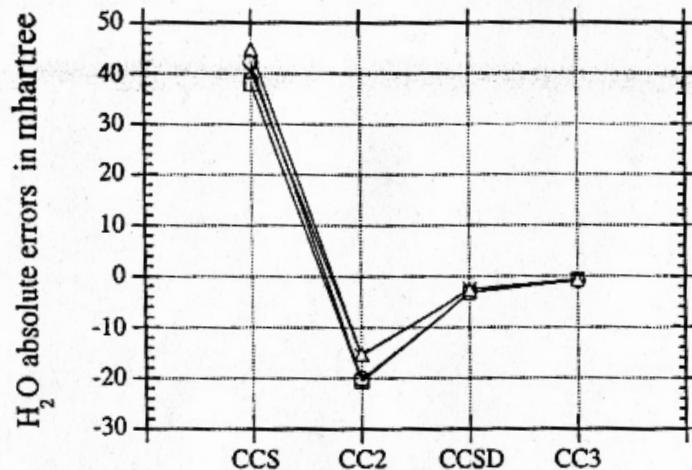


Fig. 1. The absolute deviation from the FCI excitation energies for  $H_2O$ :  $2^1A_1$  (□),  $1^1B_1$  (△),  $1^1B_2$  (×),  $1^1A_2$  (○).

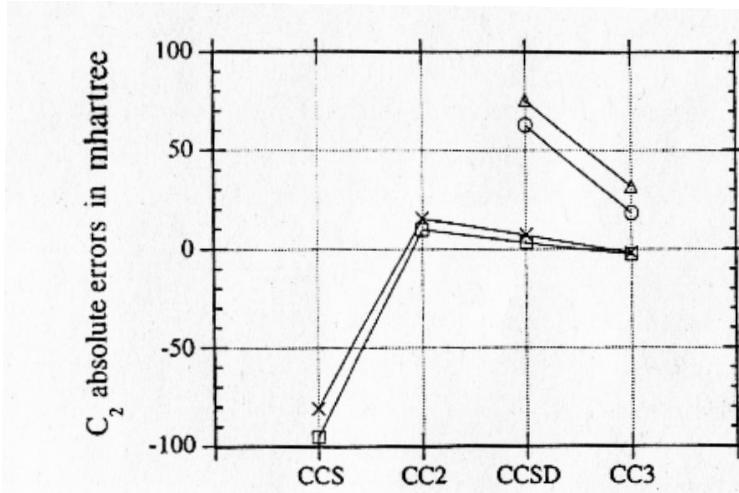
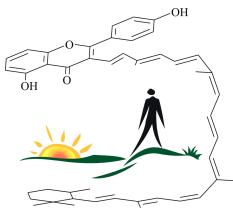


Fig. 3. The absolute deviation from the FCI excitation energies for  $C_2$ :  $1^1\Pi_u$  (□),  $1^1\Delta_g$  (△),  $1^1\Sigma_u^+$  (×),  $1^1\Pi_g$  (○).

- Errors decrease systematically (for correlated models)
- Errors may be positive or negative (different correlation cases)



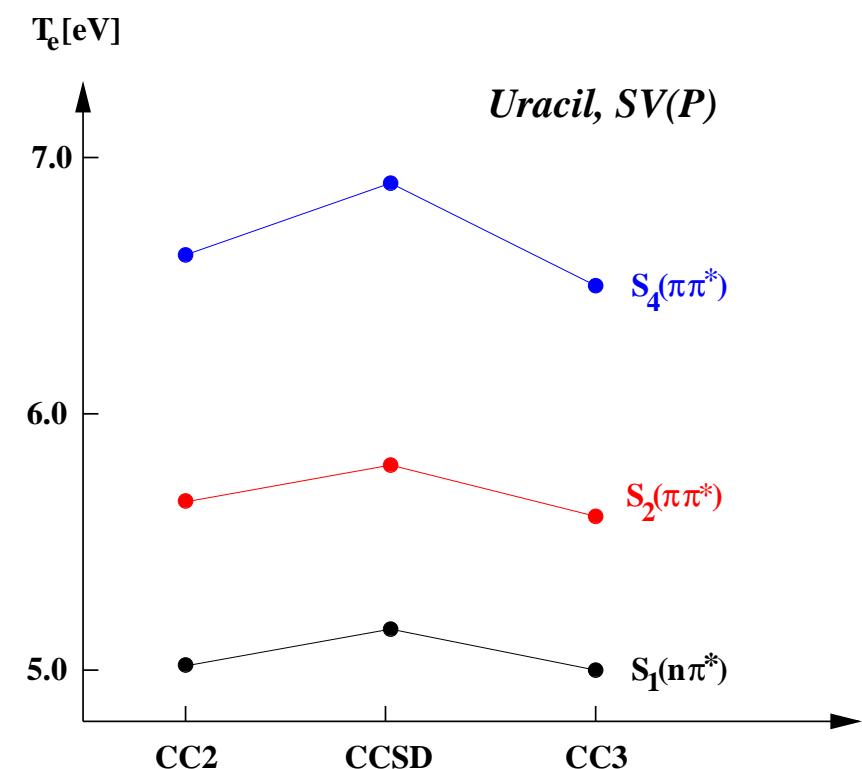
### III. Accuracy: Electron Correlation

#### The NA Bases: Uracil

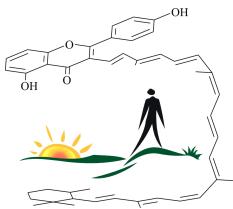
T Fleig, S Knecht, C Hättig, *J. Phys. Chem. A* **2007**, accepted.

State	RI-CC2	CC2	CCSD	CC3	$t_1$
$S_2$	5.67	5.67	5.81	5.61	91%
$S_4$	6.63	6.63	6.91	6.51	88%
$S_1$	5.02	5.01	5.17	5.00	92%

- RI is a valid approximation!
- CC2 values close to CC3
- CCSD does not necessarily give correct answer!
- Triple-excitation character increases in  $S_4$  state



Is this fortuitous or systematic?



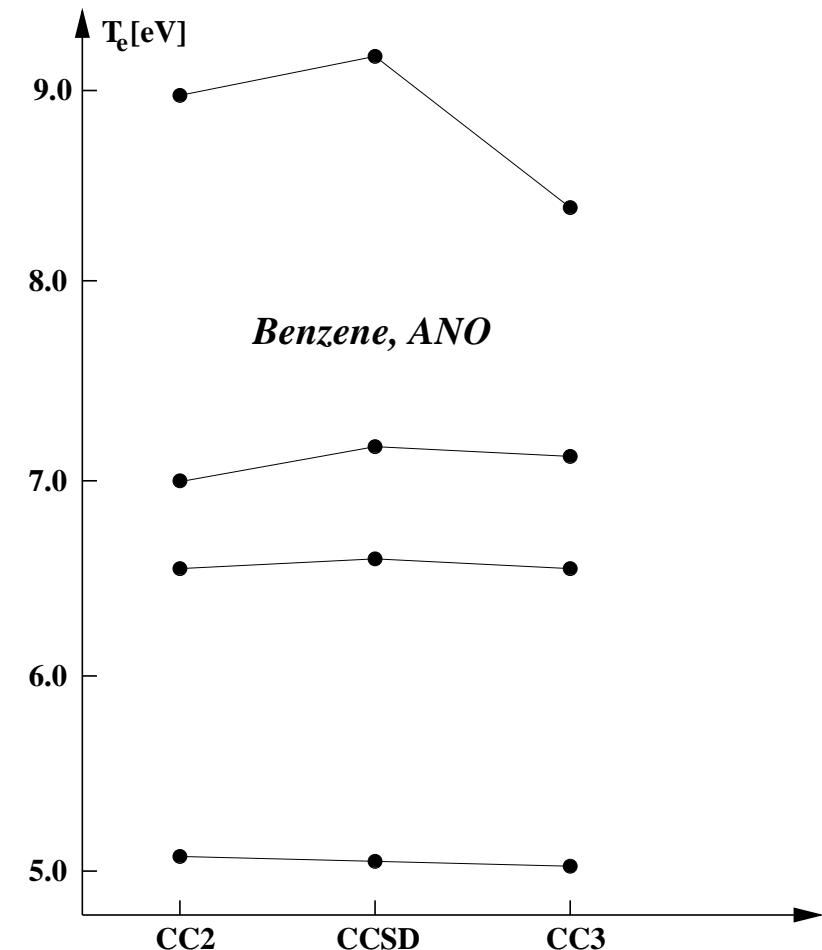
### III. Accuracy: Electron Correlation

#### A comparison: Benzene

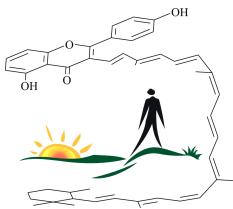
O Christiansen et al., *J. Chem. Phys.* **1996**, *105*, 6921-6939.

State	CC2	CCSD	CC3	$t_1$
$1^1B_{2u} (\pi\pi^*)$	5.27	5.19	5.08	86%
$1^1B_{1u} (\pi\pi^*)$	6.56	6.59	6.54	93%
$1^1E_{1u} (\pi\pi^*)$	7.01	7.17	7.13	93%
$2^1E_{2g} (\pi\pi^*)$	8.97	9.17	8.41	66%

- CC2 is accurate for singly-excited states
- Deviations between models are smaller
- Oscillatory behavior



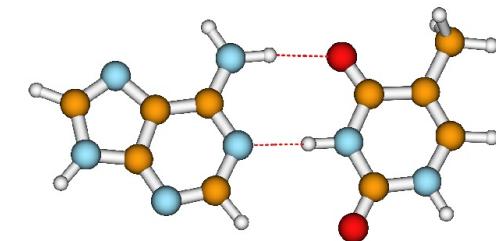
Similar trends in  $\pi$  systems



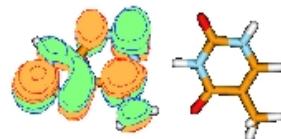
## IV. Results

### Nucleic acid base dimers

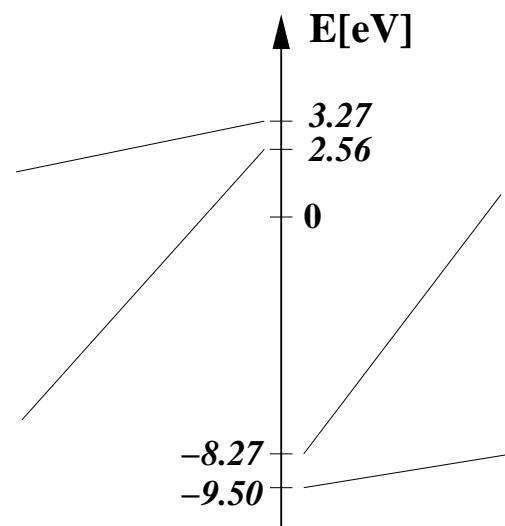
- Exploit knowledge from calibration studies
- Identify CT states in spectra
- Ground-state structures ( $C_1$ , frozen core); vertical excited-state calculations



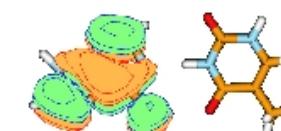
LUMO Adenine



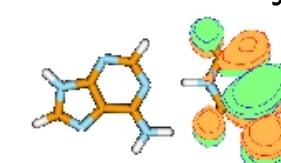
LUMO Thymine

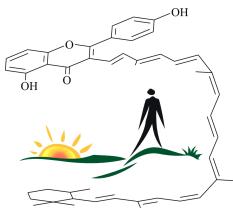


HOMO Adenine



HOMO Thymine





## IV. Results

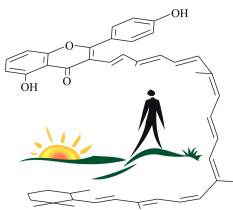
### Test: DFT performance

Adenine-Thymine (AT) base pair, Watson-Crick structure

RI-TDDFT (SVWN), TZVP			
State	System	Character	T <sub>v</sub> [eV]
S <sub>1</sub>	A→T	$\pi - \pi^*$	3.30
S <sub>2</sub>	A→T	$n - \pi^*$	3.74
S <sub>3</sub>	T→T	$n - \pi^*$	4.21
S <sub>4</sub>	T→A	$\pi - \pi^*$	4.33
S <sub>5</sub>	A→A	$n - \pi^*$	4.37
S <sub>6</sub>	T→T(A)	$\pi - \pi^*$	4.52
S <sub>7</sub>	A→A	$\pi - \pi^*$	4.53
S <sub>8</sub>	AT→T	$n/\sigma - \pi^*$	4.59
S <sub>9</sub>	A→T	$\pi - \pi^*$	4.61

RI-CC2, TZVP			
State	System	Character	T <sub>v</sub> [eV]
S <sub>1</sub>	AT→T	$n - \pi^*$	5.12
S <sub>2</sub>	A→A	$\pi - \pi^*$	5.26
S <sub>3</sub>	A→A	$\pi - \pi^*$	5.31
S <sub>4</sub>	A→A	$\pi - \pi^*$	5.41
S <sub>5</sub>	T→T	$\pi - \pi^*$	
S <sub>6</sub>	A→A	$n - \pi^*$	5.46
S <sub>7</sub>	A→A	$n - \pi^*$	5.96
S <sub>7</sub>	A→T	$\pi - \pi^*$	6.17
S <sub>8</sub>	AT→T	$n - \pi^*$	6.31
S <sub>9</sub>	AT→A	$n - \pi^*$	6.51

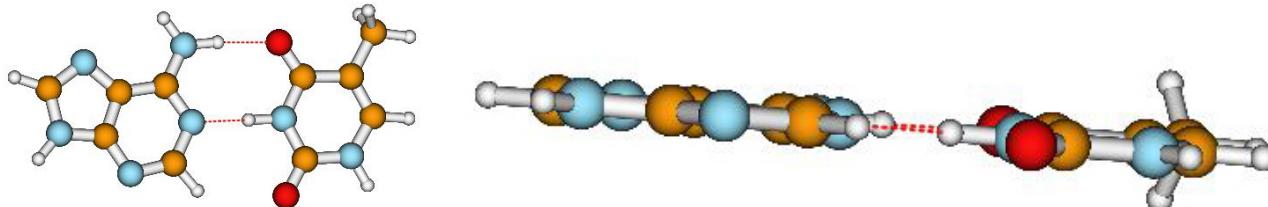
- DFT: Low-lying CT states, physically incorrect
- CC2: Spectrum corresponds to qualitative expectations



## IV. Results

### Adenine-Thymine: Structure and excited states

WC structure, opt. RI-CC2, aug-cc-pVDZ

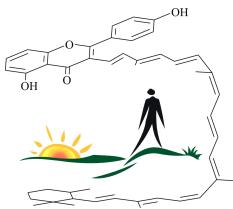


RI-CC2, aug-cc-pVDZ

RI-CC2, cc-pVDZ<sup>11</sup>

State	Excitation Type	Character (CT)	T <sub>v</sub> [eV]	T <sub>v</sub> [eV]
S <sub>1</sub>	T→T/A→T	n - π* (20 %)	4.81	5.13
S <sub>2</sub>	T→T	π - π*	4.94	5.37
S <sub>3</sub>	A→A	π - π*	5.05	5.25
S <sub>4</sub>	A→A	π - π*	5.05	5.45
S <sub>5</sub>	A→A	π - Ryd	5.29	
S <sub>6</sub>	A→A/A→T	n - π* (20 %)	5.31	5.51
S <sub>7</sub>	T→T	π - Ryd	5.47	
S <sub>8</sub>	A→A	π - Ryd	5.71	
S <sub>9</sub>	A→A/A→T	π - π* (40 %)	5.71	6.26
...				
S <sub>12</sub>	T→T/A→T	π - π* (45 %)	5.91	6.42
...				

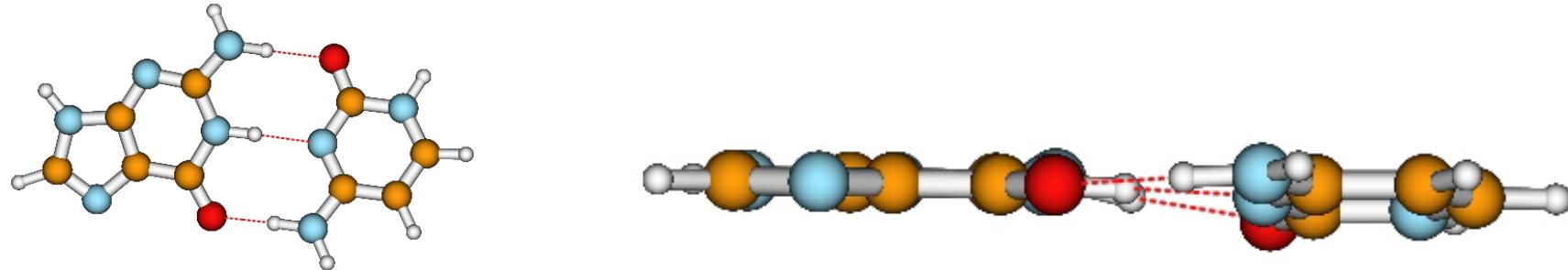
<sup>11</sup>S Perun, A L Sobolewski, W Domcke, *J. Phys. Chem. A* **2006**, *110*, 9031-9038



## IV. Results

### Guanine-Cytosine: Structure and excited states

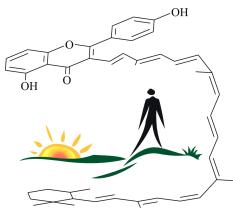
WC structure, opt. RI-CC2, aug-cc-pVDZ



RI-CC2, aug-cc-pVDZ

State	Excitation Type	Character (CT contr.)	T <sub>v</sub> [eV]
S <sub>1</sub>	G→G/G→C	π - π* (35 %)	4.61
S <sub>2</sub>	C→C	π - π*	4.77
S <sub>3</sub>	G→G	π - Ryd	4.77
S <sub>4</sub>	G→G/G→C	π - π* (35 %)	5.03
S <sub>5</sub>	G→G/G→C	π - Ryd (40 %)	5.18
S <sub>6</sub>	G→G	π - π*	5.20
S <sub>7</sub>	C→C	π - π*	5.22
S <sub>8</sub>	C→C/G→C	n - π* (20 %)	5.27
S <sub>9</sub>	G→G/G→C	π - Ryd (50 %)	5.41
...			

- Strong CT mixing
- Difficult designation

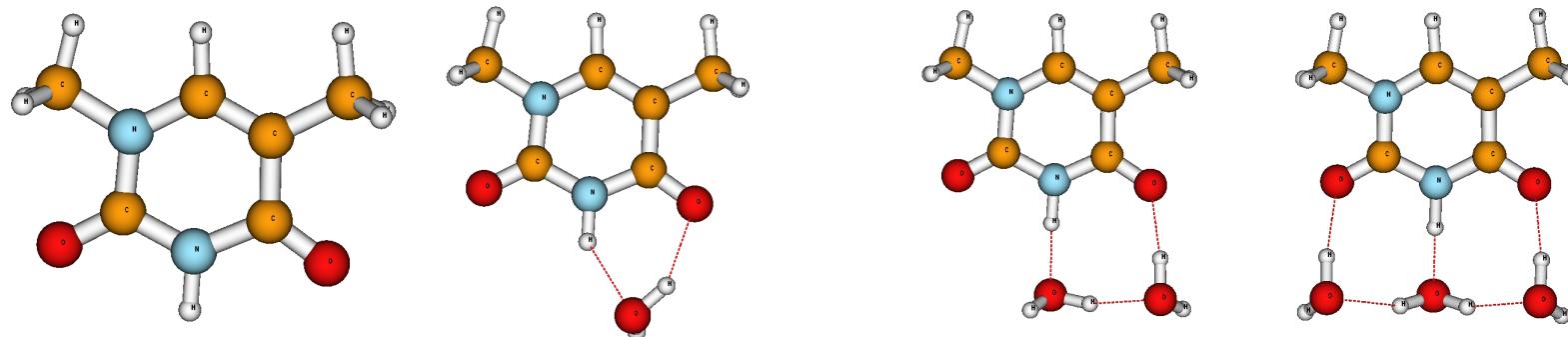


## IV. Results

### 1-Methylthymine

Investigation of an excited-state process.

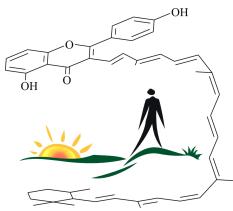
RI-CC2/cc-pVDZ optimized structures:



- Trends by adding water molecules to form a solvation shell

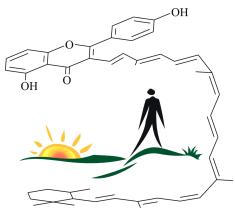


Poster by Mihajlo Etinski



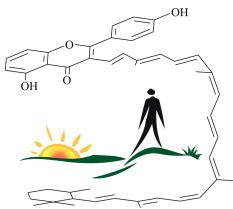
# Final Conclusions

- CC2 correlation treatment appears to be **very accurate**  
Errors << 0.4 eV
- CC2 applicable for larger molecules
- Description of excited states:  
Selection of basis set is **a crucial issue**  
(aug) leads to **significant** corrections
- In present case (NA bases and dimers):  
**Favorable** error cancellation  
Basis set: too low; Correlation: too high!
- Advocated model: **RI-CC2/aug-cc-pVDZ**



Thanks !

- Stefan Knecht (Düsseldorf)
- Christof Hättig (Bochum)
- Ove Christiansen (Aarhus)

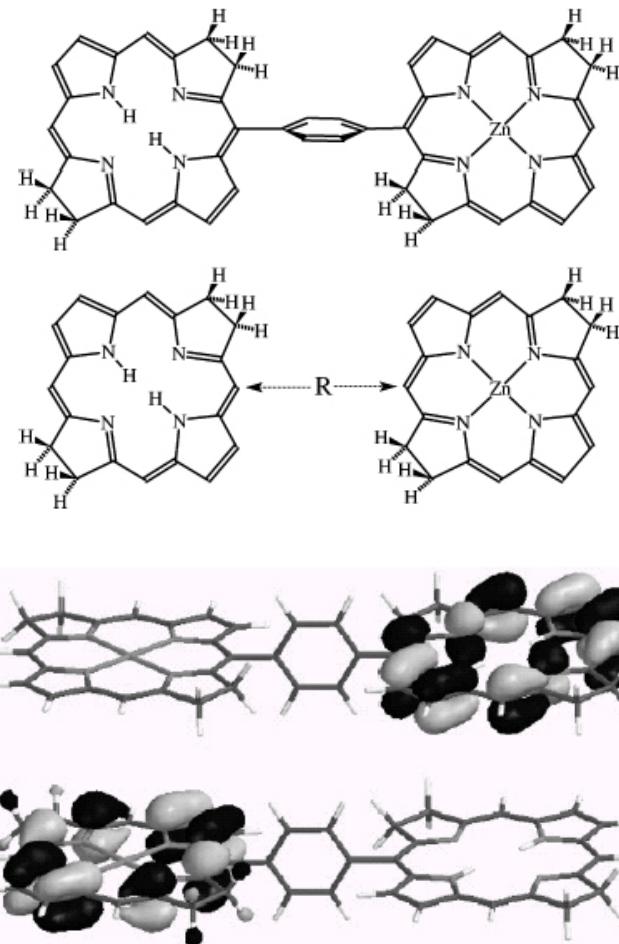


# I. Introduction and Motivation

## Charge-Transfer Problem (in DFT)

### Zinc-Bacteriochlorin-Bacteriochlorin complex

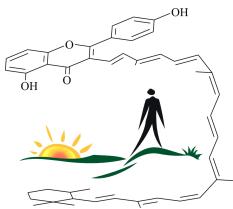
A Dreuw, M Head-Gordon, *J. Am. Chem. Soc.* **2004**, *126*, 4007-4016



(TDDFT, BLYP functional, 6-31G\*):

State	Excitation energies [eV]			Monomers	
	Full	Model	Transition	Theory	Exp.
$S_1$	1.33	1.32	$ZnBC \rightarrow BC$		
$S_2$	1.46	1.47	$BC \rightarrow ZnBC$		
$S_3$	1.86	1.90	$BC \rightarrow ZnBC$		
$S_4$	1.94	1.96	$ZnBC \rightarrow BC$		
$S_5$	2.05	2.07	$\pi - \pi^* ZnBC$	2.07	1.65
$S_6$	2.09	2.12	$\pi - \pi^* BC$	2.10	1.6
$S_7$	2.38	2.40	$\pi - \pi^* BC$	2.39	2.3
$S_8$	2.42	2.46	$\pi - \pi^* ZnBC$	2.44	2.2
$S_9$	2.43	2.42	$ZnBC \rightarrow BC$		

⇒ Spectrum ist qualitatively wrong !

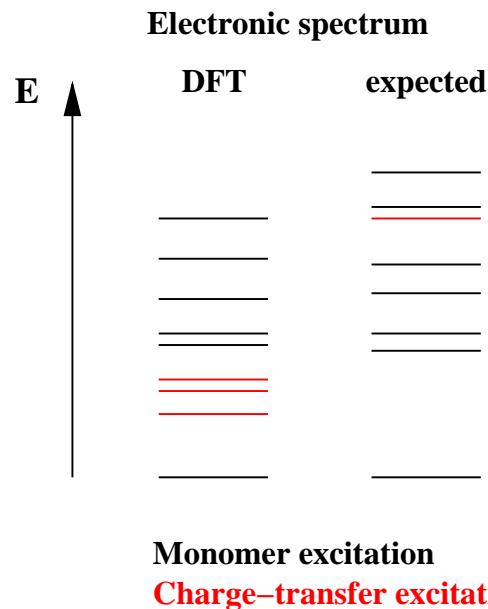


# I. Introduction and Motivation

## Charge-Transfer Problem (in DFT)

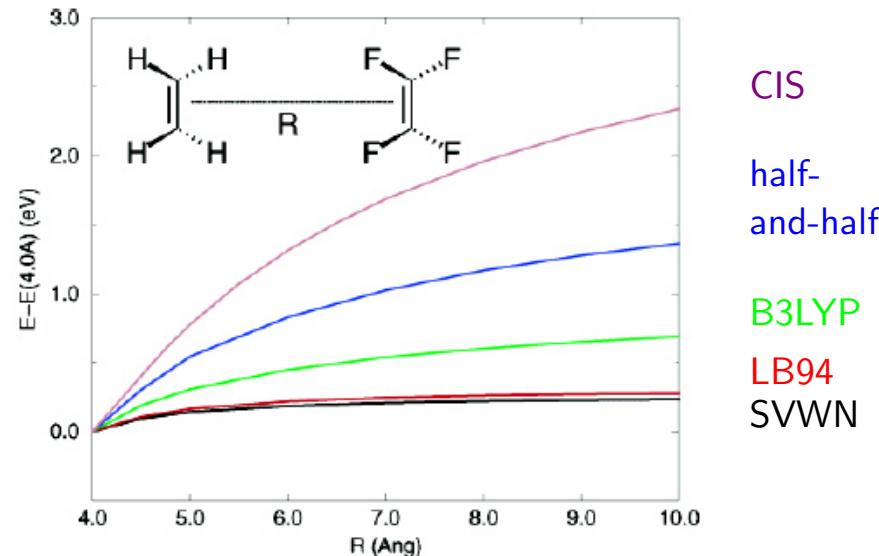
### Problem 1:

CT excitation energies are considerably too small



### Problem 2:

CT states exhibit wrong asymptotic behavior



1. A Dreuw, M Head-Gordon, *Chem. Rev.* **2005**, *105*, 4009-4037
2. A Dreuw, M Head-Gordon, *J. Am. Chem. Soc.* **2004**, *126*, 4007-4016
3. A Dreuw, J L Weisman, M Head-Gordon, *J. Chem. Phys.* **2003**, *119*, 2943-2946