

# **V23 - Stochastic Dynamics simulations of a photosynthetic vesicle**

**where bioinformatics meets biophysics**

**I** Introduction: prelude photosynthesis

**II** Process view and geometric model of a chromatophore vesicle  
Tihamér Geyer & V. Helms (Biophys. J. 2006a, 2006b)

**III** Stochastic dynamics simulations  
T. Geyer, Florian Lauck & V. Helms (J. Biotechnol. 2007)

**IV** Parameter fit through evolutionary algorithm  
T. Geyer, X. Mol, S. Blaß & V. Helms (PLoS ONE 2010)

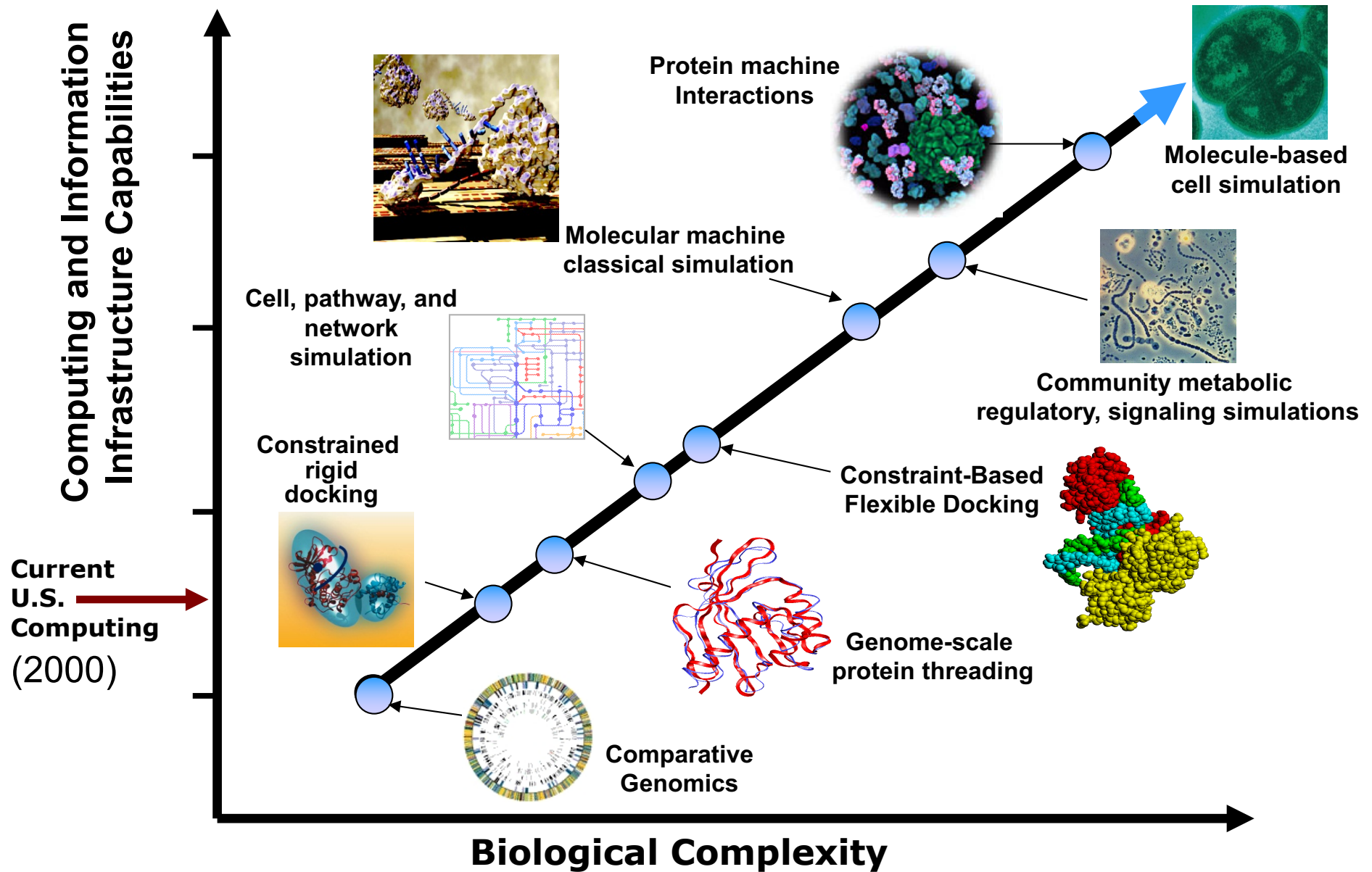
## Content of final exam (March 3, 2017)

Lecture	Slides relevant for exam
1	18-24
2	1-14, 18-20
3	16-27
4	All
5	1-32
6	24-37
7	None
8	1-18
9	1-15
10	6-18, 30
11	8-33
12	The main ideas of 1-14, 27-43

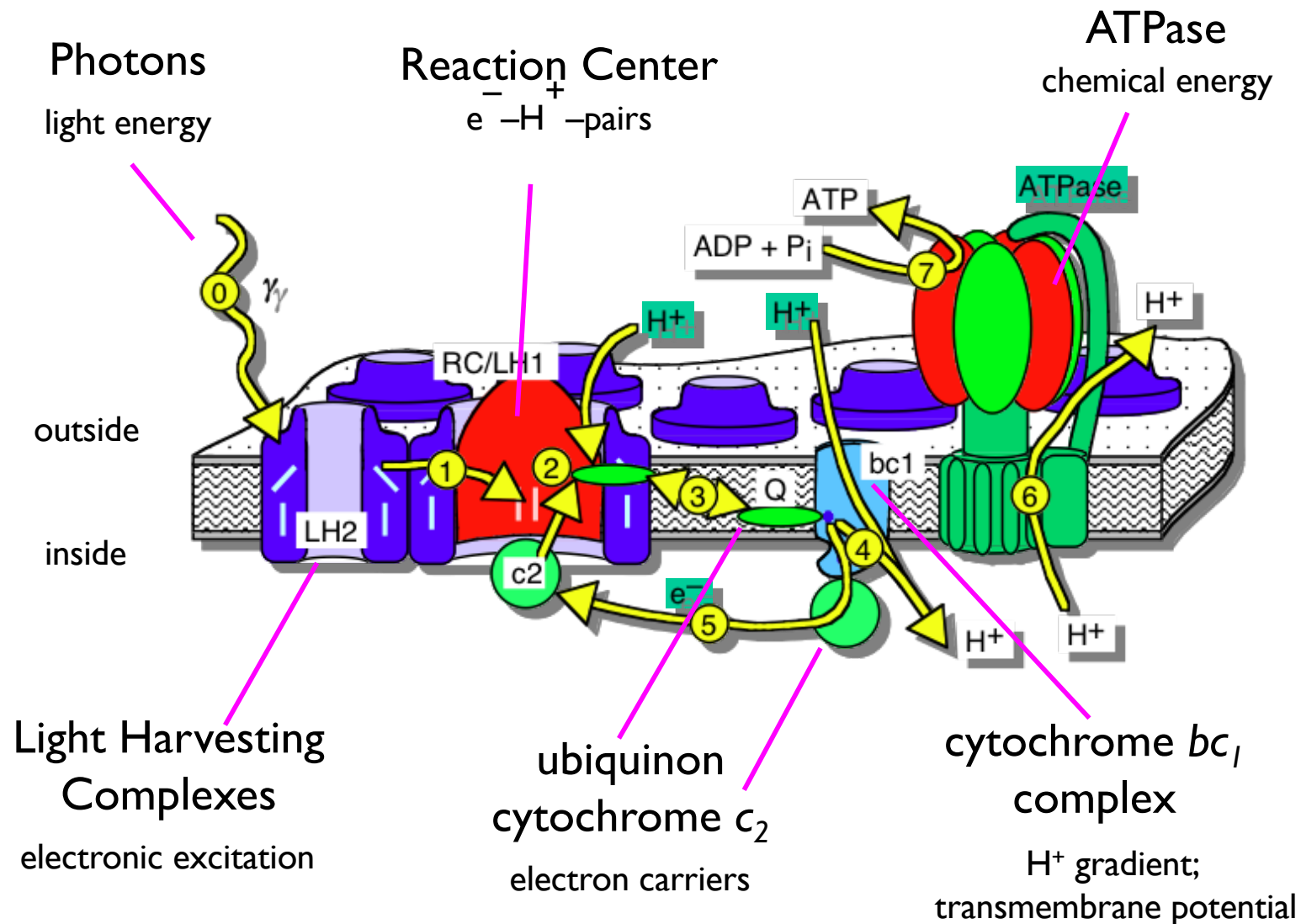
Lecture	Slides relevant for exam
13	1-30
14	1-19, 25-30, 33, 42
15	2,3, 6-8, 28, 30, 39
16	Main ideas of 1-16
17	1-34
18	1-14
19	18-41
20	17-23
21	None
22	None
23	None
24	None
25	None

**Relevant are also the assignments !  
(theoretical parts, not the programming parts)**

# "Genomes To Life" Computing Roadmap (NIH/DOE)

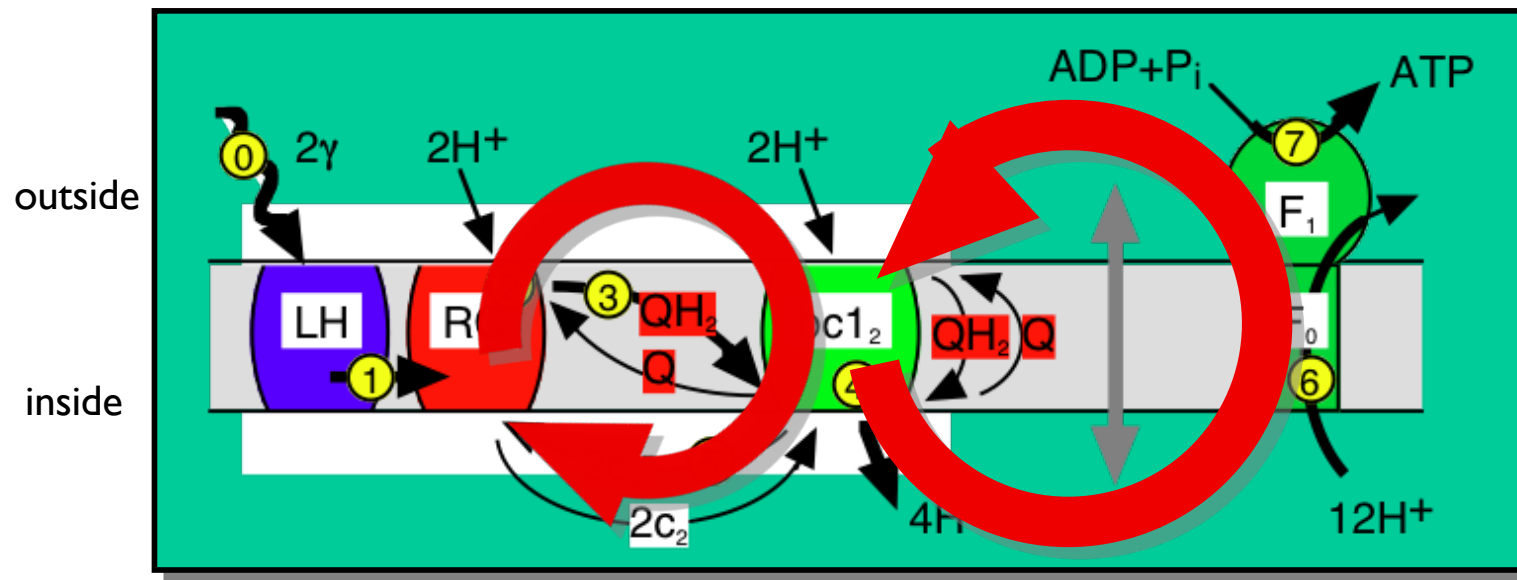
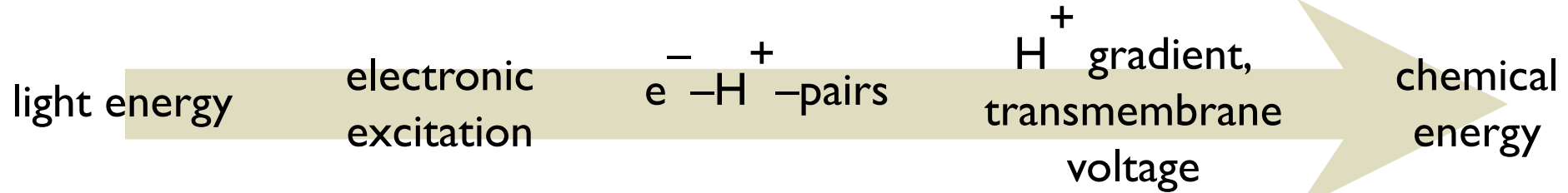


# Bacterial Photosynthesis 101



# Photosynthesis – cycle view

The conversion chain: stoichiometries must match turnovers!



2 cycles:

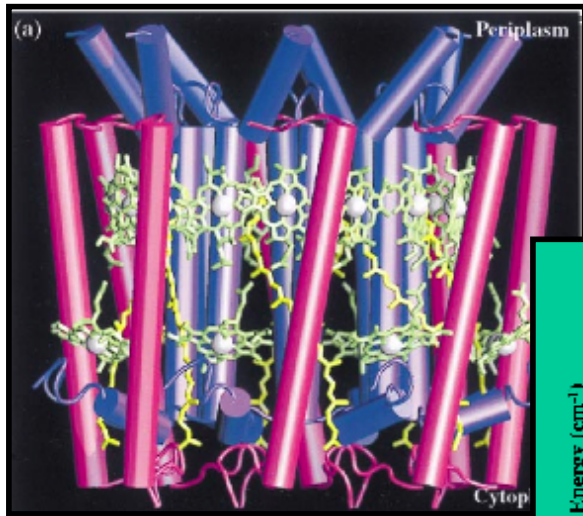
electrons

protons

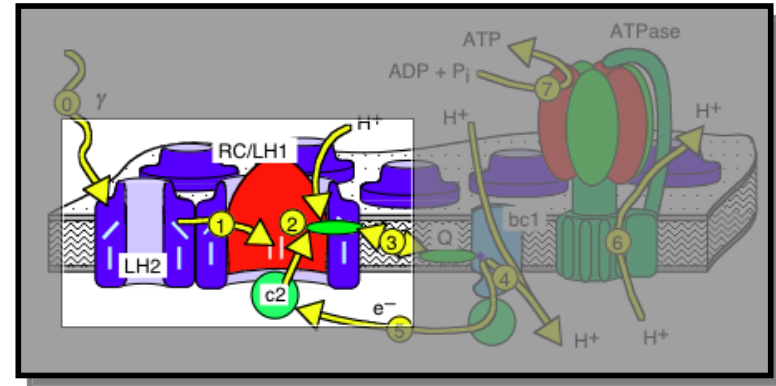
# LH1 / LH2 / RC — a la textbook

Collecting photons

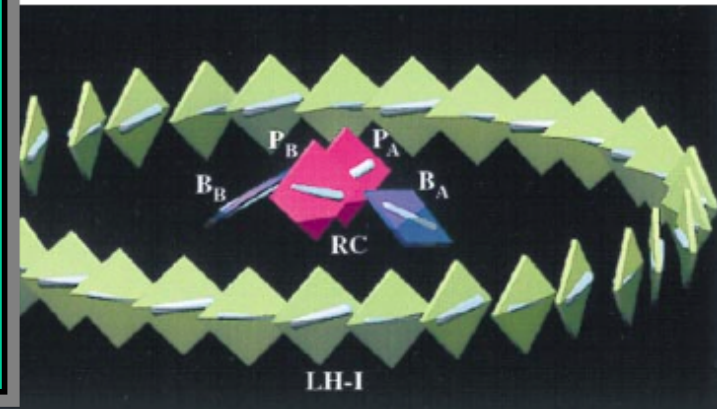
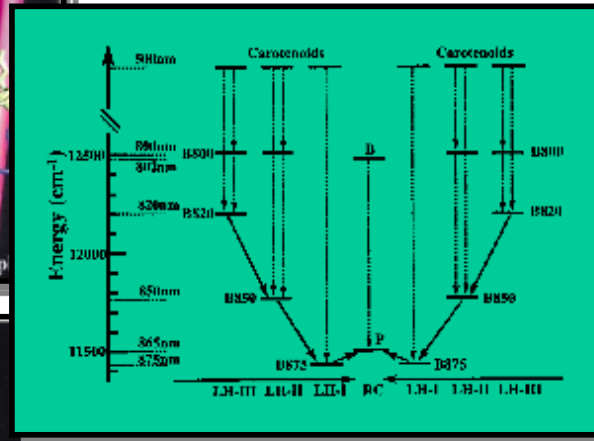
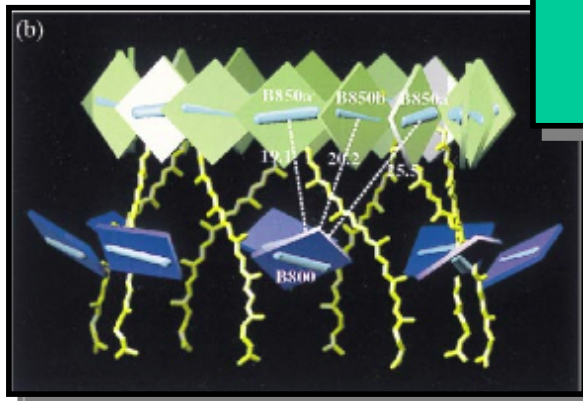
LH2: 8  $\alpha\beta$  dimers



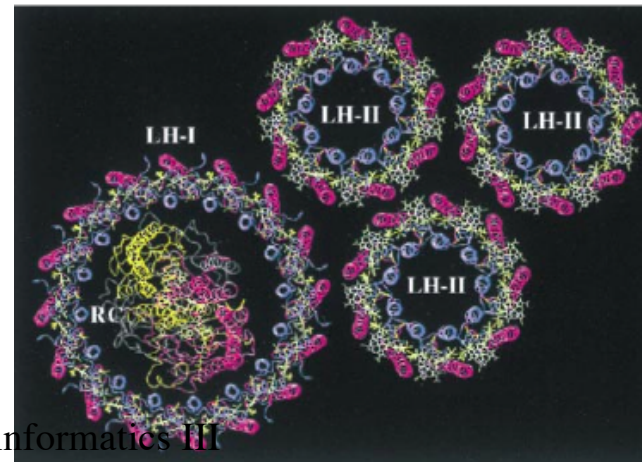
downhill transport of  
excitons  
LH2  $\rightarrow$  LH1  $\rightarrow$  RC



LH1: 16  $\alpha\beta$  dimers

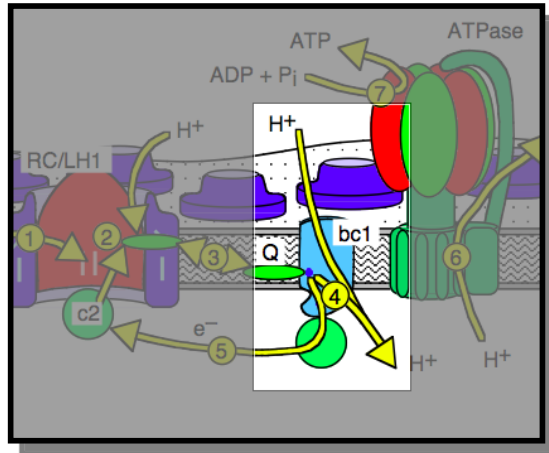


B800, B850, Car.



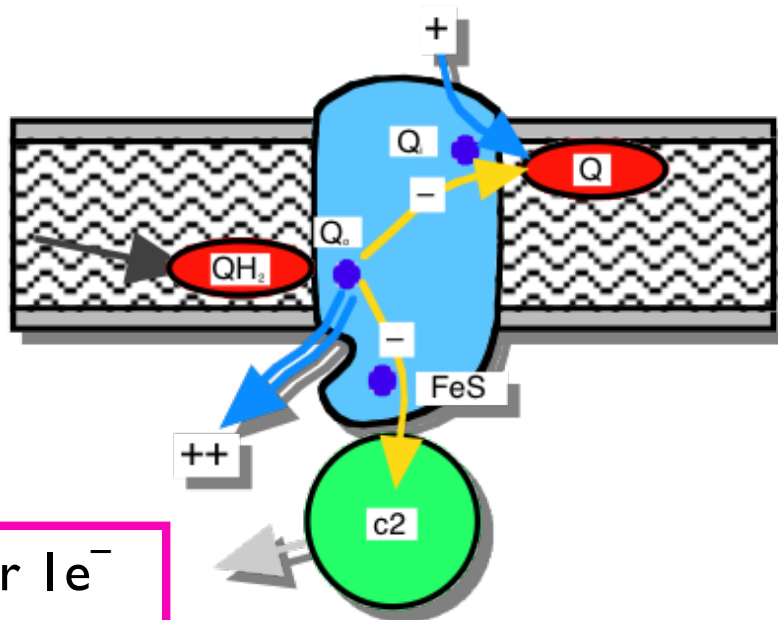


# The Cytochrome *bc*<sub>1</sub> complex

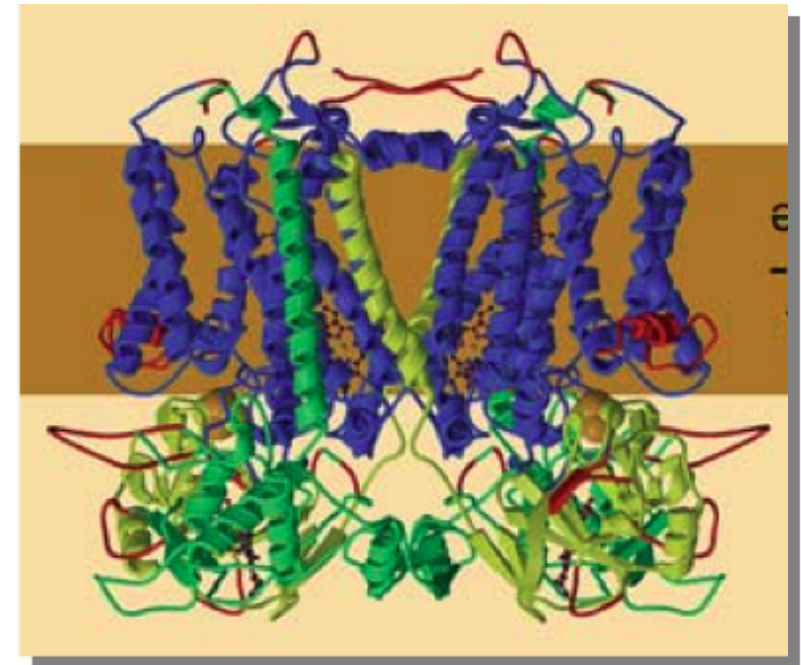


the "proton pump"

Q-cycle:



$2\text{H}^+$  per  $1\text{e}^-$

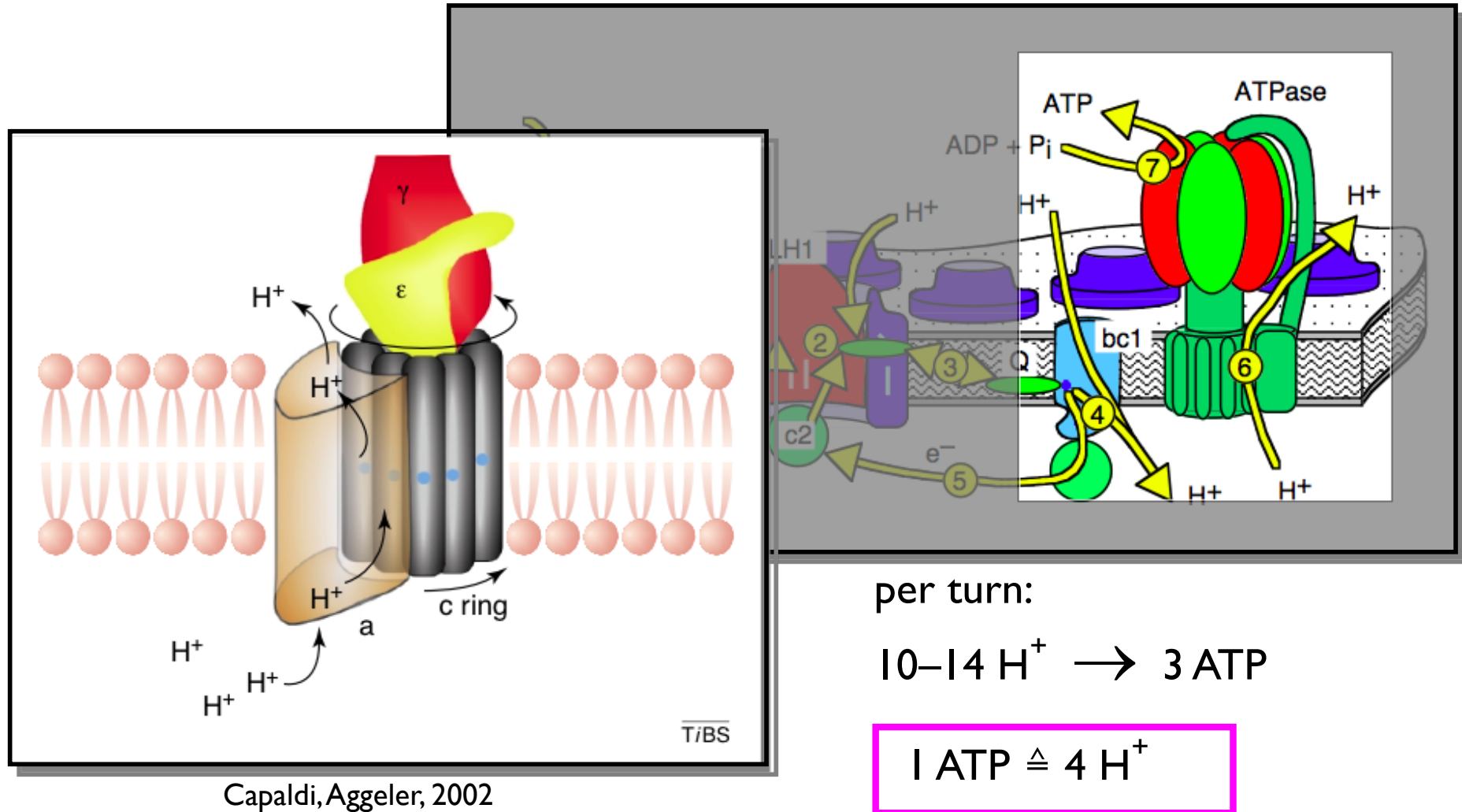


Berry, etal, 2004

X-ray structures known  
always forms a dimer

# The $F_0F_1$ -ATP synthase I

at the end of the chain: producing ATP from the  $H^+$  gradient



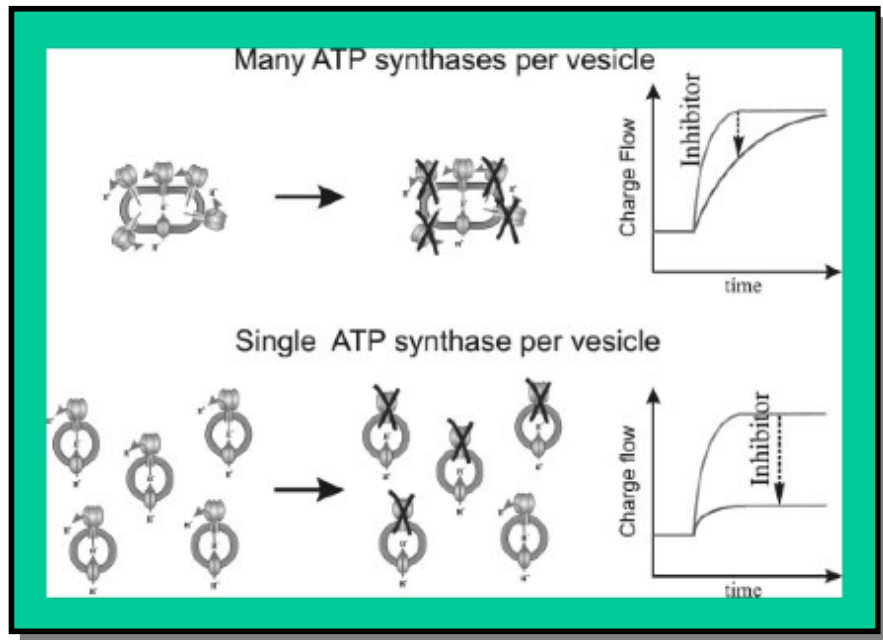


# The $F_1F_0$ -ATP synthase

"...mushroom like structures observed in AFM images..."

ATPase is "visible"

1 ATPase per vesicle

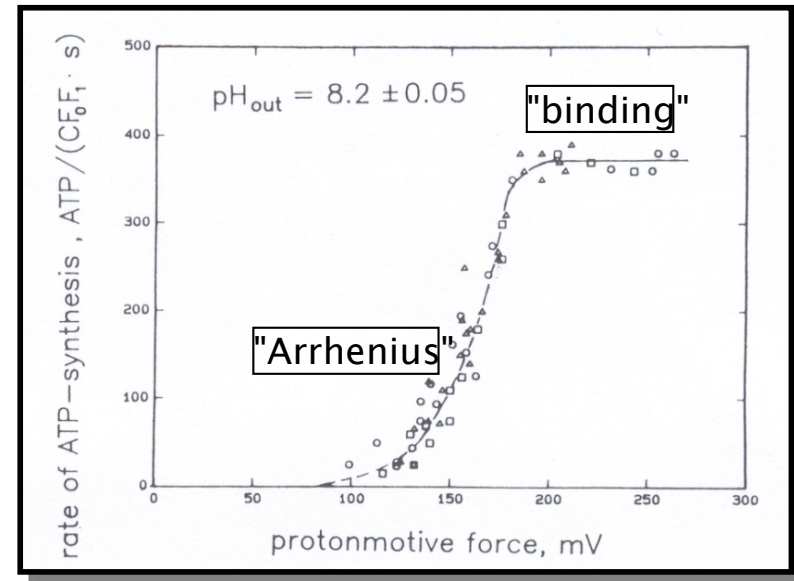


Feniouk et al, 2002

per turn: 10–14  $H^+$  per 3 ATP

$$1 \text{ ATP} \triangleq 4 H^+$$

limited throughput of the ATPase



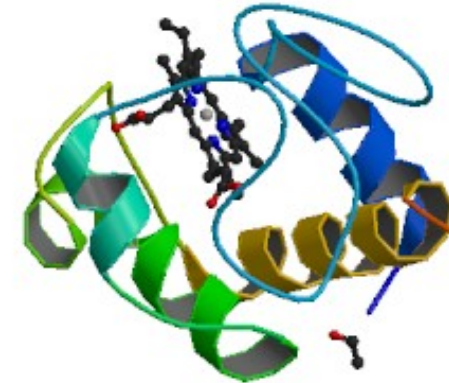
ATPase from	ATP/s	$H^+$ /s
chloroblasts	<400	1600
E. coli	<100	400

Gräber et al, 1991, 1999

# The electron carriers

**Cytochrome c:** carries electrons from  $bc_1$  to R

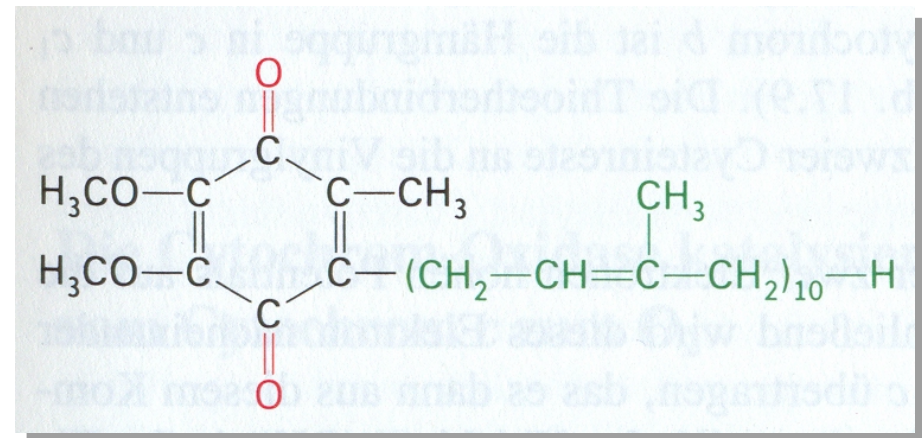
- heme in a hydrophilic protein shell
- 3.3 nm diameter, water-soluble



**Ubiquinone UQ10:**

carries electron-proton pairs  
from RC to  $bc_1$

- long (2.4 nm)  
hydrophobic  
isoprenoid tail,  
membrane-  
soluble

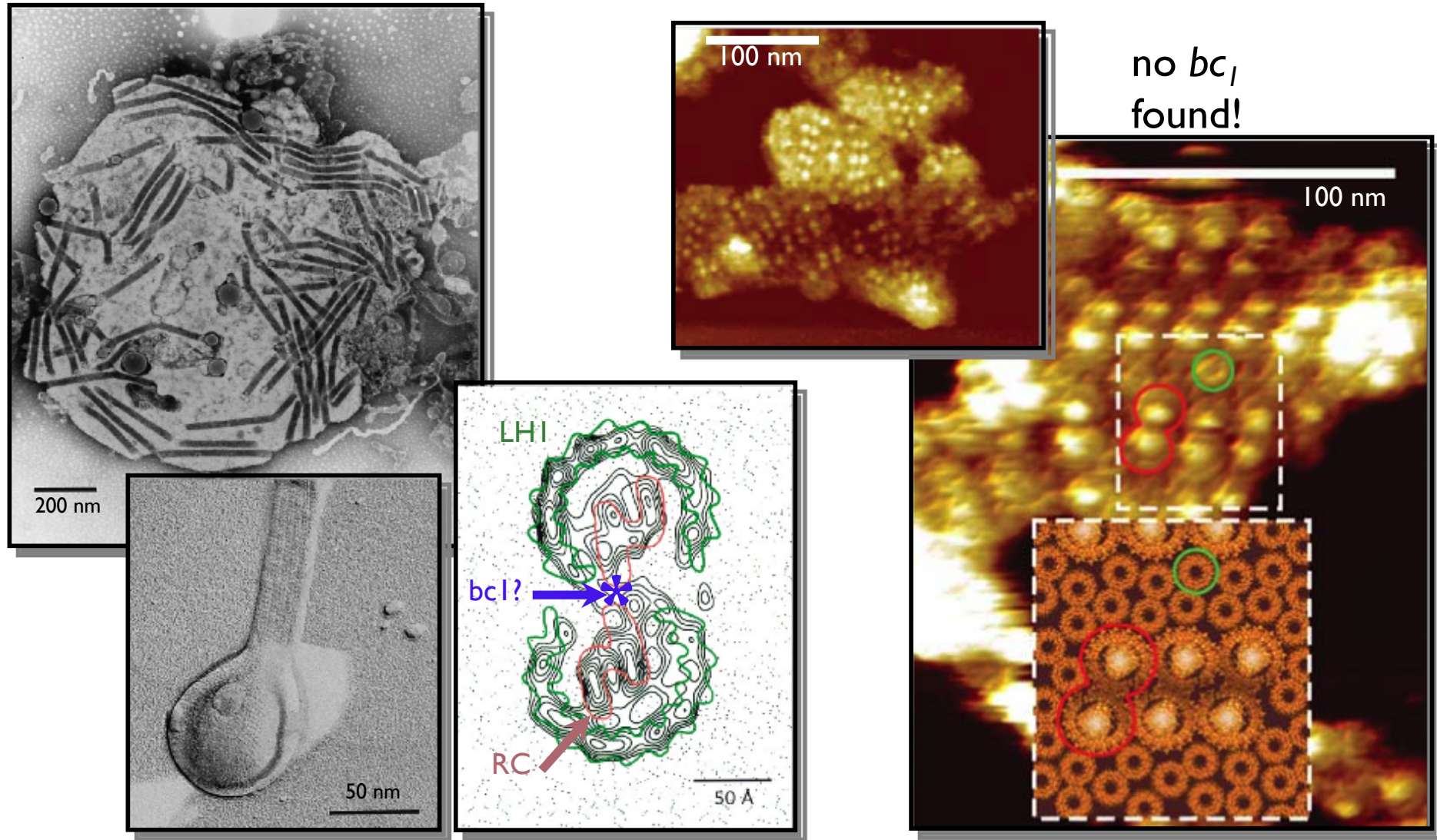


taken from Stryer

# Tubular membranes – photosynthetic vesicles where are the $bc_1$ complexes and the ATPase?

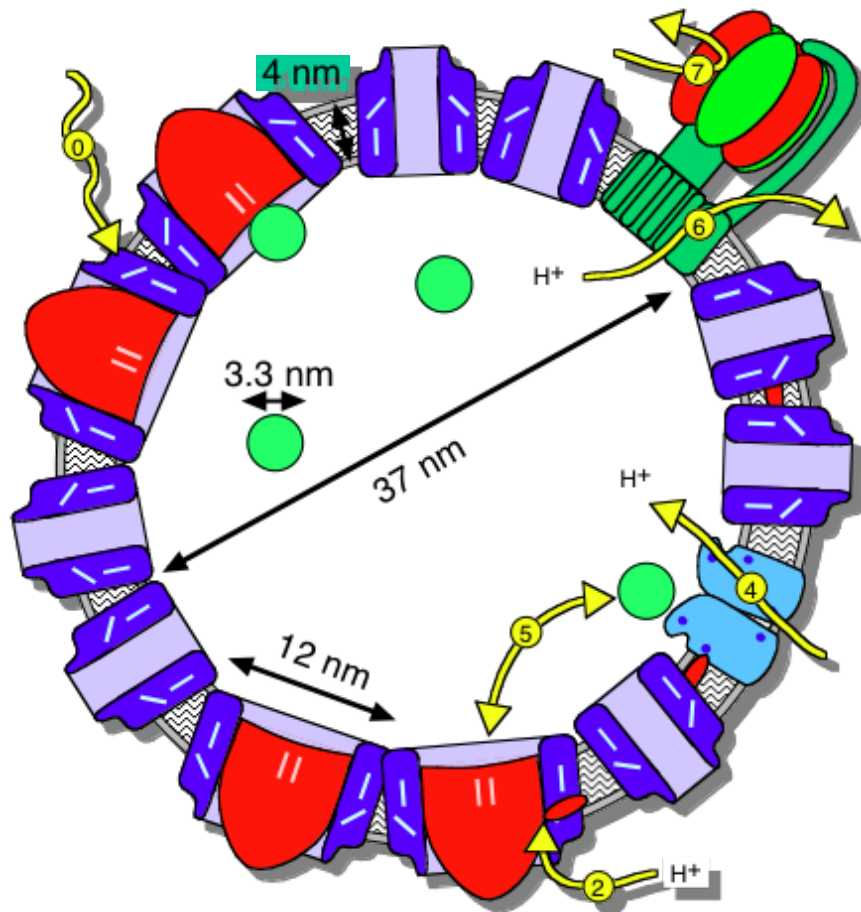
Jungas et al., 1999

Bahatyrova et al., 2004





# Chromatophore vesicle: typical form in *Rh. sphaeroides*



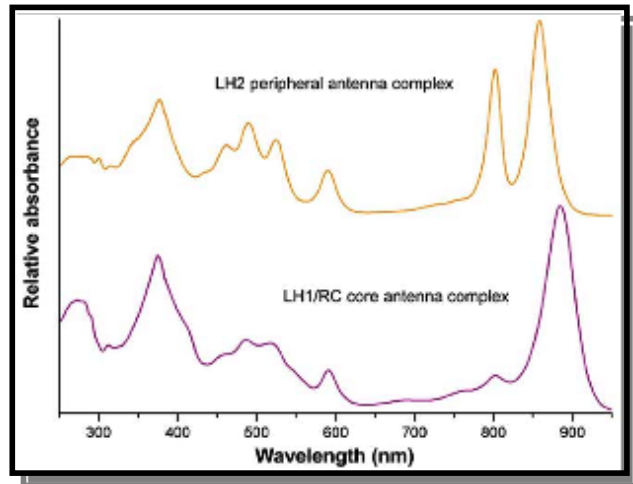
Lipid vesicles  
30–60 nm diameter  
H<sup>+</sup> and cyt c inside

average  
chromatophore surface  
vesicle, 45 nm Ø: 6300 nm<sup>2</sup>

Vesicles are really small!

# Photon capture rate of LHC's

relative absorption spectrum  
of LH1/RC and LH2

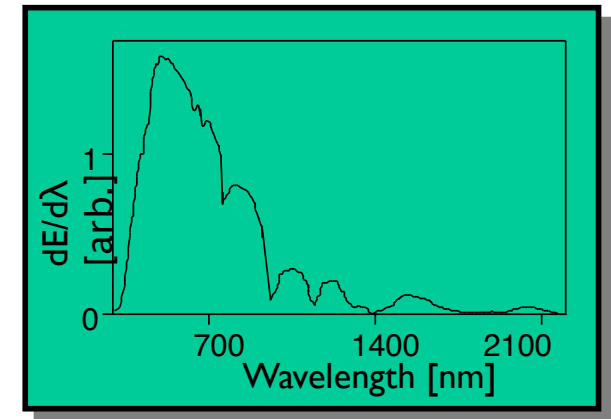


Cogdell et al, 2003

+ Bchl extinction coeff.  
normalization ( $\sigma_{\text{Bchl}} = 2.3 \text{ \AA}^2$ )

Franke, Amesz, 1995

sun's spectrum at ground  
(total: 1 kW/m<sup>2</sup>)



Gerthsen, 1985

multiply

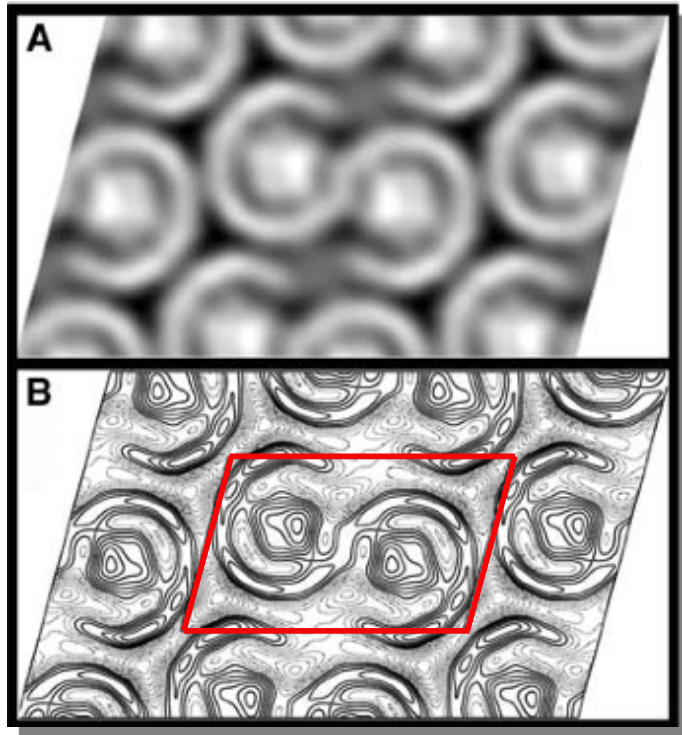
capture rate:  $0.1 \frac{\gamma}{\text{s kW Bchl}}$

typical growth condition:  
18 W/m<sup>2</sup> Feniouk et al, 2002

LH1: 16 * 3 Bchl	14 γ/s
LH2: 10 * 3 Bchl	10 γ/s

# LH1 / LH2 / RC — native

electron micrograph  
and density map



Siebert et al, 2004  
125 \* 195 Å<sup>2</sup>,  $\gamma = 106^\circ$

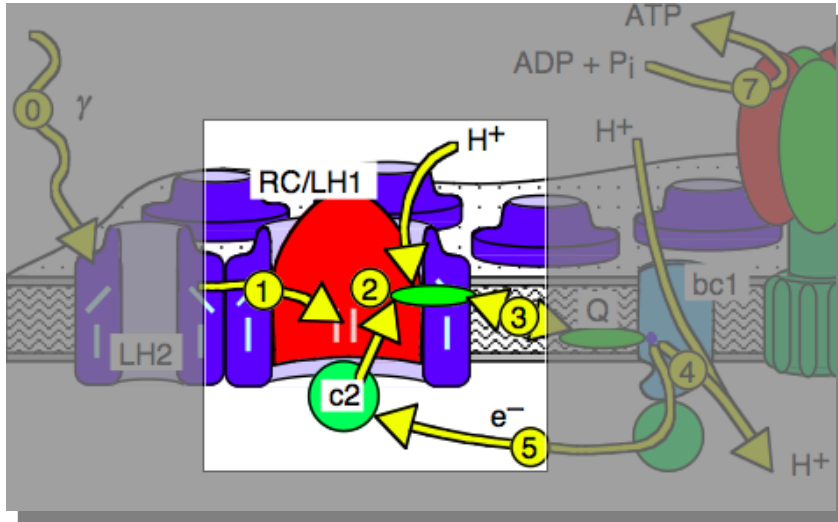
	Area per:	per vesicle (45 nm)
LH1 monomer (hexagonal)	146 nm <sup>2</sup>	
LH1 dimer	234 nm <sup>2</sup>	
LH2 monomer	37 nm <sup>2</sup>	
LH1 <sub>2</sub> + 6 LH2	456 nm <sup>2</sup>	11

Chromatophore surface  
vesicle, 45 nm Ø: 6300 nm<sup>2</sup>



# Photon processing rate at the RC

Which process limits the RCs turnover?



Unbinding of the quinol

25 ms Milano et al. 2003

+ binding, charge transfer  
 $\approx 50$  ms per quinol (estimate)

with  $2e^- H^+$  pairs per quinol

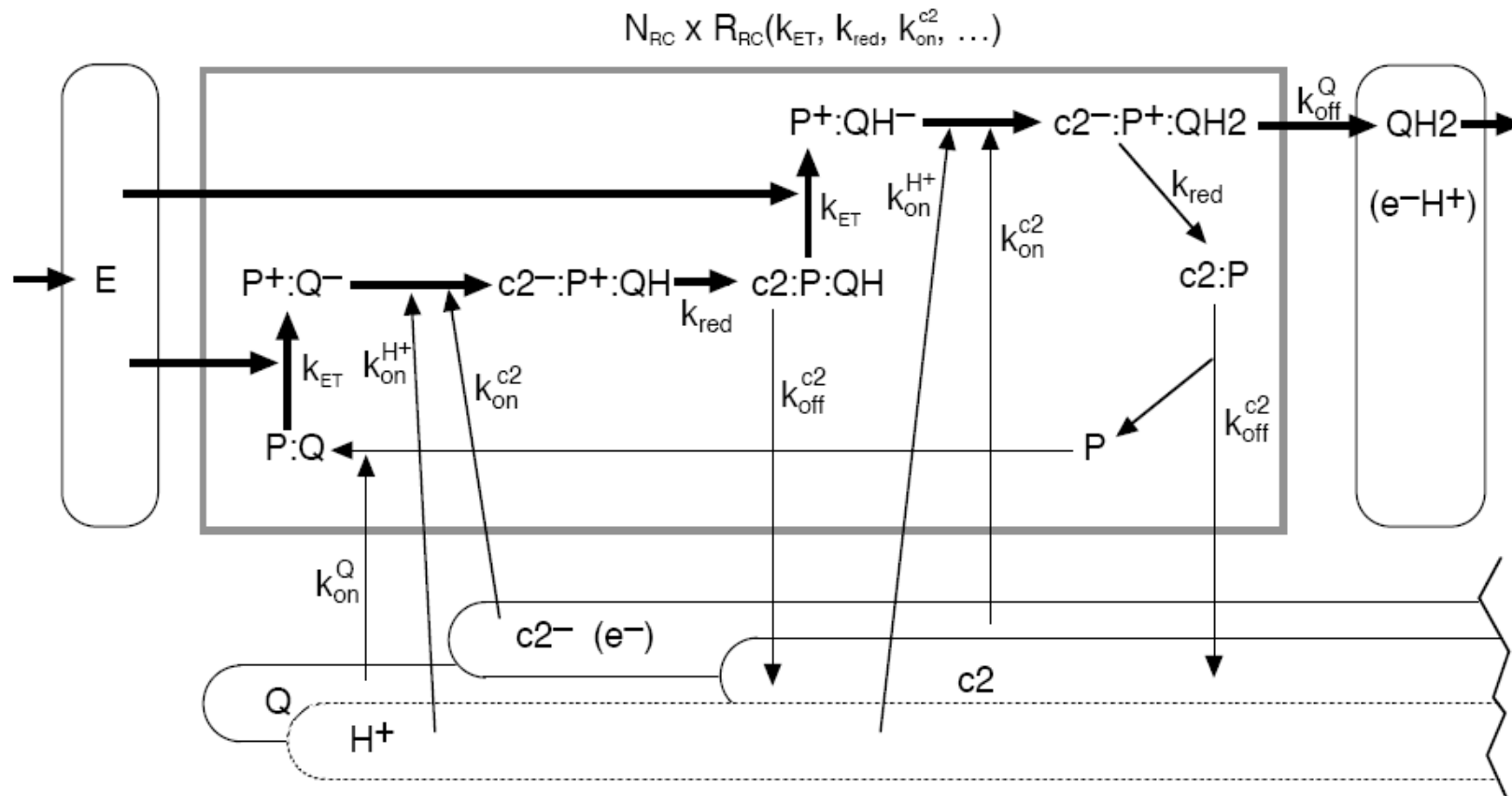
40–50  $\gamma/s$  per RC  
 $\approx 22 QH_2/s$

1 RC can serve 1 LH1  
 + 3 LH2  
 = 44  $\gamma/s$

$LH1_2 + 6 LH2 \triangleq 456 \text{ nm}^2 \rightarrow$  11 LH1 dimers including 22 RCs  
 on one vesicle

480 Q/s can be loaded @ 18 W/m<sup>2</sup> per vesicle

## Modelling of internal processes at reaction center



All individual reactions with their individual rates  $k$  together determine the overall conversion rate  $R_{RC}$  of a single RC.

Thick arrows : flow of the energy from the excitons through the cyclic charge state changes of the special pair Bchl (P) of the RC.

## Rounded rectangles : reservoirs

## **$bc_1$ Placement — Diffusional limits?**

Roundtrip times

maximal capacity of the carriers:

$$T = T_{RC} + T_{bc1} + T_{Diff}$$

### **Cytochrome $c_2$ :**

$$T_{RC} \approx 1 \text{ ms}$$

$$T_{bc1} \approx 12 \text{ ms}$$

$$T_{Diff} \approx 3 \mu\text{s}$$

$$T_{\text{round-trip}} = 13 \text{ ms} \leq 3 \text{ cyt } c \text{ per vesicle}$$

sufficient to carry  $e^-$ 's  
available: 22 cyt  $c$  per vesicle

### **Quinol:**

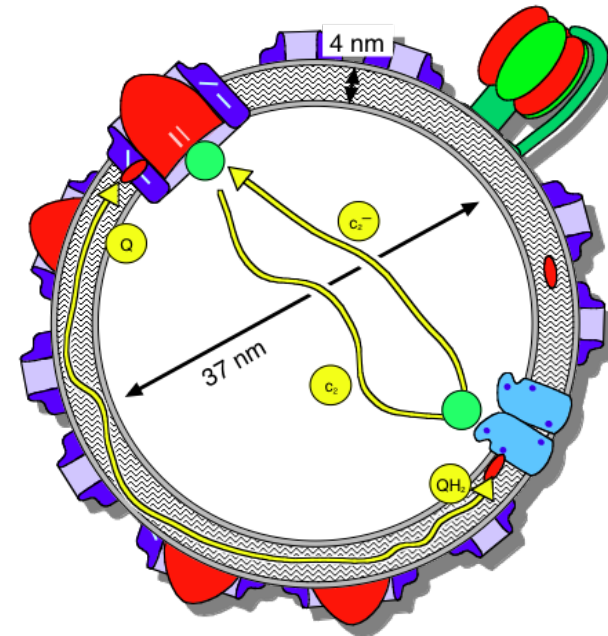
$$T_{RC} \approx 50 \text{ ms}$$

$$T_{bc1} \approx 23 \text{ ms}$$

$$T_{Diff} \approx 1 \text{ ms}$$

$$T_{\text{round-trip}} = 75 \text{ ms} \leq 7 \text{ Q per vesicle}$$

sufficient to carry  $e^-$ 's.  
available: 100 Q per vesicle

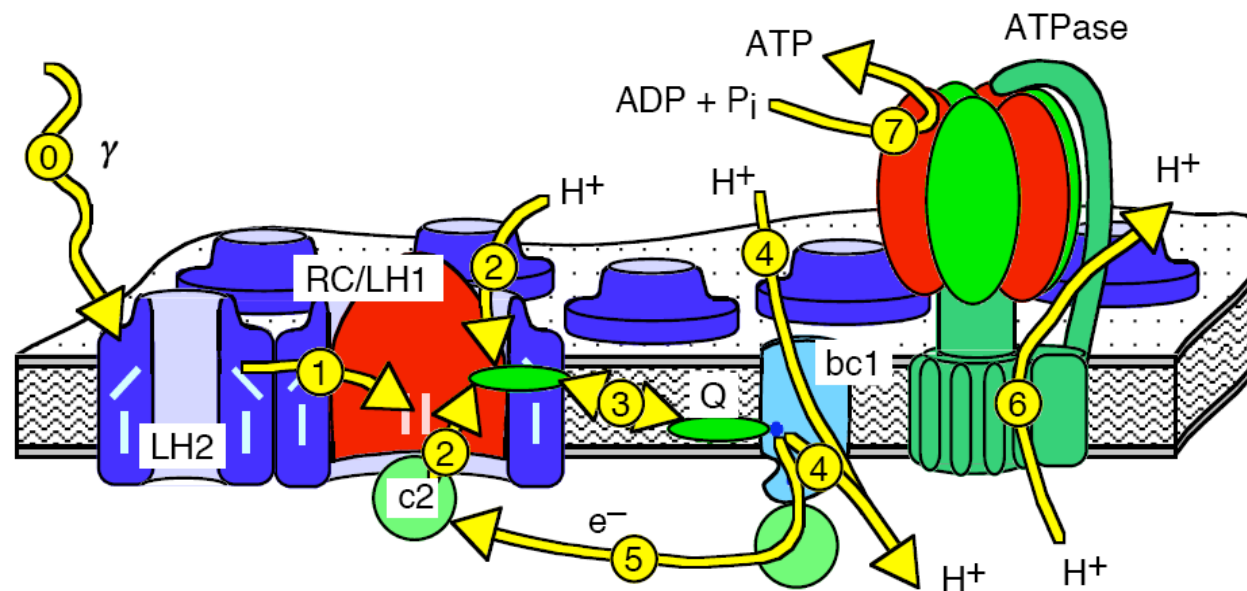


Diffusion is not limiting

→ poses no constraints  
on the position of  $bc_1$

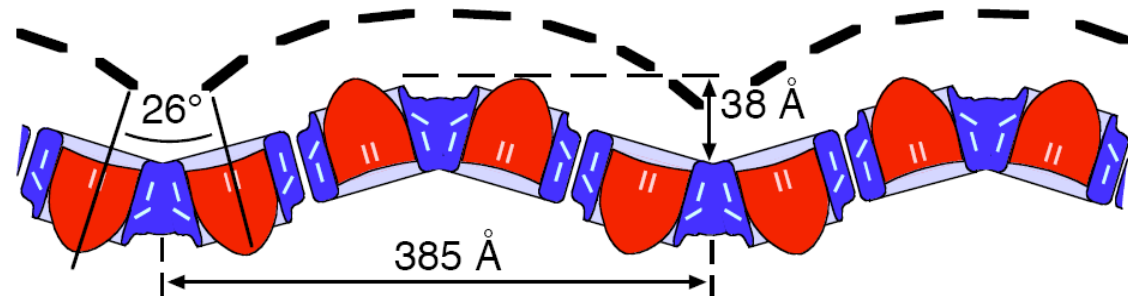
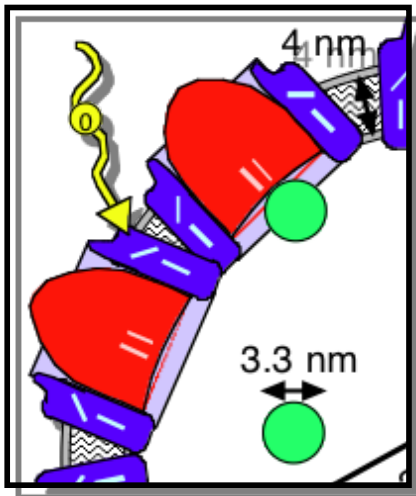
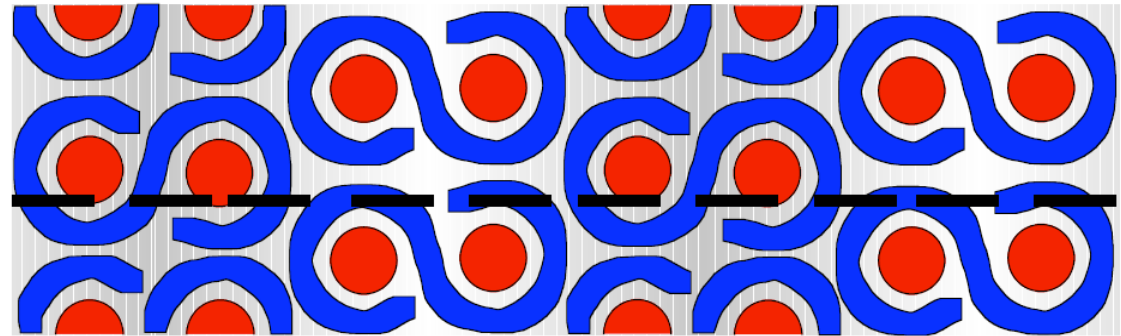
# Parameters

protein	throughput per protein (natural units)	H <sup>+</sup> equivalents per protein [1/s]	total number per avg. vesicle of 45 nm diameter	rate determined from	explained in section
LH2	10 $\gamma/s$	20	60	absorption spectra +	III A
LH1 dimer	$2 \times 14 \gamma/s$	56	10	+ light intensity of 18 W/m <sup>2</sup>	III A
RC	22 QH2/s	88	20	QH2 (un)binding	III B
bc1 dimer	$\leq 2 \times 42 c2/s$	168	3 (... 10)	measured activity at $\Delta pH = 0$	III C
ATPase	$\leq 100$ ATP/s	400	1	measured throughput	III D
cytochrome <i>c</i> <sub>2</sub>	80 e <sup>-</sup> /s	160	20	(un)binding at the bc1	V A (III B, III C)
ubiquinone	$10 \times 2(e^- H^+)/s$	40	100	(un)binding at the RC and the bc1	V A (III B, III C)



## reconstituted LH1 dimers in planar lipid membranes explain intrinsic curvature of vesicles

Drawn after AFM images of Scheuring *et al* of LH1 dimers reconstituted into planar lipid membranes.



Values fit nicely to the proposed arrangement of LH1 dimers, when one assumes that they are stiff enough to retain the bending angle of 26° that they would have on a spherical vesicle of 45 nm diameter and taking into account the length of a single LH1 dimer of about 19.5 nm.

# Proposed setup of a chromatophore vesicle

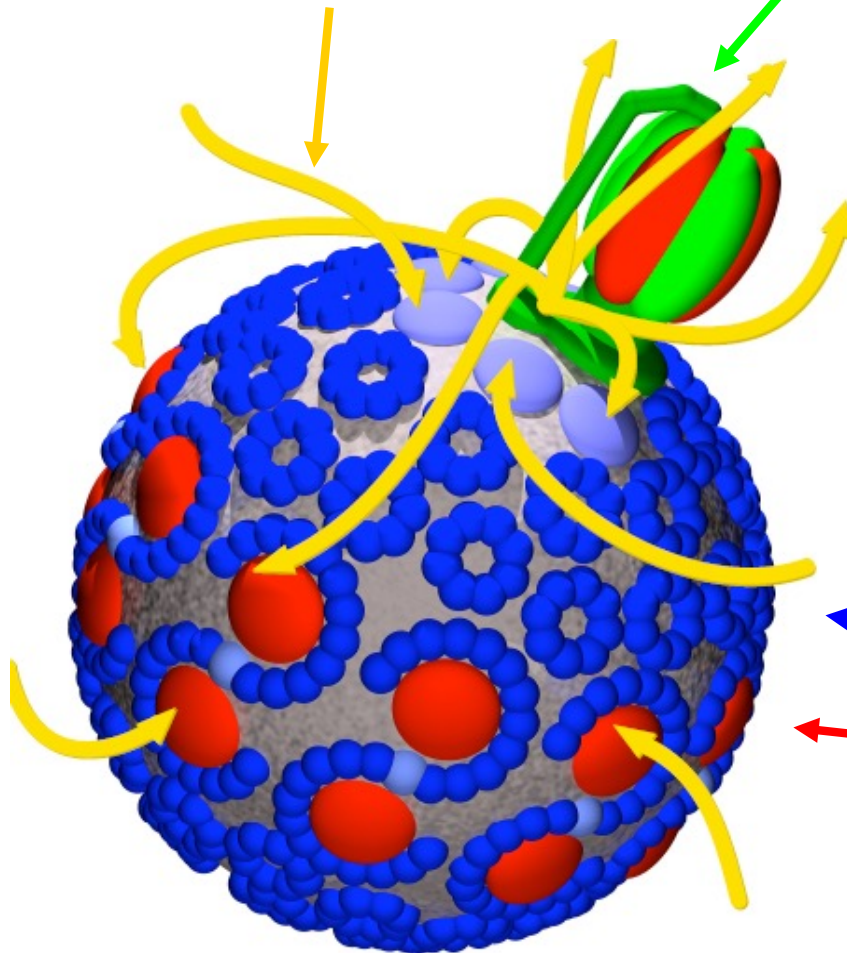
yellow arrows: diffusion of the protons out of the vesicle via the ATPase and to the RCs and bc1s.

At the „poles“

green/red: the ATPase

light blue: the bc1 complexes

Increased proton density close to the ATPase suggests close proximity of ATPase and  $bc_1$  complexes.



blue: small LH2 rings (blue)

blue/red: Z-shaped LH1/RC dimers form a linear array around the “equator” of the vesicle, determining the vesicle’s diameter by their intrinsic curvature.



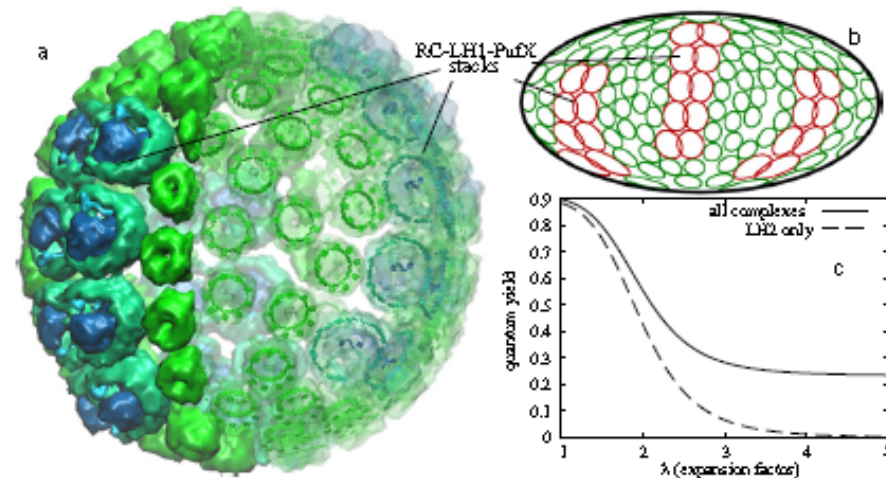
# Summary

Integrated model of binding + photophysical + redox processes  
inside of chromatophore vesicles

Various experimental data  
fit well together

Equilibrium state.

How to model  
non-equilibrium processes?



Biophysical Journal Volume 99 July 2010 67–75

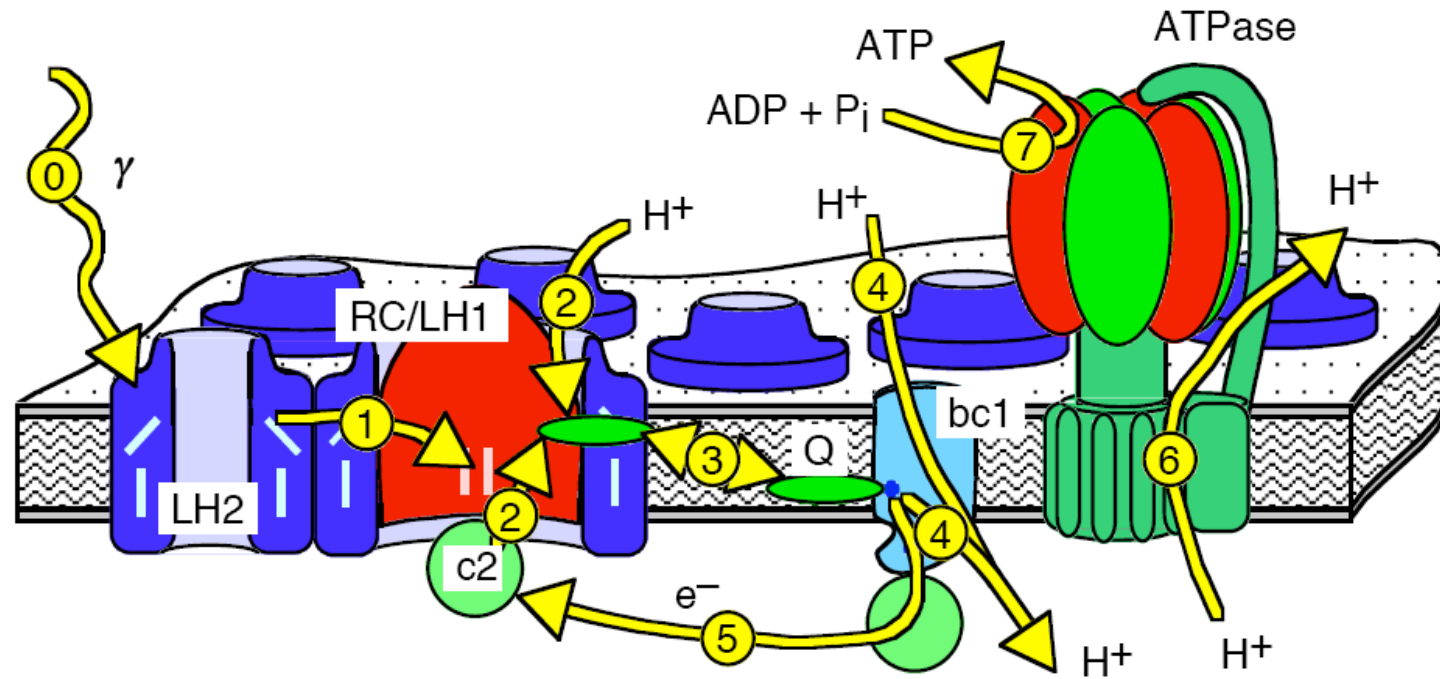
67

## Photosynthetic Vesicle Architecture and Constraints on Efficient Energy Harvesting

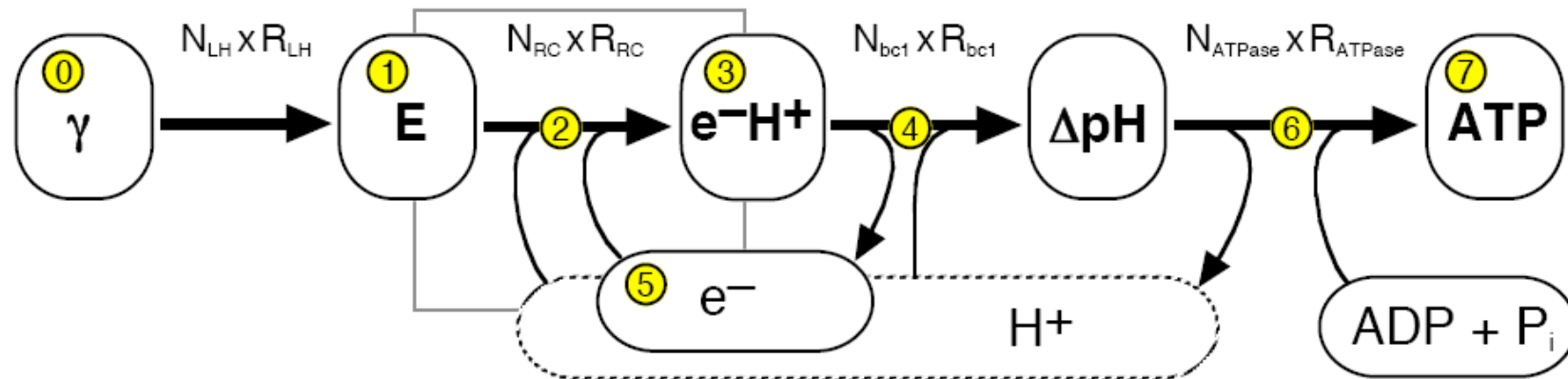
Melih Sener,<sup>†‡</sup> Johan Strümpfer,<sup>†§</sup> John A. Timney,<sup>¶</sup> Arvi Freiberg,<sup>||</sup> C. Neil Hunter,<sup>¶</sup> and Klaus Schulten<sup>†‡§\*</sup>

<sup>†</sup>Beckman Institute for Advanced Science and Technology, <sup>‡</sup>Department of Physics, and <sup>§</sup>Center for Biophysics and Computational Biology, University of Illinois at Urbana-Champaign, Urbana, Illinois; <sup>¶</sup>Department of Molecular Biology and Biotechnology, University of Sheffield, Sheffield, United Kingdom; and <sup>||</sup>Institute of Physics and <sup>\*\*</sup>Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia

# Photosynthesis: textbook view



# Viewing the photosynthetic apparatus as a conversion chain



Thick arrows : path through which the photon energy is converted into chemical energy stored in ATP via the intermediate stages (rounded rectangles).

Each conversion step takes place in parallelly working proteins. Their number  $N$  times the conversion rate of a single protein  $R$  determines the total throughput of this step.

$\gamma$  : incoming photons collected in the LHCs

**E** : excitons in the LHCs and in the RC

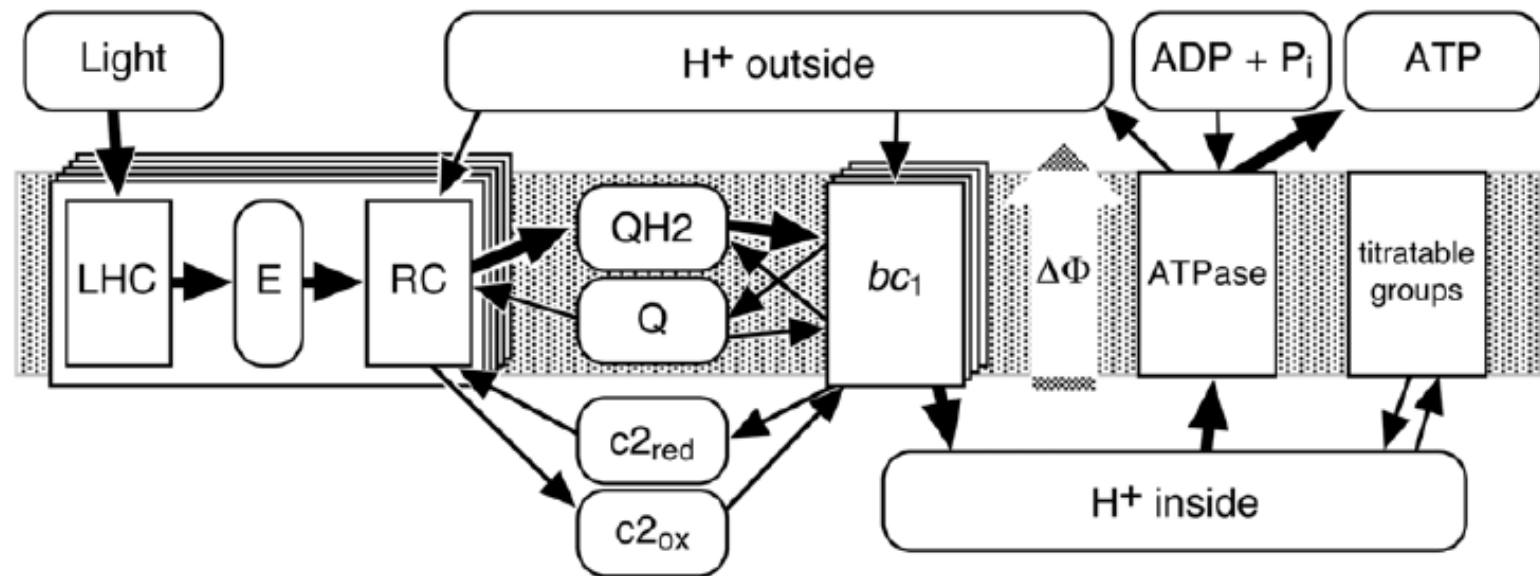
$e^-H^+$  electron–proton pairs stored on the quinols

$e^-$  for the electrons on the cytochrome  $c_2$

$pH$  : transmembrane proton gradient

$H^+$  : protons outside of the vesicle (broken outline of the respective reservoir).

# Stochastic dynamics simulations: Molecules & Pools model



Round edges: **pools** for metabolite molecules

Rectangles: protein machines are modeled explicitly as multiple copies

fixed set of parameters

integrate rate equations with stochastic algorithm

# Stochastic simulations of cellular signalling

**Traditional computational approach** to chemical/biochemical kinetics:

- (a) start with a set of coupled **ODEs** (reaction rate equations) that describe the time-dependent concentration of chemical species,
- (b) use some **integrator** to calculate the concentrations as a function of time given the rate constants and a set of initial concentrations.

Successful **applications** : studies of yeast cell cycle, metabolic engineering, whole-cell scale models of metabolic pathways (E-cell), ...

**Major problem**: cellular processes occur in very small volumes and frequently involve **very small number of molecules**.

E.g. in gene expression processes a few TF molecules may interact with a single gene regulatory region.

*E.coli* cells contain on average only 10 molecules of Lac repressor.

## Include stochastic effects

(Consequence1) → modeling of reactions as continuous fluxes of matter is no longer correct.

(Consequence2) Significant **stochastic fluctuations** occur.

To study the stochastic effects in biochemical reactions, stochastic formulations of chemical kinetics and Monte Carlo computer simulations have been used.

**Daniel Gillespie** (J Comput Phys 22, 403 (1976); J Chem Phys 81, 2340 (1977)) introduced the exact **Dynamic Monte Carlo (DMC)** method that connects the traditional chemical kinetics and stochastic approaches.



# Basic outline of the direct method of Gillespie

**(Step i)** generate a list of the components/species and define the initial distribution at time  $t = 0$ .

**(Step ii)** generate a list of possible events  $E_i$  (chemical reactions as well as physical processes).

**(Step iii)** using the current component/species distribution, prepare a probability table  $P(E_i)$  of all the events that can take place.

Compute the total probability

$$P_{tot} = \sum P(E_i)$$

$P(E_i)$  : probability of event  $E_i$  .

**(Step iv)** Pick two random numbers  $r_1$  and  $r_2 \in [0...1]$  to decide which event  $E_\mu$  will occur next and the amount of time  $\tau$  after which  $E_\mu$  will occur.

Resat et al., J.Phys.Chem. B 105, 11026 (2001)

## Basic outline of the direct method of Gillespie

Using the random number  $r_1$  and the probability table, the event  $E_\mu$  is determined by finding the event that satisfies the relation

$$\sum_{i=1}^{\mu-1} P(E_i) < r_1 P_{tot} \leq \sum_{i=1}^{\mu} P(E_i)$$

The second random number  $r_2$  is used to obtain the amount of time  $\tau$  between the reactions

$$\tau = -\frac{1}{P_{tot}} \ln(r_2)$$

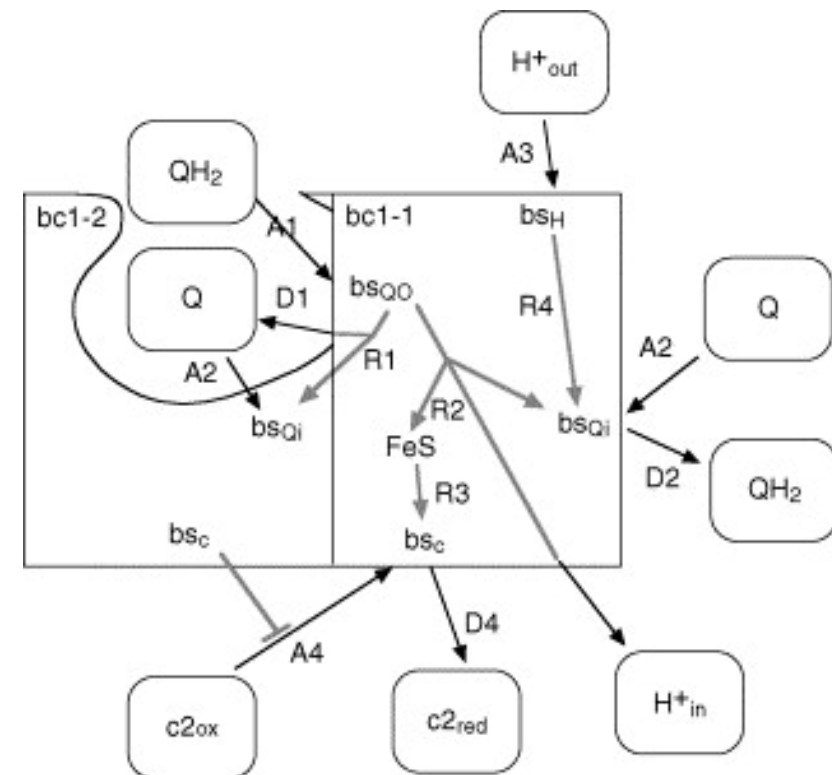
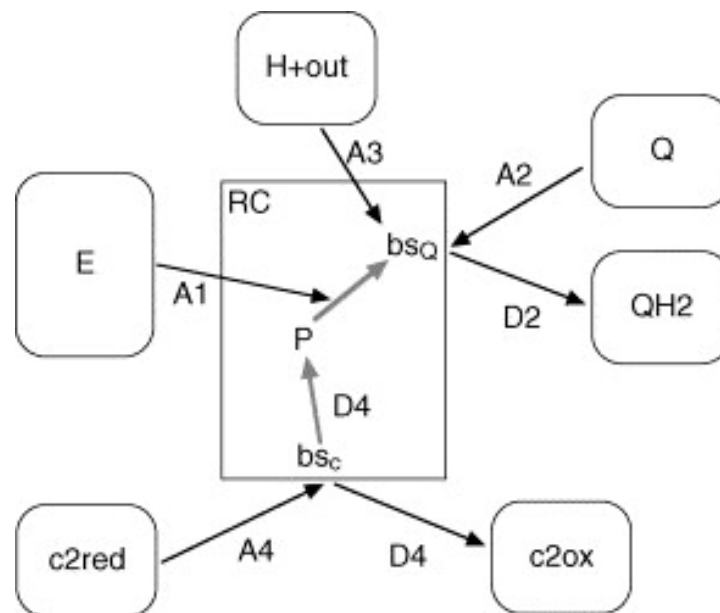
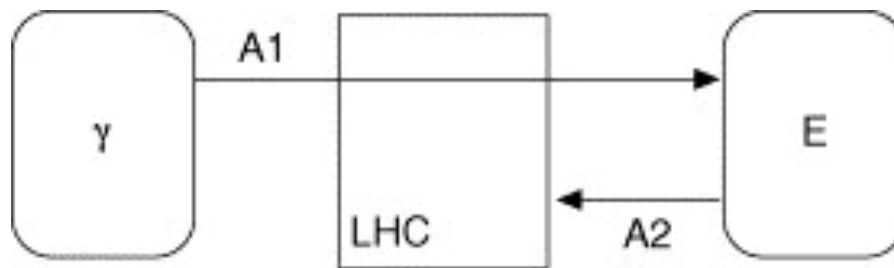
As the total probability of the events changes in time, the time step between occurring steps varies.

Steps (iii) and (iv) are repeated at each step of the simulation.

The necessary number of runs depends on the inherent noise of the system and on the desired statistical accuracy.

Resat et al., J.Phys.Chem. B 105, 11026 (2001)

# reactions included in stochastic model of chromatophore



# Stochastic simulations of a complete vesicle

Model vesicle:    12 LH1/RC-monomers  
                      1-6  $bc_1$  complexes  
                      1 ATPase

                      120 quinones  
                      20 cytochrome  $c_2$

integrate rate equations with:

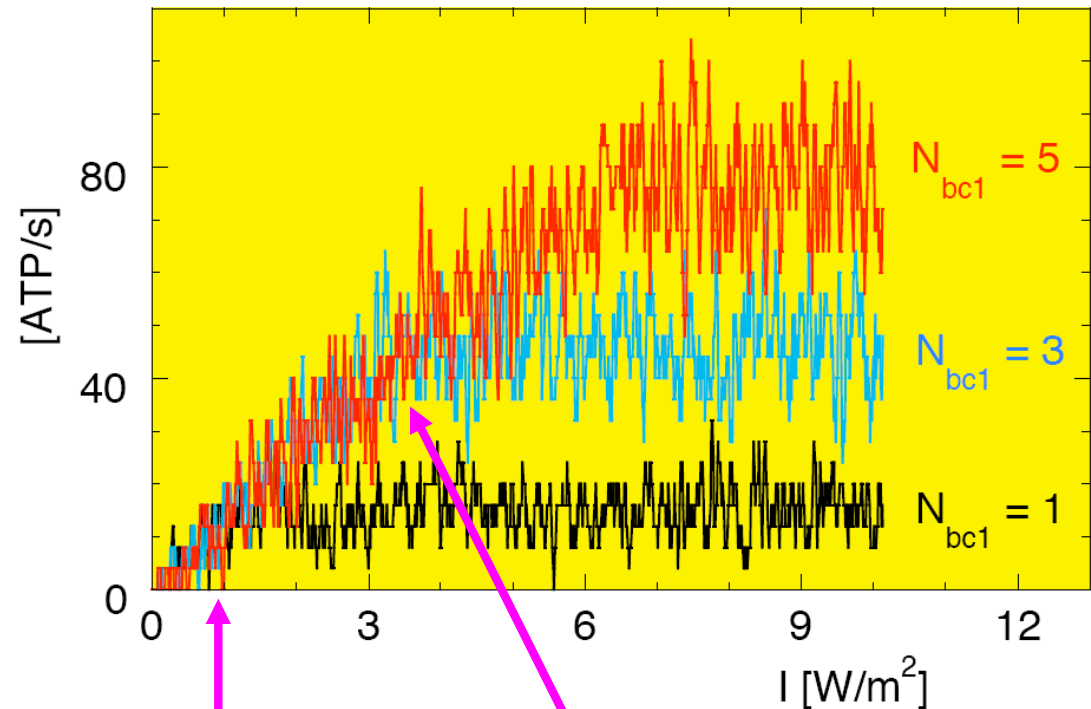
- Gillespie algorithm (associations)
- Timer algorithm (reactions); 1 random number determines when reaction occurs

simulating 1 minute real time requires 1.5 minute on one opteron 2.4 GHz proc

## simulate increase of light intensity (sunrise)

during 1 minute,  
light intensity is slowly  
increased from 0 to 10 W/m<sup>2</sup>  
(quasi steady state)

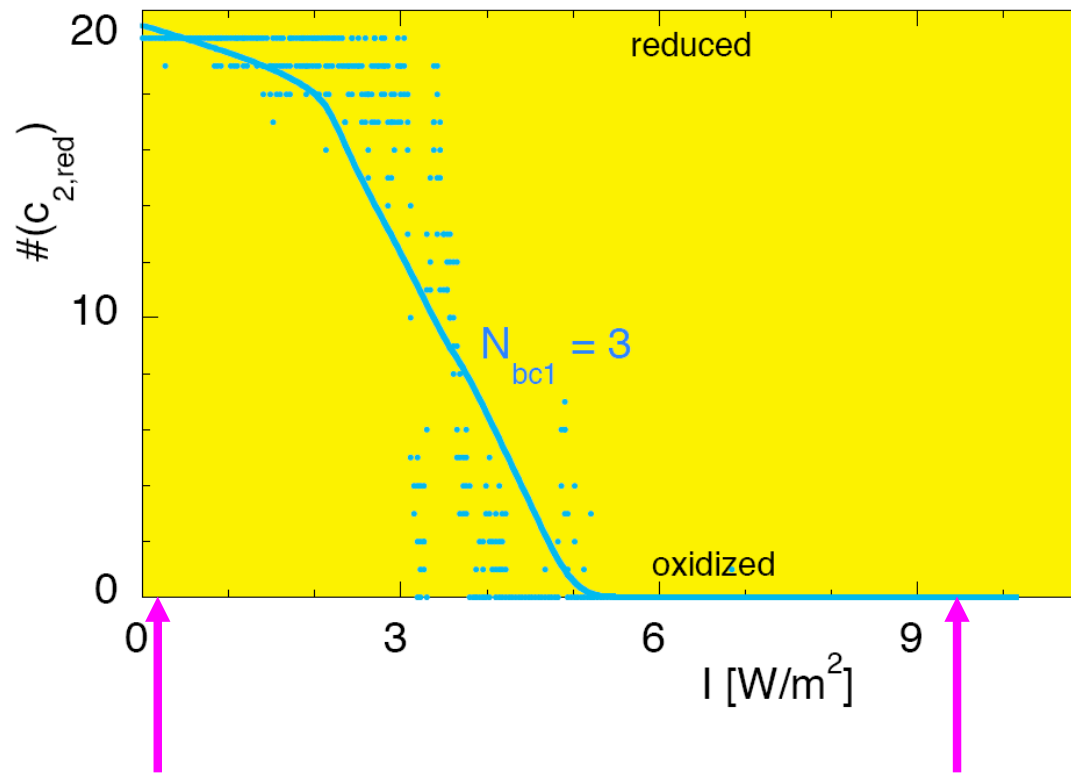
- there are two regimes
- one limited by available light
  - one limited by bc<sub>1</sub> throughput



**low light intensity:**  
linear increase of  
ATP production  
with light intensity

**high light intensity:**  
**saturation** is reached  
the later the higher the  
number of bc<sub>1</sub> complexes

## oxidation state of cytochrome $c_2$ pool

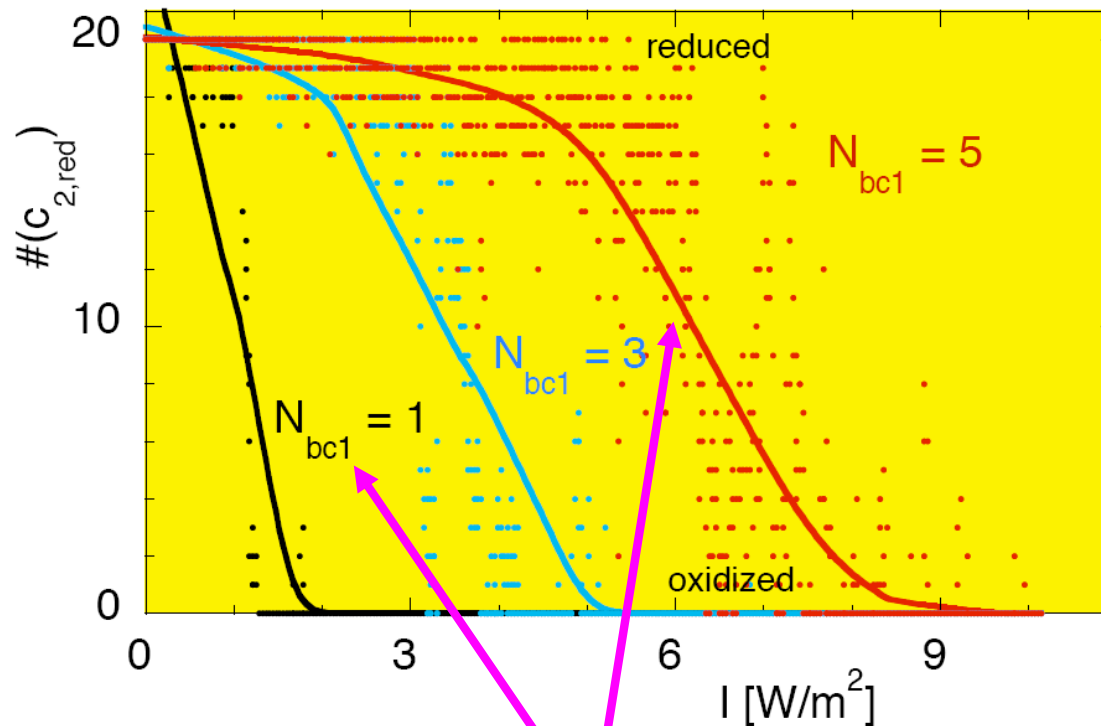


**low light intensity:**  
all 20 cytochrome  $c_2$   
are reduced by  $bc_1$

**high light intensity**  
 $RCs$  are faster than  $bc_1$ ,  
 $c_2$ s wait for electrons

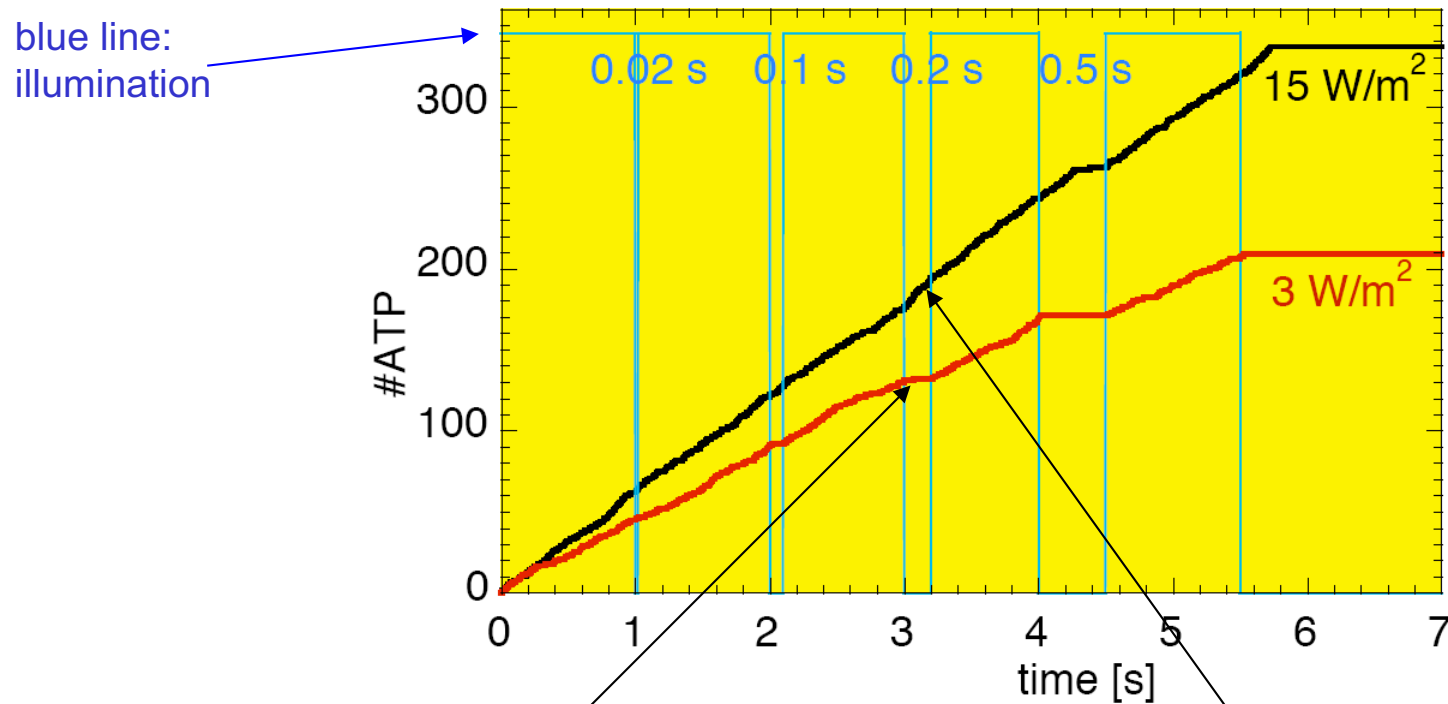


## oxidation state of cytochrome $c_2$ pool



more  $bc_1$  complexes  
can load more  
cytochrome  $c_2$ s

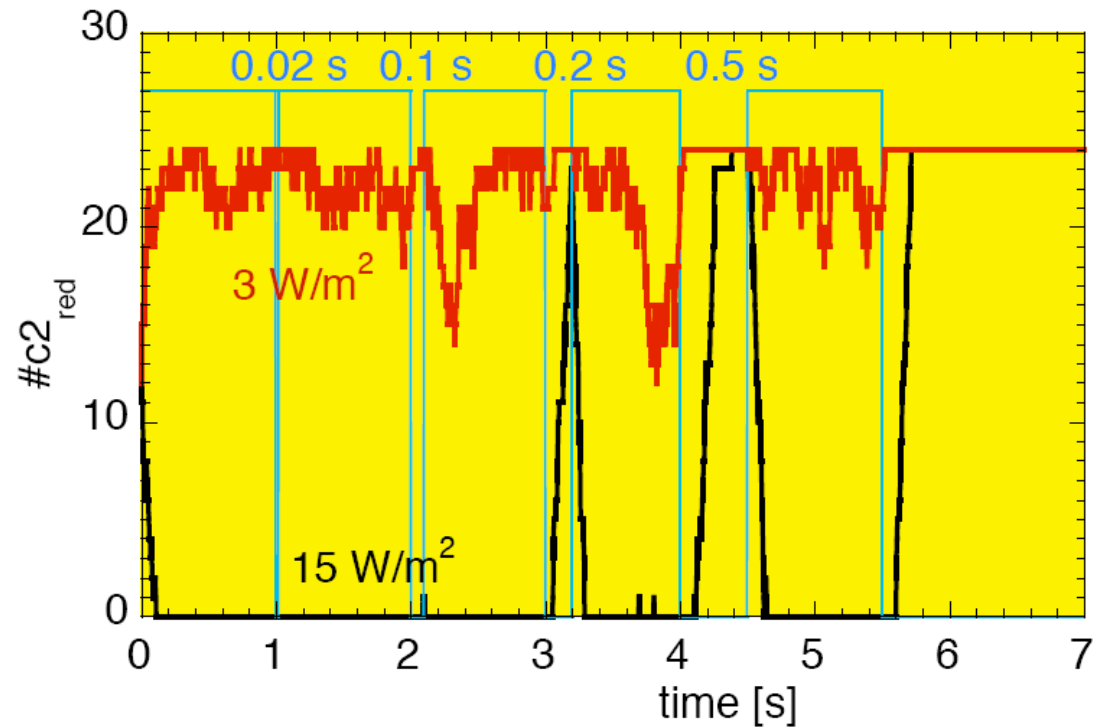
## total number of produced ATP



low light intensity: any interruption stops ATP production

high light intensity: interruptions are buffered up to 0.3 s duration

## **c<sub>2</sub> pool acts as buffer**



At high light intensity, c<sub>2</sub> pool is mainly oxidized.

If light is turned off, bc<sub>1</sub> can continue to work (load c<sub>2</sub>s, pump protons, let ATPase produce ATP) until c<sub>2</sub> pool is fully reduced.

# What if parameters are/were unknown ?

## Bridging the Gap: Linking Molecular Simulations and Systemic Descriptions of Cellular Compartments

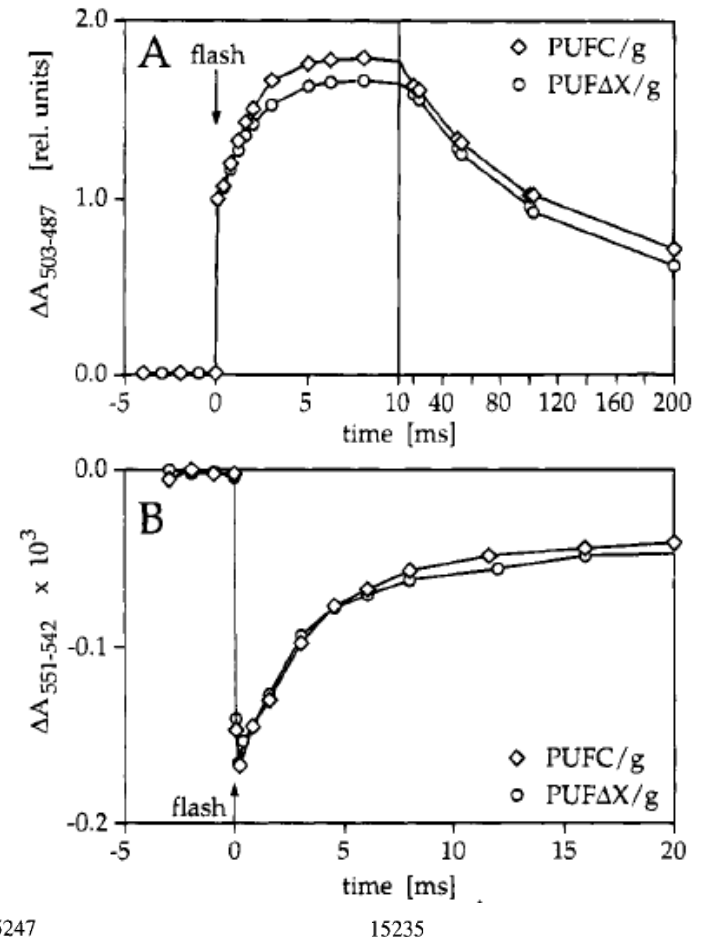
Tihamér Geyer\*, Xavier Mol, Sarah Bläß, Volkhard Helms

Center for Bioinformatics, Saarland University, Saarbrücken, Germany

PLoS ONE (2010)

choose 25 out of 45 system parameters  
for optimization.

take 7 different non-equilibrium time-resolved  
experiments from Dieter Oesterhelt lab  
(MPI Martinsried).



Biochemistry 1995, 34, 15235–15247

15235

Role of PufX Protein in Photosynthetic Growth of *Rhodobacter sphaeroides*.

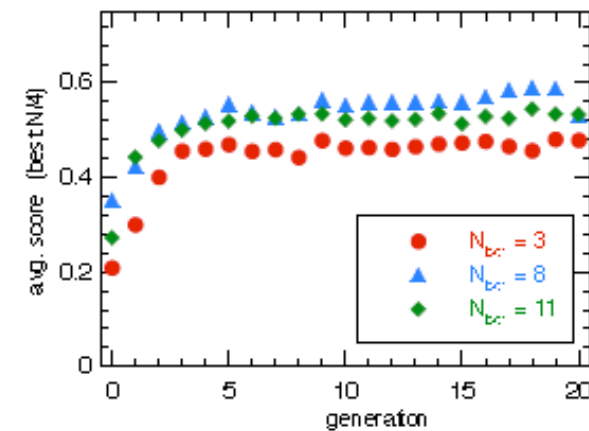
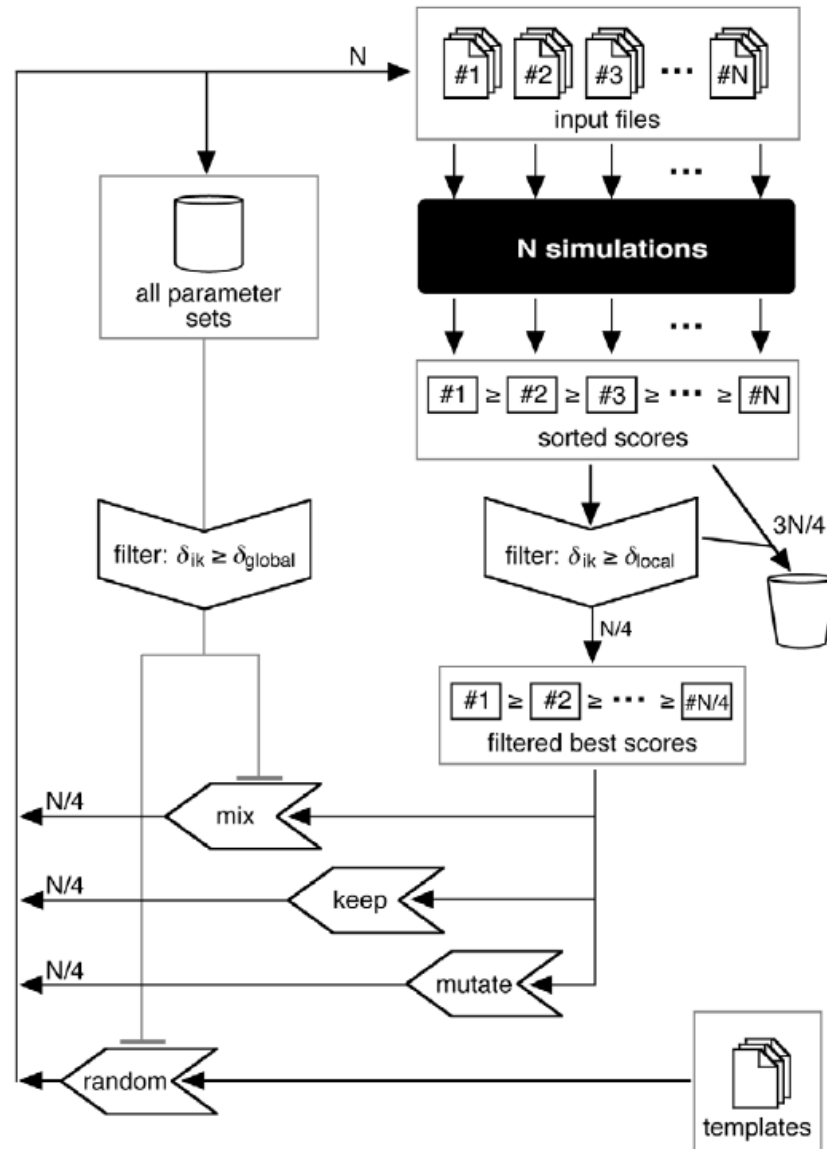
1. PufX Is Required for Efficient Light-Driven Electron Transfer and Photophosphorylation under Anaerobic Conditions<sup>†</sup>

## Parameters not optimized

Parameter	Value	Description
$bc_1::k_{on}(H^+_{out})$	$10^{10} \text{ nm}^3 \text{ s}^{-1}$	rate for proton uptake from the cytoplasm by $bc_1$
$bc_1::k_{tr}(e:Q_0 \Rightarrow FeS)$	$2.3 * 10^3 \text{ s}^{-1}$	rate for electron transfer from $Q_0$ to FeS
$bc_1::k_{tr}(e:c_1 \Rightarrow c_2)$	$10^5 \text{ s}^{-1}$	electron transfer rate from $c_1$ to bound cytochrome $c_2$
$bc_1::k_{tr}(e:Q_0 \Rightarrow b_L)$	$10^4 \text{ s}^{-1}$	electron transfer from $Q_0$ to $b_L$ heme
$bc_1::k_{tr}(e:b_L \Rightarrow b_H)$	$10^4 \text{ s}^{-1}$	electron transfer from $b_L$ to $b_H$ heme
$\Delta\Phi::V$	$2.65 * 10^4 \text{ nm}^3$	inner volume of the vesicle
$\Delta\Phi::A$	$5.28 * 10^3 \text{ nm}^2$	membrane area (Q pool „volume“)
$\Delta\Phi::C_{Hn}$	1.0 e	effective charge of a free proton in the vesicle
$\Delta\Phi::C_{Hm}$	1.0 e	effective charge of a proton on the titratable groups
$\Delta\Phi::C_{prot}$	-1.0 e	effective charge of an $e^-$ translocated through an RC
$\Delta\Phi::C_{red}$	-0.5 e	effective charge of a reduced cytochrome $c_2$
$\Delta\Phi::C_{cox}$	0.5 e	effective charge of an oxidized cytochrome $c_2$
PR::N <sub>p</sub>	80	number of titratable groups in the vesicle
PR::pK	5.0	pK of the titratable groups
$N_{core}$	10	number of dimeric core complexes (2 RC + 1 LHC)
$N_{bc1}$	10	number of cytochrome $bc_1$ complexes
$N_{ATPase}$	1	number of ATPases
$N_{c2}$	20	total number of cytochrome $c_2$
$N_Q$	200	total number of quinones

**Table S1: Model Parameters Not Included in the Optimization Process**

# Parameter optimization through evolutionary algorithm



## 25 optimization parameters

Analyze 1000 best  
parameter sets among  
32.800 simulations:

$$\langle P \rangle = \exp[\langle \log P \rangle]$$

$$\sigma^2 = \langle (\log P - \langle \log P \rangle)^2 \rangle$$

$$P_{\min} = \exp[\langle \log P \rangle - \sigma]$$

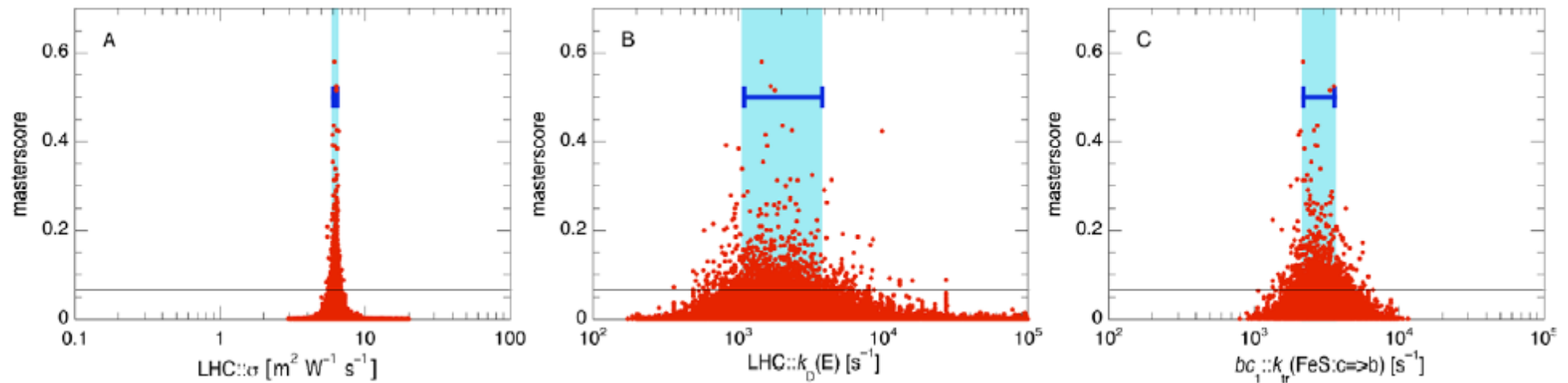
$$P_{\max} = \exp[\langle \log P \rangle + \sigma]$$

parameter	units	$\langle P \rangle$	$P_{\min} \dots P_{\max}$	$P_{\min}/P_{\max}$
LHC:: $\sigma$	$\text{m}^2 \text{W}^{-1} \text{s}^{-1}$	6.22	6.02...6.42	0.94
LHC:: $N_0$	1	1.31	0.81 ... 2.13	0.38
LHC:: $k_D(E)$	$\text{s}^{-1}$	$1.9 \times 10^3$	$(1.1 \dots 3.8) \times 10^3$	0.29
RC:: $k_{on}(E)$	$\text{s}^{-1}$	$2.4 \times 10^6$	$(1.2 \dots 4.5) \times 10^6$	0.27
RC:: $k_{on}(H^+)$	$\text{nm}^3 \text{s}^{-1}$	$1.4 \times 10^8$	$(1.3 \dots 1.6) \times 10^8$	0.81
RC:: $k_{on}(Q)$	$\text{nm}^2 \text{s}^{-1}$	$6.0 \times 10^4$	$(4.4 \dots 8.1) \times 10^4$	0.54
RC:: $k_{off}(QH2)$	$\text{s}^{-1}$	87	70...108	0.65
RC:: $k_{on}(c2red)$	$\text{nm}^3 \text{s}^{-1}$	$9.2 \times 10^5$	$(7.3 \dots 11.5) \times 10^5$	0.63
RC:: $k_{off}(c2ox)$	$\text{s}^{-1}$	$2.2 \times 10^3$	$(1.6 \dots 3.0) \times 10^3$	0.53
bc1:: $k_{on}(QH2@Qo)$	$\text{nm}^2 \text{s}^{-1}$	$1.2 \times 10^4$	$(0.79 \dots 1.7) \times 10^4$	0.46
bc1:: $k_{off}(Q@Qo)$	$\text{s}^{-1}$	28.3	26.3...30.4	0.86
bc1:: $k_{tr}(Q:Qo \rightarrow Q_i)$	$\text{s}^{-1}$	$4.9 \times 10^3$	$(3.6 \dots 6.7) \times 10^3$	0.54
bc1:: $k_{on}(Q@Qi)$	$\text{mm}^2 \text{s}^{-1}$	$6.7 \times 10^5$	$(4.5 \dots 10) \times 10^5$	0.45
bc1:: $k_{off}(QH2@Qi)$	$\text{s}^{-1}$	86	68...110	0.62
bc1:: $k_{tr}(QH2:Q_i \rightarrow Q_o)$	$\text{s}^{-1}$	$3.8 \times 10^3$	$(2.6 \dots 5.5) \times 10^3$	0.47
bc1:: $k_{on}(c2ox)$	$\text{nm}^3 \text{s}^{-1}$	$9.4 \times 10^6$	$(6.3 \dots 14) \times 10^6$	0.47
bc1:: $k_{off}(c2red)$	$\text{s}^{-1}$	$6.0 \times 10^3$	$(3.3 \dots 11) \times 10^3$	0.30
bc1:: $k_{off}(H+@Qo)$	$\text{s}^{-1}$	$2.4 \times 10^4$	$(1.3 \dots 4.3) \times 10^4$	0.30
bc1:: $k_{tr}(FeS:b \rightarrow c)$	$\text{s}^{-1}$	$3.9 \times 10^3$	$(3.1 \dots 5.1) \times 10^3$	0.61
bc1:: $k_{tr}(FeS:c \rightarrow b)$	$\text{s}^{-1}$	$2.8 \times 10^3$	$(2.2 \dots 3.6) \times 10^3$	0.61
bc1:: $k_{tr}(e:b_H \rightarrow Q_i)$	$\text{s}^{-1}$	$7.7 \times 10^3$	$(5.0 \dots 12) \times 10^3$	0.42
bc1:: $\Phi_0$	mV	102	83...114	0.73
$\Delta\Phi::U_0$	mV/e	10.3	9.5...11	0.85
$\Delta\Phi::\Delta\Phi_0$	mV/pH	10	7.6...13.7	0.55
PR::pK	1	4.84	3.9...5.9	0.66



# Sensitivity of master score

Decay rate of excitons  
in LHC



Absorption cross section  
light harvesting complex

Kinetic rate for hinge  
motion of FeS domain in  
bc1 complex

Some parameters are very sensitive, others not.

# Three best-scored parameter sets

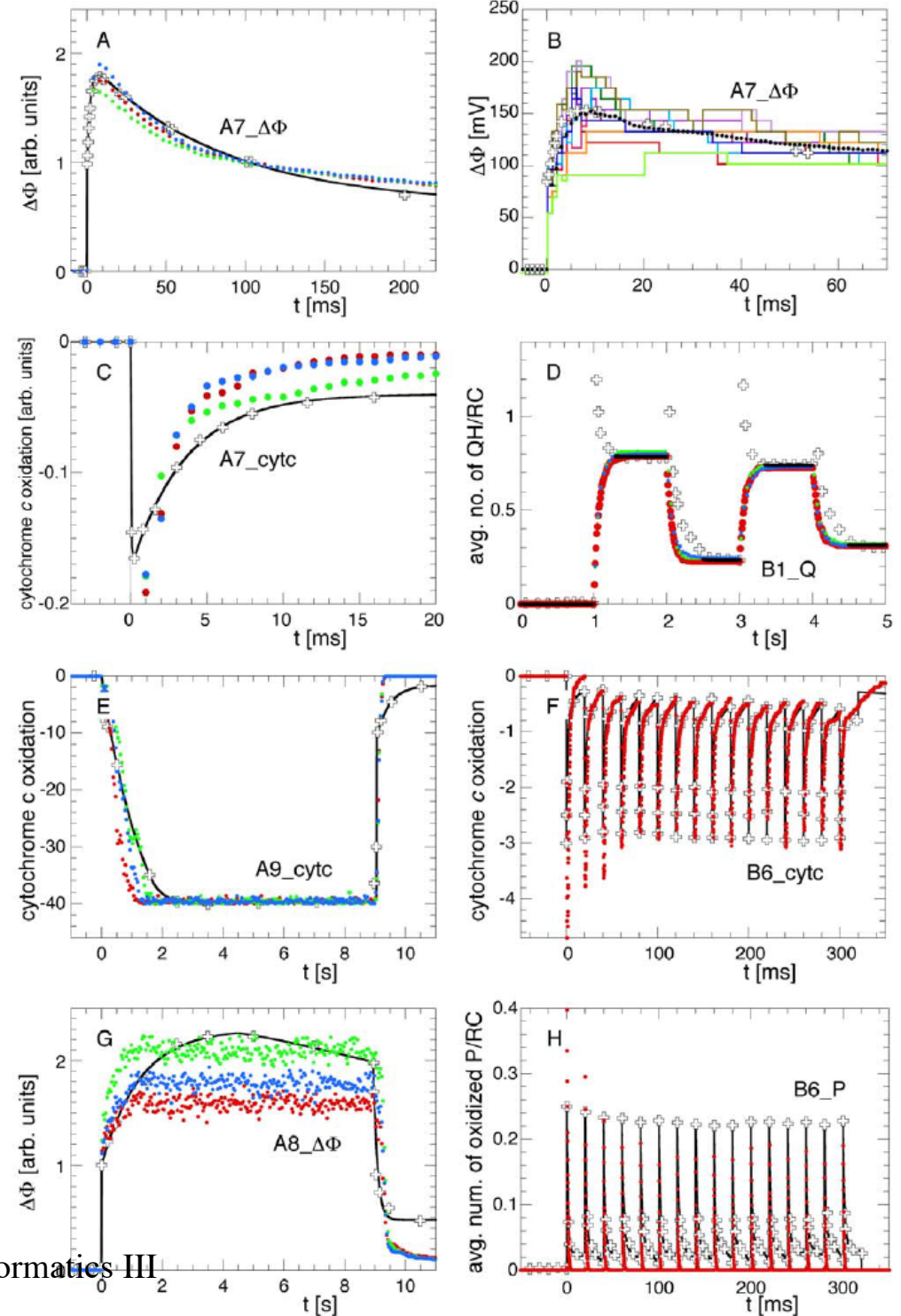
Score of individual parameter set  $i$   
for matching one experiment:

$$s_i = \frac{C_i}{\sum (x(t_i) - f(t_i))^2}$$

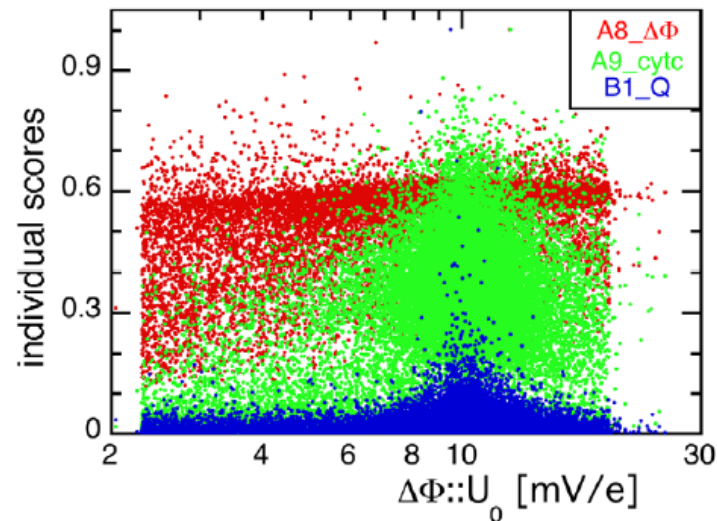
$x(t_i)$ : simulation result

$f(t_i)$ : smooth fit of exp. data

**Master score** for one  
parameter set: defined as  
product of the individual  
scores  $s_i$



# Different experiments yield different sensitivity



“importance score”:  
Sum of the sensitivities  
 $P_{min}/P_{max}$  of all relevant  
parameters

**Table 2.** Importance scores and correlation coefficients between the master score and the respective individual scores of the experimental scenarios denoting the relative importance of each of the experiments for the parameter value optimization.

experiment	A7 cytc	A7 ΔΦ	A8 ΔΦ	A9 cytc	B1 Q	B6 P	B6 cytc	BC1
importance score	4.4	7.7	5.8	9.7	3.8	5.2	8.9	5.5
correlation	0.09	0.44	0.22	0.38	0.83	0.17	0.31	0.41

The importance scores are determined as the sums of the sensitivities of all relevant parameters against the individual scores (see table S2 for all the individual values). The correlation coefficients are obtained from a linear fit of the master score against the respective individual score.

Analysis could suggest new  
experiments that would be  
most informative!

## Summary

Only 1/3 of the kinetic parameters previously known.

Stochastic parameter optimization converges robustly into the same parameter basin as known from experiment.

Two large-scale runs (15 + 17 parameters) yielded practically the same results.

If implemented as grid search, less than 2 points per dimension.

It appears enough to know 1/3 – 1/2 of kinetic rates about a system to be able to describe it quantitatively (IF connectivities are known).

## Review – algorithms / methods etc in Bioinfo III

*“There is no such thing as a free lunch”.*

Alvin Hansen, economist (1953)

There exist several “No Free Lunch Theorems” for optimization problems.

E.g. Wolpert & Macready (1997) showed:

For any search/optimization algorithm, any elevated performance over one class of problems is exactly paid for in performance over another class.

## Review – simulation / analysis methods

	Used where	Pro	Con
Enrichment methods	Annotate gene function, histone peaks, motifs	Proper statistical analysis	Not causal, mechanistic reasons remain unclear
Graph algorithms	Modules in PPI networks, PP complexes, MCDS algo for key genes in GRNs, Cut-sets in metabolic networks	<ul style="list-style-type: none"><li>- graph layout provides intuitive view of network topology,</li><li>- ILPs give optimal solutions,</li><li>- heuristic algorithms can be fast</li></ul>	<ul style="list-style-type: none"><li>- ILPs very time-consuming,</li><li>- heuristic solutions may be not accurate,</li><li>- graph algorithms suffer from noisy data</li></ul>

## Review – simulation / analysis methods

	Used where	Pro	Con
Pearson correlation	Gene co-expression, DNA co-methylation	Quantitative measure	Suffers from outliers (V21); correlations are not causal
Rank-based correlation	Gene co-expression	Avoids outlier problems	Sensitive to small variations, large variations may be condensed into small rank differences
Bayesian network	Anywhere (here: PPIs)	Integrates arbitrary data; automatic weighting of likelihoods	Not causal (but this can be included)
Boolean network	GRNs	Finite state space, understand system completely, causal	Values restricted to boolean levels (but can be generalized)



## Review – simulation / analysis methods

	Used where	Pro	Con
FBA	Metabolic networks	Gives one optimal solution	None (?)
EFMs / EPs	Metabolic networks	Full insight into metabolic capabilities of system	Already medium-sized systems have 10.000s + EFMs
ODE	Metabolic systems, Signaling systems	Quantitative, time-dependent models, simple systems can be solved analytically, simple numerical implementation	Needs many parameters, not suitable for small particle numbers
Stochastic simulations	Metabolic systems	Capture stochastic effects of few particles,	Not deterministic (different solution each time); costly