# V22 - Stochastic Dynamics simulations of a photosynthetic vesicle

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)



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



**Biological Complexity** 



# **Bacterial Photosynthesis 101**





#### **Photosynthesis – cycle view**



2 cycles: electrons protons



#### LH1 / LH2 / RC — a la textbook

# ATPase **Collecting photons** ADP + LH2: 8 $\alpha\beta$ dimers downhill transport Periplasm of excitons $LH2 \rightarrow LH1 \rightarrow RC$ LH1: 16 $\alpha\beta$ dimers RC (b) TH-ILLIAN THE BC THE THEIL THEIL LH-I DH-

B800, B850, Car.



Hu *et al*, 1998 5

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



the "proton pump"





Berry, etal, 2004

X-ray structures known always forms a dimer



# The F<sub>o</sub>F<sub>1</sub>-ATP synthase I

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





# The F<sub>1</sub>F<sub>0</sub>-ATP synthase

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

➔ ATPase is "visible" 1 ATPase per vesicle



#### limited throughput of the ATPase



Gräber et al, 1991, 1999

# The electron carriers

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

- 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





### 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

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



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





# LH1 / LH2 / RC — native

# electron micrograph and density map



Siebert et al, 2004 125 \* 195 Ų, γ = 106°

	Area per:	per vesicle (45 nm)
LH1 monomer (hexagonal)	146 nm²	
LH1 dimer	234 nm²	
LH2 monomer	37 nm²	
LH1 <sub>2</sub> + 6 LH2	456 nm²	11

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



# Photon processing rate at the RC

Which process limits the RCs turnover?



1 RC can serve 1 LH1 + 3 LH2 = 44 γ/s → 25 ms Milano et al. 2003

+ binding, charge transfer
≈ 50 ms per quinol (estimate)

with 2e- H+ pairs per quinol

→ 40–50 γ/s per RC  $\approx 22 \text{ QH}_2/\text{s}$ 

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

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*<sub>1</sub> Placement — Diffusional limits?

Roundtrip times maximal capacity of the carriers:

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

# **Cytochrome c**<sub>2</sub>:

T<sub>Diff</sub>≈3μs  $T_{pc} \approx 1 \text{ ms}$   $T_{bc1} \approx 12 \text{ ms}$  $T_{round-trip} = 13 \text{ ms} \rightarrow \leq 3 \text{ cyt c per vesicle}$ sufficient to carry e-'s available: 22 cyt c per vesicle

#### Quinol:

 $T_{_{RC}} \approx 50 \text{ ms}$   $T_{_{bc1}} \approx 23 \text{ ms}$   $T_{_{Diff}} \approx 1 \text{ ms}$ 

 $T_{round-trip} = 75 \text{ ms} \rightarrow \leq 7 \text{ Q per vesicle}$ sufficient to carry e-'s.

available: 100 Q per vesicle





Diffusion is not limiting

 $\rightarrow$  poses no constraints on the position of  $bc_1$ 

#### **Parameters**

protein	throughput	H+	total number	rate determined from	explained
	per protein	equivalents	per avg.		in section
	(natural units)	per protein	vesicle of		
		[1/s]	45 nm diameter		
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	$\leq 2  imes 42$ 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 $c_2$	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.

Geyer & Helms, Biophys J. (2006)

# **Summary 1**

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

#### Photosynthetic Vesicle Architecture and Constraints on Efficient Energy Harvesting

Melih Şener,<sup>1‡</sup> Johan Strümpfer,<sup>1§</sup> John A. Timney,<sup>9</sup> Arvi Freiberg,<sup>||</sup> Arvi Freiberg,<sup>||</sup> C. Neil Hunter,<sup>9</sup> and Klaus Schutten<sup>1‡§</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>I</sup>Institute of Physics and <sup>\*\*</sup>Institute of Molecular and Cell Biology, University of Tartu, Tartu, Estonia



67

#### **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 parallely working proteins. Their number N times the conversion rate of a single protein R determines the total throughput of this step.

- $\boldsymbol{\gamma}$  : incoming photons collected in the LHCs
- E : excitons in the LHCs and in the RC
- $e^-H^+\,$  electron–proton pairs stored on the quinols
- $e^{\scriptscriptstyle -}$  for the electrons on the cytochrome  $c_2$
- pH : transmembrane proton gradient
- H<sup>+</sup> : protons outside of the vesicle (broken outine 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)  $\rightarrow$  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*<sup>1</sup> 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)

- $\rightarrow$  there are two regimes
- one limited by available light
- one limited by  $bc_1$  throughput





# oxidation state of cytochrome c<sub>2</sub> pool





# oxidation state of cytochrome c<sub>2</sub> pool





# total number of produced ATP



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



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



At high light intensity, c2 pool is mainly oxidized.

If light is turned off, bc1 can continue to work (load c2s, pump protons, let ATPase produce ATP) until c2 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 Blaß, 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).



 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>

Wolfgang P. Barz,<sup>‡,‡</sup> Francesco Francia,<sup>||</sup> Giovanni Venturoli,<sup>||</sup> B. Andrea Melandri,<sup>||</sup> André Verméglio,<sup>⊥</sup> and Dieter Oesterhelt\*,<sup>‡</sup>

# **Parameters not optimized**

Parameter	Value	Description
bc1∷kon(H⁺ont)	10 <sup>10</sup> nm <sup>3</sup> s <sup>-1</sup>	rate for proton uptake from the cytoplasm by $bc_1$
bc1∷kt(e:Q₀=>FeS)	2.3 * 10 <sup>3</sup> s <sup>-1</sup>	rate for electron transfer from Q <sub>o</sub> to FeS
$bc_1$ :: $k_{tr}(e:c_1 \Rightarrow c_2)$	10 <sup>5</sup> s <sup>-1</sup>	electron transfer rate from $c_1$ to bound cytochrome $c_2$
$bc_1::k_{tr}(e:Q_0 \Longrightarrow b_L)$	10 <sup>4</sup> s <sup>-1</sup>	electron transfer from $Q_0$ to $b_L$ heme
$bc_1::k_{tr}(e:b_L=>b_H)$	104 s-1	electron transfer from $b_{\rm L}$ to $b_{\rm H}$ heme
ΔΦ:: <i>V</i>	2.65 * 10 <sup>4</sup> nm <sup>3</sup>	inner volume of the vesicle
∆ <b>Φ∷</b> A	5.28 * 10 <sup>3</sup> nm <sup>2</sup>	membrane area (Q pool "volume")
$\Delta \Phi$ :: $C_{Hin}$	1.0 e	effective charge of a free proton in the vesicle
ΔΦ:: <i>C</i> <b>h</b> m	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
$\Delta \Phi$ :: $C_{ared}$	-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::Np	80	number of titratable groups in the vesicle
PR∷pK	5.0	pK of the titratable groups
Ncare	10	number of dimeric core complexes (2 RC + 1 LHC)
Мы	10	number of cytochrome <i>bc</i> <sub>1</sub> complexes
NATPase	1	number of ATPases
N <sub>c2</sub>	20	total number of cytochrome $c_2$
Nq	200	total number of quinones

Table S1: Model Parameters Not Included in the Optimization Process

### Parameter optimization through evolutionary algorithm





57

# **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	< <i>P</i> >	P <sub>min</sub> P <sub>max</sub>	P <sub>min</sub> / P <sub>max</sub>
LHC::o	$m^2 W^{-1} s^{-1}$	6.22	6.026.42	0.94
LHC::No	1	1.31	0.81 2.13	0.38
LHC::k <sub>D</sub> (E)	s <sup>1</sup>	1.9 * 10 <sup>3</sup>	(1.13.8) * 10 <sup>3</sup>	0.29
$RC:k_{on}(E)$	s <sup>1</sup>	24 * 10 <sup>6</sup>	(1.24.5) * 10 <sup>6</sup>	0.27
RC::kon(H <sup>+</sup> )	nm <sup>3</sup> s <sup>1</sup>	1.4 * 10 <sup>8</sup>	(1.31.6) * 10 <sup>8</sup>	0.81
${\rm RC}$ :: $k_{\rm on}({\rm Q})$	nm² s ¹	6.0 * 10 <sup>4</sup>	(4.48.1) * 10 <sup>4</sup>	0.54
RC::k <sub>off</sub> (QH2)	s <sup>1</sup>	87	70108	0.65
$RC:k_{on}(c2red)$	nm <sup>3</sup> s <sup>-1</sup>	9.2 * 10 <sup>5</sup>	(7.311.5) * 10 <sup>5</sup>	0.63
RC::k <sub>off</sub> (c2ox)	s <sup>1</sup>	2.2 * 10 <sup>3</sup>	(1.63.0) * 10 <sup>3</sup>	0.53
bc1::kon(QH2@Qo)	nm² s ¹	1.2 * 10 <sup>4</sup>	(0.791.7) * 10 <sup>4</sup>	0.46
bc1::koff(Q@Qo)	s <sup>1</sup>	28.3	26.330.4	0.86
$bc_1 :: k_n (\mathbf{Q}; \mathbf{Q}_o - \! > \! \mathbf{Q}_i)$	s <sup>1</sup>	4.9 * 10 <sup>3</sup>	$(3.66.7) * 10^3$	0.54
bc1::kon(Q@Qi)	mm² s 1	6.7 * 10 <sup>5</sup>	(4.510) * 10 <sup>5</sup>	0.45
bc1::koff(QH2@Qi)	s <sup>1</sup>	86	68110	0.62
$bc_1::k_{tr}(QH2;Q_i - >Q_o)$	s <sup>1</sup>	3.8 * 10 <sup>3</sup>	(2.65.5) * 10 <sup>3</sup>	0.47
$bc_1$ :: $k_{on}(c2ox)$	nm³ s 1	94 * 10 <sup>6</sup>	(6.314) * 10 <sup>6</sup>	0.47
$bc_1$ :: $k_{off}$ (c2red)	s <sup>1</sup>	6.0 * 10 <sup>3</sup>	(3.311) * 10 <sup>3</sup>	0.30
bc1::koff(H+@Qo)	s <sup>1</sup>	2.4 * 10 <sup>4</sup>	(1.34.3) * 10 <sup>4</sup>	0.30
$bc_1::k_0(\text{FeS:b}=>c)$	s <sup>1</sup>	3.9 * 10 <sup>3</sup>	(3.15.1) * 10 <sup>3</sup>	0.61
<i>bc</i> 1:: <i>k</i> tr(FeS:c=>b)	s <sup>1</sup>	2.8 * 10 <sup>3</sup>	$(2.23.6) * 10^3$	0.61
$bc_1::k_t(\mathbf{e}:\mathbf{b}_{H}=>\mathbf{Q}_{i})$	s <sup>1</sup>	7.7 * 10 <sup>3</sup>	(5.012) * 10 <sup>3</sup>	0.42
$bc_1$ :: $\Phi_0$	mV	102	83114	0.73
ΔΦ::U <sub>o</sub>	mV/e	10.3	9.511	0.85
$\Delta \Phi$ :: $\Delta \Phi_0$	mV/pH	10	7.613.7	0.55
PR::pK	1	4.84	3.95.9	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:

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

 $x(t_i)$ : 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 🗚 Ф	A9 cytc	B1 Q	86 P	B6 cytc	BC1
importance score	4.4	7.7	5.8	9.7	3.8	52	8.9	55
correlation	0.09	0.44	022	038	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 52 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 2

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).

