Modeling the photosynthetic system I as complex interacting network

Montepietra D.¹, Bellingeri M.^{2,3*}, Scotognella F.^{3,4}, and Cassi D.²

¹Dipartimento di Fisica, Università di Modena e Reggio Emilia, via Campi, 213/a, 41125 Modena, Italy

²Dipartimento di Scienze Matematiche, Fisiche e Informatiche, Università di Parma, via G.P. Usberti, 7/a, 43124 Parma, Italy

³Dipartimento di Fisica, Politecnico di Milano, Piazza Leonardo da Vinci 32, 20133, Milano, Italy

⁴Center for Nano Science and Technology@PoliMi, Istituto Italiano di Tecnologia, Via Giovanni Pascoli, 70/3, 20133, Milan, Italy

* Corresponding author: michele.bellingeri@polimi.it

Abstract

In this paper we model the excitation energy transfer (EET) of the photosynthetic system I (PSI) of the common pea plant *Pisum Sativum* as complex interacting network. The magnitude of the link energy transfer between nodes-chromophores is computed by Forster Resonant Energy Transfer (FRET) using the pairwise physical distances between chromophores from the PDB (Protein Data Bank). We measure the global PSI network EET efficiency adopting well-known network theory indicators: the network efficiency (Eff) and the largest connected component (LCC). We find that when progressively removing the weak links of lower EET, the network efficiency (Eff) decreases while the EET paths integrity (LCC) is still preserved. This finding would show that the PSI is a resilient system owning a large window of functioning feasibility and it is completely impaired only when removing most of the network links. Furthermore, we perform nodes removal simulations to understand how the nodes-chromophores malfunctioning may affect the PSI functioning. We discover

that the removal of the core chlorophylls triggers the fastest decrease in the network efficiency (E_{ff}), unveiling them as the key component boosting the high EET efficiency. Our outcomes open new perspectives of research, such comparing the PSI energy transfer efficiency of different natural and agricultural plant species and investigating the light-harvesting mechanisms of artificial photosynthesis both in plant agriculture and in the field of solar energy applications.

Keywords: Complex network, photosystem I, photosynthetic network, biological network, network robustness, network attack.

1. INTRODUCTION

The ability of photosynthetic oxygenic organisms to convert light energy into chemical energy depends on a group of large membrane-bound complexes whose coordinated activity allows the capture of photons and their conversion into highly energetic molecules as NADPH and ATP through an electron transport chain [1]. The whole photosynthetic process is driven by the biological complexes possessing the light-capturing function: photosystem I (PSI) and photosystem II (PSII) [2,3]. The photosynthetic process starts with the catch of an incoming photon by PSII that utilizes it to oxidize water. The oxidation of water produces oxygen and reduces membrane-embedded quinones. The reduced quinones are then utilized by the cytochrome b_6f complex to create a proton gradient across the membrane and to reduce the small copper protein plastocyanin (PC), that is the electron donor of PSI. After an additional photon is absorbed by PSI, an oxidation takes place and the removed electron migrates to finally reduce ferredoxin (Fd). Eventually, reduced Fd will reduce NADP to NADPH that will power the Calvin cycle to produce carbohydrates [4]. The photosynthetic process is schematically depicted in Figure 1.

The light-harvesting biological complexes PSI and PSII are thought to be the fundamental (and probably the oldest) photosynthetic units [5]. Within the PSII and PSI complexes light is absorbed and transferred to a reaction center (RC) through a highly efficient Excitation Energy Transfer (EET) [6]. The two photosystems display a common structural organization and two main functional moieties: a

core complex, containing the RC where the photochemical reactions occur, and a peripheral antenna system devoted to the increase of the light harvesting capability and to the regulation of the photosynthetic process [4,7,8]

The core complexes have been well conserved during the evolution, as most of the subunits are similar in prokaryotic and eukaryotic photosystems and only a few are specific to each group. On the contrary, the peripheral antenna system displays great variability, being composed of peripheral associated membrane proteins in cyanobacteria, called phycobilisomes, and integral Light Harvesting Complex (LHC) membrane proteins in eukaryotic cells [9].

The light-harvesting process and the EET from the antenna complexes to the RC are facilitated by pigment-protein complexes (PPCs). In green plants (Viridiplantae), most of the light is absorbed by the photosynthetic pigments chlorophyll *a* and *b*, allowing efficient light harvesting and ultrafast EET amongst antenna chlorophylls, leading to the quantum and thermodynamic efficiencies which are the highest known [9,10]

When a photon is absorbed, a chlorophyll (Chl) is excited to the singlet excited state (1Chl*). In PSI this excitation energy is transferred from the antenna pigments to the primary electron donor P700, a special Chls pair located at the RC. The excitation energy is used to build up the singlet excited state P700*, leading to charge separation, where P700⁺ remains and the electron is transferred across the thylakoid membrane by a chain of electron carriers.

Other cofactors present in PSI are also involved in the processes of light capturing, charge separation and prevention of photodamage, like carotenoids (Cars). This network of pigments, whose absorption spectra spans a broad spectral range, is embedded in a protein scaffold which enforces chromophores spatial organization by holding them in a relatively rigid position and orientation [3].

While huge amount of research has been focused at understanding the PSI core electron transfer chemistry which is induced by the arrival of excitation energy at the RC [11–17], few attention has

been directed to unveil the underneath structural and topological patterns shaping the high EET efficiency of the PSI system [2,3].

The purpose of this research is the analysis of the structural organization of a plant PSI as a complex networked system for EET by means of network science viewpoint. Network science has proven to be a powerful tool for the study of complex systems from many different realms [18–26]. Basically, a network is a model composed by nodes (or vertices) and links (edges), where nodes indicate the components of the system and the links define the interactions among them [27,28]. Here we model the PSI light-harvesting system of the plant *Pisum sativum* as a complex network of nodes-chromophores describing the energy transfer links among them. First, we find that the PSI is a highly connected network with very short EET pathways from the nodes harvesting the light to the RC. Second, we discover that the capacity to perform EET in the whole network is severely impaired only by removing the major amount of EET links. This unveils the high resilience of the PSI system, which holds the capacity to perform the energy transfer in the whole network even reaching severe conditions for its feasibility. Last, we simulate different scenarios of chromophores malfunctioning by removing nodes from the network [18,29]. We find that the removal towards the core Chls is clearly more harmful in decreasing the PSI functioning, suggesting the core as the key structural component of the high efficiency EET in the PSI of the *P. Sativum*.

2. METHODS

2.1 Network theory

A generic network (or graph) $G(N,L)$ is composed by a set of N nodes and L links connecting them. The network G can be abstracted in $N \times N$ square adjacency matrix A , where the element $a_{i,j}=1$ indicates the presence of a link between node i and node j and 0 otherwise [27]. In the case G is undirected, e.g. the link have no directionality, the adjacency matrix is symmetrical with $a_{i,j}=a_{j,i}$. In a directed network the adjacency matrix is not symmetrical, the links own proper direction and $a_{i,j}=1$ stands for a link from i to j , and $a_{j,i}=1$ indicates that the link goes from j to i . In the case the links

among nodes are only present-absent, as in G above described, the network is called binary or topological or unweighted. Nevertheless, many real networks are naturally weighted displaying a large heterogeneity in the capacity of the links-connections. In fact, in weighted network links are associated with weight values that differentiate them in terms of their strength, intensity, or capacity [19,22,30–33]. Thus, a weighted network can be abstracted by a $N \times N$ weighted adjacency matrix W , where the element $w_{ij} > 0$ in the case there is link of weight w_{ij} between the node i and the node j and $w_{ij} = 0$ otherwise. Many empirical researches demonstrate that purely binary-topological models are inadequate to explain the complex structure of the real systems, and that there is a need for weighted network models to overcome the pure topological description [27,28,34]

2.2 Building the Photosystem I Network (PSI)

We build the PSI network of the common pea plant (*P. sativum*) starting from the PDB (Protein Data Bank) entry 4RKU. In the PDB, we find all available crystals of PSI belong to this species. We model the PSI as weighted directed network of N nodes-chromophores and L links-EET. The nodes represent chromophores with different functions, i.e. β -carotenes (Cars) and chlorophylls (Chls a or b). To model the PSI network we assumed that Cars and Chls can be summarized by one of their atoms. In fact, Cars are long conjugated molecules with electrons delocalized in the entirety of the molecule, making the internal processes much faster than the external ones (i.e. energy transfer) [35]; so each Car can be considered almost like a point-node object delivering energy in the system. For this reason, we choose the C_{15} carbon, located in the middle of the molecule, to model the carotenoids. We found 21 Cars molecules in the structure. For Chls we represented the molecule using its Mg atom, that is a common practice in the literature [6,36]. We found 158 Chl molecules in the PSI complex (including P700). In this way every node of the PSI network is a C or a Mg atom, representing the corresponding molecule. Modeling the PSI we obtain a network with a total number of nodes $N = 179$.

Then, we assume that the underlying mechanism of energy transfer (EET) between Chls, and between Cars and Chls is represented by Forster Resonant Energy Transfer (FRET) [37]. FRET contributes to the most of the EET acting as a good approximation to investigate energetic coupling of the

underlying spatial organization of the chromophores [38,39]. Future investigations may expand the model presented here by including the quantum mechanical dynamics to model the EET of the PSI.

The FRET efficiency (E):

$$E_{ij} = \frac{1}{1 + \left(\frac{r_{ij}}{R_0}\right)^6} \quad (1)$$

where E_{ij} is the FRET efficiency the link connecting nodes i and j , R_0 is the Forster distance between i and j , and r_{ij} is the physical distance between the nodes. All pairwise physical distances between nodes can be calculated directly from their PDB spatial coordinates. Therefore, E_{ij} represents the weight w_{ij} of the link joining node i and node j . The value of R_0 of Chls-Chls FRET is 90 Å [40,41], and we assume the same value for Cars-Chls FRET. Furthermore, in the photosynthetic process, Cars absorb and transfer lower wavelength photons to Chls, which in turn cannot transfer the excitation back to Cars but only send it to other Chls, until they reach the P700 [42]. For this reason, carotenoids will be represented by nodes with outgoing links only. On the contrary, the two Chls forming P700 will have only incoming links as they can only receive excitation; furthermore, the Chls of the P700 do not perform EET and thus no direct link connection is present between them. Last, with the sake of simplicity, our model does not consider the competition between parallel energy transfers that may induce radiative (light emission) and non-radiative (heat production) de-excitation, with a global EET efficiency decreasing.

2.3 The PSI network with the energy funnel

Light-harvesting complexes (LHCs) chlorophylls are the light entry gate of photosynthesis [43]. The different LHCs Chls are used to absorb light in spectral regions where more radiant energy is available, increasing the amount of sunlight that can be collected [44]. In the PSI, four types of Chls LHCs are found: LHC1, LHC2, LHC3, LHC4 [4]. Their schematic representation can be seen in Fig.

1. Different LHCs have different light absorption peaks, corresponding to different site energies of the Chls. A photon with a certain energy absorbed in one of these LHC will therefore be directed only to Chls with lower or equal site energies. This generates an energetic funnel, that it is used by the photosynthetic organisms to speed up and increase the EET efficiency [45]. The first step in modelling this funnel is thus the assignment of the different site energies to the LHCs, that will define the connectivity among them. We assign the site energies of the LHCs as in [38] (Table 2). Further, [38] showed that inter-LHCs EET is much less efficient than LHC-core EET. Following this finding, we assume that only adjacent LHCs can transfer energy between them. We depict the possible links among LHCs complexes in Fig. 2. In Table 1 we summarize the different nodes-chromophores and their properties.

2.4 The cut-off distance (CD)

The links of the PSI network own different FRET efficiencies as a function of the nodes pairwise distances of the PDB structure. The closer the nodes, the higher the link weight (FRET efficiency) joining them. For this reason, we can build the PSI network with different number of links as a function of their FRET magnitude. To do this, we define the cut-off distance (CD) as the distance between two nodes over which we assume the link between them absent because the associated FRET efficiency computed by Eq. 1 is too low. Thus, by lowering the CD parameter, we require shorter pairwise distance to form links among nodes. In other terms, starting from the network where all links are drawn, when lowering the CD, we operate a removal of the links of lower weight (weak links) maintaining the links of higher weight (strong links). We created several networks using the cut-off distance values: NO cut-off, 90, 80, 70, 60, 50, 40, 30, 20 and 10 Å.

2.4 Measures of network functioning

The largest connected component (LCC): The *LCC* is a widely used measure of the network functioning [18,46,47]. The *LCC* is also known as giant component and it accounts the highest number of connected nodes in the network. We compute two types of *LCC*, by considering the weakly largest connected component (*LCC_{weak}*) and strongly largest connected component (*LCC_{strong}*). The

LCC_{weak} represents the largest number of nodes in which every node can reach any other by an undirected path. Differently, the LCC_{strong} represents the largest number of nodes in which every node can reach any other node by a directed path [27].

The network efficiency (Eff): Eff is a widely used measure of network functioning evaluating how efficiently the network exchanges information [20,48,49]. Eff is based on the shortest paths (SP) notion. A path is a sequence of links connecting two nodes in the network; a binary shortest path between a pair of nodes is the minimum number of links that we have to pass travelling between the nodes [50]. Differently, the weighted shortest paths (wSP) consider both the number of links and the weight attached to them. To compute wSP, we first compute the inverse of the link weights; then we compute the weighted shortest paths as the minimum sum of the inverse link weights necessary to travel among nodes [27,34].

The average efficiency of the network is then defined:

$$Eff = \frac{1}{N \cdot (N - 1)} \sum_{i \neq j \in G} \frac{1}{d(i, j)} \quad (2)$$

where $d_{i,j}$ is the shortest path between node i and node j , and N is the total number of network nodes. Eq. 1 indicates that the longer the shortest paths among nodes, the lower the efficiency of the network. The efficiency can be computed for undirected and directed networks when considering undirected and directed shortest paths respectively. Furthermore, the network efficiency can properly evaluate binary and weighted networks; in the case the network is weighted, $d_{i,j}$ indicates the weighted shortest path between nodes i and j . Since the PSI is a weighted and directed network, we use the weighted and directed network efficiency (Eff).

The number of connected nodes to P700 (CN): we account the number of connected nodes to P700 through a directed path. This measure indicates how many nodes in the network can actually transfer energy (directly or indirectly) to the P700 reaction center. In other words, CN accounts the number of nodes-chromophores that can properly contribute to the PSI functioning.

2.5 The PSI network properties

We analysed the PSI system using well known notions from networks science. **Nodes degree (k)**: is the total number of links to the nodes [27]. We also computed the number of ingoing links, i.e. the in-degree (k_{in}) and the number of outgoing links, the out-degree (k_{out}). **Nodes Strength (S)**: is the sum of the link weights to the nodes [49]. The strength of a node is the weighted counterpart of the node degree (k). Analogously we can compute the ingoing strength (S_{in}) and the outgoing strength (S_{out}). **Nodes betweenness centrality (BC)**: is a widely used measure of nodes centrality based on shortest paths. The BC of a node is the number of the shortest paths between all nodes pairs that pass through that node [27,51]. Here we adopt both the binary betweenness centrality (BC) and the weighted version of the betweenness centrality of the nodes (BC_w). The BC_w is computed using the weighted shortest paths. The weighted betweenness centrality of a node accounts the number of weighted shortest paths from any couple of nodes passing along that node. The higher is the BC_w of a node, the higher is the number of weighted shortest paths passing along it (and the more central is the node). **δ centrality**: the node δ centrality accounts the decrease of the network efficiency (Eff) triggered by the removal of that node from the network [48]. The higher the δ , the higher the contribution of the node on the network efficiency (Eff).

2.6. The nodes removal simulations

We simulated different nodes removal (attack) processes over the PSI network [18,52]. The PSI network nodes removal may describe functionality loss of the chromophores, that can be due by aging, disease, pollution or others [3].

The nodes removal strategies are:

- **Rand**: network nodes are randomly removed. It represents the case of failures in the network and it is used as benchmark comparison of the harmful of the selective attack strategies [18,29].

- **Rand LHC:** in this strategy we randomly remove nodes in the LHC ensemble that is composed by 59 nodes.
- **Rand Core:** in this strategy we randomly removed nodes in the Core Chls ensemble that is composed by 97 nodes.
- **Rand Car:** in this strategy we randomly removed nodes in the Cars ensemble that is composed by 21 nodes.

We measured the functioning of the network during the nodes removal process using the network directed weighted efficiency (Eff) and the weakly largest connected component (LCC_{weak}). Since the random removals are stochastic processes, we average the outcomes from 10^3 simulations. The damage triggered in the system by the specific PSI nodes removal (Rand Car, Core and LHC) is then compared with the damage induced by the random removal (Rand) of the same number of nodes. In this way we can discriminate the importance of the different types of nodes-chromophores.

2.7. The robustness of the network

When nodes are removed from the PSI network, we assist to the elongation of the EET paths and consequent system functioning decrease (Fig. 2 a, b). We can assess the system damage subjected to nodes removal according to the functioning measures. The more important the nodes removed from the network, the steeper is the decrease in the network functioning measure. For example, in the Fig. 2c the red removal strategy identifies more important nodes causing a steeper decrease in the network functioning efficiency (Eff) with respect to the blue strategy. To compare the response among nodes removal strategies, we resume the outcome in a single value defined network robustness (R), depicted by the bar plots of the Fig. 3d. The value of R corresponds to the area below the curve of the system functioning measure subjected to nodes removal. We referred to R_{EFF} and R_{LCC} when the network functioning is evaluated by the Eff and LCC respectively.

3. RESULTS

3.1. The network properties

In Table 3 we resume the PSI network properties for each CD value. The number of links L , average nodes degree $\langle k \rangle$ and average nodes strength $\langle S \rangle$ linearly decrease lowering CD (Fig. 3). The nodes degree k ranges in the interval $0 < k < 270$. The nodes degree distribution is: a tri-modal with three bell-shaped distributions from NO CD to CD=50 Å, one bell-shaped distribution with most of the values approaching the mean from CD=40 Å to CD=20 Å, and it becomes a Dirac distribution where all nodes present degree 1 for CD=10 Å (Appendix, Fig. A1). The node strength distribution follows the node degree distribution patterns indicating a nodes degree-strength coupling (Fig. A2). The average link weight $\langle E_w \rangle$ shows a very slow increase lowering CD (Fig. 3); the link weight distributions are highly skewed to the left with many links presenting weight 1 and few links of lower weight (Fig. A3). The network efficiency (Eff) remains roughly constant by decreasing CD from NO CD to CD=70 Å; after this value, we find a more pronounced Eff decrease reaching the minimum for CD=10 Å (Fig. 3). LCC_{weak} spans the whole network with $N=179$ nodes up to reach CD>10 Å (Fig. 4). LCC_{strong} holds the values $N=97$ up to reach CD>10 Å (Fig. 3). We find 177 nodes connected to the P700 (CN) and this number holds constant up to reach CD=20 Å (Fig. 3). Only for CD=10 Å CN drop to *quasi* zero. It is worth noting that for LCC_{weak} , LCC_{strong} and CN measures the critical CD value is 10 Å, below which the measures quickly collapses.

3.2 The properties and connectivity of the different nodes-chromophores

The Core Chls show the highest degree and in-degree (k and k_{in}), betweenness centrality (BC and BC_w), strength (S and S_{in}) and δ centrality than all the others types of nodes-chromophores (Figs 5-6); the LHCs show the highest outgoing degree (k_{out}) and strength (S_{out}) (Figs 5-6); the P700 has the highest in-degree (k_{in} , comparable to the Core) and in-strength (S_{in}) (Figs 5-6). We find a higher number of LHCs-Core links than the number of intra-LHCs links, e.g. each LHC shares more than 1200 links with the Core Chls and *circa* 100 intra-LHCs links for CD=NO (Fig. 7). Nonetheless,

when lowering CD, we find that the number LHCs-Core links tends to decrease, while the number of intra-LHCs links holds roughly constant. For all LHCs we find that the number of LHC-Core links become similar to the number of intra-LHCs links for $CD=50 \text{ \AA}$ (Fig. 7).

3.3 Nodes removal

Carotenoids: Rand Car induces less average damage to the Eff (higher R_{Eff}) of the PSI network than the random removal control (Rand) (Fig. 8). The Rand Car shows roughly half of the R_{EFF} (50% fewer) of Rand. This difference between the R_{Eff} values decreases by lowering CD. Also with the R_{LCC} , Rand Car induces less average damage to the LCC_{weak} than Rand (Fig. A13). The difference between the R_{LCC} values of the Rand Car and Rand is less pronounced for $CD=10 \text{ \AA}$ (Fig. A13).

LHCs: Random removal of LHCs Chls (Rand LHCs) causes less damage to Eff (higher R_{EFF}) than the random removal control (Rand) for all CDs (Fig. 8). The Rand LHCs induces less average damage to the LCC_{weak} (higher R_{LCC}) of than Rand (Fig. A13). This difference is less pronounced for $CD=10 \text{ \AA}$ (Fig. A13).

Core: Random removal of Core Chls (Rand Core) causes more damage to Eff (lower R_{EFF}) than the random removal control (Rand) (Fig. 8). The R_{EFF} obtained by Rand Core is *circa* 25% lower than the one induced by the Rand. Only $CD = 10 \text{ \AA}$, Rand Core triggers greater damage than Rand. Rand Core causes more damage to LCC_{weak} than Rand (Fig. A13). This difference is less pronounced for $CD=10 \text{ \AA}$.

4. Discussion

4.1 Network properties

The Core Chls nodes show the highest degree (k) among all other types of chromophores, i.e. they share the higher number of links in the network (Figs 5-6) indicating the Core as the PSI network component playing a major role for the global energy transfer. This is clearly confirmed by their

highest betweenness centrality for all CDs (Figs 5-6). The nodes with highest betweenness are the articulation points routing the shortest and the most efficient energy transfer paths in the network. As a consequence, these findings demonstrate that the Core is the ‘highly connected’ and ‘most central’ component in the PSI network of the *P. Sativum*, corroborating previous outcomes from literature showing that the primary function of the Core Chls is to enhance the EET efficiency in the system [6,53].

Differently, the discovery that the LHCs nodes own the highest number of outgoing links (k_{out}) than all other types of nodes chromophores and the Cars nodes have only outgoing links (Figs 5-6) would indicate these types of chromophores as important light entry gate. This would confirm empirical evidences from literature, showing that Cars and LHCs have the primary function of light harvesting, and then perform energy transfer to other types of nodes throughout the outgoing links [6,35,53].

4.2 Feasibility of the PSI system

When lowering the cut-off distance (CD), shorter pairwise distance is required between nodes to perform FRET energy transfer links. In other terms, lowering the CD we progressively remove from the network the weak links of lower FRET efficiency and we maintain only the most efficient and most probable FRET links (strong links). The strong links connecting closest chromophores can be viewed as the ‘feasible FRET’ even when the system works at very low regime of functioning, such as scarce light radiation [54]. The low level of functioning may also occur after the vanishing of FRET links, such as for Chls ageing following the life of the leaf, with older PSI still possessing the higher FRET links only [55]. Therefore, CD tuning gives us information on the PSI system feasibility at different level of functioning efficiency and its resilience with regards of leaf ageing.

We discover that lowering the CD, the number of feasible energy transfer links (L) linearly decrease (Fig. 3). This decrease triggers the elongation of the EET paths in the PSI network and the consequent lowering of its efficiency Eff (Fig.3). Hence the PSI system experiences a reduction in the network efficiency, the LCC representing the maximum number of connected nodes-chromophores holds constant up to reach $CD=20$ Å. Further, the number of nodes connected by directed paths to the P700

reaction centre (CN) hold constant up to reach $CD=20 \text{ \AA}$. This would indicate that only for $CD=10 \text{ \AA}$ we assist to the network dismantle that completely impairs the PSI functioning. These outcomes give us information about the whole feasibility of the PSI network functioning, showing that system efficiency is negatively affected by lowering CD; even so, the possibility to perform FRET in the whole network is severely impaired only reaching the lowest value of the CD. The Forster energy transfer process (FRET) appears to have played the role of a design constraint shaping the evolution of photosynthetic organisms and have had profound influence in shaping their architecture [38,39,43]. Our outcomes shed light on the structural evolution of the PSI, showing how the PSI network formed by FRET links is a resilient system with a large '*window of feasibility*', holding the capacity to perform the energy transfer in the whole network even when deprived of the most of FRET links.

4.3 Connectivity among different chromophore types

When analysing the connectivity among different types of chromophores we find that for higher value of CD there is higher LHCs-Core linkage density than intra LCHs. Only decreasing CD up to $CD=50 \text{ \AA}$, the intra-LHCs and the LHCS-Core links density become comparable (Fig. 7). In other words, lowering the CD produces a faster decrease of the LCHs-Core links with small effect in decreasing the intra-LHCs links. On one hand, this indicates that LHCs are spatially closer among them than they are with the Core Chls. On the other hand, this shows that a large amount LHCs-Core links is of lower FRET efficiency. For this reason, for $CD=50 \text{ \AA}$ the network is formed by higher FRET links shared in similar number in LHCs-Core and intra-LHCs links (Fig. 7). This finding would describe a PSI system composed by a fraction of robust and highly probable energy transfer links, that we can call '*low regime links*', shared roughly in the same quantity between intra-LHCs and LHCs-Core links. These links may act as baseline mechanism supporting the system functioning for below-optimum regime, such as for example when the light harvesting is lower. Differently, when the efficiency is higher ($CD>50 \text{ \AA}$), even reaching optimal values (for $CD>80 \text{ \AA}$), a large amount of LHCs-Core links becomes likely to perform the energy transfer supporting the low regime links. We call these numerous and low FRET the '*high regime links*' and they may represent an internal mechanism for

boosting the PSI system toward higher level of functioning.

4.4 The nodes-chromophores attack

The nodes removal (attack) strategies simulations have the aim to mimic the loss of functionality of the chromophores of the PSI network, giving us information on how different network nodes contribute to the functioning efficiency of the overall system [2,3]. In normal conditions when chromophores lose their functionality they get readily substituted in the turnover process [55,56]. In the case the turnover machinery is impaired the system loses chromophores making the photosystem functionality deteriorate progressively. The decrease in the number of chromophore in the photosynthesizing organisms may be the result of genetic mutations, chemical damage (such as the loss of Chls), pollution or the changing environment of the chromophores [3]. The reduction of the Chls number (or functionality) can also be associated to natural phenomena as the yellowing of senescing leaves and autumnal colour changes in the foliage of deciduous trees, but also in seeming evergreens, such as ferns and algae [57].

Removing a node-chromophore in the PSI network would affect the system in two ways. On one hand, it removes from the system one entry gate for light, thus lowering the light harvesting capacity. On the other hand, it removes all the links relating to that node-chromophore breaking up the paths for the FRET passing into that node. The paths disruption elongates the energy transfer travel among nodes producing a decrease in the network efficiency (Eff) (Fig. 2).

We observed that removal of nodes corresponding to carotenoid molecules (Cars) induces lower average damage to the network functioning (Eff and LCC) than the random nodes removal (Fig. 8). In fact, Cars are nodes possessing only outgoing links and very low betweenness centrality (BC) not playing the role of halfway hubs for the FRET paths in the network. Thus, their removal will not induce the FRET paths disruption, resulting in a lower network efficiency decrease, outlining the small Cars importance for the structural efficiency of the energy transfer, and relegating their main role for other important functions, as broadening the spectrum of absorbable light and photo-protection [6].

Similarly, we find that random removal of Chls belonging to LHCs causes lower average damage to network functioning (*Eff* and *LCC*) (Fig. 8). As for carotenoids, LHCs nodes show low betweenness centrality (Figs 5-6), indicating their minor role for routing the shortest paths in the network and consequently the lower damage to the PSI network efficiency (*Eff*) caused by their removal. For this reason, similar to the Cars nodes the LHCs would play a major function of light entry gate as proposed by [58].

Differently, we find that the random removal of Core Chls (Rand Core) triggers the greatest decrease in the network functioning (*Eff* and *LCC*) (Fig. 8). Furthermore, the single removal of Core Chls triggers the highest decrease in the network efficiency (*Eff*) than all the other types of nodes (δ indicator in Figs 5-6). The higher damage to the system EET induced by the removal of the chlorophylls from the Core may be explained with their peculiar embedding within the structure of the PSI network. The Core Chls show the highest node betweenness centrality, unveiling them as major articulation points for the FRET shortest paths among the network nodes (Fig. 5). Their removal breaks up many higher efficiency FRET paths causing greater damage to global energy transfer efficiency, raising the Core Chls as the key network component directly involved in the increase of EET of the *P. Sativum* PSI, confirming by means of network theory previous results from experimental analysis [59,60].

5. Conclusion

In this paper we modelled the PSI of the *Pisum Sativum* as complex interacting network. First, we discovered that progressively removing the weaker FRET links decreases the network efficiency, while the nodes connectivity is still preserved. This finding would unveil a large window of feasibility of the PSI functioning, in which the photosynthetic process it is completely impaired only losing almost all the links in the network. In other terms, the PSI networked system would be highly resilient to the malfunctioning of FRET links. Second, we discovered that the Core Chls removal triggers the fastest functioning decrease, indicating the chromophores in the core of the PSI network as the main

actor boosting the EET. Since the nodes removal simulates the chromophores malfunctioning, caused for example by the senescing leaves, biological diseases, pollution chemical damage or others [57], the PSI energy transfer efficiency would be more vulnerable to the core chlorophylls functionality loss. The outcomes presented here open new promising researches, for example building and comparing the energy transfer efficiency of the PSI network of different natural and agricultural plant species, or improving the light-harvesting mechanisms of artificial photosynthesis both in plant agriculture [59,60] and in the field of solar energy applications [61].

Acknowledgments

MB and FS acknowledges financial support from Fondazione Cariplo, grant n° 2018-0979. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement No. [816313]). Many thanks to Prof. Guglielmo Lanzani whose comments greatly improved the first version of this paper.

References

1. Michael NDC. 2019 Lehninger Principles of Biochemistry, 7th. *J. Chem. Inf. Model.* (doi:10.1017/CBO9781107415324.004)
2. Baker LA, Habershon S. 2015 Robustness, efficiency, and optimality in the Fenna-Matthews-Olson photosynthetic pigment-protein complex. *J. Chem. Phys.* **143**. (doi:10.1063/1.4930110)
3. Baker LA, Habershon S. 2017 Photosynthetic pigment-protein complexes as highly connected networks: Implications for robust energy transport. *Proc. R. Soc. A Math. Phys. Eng. Sci.* **473**. (doi:10.1098/rspa.2017.0112)
4. Mazor Y, Borovikova A, Nelson N. 2015 The structure of plant photosystem i super-complex at 2.8 Å resolution. *Elife* **4**, 1–18. (doi:10.7554/eLife.07433)
5. Nelson N, Ben-Shem A. 2005 The structure of photosystem I and evolution of photosynthesis. *BioEssays* **27**, 914–922. (doi:10.1002/bies.20278)

6. Mirkovic T, Ostroumov EE, Anna JM, Van Grondelle R, Govindjee, Scholes GD. 2017 Light absorption and energy transfer in the antenna complexes of photosynthetic organisms. *Chem. Rev.* **117**, 249–293. (doi:10.1021/acs.chemrev.6b00002)
7. Fromme P, Jordan P, Krauß N. 2001 Structure of photosystem I. *Biochim. Biophys. Acta - Bioenerg.* **1507**, 5–31. (doi:10.1016/S0005-2728(01)00195-5)
8. Shen J-R. 2015 The Structure of Photosystem II and the Mechanism of Water Oxidation in Photosynthesis. *Annu. Rev. Plant Biol.* (doi:10.1146/annurev-arplant-050312-120129)
9. Caffarri S, Tibiletti T, Jennings R, Santabarbara S. 2014 A Comparison Between Plant Photosystem I and Photosystem II Architecture and Functioning. *Curr. Protein Pept. Sci.* **15**, 296–331. (doi:10.2174/1389203715666140327102218)
10. Wraight CA, Clayton RK. 1974 The absolute quantum efficiency of bacteriochlorophyll photooxidation in reaction centres of Rhodospseudomonas spheroides. *BBA - Bioenerg.* (doi:10.1016/0005-2728(74)90009-7)
11. Cheng Y-C, Fleming GR. 2009 Dynamics of Light Harvesting in Photosynthesis. *Annu. Rev. Phys. Chem.* (doi:10.1146/annurev.physchem.040808.090259)
12. Lambert N, Chen YN, Cheng YC, Li CM, Chen GY, Nori F. 2013 Quantum biology. *Nat. Phys.* (doi:10.1038/nphys2474)
13. Abramavicius D, Mukamel S. 2010 Quantum oscillatory exciton migration in photosynthetic reaction centers. *J. Chem. Phys.* (doi:10.1063/1.3458824)
14. Scholes GD, Fleming GR, Olaya-Castro A, Van Grondelle R. 2011 Lessons from nature about solar light harvesting. *Nat. Chem.* (doi:10.1038/nchem.1145)
15. Frankcombe TJ. 2015 Explicit calculation of the excited electronic states of the photosystem II reaction centre. *Phys. Chem. Chem. Phys.* (doi:10.1039/c4cp04468a)
16. Olson JM. 1981 Evolution of photosynthetic reaction centers. *BioSystems* **14**, 89–94. (doi:10.1016/0303-2647(81)90024-1)
17. Johansson M, Callaghan T V., Bosiö J, Åkerman HJ, Jackowicz-Korczynski M, Christensen TR. 2013 Rapid responses of permafrost and vegetation to experimentally increased snow cover in sub-arctic Sweden. *Environ. Res. Lett.* (doi:10.1088/1748-9326/8/3/035025)
18. Bellingeri M, Cassi D, Vincenzi S. 2014 Efficiency of attack strategies on complex model and real-world networks. *Phys. A Stat. Mech. its Appl.* (doi:10.1016/j.physa.2014.06.079)
19. Bellingeri M, Bodini A. 2013 Threshold extinction in food webs. *Theor. Ecol.* (doi:10.1007/s12080-012-0166-0)

20. Bullmore E, Sporns O. 2009 Complex brain networks: Graph theoretical analysis of structural and functional systems. *Nat. Rev. Neurosci.* (doi:10.1038/nrn2575)
21. Zanin M, Lillo F. 2013 Modelling the air transport with complex networks: A short review. *Eur. Phys. J. Spec. Top.* (doi:10.1140/epjst/e2013-01711-9)
22. Bellingeri M, Bevacqua D, Scotognella F, LU ZM, Cassi D. 2018 Efficacy of local attack strategies on the Beijing road complex weighted network. *Phys. A Stat. Mech. its Appl.* (doi:10.1016/j.physa.2018.06.127)
23. Agliari E, Barra A, Galluzzi A, Guerra F, Tantari D, Tavani F. 2015 Retrieval capabilities of hierarchical networks: From dyson to hopfield. *Phys. Rev. Lett.* (doi:10.1103/PhysRevLett.114.028103)
24. Onnela JP, Saramäki J, Hyvönen J, Szabó G, De Menezes MA, Kaski K, Barabási AL, Kertész J. 2007 Analysis of a large-scale weighted network of one-to-one human communication. *New J. Phys.* (doi:10.1088/1367-2630/9/6/179)
25. Vicente-Serrano SM, Beguería S, Lorenzo-Lacruz J, Camarero JJ, López-Moreno JJ, Azorin-Molina C, Revuelto J, Morán-Tejeda E, Sanchez-Lorenzo A. 2012 Performance of drought indices for ecological, agricultural, and hydrological applications. *Earth Interact.* (doi:10.1175/2012EI000434.1)
26. Piccardi C, Tajoli L. 2018 Complexity, centralization, and fragility in economic networks. *PLoS One* (doi:10.1371/journal.pone.0208265)
27. Boccaletti S, Latora V, Moreno Y, Chavez M, Hwang DU. 2006 Complex networks: Structure and dynamics. *Phys. Rep.* (doi:10.1016/j.physrep.2005.10.009)
28. Bellingeri M, Bevacqua D, Scotognella F, Cassi D. 2019 The heterogeneity in link weights may decrease the robustness of real-world complex weighted networks. *Sci. Rep.* (doi:10.1038/s41598-019-47119-2)
29. Holme P, Kim BJ, Yoon CN, Han SK. 2002 Attack vulnerability of complex networks. *Phys. Rev. E - Stat. Physics, Plasmas, Fluids, Relat. Interdiscip. Top.* (doi:10.1103/PhysRevE.65.056109)
30. Barrat A, Barthélemy M, Pastor-Satorras R, Vespignani A. 2004 The architecture of complex weighted networks. *Proc. Natl. Acad. Sci. U. S. A.* (doi:10.1073/pnas.0400087101)
31. Sporns O, Zwi JD. 2004 The small world of the cerebral cortex. *Neuroinformatics.* (doi:10.1385/NI:2:2:145)
32. Squartini T, Van Lelyveld I, Garlaschelli D. 2013 Early-warning signals of topological collapse in interbank networks. *Sci. Rep.* (doi:10.1038/srep03357)
33. Bellingeri M, Bodini A. 2016 Food web's backbones and energy delivery in ecosystems. *Oikos* (doi:10.1111/oik.02244)

34. Bellingeri M, Bevacqua D, Scotognella F, Alfieri R, Cassi D. 2020 A comparative analysis of link removal strategies in real complex weighted networks. *Sci. Rep.* (doi:10.1038/s41598-020-60298-7)
35. Llansola-Portoles MJ, Pascal AA, Robert B. 2017 Electronic and vibrational properties of carotenoids: From in vitro to in vivo. *J. R. Soc. Interface.* (doi:10.1098/rsif.2017.0504)
36. Sener MK, Lu D, Ritz T, Park S, Fromme P, Schulten K. 2002 Robustness and optimality of light harvesting in cyanobacterial photosystem I. *J. Phys. Chem. B* **106**, 7948–7960. (doi:10.1021/jp020708v)
37. Förster T. 1948 Zwischenmolekulare Energiewanderung und Fluoreszenz. *Ann. Phys.* (doi:10.1002/andp.19484370105)
38. Şener MK, Jolley C, Ben-Shem A, Fromme P, Nelson N, Croce R, Schulten K. 2005 Comparison of the light-harvesting networks of plant and cyanobacterial photosystem I. *Biophys. J.* **89**, 1630–1642. (doi:10.1529/biophysj.105.066464)
39. Şener M, Strümpfer J, Hsin J, Chandler D, Scheuring S, Hunter CN, Schulten K. 2011 Förster energy transfer theory as reflected in the structures of photosynthetic light-harvesting systems. *ChemPhysChem* **12**, 518–531. (doi:10.1002/cphc.201000944)
40. Hussain SA. 2009 An Introduction to Fluorescence Resonance Energy Transfer (FRET).
41. Şener M, Strümpfer J, Hsin J, Chandler D, Scheuring S, Hunter CN, Schulten K. 2011 Förster energy transfer theory as reflected in the structures of photosynthetic light-harvesting systems. *ChemPhysChem*. (doi:10.1002/cphc.201000944)
42. Scholes GD, Fleming GR. 2005 Energy Transfer and Photosynthetic Light Harvesting. In *Adventures in Chemical Physics: A Special Volume of Advances in Chemical Physics: Volume 132*, (doi:10.1002/0471759309.ch2)
43. Blankenship RE, Chen M. 2013 Spectral expansion and antenna reduction can enhance photosynthesis for energy production. *Curr. Opin. Chem. Biol.* **17**, 457–461. (doi:10.1016/j.cbpa.2013.03.031)
44. Berg J, Tymoczko J, Stryer L. 2002 *Biochemistry, 5th edition*.
45. Croce R, Van Amerongen H. 2013 Light-harvesting in photosystem i. *Photosynth. Res.* **116**, 153–166. (doi:10.1007/s11120-013-9838-x)
46. Albert R, Barabási AL. 2002 Statistical mechanics of complex networks. *Rev. Mod. Phys.* (doi:10.1103/RevModPhys.74.47)
47. Bellingeri M, Agliari E, Cassi D. 2015 Optimization strategies with resource scarcity: From immunization of networks to the traveling salesman problem. *Mod. Phys. Lett. B* (doi:10.1142/S0217984915501808)

48. Latora V, Marchiori M. 2001 Efficient behavior of small-world networks. *Phys. Rev. Lett.* (doi:10.1103/PhysRevLett.87.198701)
49. Bellingeri M, Cassi D. 2018 Robustness of weighted networks. *Phys. A Stat. Mech. its Appl.* (doi:10.1016/j.physa.2017.07.020)
50. Freeman LC. 1978 Centrality in social networks conceptual clarification. *Soc. Networks* (doi:10.1016/0378-8733(78)90021-7)
51. Newman M. 2010 *Networks: An Introduction*. (doi:10.1093/acprof:oso/9780199206650.001.0001)
52. Iyer S, Killingback T, Sundaram B, Wang Z. 2013 Attack Robustness and Centrality of Complex Networks. *PLoS One* (doi:10.1371/journal.pone.0059613)
53. Ringsmuth AK. 2014 Multiscale Analysis and Optimisation of Photosynthetic Solar Energy Systems.
54. Hankamer B, Lehr F, Rupprecht J, Mussnug JH, Posten C, Kruse O. 2007 Photosynthetic biomass and H₂ production by green algae: From bioengineering to bioreactor scale-up. In *Physiologia Plantarum*, (doi:10.1111/j.1399-3054.2007.00924.x)
55. González-Rodríguez AM, Morales D, Jiménez MS. 2002 Leaf age effects on chlorophyll, Rubisco, photosynthetic electron transport activities and thylakoid membrane protein in field grown grapevine leaves. *J. Plant Physiol.* **159**, 799–803. (doi:10.1078/0176-1617-0597)
56. Beisel KG, Jahnke S, Hofmann D, Köppchen S, Schurr U, Matsubara S. 2010 Continuous turnover of carotenoids and chlorophyll a in mature leaves of Arabidopsis revealed by ¹⁴CO₂ pulse-chase labeling. *Plant Physiol.* **152**, 2188–2199. (doi:10.1104/pp.109.151647)
57. Haag J. 2019 Molecular and Biochemical Enhancement of Chlorophyll in Sports Turf.
58. Perrine Z, Negi S, Sayre RT. 2012 Optimization of photosynthetic light energy utilization by microalgae. *Algal Res.* **1**, 134–142. (doi:10.1016/j.algal.2012.07.002)
59. Melis A. 2009 Solar energy conversion efficiencies in photosynthesis: Minimizing the chlorophyll antennae to maximize efficiency. *Plant Sci.* (doi:10.1016/j.plantsci.2009.06.005)
60. Polle JEW, Kanakagiri S, Jin ES, Masuda T, Melis A. 2002 Truncated chlorophyll antenna size of the photosystems - A practical method to improve microalgal productivity and hydrogen production in mass culture. In *International Journal of Hydrogen Energy*, (doi:10.1016/S0360-3199(02)00116-7)
61. Gust D, Moore TA, Moore AL. 2009 Solar fuels via artificial photosynthesis. *Acc. Chem. Res.* (doi:10.1021/ar900209b)

Tables

Table 1: Summary of the different nodes-chromophores of *P. Sativum* PSI network.

Node Type	Acronyms	Number	Biological function	Link direction
Carotenoids	Cars	21	Photodamage protection; Lower wavelength absorption; Energy transfer	Outgoing
Core Chlorophylls	Core Chls	97	Energy transfer	Outgoing + ingoing
Ligh Harvesting Complexes Chlorophylls	LHCs	59	Light harvesting; Wider harvested light spectrum; Energy transfer	Outgoing + ingoing
P700	P700	2	Reaction center of PSI, special pair of chlorophylls	Ingoing

Table 2: Site energies related to different types of node-chromophores.

NODE IDENTITY	SITE ENERGY (nm)
Cars	400 - 550
LHC4	676.2
LHC3	676.4
LHC1	677.6
LHC2	678.6
Core Chls	678.8
P700	700.0

Table 3: PSI network features for each CD value. $\langle k \rangle$ average nodes degree, $\langle k_{out} \rangle$ average nodes out-degree and $\langle k_{in} \rangle$ average nodes in-degree; L total number of links; $\langle S \rangle$ average node strength, $\langle S_{in} \rangle$ average in-strength, $\langle S_{out} \rangle$ average out-strength; Ew average link weights; Eff global network efficiency; LCC_{weak} weakly largest connected component; LCC_{strong} strongly largest connected component; CN number of connected nodes to the P700.

CD	$\langle k \rangle$	$\langle k_{out} \rangle$	$\langle k_{in} \rangle$	L	$\langle S \rangle$	$\langle S_{out} \rangle$	$\langle S_{in} \rangle$	Ew	Eff	LCC_{weak}	LCC_{strong}	CN
NO	190	95	95	11924	151.3	75.65	75.65	0.7972	0.71	179	97	177
90 Å	158	79	79	9409	141.68	70.84	70.84	0.8984	0.71	179	97	177
80 Å	141	70	70	8279	131.79	65.9	65.9	0.9348	0.71	179	97	177
70 Å	122	61	61	7028	117.32	58.66	58.66	0.9640	0.68	179	97	177
60 Å	100	50	50	5681	98.55	49.28	49.28	0.9832	0.64	179	97	177
50 Å	76	38	38	4181	75.54	37.77	37.77	0.9939	0.58	179	97	177
40 Å	53	27	27	2828	53.3	26.65	26.65	0.9982	0.50	179	97	177
30 Å	32	16	16	1642	31.74	15.87	15.87	0.9996	0.40	179	97	177
20 Å	12	6	6	676	12.48	6.24	6.24	1.0000	0.26	179	97	175
10 Å	1	1	1	74	1.39	0.69	0.69	1.0000	0.01	6	4	2

Figures

Figure 1: *a)* Schematic representation of the different protein complexes and the mechanisms involved in the photosynthetic process. In order, from left to right we find: OEC (Oxygen Evolving Complex), Photosystem II (PSII), Plastoquinone pool (PQ), b_6f cytochrome (b_6f), Plastocyanin (PC), Photosystem I (PSI), Ferredoxin (Fd), Ferredoxin NADP(+) Reductase (FNR) and ATP-synthase. *b)* Representation of the PSI network nodes embedded in the 3D protein structure. Nodes color indicates the different PSI network chromophores as in the right figure scheme (b). The arrows show the links between LHCs and other Chls with their directionality. *c)* Schematic representation of the different nodes of the PSI network and their connectivity. Car (orange) can only be starting nodes, thus they have only outgoing edges. We stress that every node can be the starting point of EET, but Car can only have outgoing links because of their higher site energy. Core Chls (green) can have both outgoing and ingoing edges. On the other hand, Chls belonging to different LHCs (LHC1 in blue, LHC2 in magenta, LHC3 in cyan and LHC4 in purple) can only transfer energy within the same LHC or to Chls with lower energies. Finally, P700 (red) is the point of arrival of the FRET process and only possesses ingoing edges. The two P700 chlorophylls are not directly connected.

Figure 2: example network where nodes-chromophores perform EET toward the RC (a); the network in (a) after the removal two green nodes, where we can see the elongation of the EET paths and the consequent network functioning decrease; functioning efficiency measure (Eff) as a function of the fraction of nodes removed (q) for two hypothetical different attack strategies (c); the red strategy triggers a faster efficiency (Eff) decrease than the blue strategy and the network robustness area below the red curve is lower than the one below the blue curve. The robustness (R_{Eff}) of the efficiency computed for the two strategies (d). The bar robustness R_{Eff} corresponds to the area below the curve in the left chart (c); R_{Eff} value is normalized by the max value, e.g. by R_{Eff} of the blue strategy.

Figure 3: Network properties as a function of CD. average nodes degree $\langle k \rangle$, average nodes strength $\langle s \rangle$, largest connected component LCC , number of links L , network efficiency Eff , average link weight Ew , CN number of connected nodes to the P700.

Figure 4: The normalized LCC_{weak} size and the representation of the various LCC_{weak} of the PSI network found at different cut-off distances (CD) as seen in the underlying protein structure. The nodes are colored while they appear in the LCC according to the scheme colors used in Fig.1.

Figure 5: Node types vs nodes centrality feature for the PSI network features for no cut-off (CD=NO). Keys are: k degree of the nodes, k_{out} out-degree, k_{in} in-degree; BC binary directed betweenness centrality; BC_w binary directed betweenness centrality; S node strength, S_{in} ingoing node strength, S_{out} outgoing node strength; δ decrease in the directed network efficiency after single node removal.

Figure 6: Node centrality features for the different types of PSI network chromophores as a function of the CD. Keys are: k degree of the nodes, k_{out} out-degree, k_{in} in-degree; BC binary directed betweenness centrality; BC_w binary directed betweenness centrality; S node strength, S_{in} ingoing node strength, S_{out} outgoing node strength; δ decrease in the directed network efficiency after single node removal.

Figure 7: Number of links L among different types of nodes-chromophores as function of CD.

Figure 8: PSI network robustness (R_{eff}) under the random removal (Rand) and the intentional random nodes removal strategies for the different CD. The intentional random removals: top row Cars, half row LHCs and bottom row Core Chls.