

# Lawrence Berkeley National Laboratory

## Recent Work

### **Title**

Spectroscopic study of the light-harvesting protein C-phycoerythrin associated with colorless linker peptides

### **Permalink**

<https://escholarship.org/uc/item/4vf3568x>

### **Author**

Pizarro, Shelly A.

### **Publication Date**

2000-05-12



# ERNEST ORLANDO LAWRENCE BERKELEY NATIONAL LABORATORY

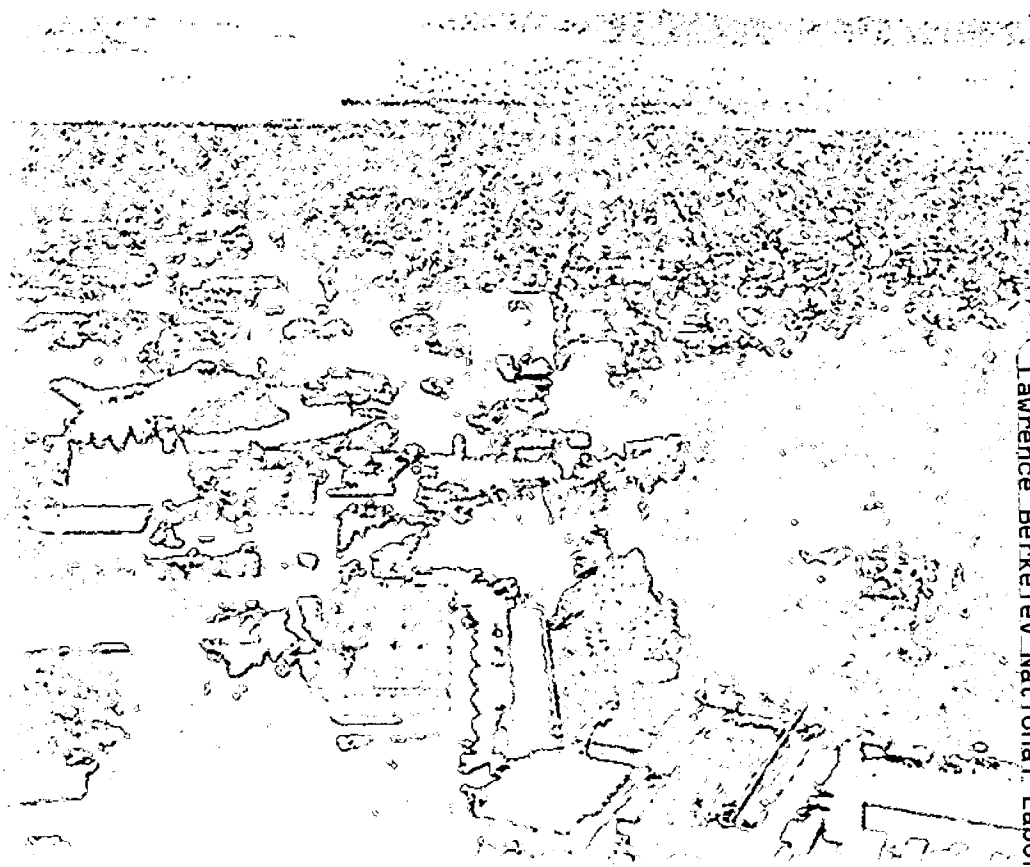
## Spectroscopic Study of the Light-Harvesting Protein C-Phycocyanin Associated with Colorless Linker Peptides

Shelly A. Pizarro

Physical Biosciences Division

May 2000

Ph.D. Thesis



Lawrence Berkeley National Laboratory

REFERENCE COPY  
Does Not  
Circulate

Bldg. 50 Library - Ref.

Copy 1

LBNL-45998

## **DISCLAIMER**

This document was prepared as an account of work sponsored by the United States Government. While this document is believed to contain correct information, neither the United States Government nor any agency thereof, nor the Regents of the University of California, nor any of their employees, makes any warranty, express or implied, or assumes any legal responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by its trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof, or the Regents of the University of California. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof or the Regents of the University of California.

**Spectroscopic Study of the Light-Harvesting Protein  
C-Phycocyanin Associated with Colorless Linker Peptides**

Shelly Ann Pizarro  
Ph.D. Thesis

Department of Chemistry  
University of California, Berkeley

and

Physical Biosciences Division  
Ernest Orlando Lawrence Berkeley National Laboratory  
University of California  
Berkeley, CA 94720

May 2000

**Spectroscopic Study of the Light-Harvesting Protein C-Phycocyanin  
Associated with Colorless Linker Peptides**

by

**Shelly Ann Pizarro**

B.A. University of Pennsylvania 1991

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

**GRADUATE DIVISION**

of the

**UNIVERSITY OF CALIFORNIA, BERKELEY**

Committee in charge:

Professor Kenneth Sauer, Chair

Professor Richard Mathies

Professor Alexander N. Glazer

Spring 2000

LBNL Report # 45998

## Abstract

### Spectroscopic Study of the Light-Harvesting Protein C-Phycocyanin

Associated with Colorless Linker Peptides

by

Shelly Ann Pizarro

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Kenneth Sauer, Chair

The phycobilisome (PBS) light-harvesting antenna is composed of chromophore-containing biliproteins and 'colorless' linker peptides and is structurally designed to support unidirectional transfer of excitation energy from the periphery of the PBS to its core. The linker peptides have a unique role in this transfer process by modulating the spectral properties of the associated biliprotein. There is only one three-dimensional structure of a biliprotein/linker complex available to date (APC/L<sub>C</sub><sup>7,8</sup>) and the mechanism of interaction between these two proteins remains unknown. This study brings together a detailed spectroscopic characterization of C-Phycocyanin (PC)-linker complexes (isolated from *Synechococcus* sp. PCC 7002) with proteomic analysis of the linker amino acid sequences to produce a model for biliprotein/linker interaction.

The amino acid sequences of the rod linkers [L<sub>R</sub><sup>8,9</sup>, L<sub>R</sub><sup>32,3</sup> and L<sub>RC</sub><sup>28,5</sup>] were examined to identify evolutionarily conserved regions important to either the structure or function of this protein family. Although there is not one common homologous site among all the linkers, there are strong trends across each separate subset (L<sub>C</sub>, L<sub>R</sub> and

$L_{RC}$ ) and the N-terminal segments of both  $L_R^{32.3}$  and  $L_{RC}^{28.5}$  display multiple regions of similarity with other linkers. Predictions of the secondary structure of  $L_R^{32.3}$  and  $L_{RC}^{28.5}$ , and comparison to the crystal structure of  $L_C^{7.8}$ , further narrowed the candidates for interaction sites with the PC chromophores.

Measurements of the absorption, fluorescence, CD and excitation anisotropy of PC trimer,  $PC/L_R^{32.3}$ , and  $PC/L_{RC}^{28.5}$ , document the spectroscopic effect of each linker peptide on the PC chromophores at a series of temperatures (298 to 77 K). Because  $L_R^{32.3}$  and  $L_{RC}^{28.5}$  modulate the PC trimer spectral properties in distinct manners, it suggests different chromophore-interaction mechanisms for each linker. The low temperature absorbance spectrum of the PC trimer is consistent with an excitonic coupling interaction between neighboring  $\alpha 84$  and  $\beta 84$  chromophores. Association with  $L_R^{32.3}$  does not greatly alter this band shape but the absorbance of the  $PC/L_{RC}^{28.5}$  complex is dramatically different. This indicates that  $L_{RC}^{28.5}$  is disrupting the  $\alpha 84$  -  $\beta 84$  relation established in the PC trimer. From these, and other polarized spectroscopy measurements, we conclude that both  $L_R^{32.3}$  and  $L_{RC}^{28.5}$  affect the spectral properties of the terminally emitting PC trimer chromophore ( $\beta 84$ ), and that  $L_{RC}^{28.5}$  is additionally perturbing the relationship between the  $\alpha 84$  and  $\beta 84$  chromophores to either disrupt or enhance their coupling interaction. The linker can perturb the PC chromophores through either specific aromatic residues or a concentration of electrostatically charged residues. Structurally, the linker disrupts the  $C_3$  symmetry of the associated biliprotein and this asymmetric interaction can serve to guide the transfer of excitation energy in one direction.

*Everyone has a without whom not,  
mine is surely Tom.*



## Table of Contents

<b>Chapter 1. Introduction.....</b>	<b>1</b>
Historical overview.....	1
PBS Morphology.....	2
C-Phycocyanin.....	4
Linker peptides.....	5
PBS Energetics.....	7
Project Description.....	11
References for Chapter 1.....	12
Figures for Chapter 1.....	18
<b>Chapter 2. Materials &amp; Methods .....</b>	<b>21</b>
Biological Preparations:.....	21
<i>Synechococcus</i> sp. PCC 7002 cultures.....	21
Isolation and purification of PC complexes .....	21
Spectroscopic Methods.....	23
Steady state measurements.....	23
Measurements using polarized light.....	25
Low temperature measurements.....	27
Analytical Methods.....	29
Primary structure analysis - amino acid sequences.....	30
Protein secondary structure predictions.....	33
RASMOL molecular viewer program.....	40
References for Chapter 2.....	42
Tables for Chapter 2 .....	44
Figures for Chapter 2.....	45
<b>Chapter 3. Analysis of linker homology and structural predictions.....</b>	<b>47</b>
Introduction.....	47
Results .....	48
Primary structure analysis.....	48
Secondary structure predictions .....	52
Discussion.....	53
Primary structure analysis.....	53
Secondary structure predictions .....	59
Tertiary Structure Analysis.....	63
Examination of the APC/L <sub>C</sub> <sup>7,8</sup> interface region.....	64
Comparison of APC and PC tertiary structures .....	68
Comparison of linker characteristics.....	71

Concluding remarks.....	74
Summary observations.....	77
References for Chapter 3.....	78
Tables for Chapter 3.....	80
Figures for Chapter 3.....	97
<b>Chapter 4. Steady State Spectroscopy.....</b>	<b>112</b>
Introduction .....	112
Results.....	113
Absorbance .....	113
Circular dichroism.....	115
Polarized excitation anisotropy.....	116
Fluorescence .....	117
Fluorescence emission at different excitation wavelengths.....	119
Fluorescence excitation at different emission wavelengths.....	120
Discussion.....	121
PC monomer.....	121
PC trimer.....	124
PC/linker complexes.....	130
Kennard-Stepanov analysis .....	137
Summary observations.....	149
References for Chapter 4.....	150
Tables for Chapter 4.....	155
Figures for Chapter 4.....	158
Appendix 4-1- Simulations of Absorbance Spectra.....	188
<b>Chapter 5. Concluding Remarks and Future Directions.....</b>	<b>202</b>
Proposed Model for Biliprotein/Linker Interaction.....	202
Future Directions.....	207
References for Chapter 5.....	210
Figure for Chapter 5.....	211

## Abbreviations & Symbols

$\lambda$	wavelength
$\lambda_{EM}$	emission wavelength
$\lambda_{EX}$	excitation wavelength
APC	Allophycocyanin
BLAST	Basic Local Alignment Search Tool
kDa	kiloDalton or molecular weight
$L_C^{7.8}$	core linker polypeptide with molecular weight of 7.8 kDa
$L_{RC}^{28.5}$	rod-core linker polypeptide with molecular weight of 28.5 kDa
$L_R^{8.9}$	rod linker polypeptide with molecular weight of 8.9 kDa
$L_R^{32.3}$	rod linker polypeptide with molecular weight of 32.3 kDa
nm	nanometer
PBS	phycobilisome
PC	C-Phycocyanin
pcb	phycocyanobilin chromophore
PE	Phycoerythrin
PEC	Phycoerythrocyanin
ps	picosecond
RT	room temperature
SDS-PAGE	sodium dodecyl sulfate - polyacrylamide gel electrophoresis

## Acknowledgements

This dissertation represents a collective effort, and I am happy to have an opportunity to acknowledge and thank all of the people who have made it possible as well as enjoyable. First and foremost, I am grateful to Ken Sauer for his ready acceptance, unfailing support, and ceaseless enthusiasm for science and learning. Along the twisting road of failed and successful experiments, he heroically provided a steady source of ideas and optimism.

Special thanks go to the students and postdocs I have known in the Sauer/Klein group during the past seven years. I am delighted to have been part of an evolving mix of cultures, opinions, and personalities. I am indebted to so many of these people for their help with technical difficulties (from big to small crises!), their encouragement during tough times, and their friendship. Henk Visser and Karen MacFarlane deserve special mention for supplying me with a constant supply of coffee breaks, encouragement, and inspiration during the long writing process.

The EXAFS part of the group has adopted me during the past two years and, although that work is not included in this dissertation, I am very grateful to have experienced the joys and tribulations of team research. In particular, I would like to thank Mel Klein and Vittal Yachandra for the invaluable scientific and practical perspective that they have provided on my XAS data as well as their timely intervention in my graduate career. Their firm reassurance and loyal support were crucial factors in motivating me to continue towards the finish line.

Specific aspects of the project were made possible by a number of individuals and I would like to acknowledge and thank them in the following section. The Calvin lab

contains a cooperative network of dedicated scientists, and I have greatly benefited from the advice and help that LiShar Huang and Hisao Yokota provided on SDS-PAGE and protein isolation (these techniques are both science and art). Don Bryant (Penn State) and the Glazer lab group (UC Berkeley MCB) kindly gave occasional advice on bacteria culture techniques, and even provided fresh cultures when my resuscitation efforts were unsuccessful. The Tinoco lab group (UC Berkeley Chemistry) generously allowed me free rein of their CD instrument. The folks at the LBNL instrument repair shop were an important source of aid with the nuances of electronics, and performed rapid repairs of broken equipment. Finally, none of this work would have been possible without Mary Talbot's tireless care, maintenance, and guidance with the AVIV and Spex spectrometers. I am grateful for the time, energy, and friendship she has shared with me.

In closing, I would like to thank the "nonscientists" in my life for not only putting up with me but also for helping me maintain my perspective and my sanity. I am lucky to be loved and supported by a wonderful network of family and friends; although they do not comprehend my work, they have nevertheless cheered me every step of the way. Lastly, I am profoundly grateful for my husband's love, patience and never ending faith in me – these have been the keystones to my success and happiness.

## Chapter 1. Introduction

Photosynthetic organisms use a system of macromolecular protein-pigment complexes to convert solar energy into chemical energy necessary for carbohydrate synthesis. The initial step in this process is the absorption of light and subsequent excitation transfer to the photosynthetic reaction center. Light-harvesting systems need to have a large cross-section of absorption wavelengths, adaptability to different light conditions, stability for prolonged illumination and the ability to rapidly transfer the collected energy with minimum dissipation. The cyanobacterial light-harvesting complexes, or phycobilisomes (PBS), meet these requirements admirably with a quantum efficiency >95% and energy transfer rates on a nanosecond to femtosecond timescale.

The PBS is a large ( $5-15 \times 10^6$  Da) assembly of phycobiliproteins and colorless peptides attached to the stromal side of the thylakoid membrane. The phycobiliproteins possess covalently bound chromophores and are particularly amenable to spectroscopic studies for a number of reasons. First, they contain a lower pigment/protein ratio than the light-harvesting systems of higher plants (1 bilin/110 amino acid residues versus 1/15 in higher plants)<sup>1</sup> which facilitates resolution of the spectral characteristics of individual chromophores. Secondly, they are highly water-soluble and can be readily isolated without disruption of their structural and functional properties.<sup>2</sup> Finally, a number of phycobiliprotein structures have been determined by x-ray crystallography<sup>3</sup> and this information permits structure-function relationships to be formed.

Extensive experimental work during the last two decades concerning spectral, biochemical and structural properties has resulted in a detailed picture of PBS architecture and energy transfer dynamics. Steady-state and time-resolved spectroscopic

techniques have been used to study the energy transfer characteristics of PBS and their individual phycobiliproteins. Except for some ultrafast (fs) components, the experimental results support a general model based on Förster's theory<sup>4,5</sup> of inductive-resonance energy transfer. The observed ultrafast elements are suspected to arise from excitonically coupled chromophores.<sup>6-9</sup>

This study focuses on the steady-state spectroscopy of the phycobiliprotein C-phycoyanin (PC) and associated linker peptides, isolated from the marine cyanobacterium *Synechococcus* sp. PCC 7002 also known as *Agmenellum quadruplicatum*. This chapter includes a review of PBS morphology and PBS energetics, then concludes with details of the project.

### **PBS morphology**

A number of excellent reviews<sup>2,10-13</sup> on phycobilisomes and phycobiliproteins available in the literature were used to assemble the following descriptive summary. PBS are large (5 - 15 x 10<sup>6</sup> Da) multiprotein organelles located on the stromal side of the thylakoid membrane. Although they primarily transfer excitation to the photosystem 2 complexes<sup>8</sup> imbedded in the thylakoid membrane, it has been shown that transfer to photosystem 1 is also possible.<sup>14,15</sup> The size and shape of PBS differs among species<sup>16</sup> and are dependent on environmental growth conditions such as light intensity,<sup>17,18</sup> wavelength,<sup>19-22</sup> and nutrient availability.<sup>23-27</sup> Each PBS is built from two main functional groups of polypeptides: (1) pigmented phycobiliproteins (henceforth known as 'biliproteins') and (2) 'colorless' linker polypeptides.<sup>28,29</sup> The biliproteins found in blue-green algae (cyanobacteria) are divided into four classes according to their maximal

absorption wavelength: allophycocyanins (APC,  $\lambda_{A \max} = 650-655$  nm), phycocyanin (PC,  $\lambda_{A \max} = 615-640$  nm), phycoerythrin (PE,  $\lambda_{A \max} = 565-575$  nm) and phycoerythrocyanin (PEC,  $\lambda_{A \max} = 575$  nm).

The most commonly found PBS shape is described as ‘hemidiscoidal’ with typical dimensions of 70 nm across the base, 30-50 nm in height and 14-17 nm width.<sup>28,30,31</sup> The PBS possesses two morphologically different domains: a “core” structure attached to the thylakoid membrane, and extending cylindrically shaped “rods” (Figure 1-1). The core contains different forms of the biliprotein APC while the rods contain PC, PE and, in some cases of filamentous cyanobacteria,<sup>32</sup> PEC at the most distal end. The core is composed of either two or three stacked cylindrical APC assemblies, and the number of rods attached to this core varies from six to ten depending on the organism. Despite these subtle morphological differences in the general phycobilisome structure, each of the cylindrical biliprotein ‘disks’ has the same dimensions. The core cylinders are composed of four connected disks of 11 nm diameter and 3.5 nm thickness while the rod cylindrical disks have a thickness of 6 nm.<sup>31</sup> The length of each rod varies from 12 to 36 nm according to organism and growth conditions.<sup>31,33</sup> Electron microscopy and x-ray crystallography of the rod disks have shown that each 6 nm-thick disk is actually an assembly of two face-to-face 3 nm disks which are the fundamental building blocks of the peripheral rods.<sup>34-37</sup> Each 3 nm disk is composed of three monomers containing two different subunits ( $\alpha$  and  $\beta$ ) arranged in a C3-symmetrical array around a central cavity ~3.5 nm in diameter. In most biliproteins, this central cavity is assumed to be occupied by linker peptides, but some cyanobacteria and red algae possess a PE hexamer with a third type of biliprotein subunit ( $\gamma$ ) in the central cavity.<sup>38-41</sup> The core



structure contains, in addition to APC and a small colorless linker peptide, two copies of a large (70 - 128 kDa) chromophore-containing linker labeled  $L_{CM}$  also known as the 'anchor protein'.<sup>42</sup>

The crystal structure of the first biliprotein was resolved in 1985.<sup>35,36</sup> Atomic resolution and refinement of the structure of PC<sup>43,44</sup> was quickly followed by other biliprotein structures: PEC,<sup>45,46</sup> PE,<sup>37,41,47</sup> APC<sup>48</sup> and most recently APC associated with its  $L_C$ <sup>7,8</sup> linker polypeptide.<sup>49</sup> The structures of these biliproteins reveal that they possess nearly isomorphous backbone structures despite the fact that different biliproteins have to accommodate different numbers of structurally distinct chromophores at various protein sites. The basic relationship between all biliproteins is also reflected in the amino acid sequences. Sequence identity is high across species for APC (>80% homology),<sup>50</sup> PC (>70%)<sup>51</sup> and between PEC and PC (~60%).<sup>52,53</sup> The spatial arrangement of the multiple  $\alpha$ -helices in biliprotein structures along with their homologous chromophore binding sites has suggested a phylogenetic relationship between biliproteins and myoglobin.<sup>35,50</sup>

### *C-Phycocyanin (PC)*

Like all other biliproteins, the PC trimer aggregate is water soluble and has a toroidal shape (Figure 1-2). Each 34 kDa monomer has 18  $\alpha$ -helices connected by short loop segments,<sup>43,44</sup> and a number of amino acid residues at the chromophore and subunit binding sites are highly conserved across species.<sup>3,44</sup> The brilliant colors of the biliproteins originate from covalently bound, linear tetrapyrrole chromophores. These chromophores can assume several conformations in solution due to variable configurations at the methine and methylene bridges between the pyrroles. In *Synechococcus* sp. PCC 7002, each PC monomer contains three phycocyanobilin chromophores attached to the protein via a single

thioether linkage to a cysteine residue. The chromophores are named according to monomer subunit and attached residue position as follows:  $\alpha 84$ ,  $\beta 84$  and  $\beta 155$ . The  $\alpha 84$  and  $\beta 84$  chromophores are attached to cysteine residues located in  $\alpha$ -helices close to the central cavity while the  $\beta 155$  chromophore is attached to a cysteine located in one of the loop segments between helices on the periphery of the protein (Figure 1-2). The  $\beta 84$  chromophore is located closest to the central cavity of the PC trimer, where it remains partially exposed to the solvent. The shortest interchromophore distance in the PC trimer is found between  $\alpha 84$  and the  $\beta 84$  from a neighboring monomer (20.8 Å). All of the chromophores are maintained in an extended conformation through interactions with the protein<sup>44,54</sup> and their spectral properties differ according to conformation as well as the surrounding environment. Both  $\alpha 84$  and  $\beta 84$  binding pockets are well-conserved across species and biliproteins,<sup>37</sup> especially aspartate residues that provide anchoring stability for rings B and C in all phycobilins.<sup>55</sup> In addition, the structure of the  $\beta 84$  chromophore itself also appears to be highly conserved.<sup>12</sup> Figure 1-3 presents the native conformations of each PC chromophore type and it shows that, although the geometric conformation of the pyrrole rings is very similar among all three, there is great variability in the position of the propionate side chains. The interaction of these propionate groups with nearby charged amino acids in the binding pocket has been speculated to play a major role in determining the spectroscopic properties of the chromophore.<sup>56,57</sup>

### *Linker peptides*

Each PBS requires five to nine different unpigmented or 'colorless' peptides for complete assembly, and these peptides account for 10-20% of the total PBS mass. It is

assumed that these peptides are localized in the central cavity of the biliprotein hexamers and that they connect one hexamer with its neighbor. Because of this function in joining together biliproteins, they were denoted as 'linker peptides'<sup>58,59</sup> and Glazer proposed a systematic nomenclature<sup>10</sup> to differentiate them according to molecular weight and PBS location. For example,  $L_C^{7.8}$  refers to a linker of 7.8 kDa mass located in the core (C) domain. Linker polypeptides can be divided into four groups according to PBS location: (1)  $L_R$  polypeptides (8-35 kDa) are involved in the rod assembly, (2)  $L_C$  linkers (8-11 kDa) are found in the core structure, (3)  $L_{RC}$  linkers (25-35 kDa) mediate the rod-core attachment, and (4)  $L_{CM}$  polypeptides (75-120 kDa), which contain one chromophore, are involved in core assembly and attachment to the thylakoid membrane. There is a significant amount of sequence homology among these linker peptides not only with each other (discussed in Chapter 3) but also with the biliproteins (30-40%).<sup>50</sup>

The protein surface of linker polypeptides does not exhibit a hydration envelope typical of globular proteins,<sup>13</sup> and the high tendency of linkers to aggregate demonstrates hydrophobic behavior. Linker peptides are expected to be positively charged at physiological pH since their calculated isoelectric points are typically greater than pH 9. In contrast, the biliproteins are extremely water-soluble and carry significant negative charge at physiological pH. These observations suggest that the association between linker polypeptides and biliproteins is driven by a combination of hydrophobic and charge-charge interactions.<sup>13</sup> The only available crystal structure for a biliprotein/linker complex, APC/ $L_C^{7.8}$  from *M. lamosus*, shows that the  $L_C^{7.8}$  linker interacts with two monomers of the trimeric APC complex and binds via multiple charged, polar, and hydrophobic contacts to the biliprotein chains.<sup>49</sup>  $L_C^{7.8}$  interacts with two of the three  $\beta$ 84

chromophores located in the APC central cavity by different methods. In one case, it induces a change in the chromophore's physical conformation while in the second case, it alters the surrounding protein environment with a wealth of positively charged arginine residues.

Besides providing structural support, the linkers also modulate energy transfer interactions within and across each domain. Linker association typically produces long-wavelength shifts of varying magnitudes in the absorption and fluorescence maximum positions of the associated biliprotein.<sup>2,12,60-62</sup> The location of linkers within the central cavity of the torus-shaped biliprotein hexamers or trimers can produce strong chromophore/linker interactions with the central biliprotein chromophores. The resulting arrangement of short- and long-wavelength absorbing biliprotein/linker complexes support a model of unidirectional transfer of excitation energy from the periphery of the PBS to the core.

### **PBS energetics**

Each PBS contains 300-800 covalently bound chromophores<sup>50</sup> which absorb light in the range of 450-665 nm and extend photosynthetic light-harvesting to spectral regions that are minimally active for chlorophylls. The PBS assembly has a quantum efficiency > 95% for transferred excitation energy emerging as APC fluorescence.<sup>5,34</sup> After PBS excitation, fluorescence arises from photosystem 2 in 150 ps if the rods contain PE and in 120 ps if the rods contain only PC.<sup>63</sup> Kinetic and spectroscopic studies on the energy transfer of individual PBS biliproteins showed that simple emission-reabsorption mechanisms are not able to explain such efficiency.<sup>12,64-66</sup>

Energy transfer within the PBS was first demonstrated in 1973.<sup>67</sup> Since then this phenomenon has been studied in both PBS and various individual biliprotein aggregates using a number of steady-state and time-resolved spectroscopic techniques. From these studies, a general picture of the energy transfer process has emerged. The arrangement of biliproteins absorbing at different wavelengths indicates directional energy transfer and early investigations (summarized in reference 68) resulted in a scheme of stepwise 'downhill' energy transfer from the outer parts of the rod (PE → PC), to the core (APC), and then to photosystem 2. The unidirectional nature of this energy transfer between heterogeneous components of the PBS is a consequence of the energy difference of absorption between PE (PEC), PC, APC and the modulation of the spectroscopic properties of these biliproteins by the different linker peptides.<sup>12</sup>

Within a single trimer disk, the principle of sensitizing (s) and fluorescing (f) chromophores was developed. Time-resolved fluorescence measurements show that the excitation energy absorbed by the chromophores at the periphery of the disc is rapidly localized on the centrally located bilins.<sup>69-71</sup> In PC, the  $\beta$ 155 and  $\alpha$ 84 chromophores are considered to be s-chromophores while  $\beta$ 84 is the f-chromophore.<sup>12,72,73</sup> To maintain an efficient system of energy transfer, the transfer time must be faster than the fluorescence lifetime of the chromophores. Typical fluorescence lifetimes for these chromophores are 3-10 ns, while transfer times have been determined to be 8-20 ps inside hexamers and trimers, 24-120 ps between rod disks,<sup>50,72</sup> and ~150 ps for transfer from APC to the APC-B/L<sub>CM</sub> terminal emitters.<sup>74</sup> Detailed calculations became possible after elucidation of various biliprotein crystal structures. The strongest interaction and fastest energy transfer is expected to occur between chromophores at the shortest distances to each other. As the

biliproteins are assembled into higher aggregates, the interchromophore distances with new neighbors are shorter and this provides new pathways for energy transfer. In monomeric PC, the shortest distance is between the two  $\beta$ -chromophores (34 Å), while in the trimer it occurs between the  $\alpha$ 84 and  $\beta$ 84 chromophores of adjacent subunits (20.8 Å).<sup>43</sup> Various research groups<sup>1,70,71,73,75</sup> have proposed that the fast and directed transfer exhibited in PBS biliproteins can be adequately explained using Förster's inductive resonance theory,<sup>4,5</sup> but the interpretation of some ultrafast kinetic components as excitonic interactions is under debate.

Based on the crystal structure of PC and on the resolved individual absorption spectra of the  $\alpha$ 84,  $\beta$ 84 and  $\beta$ 155 chromophores,<sup>72,75</sup> the individual transfer rates within PC monomers, trimers and hexamers were calculated.<sup>7</sup> A model was derived which consisted of a combination of exciton interaction between  $\alpha$ 84 and  $\beta$ 84 chromophores of neighboring subunits and Förster transfer steps. One of the predictions of this model was an ultrafast (330-370 fs) excited state decay in trimers and hexamers which should be absent in monomers. Such a fast interaction has been detected in PC trimers (~500 fs) and is correspondingly absent in PC monomers.<sup>66</sup> The results have been interpreted as rapid Förster energy transfer between neighboring  $\alpha$ 84 and  $\beta$ 84 chromophores which are close to each other only in trimer and hexamer PC aggregates. A similar ultrafast kinetic component has been observed in trimeric APC (~400 fs) but, unlike the PC component, it was not associated with a polarization decay component.<sup>76</sup> Recent femtosecond pump-probe anisotropy measurements on APC and PC trimers<sup>8,9</sup> exhibit multiexponential decays that provide significantly shorter time constants (APC = 10-30 fs and 280 fs; PC = 20-60 fs and ~700 fs). Because both systems display initial anisotropy values (APC =

0.58-0.7; PC = 0.47) that are higher than the theoretical maximum (0.4) for a system of uncoupled chromophores, the authors propose the existence of an exciton state for the  $\alpha 84/\beta 84$  chromophore pair. The shorter rate constant is then interpreted as exciton state relaxation, while the longer fs component can be due to either localization of the excitation on one of the chromophores or Förster energy transfer between the two chromophores. A clear-cut decision between relaxation of exciton states and Förster energy transfer for these ultrafast processes is impossible to make at this point.

The effect of associated linker peptides on the biliprotein kinetics is relatively unexplored. Steady-state absorbance and fluorescence measurements show that linker peptides affect the spectral properties of the associated biliprotein by inducing long-wavelength shifts of differing magnitude.<sup>2,12,60-62</sup> The absorption/fluorescence maximum position of APC is red-shifted  $\sim 3$  nm when associated with  $L_C$ <sup>7,8,77,78</sup>. The recently resolved crystal structure of APC/ $L_C$ <sup>7,8</sup> shows that the linker is interacting with only two of the three available APC monomers and that one of the  $\beta 81$  chromophores is significantly altered in conformation as a consequence of the associated linker.<sup>49</sup> Time-resolved fluorescence decay measurements of the APC/ $L_C$ <sup>7,8</sup> aggregate show that there is no appreciable difference between the resolved kinetic rates of the APC and APC/ $L_C$ <sup>7,8</sup> complex on a picosecond timescale.<sup>77,78</sup> Femtosecond measurements are needed to determine whether the presence of the linker is disruptive to the speculated  $\alpha 84/\beta 84$  excitonic coupling. Although the APC/ $L_C$ <sup>7,8</sup> aggregate is the only structural example available, it may not be prototypical of biliprotein/linker complexes found in the PBS rods. This linker is expected to be less essential to PBS function than the  $L_R$  or  $L_{RC}$

linkers because a mutant *Synechococcus* sp. PCC 7002 lacking  $L_C^{7,8}$  was able to grow (albeit more slowly than wild type) with no appreciable difference in PBS structure or function.

### **The Project**

This study focuses on the linker polypeptides  $L_R^{8,9}$ ,  $L_R^{32,3}$ , and  $L_{RC}^{28,5}$  found in the rod structure of cyanobacterium *Synechococcus* sp PCC 7002. Because these linker peptides modulate the channels of energy transfer within the PBS, biliprotein/linker complexes are interesting systems to study spectroscopically. The goal of this project is to elucidate the critical components of the linker peptide responsible for modulation of the spectral shifts observed in the PC/linker complexes. Measurements of the absorption, fluorescence and anisotropy of PC/linker complexes document the spectroscopic effect of the linker peptides on the PC chromophores at different temperatures (Chapter 4). These spectroscopic measurements, combined with a detailed analysis of the amino acid sequences of each linker peptide (Chapter 3), are used to produce a model of biliprotein/linker interaction presented in Chapter 5.



## References for Chapter 1

1. Debreczeny, M. P. (1994) Ph. D. Thesis, LBL-35672, University of California, Berkeley, CA.
2. Glazer, A.N. (1988) In Methods in Enzymology, eds. Packer, L. and Glazer, A.N., Academic Press, New York, NY, Vol 167, pp 291-312.
3. Betz, M. (1997) *Biol. Chem.*, 378:167-176.
4. Förster, T. (1965) In Modern Quantum Chemistry: Part III, ed. Sinanoglu, O., Academic Press, New York, NY, pp 93-137.
5. Förster, T. (1967) In Comprehensive Biochemistry, eds. Florkin, M. and Stotz, E.H., Elsevier, Amsterdam, Vol 22, pp 93-137.
6. Holzwarth, A.R., Bittersmann, E., Reuter, W., Wehrmeyer, W. (1990) *Biophys. J.* 57:133-145.
7. Sauer, K., Scheer, H. (1988) *Biochim. Biophys. Acta* 936:157-170.
8. Riter, R.E., Edington, M.D., Beck, W.F. (1997) *J. Phys. Chem.* 101:2366-2371.
9. Edington, M.D., Riter, R.E., Beck, W.F. (1995) *J. Phys. Chem.* 99:15699-15704.
10. Glazer, A.N. (1985) *Ann. Rev. Biophys. Chem.* 14:47-77.
11. MacColl, R., Guard-Friar, D. (1987) Phycobiliproteins, CRC Press, Boca Raton, FL.
12. Glazer, A.N. (1989) *J. Biol. Chem.* 264:1-4.
13. Sidler, W.A. (1994) In The Molecular Biology of Cyanobacteria, ed. Bryant, D.A., Kluwer Academic Publishers, The Netherlands, pp 139-216.
14. Mullineaux, C.W. (1992) *Biochim. Biophys. Acta* 1100:285-292.
15. Glazer, A.N., Gindt, Y.M., Chan, C.F., Sauer, K. (1994) *Photosynthesis Research* 40:167-173.

16. Gantt, E. (1980) *Inter. Rev. Cytol.* 66:45-80.
17. Öquist, G. (1974) *Plant Physiol.* 30:38-40.
18. Lönneborg, A., Lind, L.K., Kalla, S.R., Gustafsson, P., Öquist, G. (1985) *Plant Physiol.* 78:110-114.
19. Ohki, K., Fujita, Y. (1992) *J. Phycol.* 28:803-808.
20. Bogorad, L. (1975) *Annu. Rev. Plant Physiol.* 26:369-401.
21. Tandeau de Marsac, N. (1983) *Bull. Inst. Pasteur* 81:201-254.
22. Grossman, A.R. (1990) *Plant Cell Environ.* 13:651-666.
23. Foulds, I.J., Carr, N.G. (1977) *FEMS Microbiol. Lett.* 2:117-119.
24. Wood, N.B., Haselkorn, R. (1980) *J. Bacteriol.* 141:1375-1385.
25. Yamanaka, G., Glazer, A.N. (1980) *Arch. Microbiol.* 124:39-47.
26. Allen, M.M. (1984) *Annu. Rev. Microbiol.* 38:1-25.
27. Collier, J.L., Grossman A.R. (1992) *J. Bacteriol.* 174:4718-4726.
28. Glazer, A.N. (1984) *Biochim. Biophys. Acta* 768:29-51.
29. Tandeau de Marsac, N., Cohen-Bazire, G. (1977) *Proc. Natl. Acad. Sci. USA* 74:1635-1639.
30. Mörschel, E., Wehrmeyer, W. (1977) *Arch. Microbiol.* 113:83-89.
31. Bryant, D.A., Guglielmi, G., Tandeau de Marsac, N., Castet, A.M., Cohen-Bazire, G. (1979) *Arch. Microbiol.* 123:113-127.
32. Bryant, D.A. (1982) *J. Gen. Microbiol.* 128:835-844.
33. Glazer, A.N., Williams, R.C., Yamanaka, G., Schachman, H.K. (1979) *Proc. Natl. Acad. Sci. USA* 76:6162-6166.
34. Bryant, D.A., Glazer, A.N., Eiserling, F.A. (1976) *Arch. Microbiol.* 110:61-75.

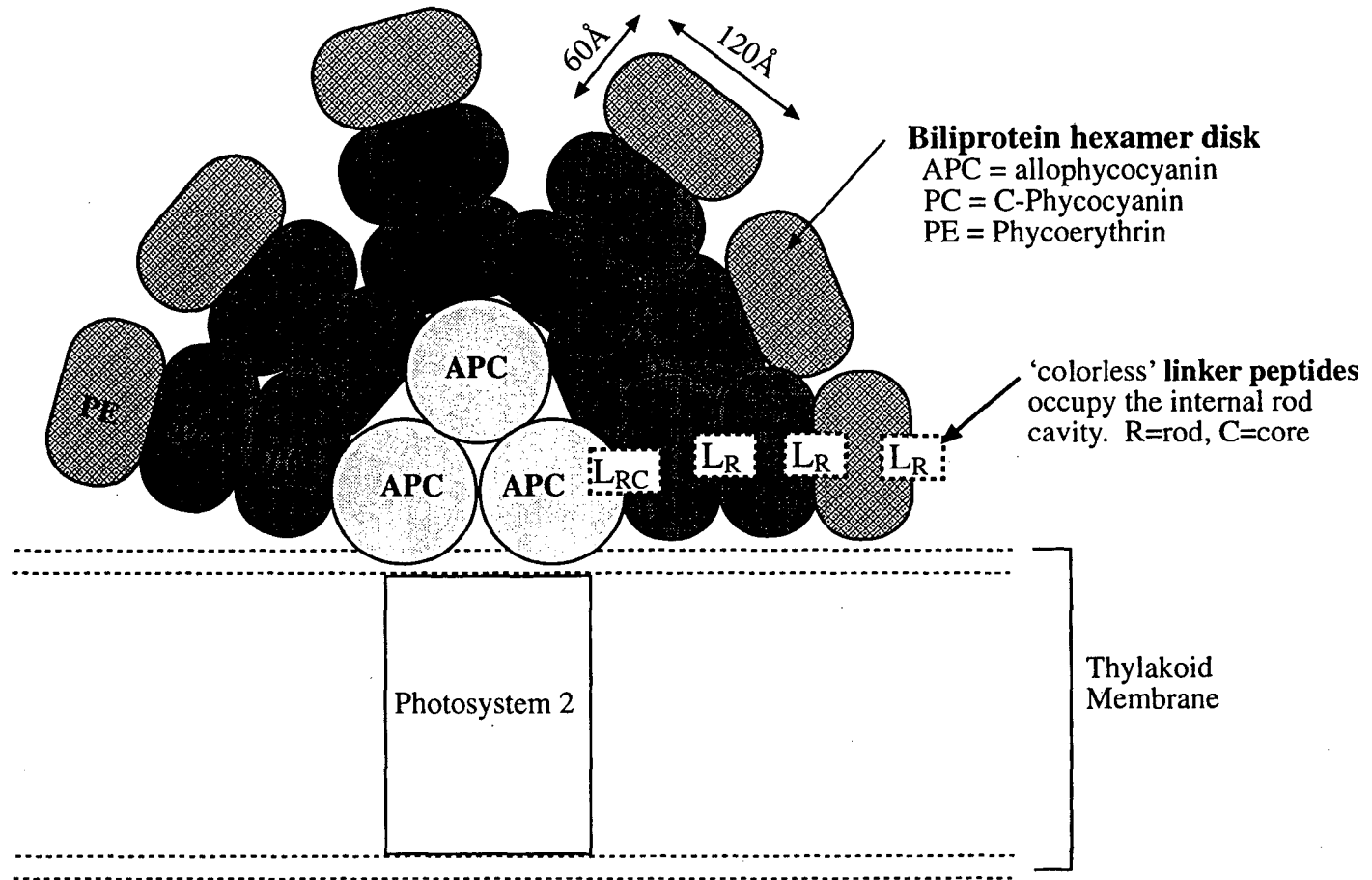
35. Schirmer, T., Bode, W., Huber, R., Sidler, W., Zuber, H. (1985) *J. Mol. Biol.* 184:257-277.
36. Schirmer, T., Huber, R., Schneider, M., Bode, W., Miller, M., Hacert, M.L. (1986) *J. Mol. Biol.* 188:651-676.
37. Ficner, R., Lobeck, K., Schmidt, G., Huber, R. (1992) *J. Mol. Biol.* 228:935-950.
38. Willbanks, S.M., Glazer, A.N. (1993) *J. Biol. Chem.* 268:1236-1241.
39. Koller, K.P., Wehrmeyer, W. (1977) *Arch. Microbiol.* 104:255-261.
40. Glazer, A.N., Hixson, C.S. (1977) *J. Biol. Chem.* 252:32-42.
41. Ficner, R., Huber, R. (1993) *Eur. J. Biochem.* 218:103-106.
42. Gantt, E. (1988) In Light-Energy Transduction in Photosynthesis: Higher Plants and Bacterial Models, eds. Stevens Jr., S.E., and Bryant, D.A., American Society of Plant Physiologists, Rockville, MD, pp 91-101.
43. Schirmer, T., Bode, W., Huber, R. (1987) *J. Mol. Biol.* 196:677-695.
44. Duerring, M., Schmidt, G.B., Huber, R. (1991) *J. Mol. Biol.* 217:577-592.
45. Rübéli, R., Schirmer, T., Bode, W., Sidler, W., Zuber, H. (1985) *J. Mol. Biol.* 186:197-200.
46. Duerring, M., Huber, R., Bode, W., Rübéli, R., Zuber, H. (1990) *J. Mol. Biol.* 211:633-644.
47. Chang, W., Jiang, T., Wan, Z., Zhang, J., Yang, Z., Liang, D. (1996) *J. Mol. Biol.* 262:721-731.
48. Brejc, K., Ficner, R., Huber, R., Steinbacher, S. (1995) *J. Mol. Biol.* 249:424-440.
49. Reuter, W., Wiegand, G., Huber, R., Than, M.E. (1999) *Proc. Nat. Acad. Sci. USA*, 96:1363-1368.

50. Zuber, H. (1987) In The Light Reactions, ed. Barber, J., Elsevier Biomedical, Amsterdam, pp 157-259.
51. de Lorimier, R., Wilbanks, S.M., Glazer, A.N. (1993) *Plant Molec. Biol.* 21:225-237.
52. Eberlein, M., Kufer, W. (1990) *Gene* 94:133-136.
53. Swanson, R.V., de Lorimier, R., Glazer, A.N. (1992) *J. Bacteriol.* 174:2640-2647.
54. Scheer, H. (1981) *Angew. Chem.* 93:230-250.
55. Rüdiger, W. (1994) In Progress in Phycological Research, eds. Round, F.E. and Chapman, D.J., Biopress Ltd., Bristol, Vol 10, pp 97-136.
56. Scharnagl, C., Schneider, S. (1989) *J. Photochem. Photobiol. B: Biol.* 3:603-614.
57. Scharnagl, C., Schneider, S. (1991) *J. Photochem. Photobiol. B: Biol.* 8:129-157.
58. Lundell, D.J., Williams, R.C., Glazer, A.N. (1981) *J. Biol. Chem.* 256:3580-3592.
59. Zilinskas, B.A., Howell, D.A. (1983) *Plant Physiol.* 71:379-387.
60. Glazer, A.N. (1987) In The Cyanobacteria, eds. Fay, P. and Van Baalen, C., Elsevier, Amsterdam, pp 69-88.
61. Rübéli, R., Zuber, H. (1988) In Photosynthetic Light-Harvesting Systems, eds. Scheer, H. and Schneider, S., Walter de Gruyter, Berlin, pp 61-70.
62. Glauser, M., Sidler, W., Zuber, H. (1993) *Photochem. Photobiol.* 57:344-351.
63. Yamazaki, I., Mimuro, M., Murao, T., Yamazaki, T., Yoshihara, K., Fujita, Y. (1984) *Photochem. Photobiol.* 39:233-240.
64. Scheer, H. (1986) In Encyclopedia of Plant Physiology: Photosynthesis III, eds. Staehlin, L.A. and Arntzen, C.J., Academic Press, New York, NY, Vol 19, pp 336-372.

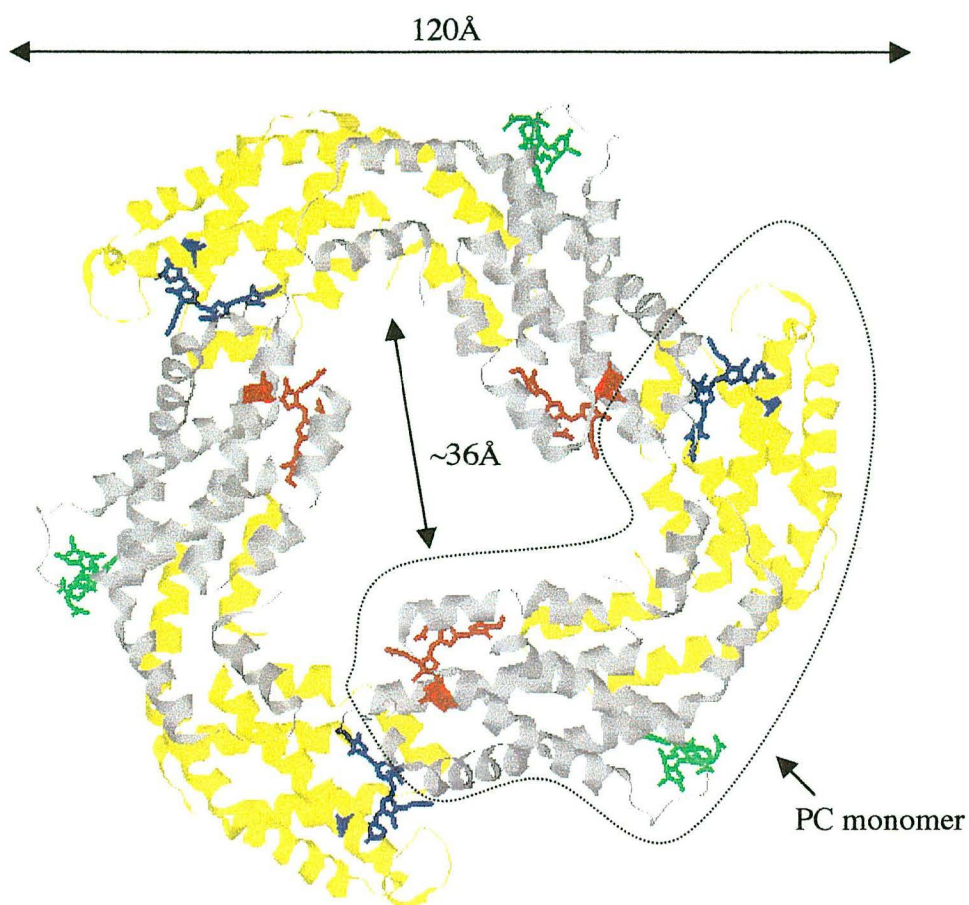
65. Wehrmeyer, W. (1990) In Experimental Phycology I, eds. Wiessner, W., Robinson, D.G. and Starr, R.C., Springer Verlag, Berlin, pp 158-172.
66. Gillbro, T., Sharkov, A.V., Kryukov, I.V., Khoroshilov, E.V., Kryukov, P.G., Fisher, R., Scheer, H. (1993) *Biochim. Biophys. Acta* 1140:321-326.
67. Gantt, E., Lipschultz, C.A. (1973) *Biochim. Biophys. Acta* 292:858-861.
68. Scheer, H. (1982) In Molecular Biology, Biochemistry and Biophysics, ed. Fong, F.K., Springer-Verlag, Berlin, Vol 35, pp 7-45.
69. Mimuro, M., Füglistaller, P., Rübli, R., Zuber, H. (1986) *Biochim. Biophys. Acta* 848:155-166.
70. Debreczeny, M.P., Sauer, K.H., Zhou, J., Bryant, D.A. (1995) *J. Phys. Chem.* 99:8412-8419; Debreczeny, M. P. (1994) Ph. D. Thesis, LBL-35672, University of California, Berkeley, CA.
71. Debreczeny, M.P., Sauer, K.H., Zhou, J., Bryant, D.A. (1995) *J. Phys. Chem.* 99:8420-8431; Debreczeny, M. P. (1994) Ph. D. Thesis, LBL-35672, University of California, Berkeley, CA.
72. Siebzehnrübl, S., Fischer, R., Scheer, H. (1987) *Z. Naturforsch.* 42c:258-262.
73. Mimuro, M., Rübli, R., Füglistaller, P., Zuber, H. (1986) *Biochim. Biophys. Acta* 851:447-456.
74. Feis, A., Friedrich, J., Gottschalk, L., Scheer, H. (1992) *J. Phys. Chem.* 96:6087-6089.
75. Fischer, R., Siebzehnrübl, S., Scheer, H. (1988) In Photosynthetic Light-Harvesting Systems, eds. Scheer, H. and Schneider, S., Walter de Gruyter, Berlin, pp 71-76.

76. Sharkov, A.V., Kryukov, I.V., Khoroshilov, E.V., Kyukov, P.G., Fischer, R., Scheer, H., Gillbro, T. (1992) *Chem. Phys Lett.* 191:633-638.
77. Maxson, P. (1988) Ph. D. Thesis, LBL-26163, University of California, Berkeley, CA, 45-76.
78. Sharkov, A.V., Kyukov, I.V., Khoroshilov, E.V., Kryukov, P.G., Fischer, R., Scheer, H., Gillbro, T. (1992) *Chem. Phys. Lett.* 191:633-638.

**Figure 1-1.** Schematic representation of a hemidiscoidal phycobilisome light-harvesting antenna



**Figure 1-2.** Schematic representation of PC trimer (adapted from reference 44)

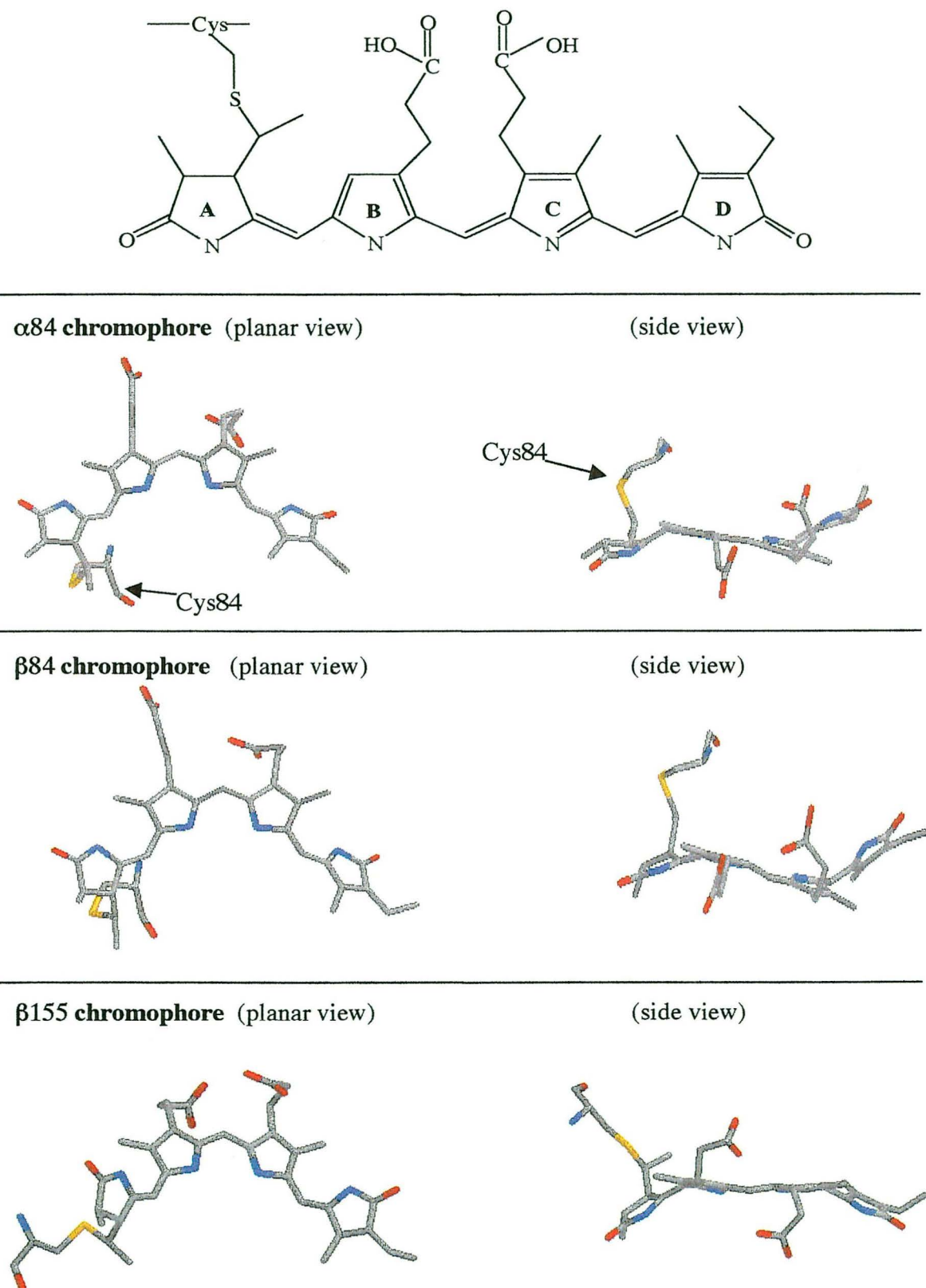


Legend	
<span style="display:inline-block; width:15px; height:10px; background-color:yellow; border:1px solid black;"></span>	$\alpha$ -subunit
<span style="display:inline-block; width:15px; height:10px; background-color:grey; border:1px solid black;"></span>	$\beta$ -subunit
<span style="display:inline-block; width:15px; height:10px; background-color:blue; border:1px solid black;"></span>	$\alpha$ 84 chromophore
<span style="display:inline-block; width:15px; height:10px; background-color:red; border:1px solid black;"></span>	$\beta$ 84 chromophore
<span style="display:inline-block; width:15px; height:10px; background-color:green; border:1px solid black;"></span>	$\beta$ 155 chromophore

Chromophore distances	
$\alpha$ 84 – $\beta$ 84*	~20Å
$\beta$ 84 – $\beta$ 155	34Å
$\alpha$ 84 – $\beta$ 155	50Å



**Figure 1-3.** Chemical structure representations of phycocyanobilin chromophores bound by thioether link to PC cysteine residue at tetrapyrrole ring A



## Chapter 2. Materials & Methods

### Biological Preparations

#### *Synechococcus* sp. PCC 7002 cultures

The cyanobacterium *Synechococcus* sp. PCC 7002 was grown at room temperature in 1 L aliquots of Medium A<sup>+</sup> (Table 2-1) supplemented with vitamin B-12. The cultures were constantly stirred and maintained in an atmosphere of [95% N<sub>2</sub>/5% CO<sub>2</sub>] under white fluorescent light.<sup>1</sup> A typical batch of wild type strain took 7-10 days to reach confluency. It was then harvested by centrifugation in a Sorvall GS-3 rotor (Dupont Instruments) at 3500 x g for 10 min, resuspended in a minimal amount of 5 mM Na phosphate pH 7 buffer and stored frozen.

#### *Isolation and purification of PC complexes*

All buffers contained 1 mM sodium azide as a preservative, and all procedures were conducted at 4 °C unless noted otherwise. Isolation of PC complexes from *Synechococcus* sp. PCC 7002 was performed using either previously harvested cells that had been stored in the freezer or freshly harvested batches. The procedure described below was adapted from that developed by Yu *et al.*<sup>2</sup> A French press was used (~22 GPa) to break apart the cell membranes. The water soluble phycobilisomes were separated from membrane fragments and other precipitating pigments using a series of successive centrifugation spins on the resulting supernatant: 35,000 x g (Sorvall SS-34 rotor) for 30 min, 300,000 x g (Beckman Ti-60 rotor) for 1 hour.

To separate the PC and APC, the blue supernatant was dialyzed in 1 L of 1 mM K phosphate/100 mM NaCl pH 7 buffer solution for 4 hours then applied to a Biogel HTP

(Biorad) column (2.5 x 10 cm) previously equilibrated with the same buffer. PC was eluted with 50 mL of 35 mM K phosphate/100 mM NaCl pH 7 buffer solution. The majority of the APC remained bound to the column until "cleaned" with 50 mL of 100 mM K phosphate/100 mM NaCl pH 7 buffer solution and subsequently discarded. The absorbance of the elution fractions was tested, and the fractions containing PC were pooled and dialyzed versus 5 mM K phosphate pH 7 buffer solution overnight with one buffer change after 2-4 hours.

The PC solution was applied to a DEAE cellulose (DE-52) (Whatman) column (5 x 5 cm) previously equilibrated with the same buffer as the dialyzed PC, and separation of the PC aggregates was achieved using a linear gradient of 5 mM to 125 mM K phosphate pH 7 buffer solutions (total volume = 400 mL). PC/linker complexes were found in the earlier eluting fractions (~ 40 mM K phosphate pH 7) as determined by SDS-PAGE and absorbance measurements and were pooled separately from the other PC fractions. The two PC pools were separately concentrated using an Amicon ultrafiltration system with PM30 membranes, dialyzed with 50 mM K phosphate pH 7 buffer solution for at least 4 hours, then stored at 4 °C.

Linear sucrose gradients (5%-15% in 50 mM K phosphate pH 7, total volume = 5 mL) prepared with high purity sucrose (Aldrich) were used to separate the PC/linker complexes, PC/L<sub>R</sub><sup>32.3</sup> and PC/L<sub>RC</sub><sup>28.5</sup>. About 100 µL of the concentrated PC/linker solution was loaded onto each gradient, then spun at 200,000 x g (Beckman SW 50.1 swinging bucket rotor) for 14 hours at 20 °C. The two resulting blue layers were carefully extracted with a syringe and used in spectral measurements without further purification. SDS-PAGE and absorbance measurements revealed that the bottom layer

contains predominantly the PC/L<sub>RC</sub><sup>28.5</sup> complex, and the top layer, the PC/L<sub>R</sub><sup>32.3</sup> complex.

## **Spectroscopic Methods**

### *Steady state measurements*

All measurements were conducted in 4 mL disposable clear polypropylene cuvettes with 1 cm path length. Absorption traces were gathered on an AVIV 14DS UV-VIS-NIR Spectrophotometer (AVIV, Lakewood, NJ) and were corrected with the corresponding buffer “blank” subtracted as baseline. Samples were diluted to a maximum absorbance of 0.1 in the visible region, and then used for fluorescence measurements. Excitation and emission spectra were gathered using a Spex Fluorolog spectrophotometer (Spex ISA, Edison, NJ) with detection at 90° from excitation and using a reference photodiode in the excitation monochromator compartment to correct for variations in lamp intensity. Other corrections were made to compensate for the photomultiplier tube detector wavelength dependent sensitivity and, in the case of polarization data, for monochromator bias towards light of parallel orientation. Wavelength calibration of the emission monochromator was performed using a standard sodium lamp in the sample compartment. A systematic mechanical error in the monochromator motor leads to increasingly bigger discrepancies in wavelength position traveling in either direction from the calibrated sodium line. Corrections due to this mechanical error are very important for the Kennard-Stepanov analysis and they were performed using an interpolation program written by M. Talbot. It accepts the experimental data, which contains fluorescence intensity values spaced at regular

wavelength intervals (1.0 nm), and compresses it to the true wavelength values determined from the scan of a sodium standard lamp. In performing this compression, the program interpolates the measured intensities resulting in minor smoothing of the data. Over the region of interest in these experiments (450-800 nm), the worst deviation from the correct wavelength is  $\leq 0.3$  nm.

Samples were excited with approximately 4 nm bandwidth of light, and the emission bandwidth was typically  $< 2$  nm. Scanning step size was 1 nm and dwell time was set at 1-5 seconds per step depending on protein concentration and experimental procedure. For example, the polarized excitation traces benefited greatly from longer averaging time at each step.

Samples were prepared in the appropriate buffer solutions (shown below) with maximal absorbance 0.1. Emission scans of each PC complex were first taken with the excitation wavelengths set at the maximum absorbance for each complex as follows:

PC monomer	1 M KSCN/5 mM K phosphate pH 7	$\lambda_{EX}=616$ nm
PC trimer	50 mM K phosphate pH 7	$\lambda_{EX}=622$ nm
PC/L <sub>RC</sub> <sup>28.5</sup>	50 mM K phosphate pH 7	$\lambda_{EX}=637$ nm
PC/L <sub>R</sub> <sup>32.3</sup>	50 mM K phosphate pH 7	$\lambda_{EX}=624$ nm.

Each of these emission scans was then repeated with the excitation set at 585 nm to eliminate the scattering from the excitation lamp in presenting the spectra. The emission bands for each complex were identical at these two excitation wavelengths. Excitation spectra were taken with the emission wavelength set at 650-660 nm.

### *Measurements using polarized light*

Circular dichroism (CD) measurements for each PC complex (prepared in the appropriate buffer as described above) were made using a Model J600 Spectropolarimeter (Japan Spectroscopic Co.). This instrument is equipped with a 450 W Xenon arc lamp as a light source, double monochromators for precise wavelength selection, two crystal prisms to produce linearly polarized light, a CD modulator to convert the output from the prisms into circularly polarized light, and finally a photomultiplier tube to detect the output signal. Data were gathered with the following parameters: 1 nm bandwidth, 20 mdeg sensitivity, time constant of 2 seconds at each 0.2 nm step. A total of 8-10 scans were averaged for each sample. CD scans of the respective solution buffers were separately acquired and subtracted from each sample's CD spectrum. In addition, the spectra were corrected against an average baseline profile of the empty cuvette taken before and after each scan. The instrument software automatically calculates the molecular ellipticity  $[\theta]$  or CD magnitude, experimentally determined as:

$$\begin{aligned} [\theta] &= 3300\Delta\epsilon \\ \Delta\epsilon &= (\epsilon_L - \epsilon_R) \\ &= \frac{1}{LC} \log\left(\frac{I_R}{I_L}\right) \end{aligned}$$

where  $L$  is the cell length,  $C$  is the sample concentration expressed in moles/liter of protein residues,  $\epsilon$  is the extinction coefficient and  $I_R$  and  $I_L$  correspond to the right- and left-circularly polarized light intensities detected by the photomultiplier tube.

Polarized excitation fluorescence measurements of each PC complex were made on the Spex Fluorolog described above equipped with polarizing prisms installed in the

excitation and emission ports of the sample compartment as shown in Figure 2-1.  $I_{\text{parallel}}$  denotes the emission intensity when the two prisms are oriented along the same vertical axis and  $I_{\text{perpendicular}}$  represents intensity measured with a  $90^\circ$  difference in the axial orientation of the prisms. Experimentally, this translates to four possible axial combinations of the excitation and emission port polarizing prisms: VV, HH, VH, and HV where V represents a vertical orientation and H is a horizontal orientation (Figure 2-1). A homogenized scattering solution made with non-dairy creamer at a very dilute concentration was used to calibrate the prism positions at maximal (VV) and minimal intensity (HV). As an excitation scan was taken, the positions of the polarizers were moved automatically through the four different possible configurations at each wavelength position (Figure 2-1). The end result is four separate excitation scans for a single sample, each one taken at a different polarizer configuration (VV, HH, HV, VH). To improve signal-to-noise at each polarizer configuration, five to eight scans were averaged before calculating the anisotropy. For PC monomers, PC trimers, and PC/L<sub>R</sub><sup>32.3</sup> the emission wavelength used in these polarized excitation scans was  $\lambda_{\text{EM}} = 647 \text{ nm}$  while for PC/L<sub>RC</sub><sup>28.5</sup>,  $\lambda_{\text{EM}} = 654 \text{ nm}$ . The anisotropy is typically calculated as:

$$r = \frac{I_{\text{parallel}} - I_{\text{perpendicular}}}{I_{\text{parallel}} + 2I_{\text{perpendicular}}}$$

however, a correction must be applied to the  $I_{\text{parallel}}$  and  $I_{\text{perpendicular}}$  terms due to the polarization bias of the monochromator gratings favoring light in a parallel orientation.<sup>3,4</sup> This instrument bias is characterized by the ratio  $N = (\text{HH}/\text{HV})$  which is dependent on wavelength.<sup>4</sup> Consequently  $N$  is used to lower the measured VV intensity to its true  $I_{\text{parallel}}$  value and the fluorescence anisotropy is experimentally determined as:

$$r = \frac{N(VV) - (VH)}{N(VV) + 2(VH)}$$

### *Low temperature measurements*

Low temperature measurements of the absorbance and fluorescence were taken using an Oxford liquid nitrogen cryostat (Model Optistat DN, Oxford Instruments, Bedford, MA) pumped to at least  $10^{-4}$  Torr with a turbopump (Model V-70, Varian vacuum products, Lexington, MA). The typical vacuum obtained at room temperature was  $10^{-6}$  Torr. A schematic diagram of the cryostat is shown in Figure 2-2. Disposable polystyrene cuvettes (1 cm square) were placed in a specialized sample holder whose temperature was controlled by heat exchange with the cryostat heater assembly or gaseous flow of liquid nitrogen stored in a reservoir. The heater was driven by a controller unit that could be set manually or through a computer interface. To obtain a more accurate measure of temperature at the sample, a T-type thermocouple connected to an Autotune Temperature Controller unit (Model CN132, Omega Engineering Inc., Stamford, CT) was inserted directly into the sample cuvette ~3 mm from the sample (but not in it, Figure 2-2). The cryostat was equipped with quartz port windows sealed with indium wire that was replaced after ~10 days of operation or when a leak was detected.

Samples were prepared at low concentrations (~25  $\mu\text{g/mL}$ , absorbance maximum 0.1) in the following buffers and were placed under house vacuum for 20-30 min before inserting into the cryostat to eliminate bubbles that would affect the clear glass formed at low temperature. PC monomers were prepared in a 75% glycerol/5 mM K phosphate pH 7 buffer while all the other PC complexes were prepared in 3 M sucrose (Sigma, > 99.5%



purity). Room temperature absorption spectra were routinely taken before and after freezing to verify that the freezing process had not damaged the proteins.

Once the desired temperature was reached by the cryostat heater controller, the thermocouple located at the sample cuvette showed that 10- 30 min was needed for the sample to equilibrate to the same temperature with the larger delays occurring at the lower temperatures. The lowest measured temperature at the sample cuvette was 84 K even after 3 hours of equilibration time was allowed. The built-in temperature detector in the Oxford cryostat is actually located 2-3 cm above the sample cuvette (Figure 2-2) and it registered 77 K; the poor conduction of this temperature to the sample compartment is a flaw of the instrument design. Fluctuations in the sample temperature were in the order of  $\pm 0.3$  K during the scan. Absorbance traces of each buffer were taken separately at each temperature and then were subtracted from those of the protein samples for baseline correction. There was no cuvette in the reference compartment of the absorbance spectrometer.

Before initiating the cooling process, the position of the cryostat and thus the sample holder was adjusted to the maximum transmittance observed at 700 nm. Typically, the cryostat alone allowed ~60% transmittance at 700 nm, and once a blank buffer sample was introduced this decreased to ~50% transmittance. When ambient humidity was present in the sample space, water droplets deposited on the inside windows as the sample was cooled. This produced a fair amount of scattering. To minimize this humidity, pump/purge cycles of dry nitrogen using a small mechanical vacuum pump were performed after each sample insertion and at times (during the rainy season) the entire cryostat would be enveloped in a disposable glove bag filled with

house nitrogen. The use of this glove bag was particularly advantageous during sample exchanges at low temperatures to prevent moist air from condensing inside the cool sample space. The AVIV sample compartment and that of the fluorimeter were also continuously purged with dry nitrogen.

For low temperature fluorescence measurements, the Spex fluorimeter sample cavity was removed to accommodate the Oxford cryostat and homemade holder (made by H. Visser). The apparatus was covered with a black cloth and the mirrors were set for right angle detection. The emission bandwidth was  $< 2.0$  nm and a red filter (Kodak wratten filter #26) was used in the emission pathway to minimize detection of first harmonic wavelengths. For excitation scans (450-700 nm) the emission monochromator was set at 655 nm, and for emission scans (590-800 nm) the excitation monochromator was set at  $\lambda_{\text{EX}} = 585$  nm. Red-edge excitation scans were taken at  $\lambda_{\text{EX}} = 660$  nm instead of 585 nm. The emission signal when  $\lambda_{\text{EX}} = 660$  nm was very weak and thus showed a higher amount of background noise; the monochromator excitation slits were opened to 2.0 nm to compensate.

### **Analytical methods**

All of the data analysis presented in this work was performed using commonly available software programs. The Kennard-Stepanov calculations and spectra presented in Chapter 4 were performed using either Excel 98 (Microsoft Corp., Redmond, WA) or Igor 2.6 (Wavemetrics, Lake Oswego, OR). The programs presented in Chapter 3 for the linker homology studies were accessed via the internet and are described below.

### *Primary Structure Analysis - Amino Acid Sequences*

The analytical software tools used for evaluating the primary and secondary structures of the linker peptides were available at no cost through the Expert Protein Analysis System (ExPASy) of the Swiss Institute of Bioinformatics:

[<http://expasy.hcuge.ch/www/tools.html>].

Interactive access was made using Netscape Navigator and the results (presented in Chapter 3) were received via electronic mail. The programs can be collectively accessed through the ExPASy site or directly at their own individual internet locations as noted throughout this description. The published amino acid sequences of the linker peptides (L<sub>R</sub><sup>8.9</sup>, L<sub>R</sub><sup>32.3</sup>, L<sub>RC</sub><sup>28.5</sup>) found in the phycobilisome rods of *Synechococcus* sp. PCC 7002 were subjected to all of the programs described below. The primary structure of the core linker L<sub>C</sub><sup>7.8</sup> from *Mastigocladus laminosus* was also included where appropriate to test method accuracy.

SWISS-PROT Sequence Database      [<http://expasy.hcuge.ch/cgi-bin/sprot-search-de>]

The amino acid sequences used were obtained from the SWISS-PROT sequence database maintained by the European Bioinformatics Institute (EBI). The database contains information from many sources, such as the European Molecular Biology Laboratory (EMBL), the Nucleotide sequence database, the Brookhaven Protein Data Bank (PDB), the NIH Protein Information Resource Center (PIR) and is cross-referenced to other gene, protein and nucleotide information data banks. For simplicity in repeatedly accessing data, the accession (AC) identification numbers for all *Synechococcus* sp. PCC 7002 proteins analyzed are as follows:

<u>Protein</u>	<u>Gene</u>	<u>AC Number</u>
PC $\alpha$ subunit	cpcA	P03943
PC $\beta$ subunit	cpcB	P03944
L <sub>RC</sub> <sup>28.5</sup>	cpcG	Q05238
L <sub>R</sub> <sup>32.3</sup>	cpcC	Q05237
L <sub>R</sub> <sup>8.9</sup>	cpcD	P31966

*Search for primary structure homology with other proteins*

Two different programs were used to perform homology searches for each of the linker protein sequences against a non-redundant combination of independent databases: PDB (Brookhaven National Laboratory), PIR (Georgetown University), PRF, EMBL and SWISS-PROT.

The Fast-A3 program [<http://www.ebi.ac.uk/htbin/fasta.py?request>] is based at the EMBL and uses an algorithm developed by Pearson and Lipman at the University of Virginia<sup>5,6</sup> to search for similarities between one sequence (the query) and any group of sequences. This method first identifies the ten best regions of similarity between the query sequence and each sequence from the search set and sets up a matrix. These "best" regions are then rescored using a scoring matrix that allows conservative replacements of amino acids, ambiguity symbols, and runs of shorter identities (less than ten regions). In the third step, the program checks to see if some of these initial highest-scoring diagonals can be joined together. Finally, the set of sequences with the highest scores is aligned to the query sequence for graphical display. The output has three forms, all organized by score from highest to lowest: (1) a histogram, along with standard deviations, showing

the overlapping regions of homology between the query sequence and the search results without specific identity of each amino acid, (2) a list of the matching result sequences identified by AC numbers and organism/protein information and (3) individual alignments of each matching result sequence with the query sequence.

For each amino acid sequence, the top 100 matches in the database were routinely requested to include homology matches to proteins that were not related to the family of cyanobacterial proteins. However such matches frequently occurred over small regions of very large proteins and were thus statistically insignificant. To avoid lengthy result outputs, Fast-A3 automatically displayed only the region of overlap between two aligned sequences and not the entire sequences.

The second program used for homology searches was the “Basic Local Alignment Search Tool” (BLAST) [<http://expasy.hcuge.ch/cgi-bin/BLASTEPFL.pl?>] maintained by the École Polytechnique de Lausanne in Switzerland. BLAST has the advantage of searching not only for direct sequence matches but also for local sequence similarities as well as multiple regions of homology.<sup>7</sup> It also provides a threshold parameter for background “noise”, the number of random hits that can be expected from using a database of a particular size. The disadvantage is that, rather than entering AC numbers for your query sequence, the interface requires that the entire protein sequence, formatted in single letter amino acid codes, be entered. One can choose to define a set of parameters to restrict the search, but for the most part I used the default set of parameters and selected the main peptide database as SWISS-PROT to compare the results with those of the Fast-A3 program. BLAST displays results as a group of result sequences

versus the original query sequence with graphical simulated lines aligned below the query for regions of homology or by specifically listing the amino acid residues.

#### MultAlin program -Multiple Sequence Alignment of Linker Peptides

The amino acid sequences of each linker were compared against each other using the MultAlin program at Institut National de la Recherche Agronomique (INRA) [<http://www.toulouse.inra.fr/multalin.html>]. MultAlin creates a multiple sequence alignment from a group of related sequences using progressive pairwise alignments.<sup>8</sup> The protein sequences are aligned with postulated gaps so that similar residues are juxtaposed. A positive score is attached to the alignment of identical amino acids, conservative substitutions for similar amino acids, and a penalty is given for gaps or non-conservative substitutions. Multiple iterations optimize the total score, evaluate all possible alignments and allow for any length gap at any position. The output of the two aligned sequences can be viewed either in the form of a text file or a GIF file which one can manipulate for color, size, consensus level and comments. In addition, the output is used as a starting point for secondary structure prediction.

#### *Protein Secondary Structure Predictions*

A variety of programs are available for secondary structure prediction from the amino acid sequence of a protein and their accuracy is typically ~70%. The recently available JPred consensus method [<http://circinus.ebi.ac.uk:8081/submit.html>] automatically runs and evaluates the results from six representative algorithms, and reports the individual as well as the consensus predictions.<sup>9</sup> The individual programs are

described below. The consensus prediction is based on a simple majority combination of the four currently most accurate methods (NNSSP, DSC, PREDATOR and PHD) and boasts an accuracy of 72.9% in the prediction of three conformational states ( $\alpha$ -helix, extended  $\beta$ -strand and random loop). Two other methods (MULPRED and ZPRED) are also automatically activated by JPred which are not included in the consensus evaluation but are supplied for comparison.

### MULPRED

MULPRED is a consensus algorithm based on the combination of different single sequence prediction methods.<sup>10-14</sup> Although descriptive literature is unavailable, the authors of JPred demonstrate that this approach is not so accurate as current multiple sequence methods.<sup>15</sup> MULPRED is incorporated into JPred and cannot be individually accessed; its prediction results are provided for comparative purposes only.

### ZPRED

[<http://kestrel.ludwig.ucl.ac.uk/zpred.html>]

ZPRED predicts secondary structure and active sites based on the conservation of amino acids between the query sequence and a family of homologous proteins.<sup>16</sup> It recognizes a sequence relationship between a segment of the polypeptide chain of the unknown structure with a sequence and conformation database from the known structures. In other words, the ZPRED algorithm is built on the observation that when sequences are aligned, the regions of insertions and low conservation of amino acids tend to occur in the random loop regions and not in the  $\alpha$ -helix or  $\beta$ -sheet structures.<sup>17</sup> The accuracy for a three-state prediction ( $\alpha$ -helix,  $\beta$ -sheet, loop) ranges between 59% and

63%. It begins with an alignment of the multiple homologous sequences and produces a standard Robson prediction<sup>10</sup> for each amino acid residue using an evaluation interval of 17 residues along the sequence. The prediction for each residue is averaged over the entire family of aligned homologous structures. A conservation number is assigned to each residue of the query sequence based on the strength of its homology with the other aligned homologous proteins; an identical match produces the highest score while a penalty is incurred for chemically different amino acids. The conservation number is averaged over three residues and the resulting values are used to weight the previously obtained averaged secondary structure prediction. More accurate conservation number scores would be possible through improvements to the alignment procedure as well as from a better evaluation method of the similarity in chemical properties of the amino acids. A second feature of the ZPRED algorithm is the identification of active sites using the assumption that residues conserved in loop regions are more significant for protein function than those conserved in  $\beta$ -sheets or  $\alpha$ -helices. Loop regions are more mobile and make poor scaffolding for the overall protein structure so from an architectural point of view there is no reason to evolutionarily conserve particular sequences unless these are also involved in function. JPred provides both the calculated ZPRED conservation number score and its resulting secondary structure prediction.

#### Discrimination of Secondary Structure Class (DSC)

[[http://bonsai.lif.icnet.uk/dsc/dsc\\_read\\_align.html](http://bonsai.lif.icnet.uk/dsc/dsc_read_align.html)]

The aim of DSC was to decompose secondary structure prediction into its basic concepts and then use simple linear statistical methods to combine these concepts to



produce an understanding of the folding process.<sup>18</sup> To date, DSC achieves the highest three-state accuracy ( $\alpha$ -helix,  $\beta$ -strand, coil) observed (70.1%) for a non neural-network approach. It begins with a simple single sequence prediction (over a window segment of 17 amino acid residues) based on the propensity of particular residues for particular secondary structures.<sup>10-14</sup> It then applies a linear discrimination function based on the following factors: residue position, hydrophobic regions, insertion and deletion areas in multiple aligned homologous sequences (typically occur in loop regions), and conservation of  $\alpha$ -helix and  $\beta$ -strand in homologous sequences. The resulting prediction for each residue is “smoothed” over 2 to 5 residues and then refined by calculating the expected ratio of  $\alpha$ -helix and  $\beta$ -strand structures as well as the frequency of certain residues (weighting factors). A second linear discrimination function is applied and then a filter for physically unlikely sequences of conformations is performed before delivering the final prediction. DSC accuracy differs according to sequence chain-length and its most accurate results are found for short (< 90 residues) or medium-length chains (90-170 residues). For longer chains, PHD, a neural-network approach described further below, is found to be significantly more accurate than DSC.

#### Nearest Neighbor Secondary Structure Prediction (NNSSP) [<http://genomic.sanger.ac.uk>]

NNSSP uses nearest-neighbor algorithms along with multiple sequence alignment to predict a three-state system of secondary structures ( $\alpha$ -helix,  $\beta$ -strand and coil) with an accuracy maximum of 72.2%.<sup>19</sup> The basic idea of the nearest-neighbor approach is the prediction of the secondary structure state of the central residue of a test segment, based on the secondary structure of homologous segments from proteins with known three-

dimensional structure. Restricting the number of homologous sequences to those most closely associated with the test sequence improves the accuracy of the method. NNSSP uses a scoring system for each residue that combines a sequence similarity matrix with a local structural environment scoring method<sup>20</sup> developed from a database of proteins with known three-dimensional structure. The database information allows the assignment of “environment classes” to each residue separately and as a group according to its local structure features, i.e., secondary structure state, solvent accessibility, polarity, and residue position. The most crucial element of the algorithm is what combination of environmental classes and similarity matrix to use when setting up the scoring table. For example, the method accuracy was improved by creating a separate environmental class for the N- and C- terminal ends of  $\alpha$ -helices and  $\beta$ -strands, which are usually difficult prediction areas.

Once each residue of the test sequence has been assigned a structural state, the prediction is refined by considering the structural state of its nearest-neighbors in the sequence segment (window size tested with 17 to 23 residues) as well as the predictions for the same aligned segment in the homologous sequences. Decisions are made on a simple majority vote and some filtering rules are applied to avoid unrealistic cases of very short helices and strands. A weighting parameter based on database statistics of secondary structure content in globular proteins was used to improve the accuracy of  $\beta$ -strand prediction from 42% to 65%, but even so the NNSSP algorithm is more accurate in predicting  $\alpha$ -helices (72.8%) than  $\beta$ -strands (66.6%).

### PREDATOR-secondary structure prediction from single or multiple sequences

[[http://www.embl-heidelberg.de/cgi/predator\\_serv.pl](http://www.embl-heidelberg.de/cgi/predator_serv.pl)]

PREDATOR is a three state ( $\alpha$ -helix,  $\beta$ -strand, coil) secondary structure prediction routine with a 68% accuracy for a single sequence.<sup>21,22</sup> The program can accept either a single sequence or a set of multiple sequences generated by a multiple alignment routine such as MultAlin described above. A single sequence input will prompt a search of large protein sequence databases for related members followed by a multiple alignment of the homologous sequences to produce a secondary structure prediction for the query sequence. The prediction technique takes into account the position and structural type of each amino acid residue individually and in relation to its nearest neighbors. A window average of predicted structure states over the length of 13 residues is applied for each individual residue. Filtering/cleaning rounds are completed to eliminate assignment conflicts between window segments before the final prediction is made. The resulting output shows the predicted secondary structure state for each residue as well as its calculated reliability index.

### PredictProtein method using a profile network prediction Heidelberg (PHD)

[<http://www.embl-heidelberg.de/predictprotein/predictprotein.html>]

PHD is composed of a series of programs that were developed at the EMBL with the purpose of improving secondary structure predictions using evolutionary information gathered from multiple sequence alignments or a multi-level system of neural networks. PHDsec<sup>23-25</sup> was used for secondary structure prediction, PHDacc<sup>26</sup> was designed to determine solvent accessibility while PHDhtm<sup>27</sup> searches for helical transmembrane

regions. At 72.2% three-state ( $\alpha$ -helix,  $\beta$ -strand, coil) accuracy, PHD is considered the most accurate secondary structure method currently available. Evaluated on the same data set of proteins, PHD has been rated at 10% higher three-state accuracy than methods using only single sequence information such as the previously discussed MULPRED, and at more than 6% higher than a method using alignment information based only on database statistics.<sup>28</sup>

The first step in the PHD analysis is to generate a multiple sequence alignment with proteins homologous to the test sequence. The second step involves feeding the alignment into a neural network system. Correctness of the multiple sequence alignment is as crucial for prediction accuracy as that the alignment contains a broad spectrum of homologous sequences. The PHD methods then process the input information on multiple levels. The first level is a feed-forward neural network (sequence-to-structure) consisting of two contributions: one from the original sequence (local) taken from a window of 13 adjacent residues and one from the family of homologous sequences (global). The output from this first level network is the structural state of the residue at the center of the input window. A second level structure-to-structure network is then applied to introduce a correlation between adjacent residues so that predicted lengths of  $\alpha$ -helices and  $\beta$ -strands are similar to those observed in three-dimensional protein structures. The next procedure involves weighting the prediction results to compensate for low accuracy  $\beta$ -strand prediction and a comparison is made per residue between the pre-weighted result and the weighted output and the prediction with the higher probability score is propagated to the next step. The final step is a simple filter to eliminate physically unrealistic predictions.

The final PHD output includes the secondary structure predictions for each residue as well as its calculated reliability. These reliability values for each structural class were plotted on a per residue basis to compare each region against all of the predicted structural possibilities. For example, if there is a strong probability for a region of 10 amino acids as an  $\alpha$ -helix, then all other structure types should be minimal in this same region; whereas if the region can be classified as either a  $\beta$ -strand or random coil, the two structural classes will have comparable probability values.

#### *RASMOL molecular viewer program*

The RASMOL program was originally developed as an interactive molecular structure viewer compatible with MacIntosh software.<sup>29</sup> It was one of the first successful graphic tools available for use on a small processor unit or personal computer. Various research computing groups around the world have modified the original programming code to add user-friendly interaction features and to produce versions that are more scientifically accurate (restrict bond lengths, angles etc.). RASMOL is now widely available along with instruction manuals in different computing platforms. The version used in this thesis was obtained through the UC-Berkeley Modular Chemistry Consortium website [<http://mc2.cchem.berkeley.edu>] and is called RasMac-ucbv1.3. We have used it to display various views of the crystal structures of the *Synechococcus* sp. PCC 7002 PC trimer and monomer units as well as the *Mastigocladus laminosus* APC trimer and  $L_C^{7.8}$  linker.

The crystal structures in the SWISS-PROT database could be downloaded directly into RASMOL. There is a number of structural display options available and a palette of

colors that can be chosen through either a command window or the drag/click menus.

Typically, the bulk of the protein is displayed in helical grey ribbons and regions of interest are accented by using the "select" and "color" commands in the RASMOL control window.

## References for Chapter 2

1. Gindt, Y.M., Zhou, J., Bryant, D.A., Sauer, K. (1992) *J. Photochem. Photobiol. B* 15:75-89.
2. Yu, M.H., Glazer, A.N., Williams, R.C. (1981) *J. Biol. Chem.* 256:13130-13136.
3. Lackowicz, J.R. (1983) Principles of Fluorescence Spectroscopy, Plenum Press, New York, NY, pp 111-150.
4. Maxson, P. (1988) Ph. D. Thesis, LBL-26163, University of California, Berkeley, CA.
5. Pearson, W.R., Lipman, D.J. (1988) *Proc. Natl. Acad. Sci. USA* 85:2444-2448.
6. Pearson, W. (1990) In Methods in Enzymology, Academic Press, San Diego, CA, Vol 183, pp 63-98.
7. Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J. (1990) *J. Mol. Biol.* 215:403-410.
8. Corpet, F. (1988) *Nucl. Acids Res.* 16:10881-10890.
9. Cuff, J.A., Clamp, M.E., Siddiqui, A.S., Finlay, M., Barton, G.J. (1998) *Bioinformatics* 14:892-893.
10. Garnier, J., Osguthorpe, D.J., Robson, B. (1978) *J. Mol. Biol.* 120:97-120.
11. Chou, P.Y., Fasman, G.D. (1974) *Biochem.* 13:211-222.
12. Lim, V.I. (1974) *J. Mol. Biol.* 88:873-894.
13. Rose, G.D. (1978) *Nature* 272:586-591.
14. Wilmot, A.C.M., Thornton, J.M. (1988) *J. Mol. Biol.* 203:221-232.
15. Cuff, J.A., Barton, G.J. (1999) *Proteins: Structure, Function and Genetics* 34:508-519.
16. Zvelebil, M., Barton, G., Taylor, W., Sternberg, M. (1987) *J. Mol. Biol.* 195:957-961.

17. Greer, J. (1981) *J. Mol. Biol.* 153:1027-1042.
18. King, R.D., Sternberg, M.J.E. (1996) *Prot. Sci.* 5:2298-2310.
19. Salamov, A.A., Solovyev, V.V. (1995) *J. Mol. Biol.* 247:11-15.
20. Bowie, J.U., Luthy, R., Eisenberg, D. (1991) *Science* 253:164-170.
21. Frishman, D., Argos, P. (1996) *Prot. Eng.* 9:133-142.
22. Frishman, D., Argos, P. (1997) *Proteins* 27:329-335.
23. Rost, B., Sander C. (1993) *Proc. Natl. Acad. Sci. USA* 90:7558-7562.
24. Rost, B., Sander C. (1993) *J. Mol. Biol.* 232:584-599.
25. Rost, B., Sander C. (1994) *Proteins* 19:55-72 for evaluation of accuracy.
26. Rost, B., Sander C. (1994) *Proteins* 20:216-226.
27. Rost, B., et al (1995) *Prot. Sci.* 4:521-533.
28. Levin, J.M., Pascarella, S., Argos, P., Garnier, J. (1993) *Prot. Eng.*, 6:849-854.
29. Sayle, R.A., Milner-White, E.J. (1995) *TIBS* 20:374-376.

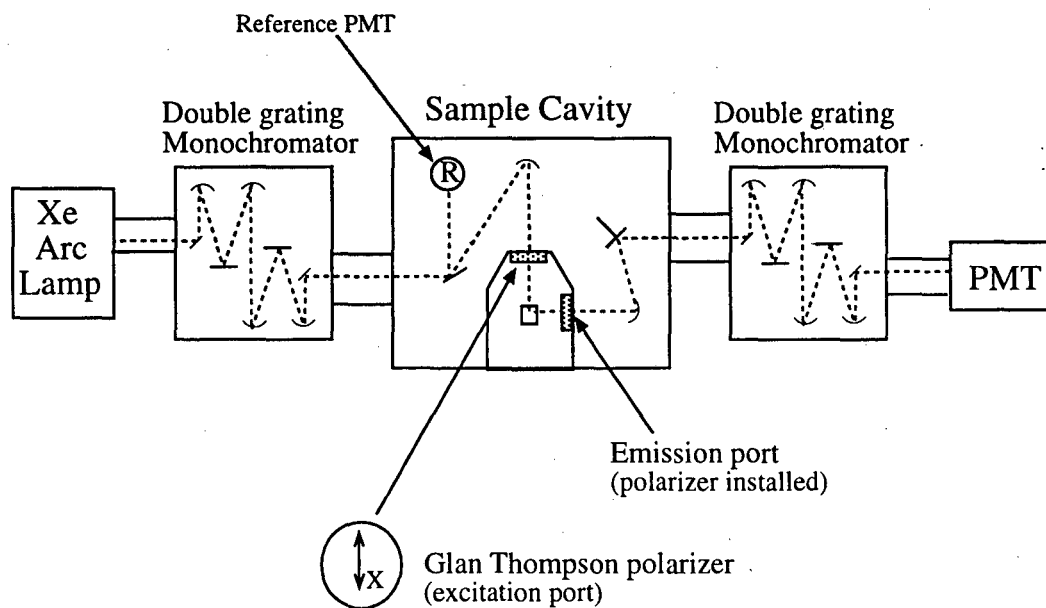


Table 2-1. Medium A+ Recipe

Chemical	Concentration of Stock (g/L)*	Volume(mL) of Stock per 1L of Medium	Final concentration (g/L)*
NaNO <sub>3</sub>	500	2	1.0
NH <sub>4</sub> Cl	150	1	0.15
K <sub>2</sub> HPO <sub>4</sub> •3H <sub>2</sub> O	50	1	0.05
CaCl <sub>2</sub> •2H <sub>2</sub> O	230	1	0.23
<u>Salt Solution</u>		100	
MgSO <sub>4</sub>	24.4		2.44
NaCl	180		18
KCl	6		0.6
Na <sub>2</sub> EDTA	30	1	0.03
1M Tris, pH 8	1 M	8.5	8.5 mM
FeCl <sub>3</sub> •6H <sub>2</sub> O	4	1	4 mg/L
<u>Trace Metals</u>		1	
H <sub>3</sub> BO <sub>3</sub>	34.2		34.2 mg/L
MnCl <sub>2</sub> •4H <sub>2</sub> O	4.3		4.3 mg/L
ZnSO <sub>4</sub> •7H <sub>2</sub> O	0.63		0.63 mg/L
Na <sub>2</sub> MoO <sub>4</sub> •2H <sub>2</sub> O	3.9		3.9 mg/L
CuSO <sub>4</sub> •5H <sub>2</sub> O	0.03		30 µg/L
CoCl <sub>2</sub> •6H <sub>2</sub> O	0.0122		12.2 µg/L
Vitamin B-12	50 mg/L	0.1	5 µg/L

\* units are g/L unless specified otherwise

**Figure 2-1.** Optical schematic of Spex Fluorolog used in fluorescence measurements  
Shown here with polarizing prisms installed



**Four possible experimental orientations**

$$\begin{array}{c} \updownarrow \\ \text{X} \end{array} + \begin{array}{c} \updownarrow \\ \text{M} \end{array} = \text{VV}$$

$$\begin{array}{c} \leftarrow \rightarrow \\ \text{X} \end{array} + \begin{array}{c} \leftarrow \rightarrow \\ \text{M} \end{array} = \text{HH}$$

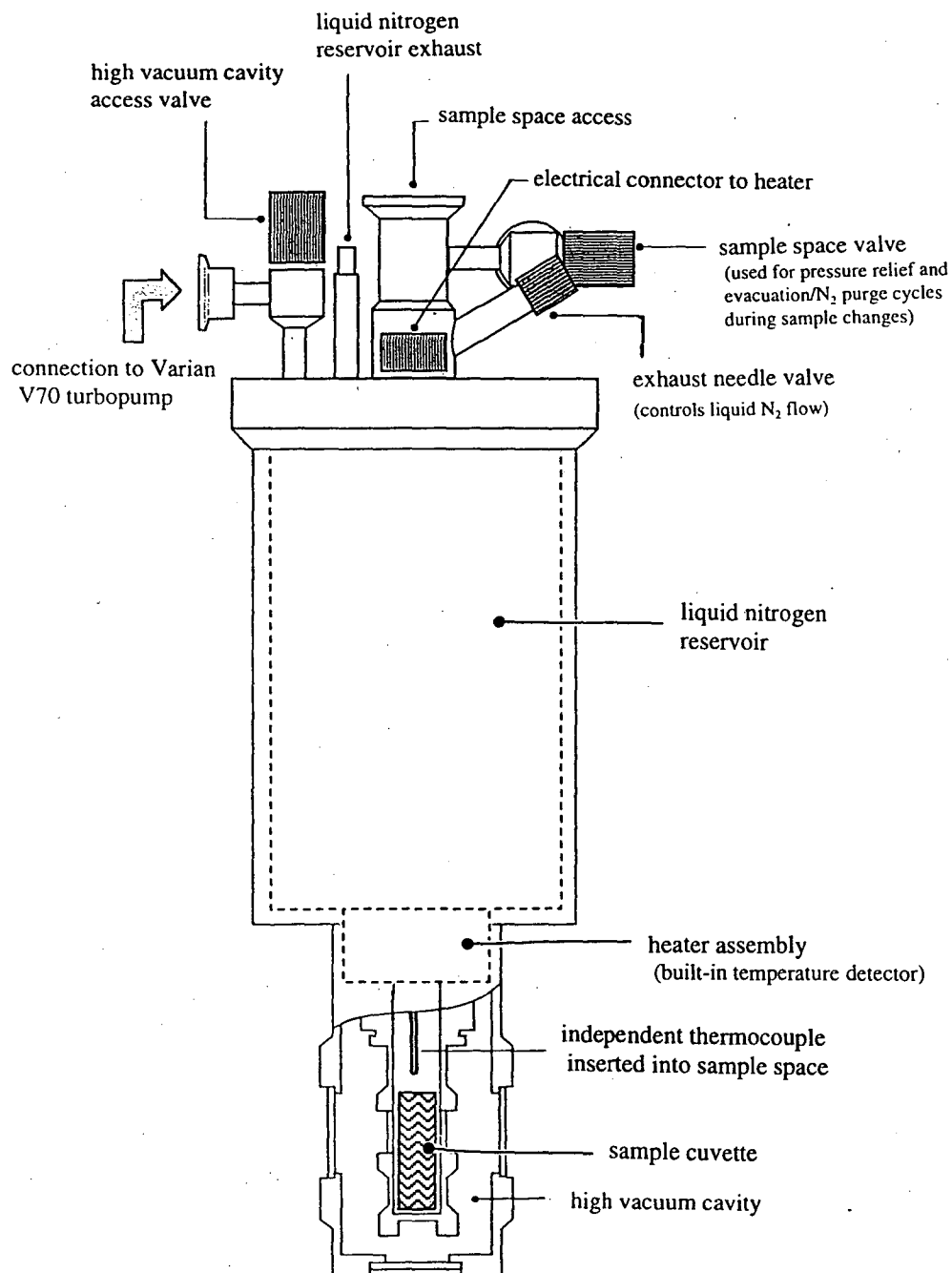
$$\begin{array}{c} \leftarrow \rightarrow \\ \text{X} \end{array} + \begin{array}{c} \updownarrow \\ \text{M} \end{array} = \text{HV}$$

$$\begin{array}{c} \updownarrow \\ \text{X} \end{array} + \begin{array}{c} \leftarrow \rightarrow \\ \text{M} \end{array} = \text{VH}$$

Experimental correction needed (adapted from reference 4):

$$\frac{I_{\text{parallel}}}{I_{\text{perpendicular}}} = \frac{\text{VV} * \text{HV}}{\text{VH} * \text{HH}}$$

**Figure 2-2.** Schematic diagram of Oxford Optistat<sup>DN</sup> liquid nitrogen cryostat  
(adapted from Oxford instruction manual)



## Chapter 3. Analysis of linker homology and structural predictions

### Introduction

The colorless linker peptides of *Synechococcus* sp. PCC 7002 are an integral part of the phycobilisome (PBS) assembly. Besides providing structural support, the rod linkers modify the spectral properties of the resident biliprotein chromophores thereby affecting the flow of energy transfer. Until recently, there was very little structural information available for the linker peptides. In the PBS rods, the linkers are assumed to occupy the ~35 Å-wide "cavity" at the center of the biliprotein hexamer unit. The C<sub>3</sub> symmetry of the C-phycocyanin (PC) trimer provides the 'single' linker peptide with three identical binding sites, and it is this lack of overall symmetry for the PC/linker complex that hindered resolution of crystallographic data. The first resolved crystal structure of a biliprotein/linker complex has now been published; it is the core biliprotein allophycocyanin (APC) with its small linker peptide L<sub>C</sub><sup>7.8</sup> isolated from *Mastigocladus laminosus*.<sup>1</sup> This structure has made it possible to pinpoint the specific physical interaction sites between the linker peptide and the APC-bound chromophores.

The structure of APC is very similar to that of PC, and all chromophores found in these two biliproteins are chemically identical phycocyanobilins named according to subunit and residue position. The goal of the analysis presented in this chapter is to examine the available primary structure of the three rod linkers found in *Synechococcus* sp. PCC 7002 (L<sub>R</sub><sup>8.9</sup>, L<sub>R</sub><sup>32.3</sup>, L<sub>RC</sub><sup>28.5</sup>) and produce a model of interaction for PC/linker aggregates. The APC/L<sub>C</sub><sup>7.8</sup> structure will be used as a guide in drawing analogous conclusions for PC/linker interactions. The chapter is divided into three parts. First, an analysis of the inherent primary structure properties of each linker peptide sequence is

presented. This includes examination of electrostatic charge content, hydrophobic regions and sequence segments homologous with other proteins. Secondly, predictions of the secondary structure of each linker sequence are calculated and compared to the known  $L_C^{7,8}$  structure. The results of this primary and secondary structure analysis of the linkers is discussed and summarized, followed by a comparative analysis of the APC and PC tertiary structures for possible linker binding sites.

## Results

### *Primary Structure Analysis*

All amino acid sequences used were obtained from the SWISS-PROT database, and those of the rod linker peptides are shown in Table 3-1 at the end of this chapter. Statistical properties calculated from each linker's amino acid sequence are reported in Table 3-2. Of the three rod linkers in this study,  $L_R^{8,9}$  has the lowest calculated isoelectric point ( $pI = 5.42$ ), with a value comparable to that of the PC  $\alpha$  ( $pI = 5.31$ ) and  $\beta$  ( $pI = 5.58$ ) subunits. In contrast, the core linker  $L_C^{7,8}$  has an extremely high  $pI$ . Table 3-2 also shows that the content of electrostatically charged amino acids in each linker is different between the rod linkers and the  $L_C^{7,8}$  linker. While all the rod linkers have an almost equal balance of positively and negatively charged amino acids,  $L_C^{7,8}$  has a 3:1 majority of positively charged residues to negatively charged ones. On average, the content of charged residues in the rod linkers comprised ~22% of the total amino acid content. The most abundant charged amino acid in the rod linker peptides is the positively-charged arginine. It ranks second only in  $L_R^{8,9}$ , where glutamate is more abundant, resulting in a slightly more negatively charged content. The other two rod

linkers,  $L_R^{32.3}$  and  $L_{RC}^{28.5}$ , have overall electrostatic charges that are either neutral or slightly positive.

Two algorithms, Fast-A3 and BLAST (described in Chapter 2), were used to perform each homology search and both yielded comparable results; only the Fast-A3 results are presented here. Tables 3-3A to 3-3D summarize the homology search results for each linker sequence; these are organized according to gene name, in order of decreasing similarity with the specified linker. A code key of gene names and organisms is provided to identify each homologous protein. The homology score next to each sequence represents the similarity between the query sequence and a homologous protein over the entire length of the sequence; the algorithm evaluates each set of residues and determines the score from the number of identical matches as well as matches between amino acids with similar properties. On a separate graph below each table, we have noted the “conservation” across species for the protein family of the query sequence by plotting the number of times a residue is repeated (i.e., if a histidine appears at the same position in 5 out of 8 sequences, then it is 63% conserved) versus the residue position. This helps in identifying evolutionary trends across the sequence length.

Generally, each linker attains the highest homology score with proteins of the same type in different cyanobacterial species, i.e.,  $L_C^{7.8}$  has the highest homology score with other  $L_C$  linkers while  $L_{RC}^{28.5}$  is most similar to other  $L_{RC}$  linkers. No significant region is common among all three rod linkers, although  $L_{RC}^{28.5}$  and  $L_R^{8.9}$  each share some sequence homology with separate regions of  $L_R^{32.3}$  (for quick reference see schematic in Figure 3-1, details are shown in Figure 3-2). The sequence for  $L_C^{7.8}$  shares

a significant region of similarity with  $L_R^{8.9}$  near its C-terminus that is also homologous to  $L_R^{32.3}$  and  $L_{RC}^{28.5}$  albeit to a lesser extent (Figure 3-3).

The results from the Fast-A3 search for homologous proteins with the  $L_R^{8.9}$  sequence are shown in Table 3-3A. There is good homology with other rod-capping linkers (code = PYS1, average homology score 43%) over the entire sequence length. The conservation graph at the bottom of the table shows that these similarities are concentrated in the central sequence region where the average conservation across species is 71% (residues 12-70 only). There is also some homology between  $L_R^{8.9}$  and other rod linkers (PYR1, PYR2, PYR3, 31% average homology) that can be as high as 41% for segments of ~67 residues and 48% for segments of ~33 residues. In addition, there are small areas of  $L_R^{8.9}$  homology with two sets of proteins of a different type. The first set includes homology matches to core linker peptides (PYC1). Although not included in Table 3-3A because their homology scores were < 25%, these  $L_C$  linkers showed an average homology match of ~45% with short sections (~30 amino acids) of  $L_R^{8.9}$ . In particular, the homology between  $L_R^{8.9}$  and the recently structurally resolved  $L_C^{7.8}$  was 50% over a region of 28 amino acids and was the highest value found in comparing  $L_C^{7.8}$  with the three rod linkers in this study (Figure 3-2). The second set of non-rod-linker type homologues for  $L_R^{8.9}$  is comprised of the ferredoxin:NADP<sup>+</sup> oxidoreductase (FNR) enzyme. The average homology match is good (52%), but it corresponds to a small region near the N-terminus of the FNR protein (69 out of 402 residues, i.e., 17% of the total sequence), which is part of a network of  $\beta$ -strands and loop segments and thus cannot be used in isolation as an accurate structural template for  $L_R^{8.9}$ .

Table 3-3B summarizes the Fast-A3 homology search results for the  $L_R^{32.3}$  linker. It has good homology across species with linkers of the same type (PYR1, 57% average homology score). The distribution of conserved amino acids is skewed, with a higher degree of conservation towards the N-terminus end; residues 1-180 have an average conservation value of 79%, while the residues in the remainder of the sequence have an average conservation value of 55%. In addition to the matching region with the  $L_R^{8.9}$  linker already noted, the  $L_R^{32.3}$  sequence also has a homologous area near its N-terminus with the  $L_{RC}^{28.5}$  linker, although it is not as strong (35%) as the  $L_R^{8.9}$  comparison (49%). A surprising number of homology matches were found between  $L_R^{32.3}$  and the core anchor linker peptide  $L_{CM}$  (APCE, 28% average homology score). The  $L_{CM}$  region that is homologous with  $L_R^{32.3}$  is classified as REP for “repeat”; a region found in multiple copies that is responsible for joining together APC trimers.<sup>2</sup>

The Fast-A3 homology search results for the  $L_{RC}^{28.5}$  linker are presented in Table 3-3C. The highest homology is once again found among the same type linkers across species (PYG1/PYG, 50% average homology score), while lower homology values are found with other types: 26% average homology with rod linkers (PYR1, PYR5, PYR6) and 27% for matches to  $L_{CM}$  anchor peptides (APCE). The latter similarity matches are to the same  $L_{CM}$  REP region that was also homologous with  $L_R^{32.3}$ . It is the N-terminal region of  $L_{RC}^{28.5}$  that is homologous with both the rod linkers and the  $L_{CM}$  polypeptides. Like in  $L_R^{32.3}$ , the N-terminal region of  $L_{RC}^{28.5}$  also has a higher degree of conserved residues with an average conservation value of 68% for the first 180 amino acids and 40% for the rest of the sequence.



Lastly, a homology search was performed for the  $L_C^{7.8}$  sequence (Table 3-3D). The  $L_C^{7.8}$  linker is very well-conserved across species with other core linkers (PYC1, average homology score 86%). It has some homology with rod linkers (PYS1, PYR1, PYR2, PYR3, average homology score 25%) and with FNR proteins (24%). Although the same  $L_C^{7.8}$  segment (near its C-terminus) displayed some similarity to C-terminal regions of  $L_R^{32.3}$  and  $L_{RC}^{28.5}$ , the closest association to the entire sequence is found with  $L_R^{8.9}$ ; each rod linker sequence is separately aligned with  $L_C^{7.8}$  in Figure 3-3.

#### *Secondary structure predictions*

The JPred consensus method of protein secondary structure prediction for three structural states ( $\alpha$ -helix,  $\beta$ -strand, loop) was applied to each linker sequence. This method is comprised of six discrete prediction algorithms (DSC, MULPRED, ZPRED, NNSSP, PREDATOR and PHD), and the results from each one are reported along with the consensus prediction reached through a simple majority vote. Figure 3-4A presents the secondary structure predictions for  $L_C^{7.8}$  ( $\alpha$ -helix = H,  $\beta$ -strand = E, random loop = -). Comparing the known secondary structure of  $L_C^{7.8}$  with the JPred consensus result shows that the structure assignments of 85% of the residues are correctly predicted. This is higher than the published average accuracy of JPred (72.9%). Individually, the most accurate method for this protein was PHD (83.6%) while the worst was ZPRED (61.1%). The prediction reliability per amino acid residue for the PHD prediction of each structural class is also plotted in Figure 3-4A.

Figures 3-4B to 3-4D present the abridged prediction results (JPred consensus and PHD predictions) for the rod linker peptides. The consensus prediction for  $L_R^{8.9}$  (Figure

3-4B) contains three  $\beta$ -strands (spanning residues R14-Q21, V42-V46, I65-K69) and one  $\alpha$ -helix (M51-L61). PHD predicts two additional  $\beta$ -strand regions (Q4-N7, A33-R37) that were not included by the consensus method. The first one has a very low reliability value (< 20%), but the second one is quite strong (maximum reliability = 80%). In contrast to  $L_R^{8.9}$  and  $L_C^{7.8}$ , the secondary structure consensus predictions for  $L_R^{32.3}$  and  $L_{RC}^{28.5}$  have a wealth of  $\alpha$ -helical regions.  $L_R^{32.3}$  is predicted to have 11  $\alpha$ -helices and 5 small  $\beta$ -strands (Figure 3-4C), while  $L_{RC}^{28.5}$  is expected to contain 7  $\alpha$ -helices and 4  $\beta$ -strands (Figure 3-4D). The regions of shared homology between the rod linker peptides are indicated in the figures. Instead of using a graph for the  $L_R^{32.3}$  and  $L_{RC}^{28.5}$  linker predictions, the PHD reliability scores are noted below each residue as a single number (0 to 9, 9=highest reliability).

## Discussion

### *Primary structure analysis*

There are many factors that contribute to a protein's structure and function. The unique combination of amino acids in the primary structure of a protein suggests specific interactions as well as general properties characteristic to that protein. Our first step in analyzing the primary structure of each linker was a calculation of its isoelectric point (pI). Across cyanobacterial species, linker peptides typically have high calculated pI values and are thus expected to be positively charged peptides at the neutral pH 7.0. This leads to hydrophobic behavior such as aggregation and precipitation when dissociated from their biliprotein binding site. In contrast, the biliproteins are highly water soluble and have more neutral pI values.

Of the three rod linkers in this study, L<sub>R</sub><sup>8.9</sup> has a number of characteristics that subtly set it apart from the other two. Comparing the total content of amino acid types in each linker, L<sub>R</sub><sup>8.9</sup> has the highest percentage of polar and charged residues, the lowest percentage of aromatic and hydrophobic residues, and the lowest calculated pI (Table 3-2). It is the only linker peptide in this set of rod linkers that contains a concentrated sequential area of polar and charged residues near its N-terminus (residues 24-33), and it could be preferentially exposed to the solvent. The other two linkers have smaller (~5 residues long) segments of polar and charged residues scattered throughout their sequences. These general properties of the amino acid sequence may indicate a difference in the biliprotein binding mechanism of L<sub>R</sub><sup>8.9</sup> and the other two linkers. The aromatic and negatively charged residues of L<sub>R</sub><sup>8.9</sup>, necessary for strong electrostatic interactions, are not well conserved and, although it has a high content of polar amino acids, only 18% of these are directly conserved across species. Commonly known as a “capping” linker, L<sub>R</sub><sup>8.9</sup> is believed to occupy the position at the end of the phycobilisome rod. Because only one coding gene has been found for this linker type in *M. laminosus*,<sup>3,4</sup> it is assumed that L<sub>R</sub><sup>8.9</sup> caps the rod regardless of the terminal biliprotein present: PC at high light intensity and phycoerythrocyanin (PEC) at low light intensity.<sup>5</sup> Unfortunately, during the protein isolation procedure from cyanobacterial cells, a PC/L<sub>R</sub><sup>8.9</sup> complex is not recovered; L<sub>R</sub><sup>8.9</sup> is thought to dissociate from the PBS rod in the low buffer solutions and is lost in the purification procedure.<sup>6</sup> This behavior suggests an ability for L<sub>R</sub><sup>8.9</sup>, in removing itself from the biliprotein binding site as well as a recognition of multiple biliproteins, that is not shared with the other linkers.

$L_R^{8.9}$  also shares some similarity with the sequence of the FNR enzyme.

Although FNR is found to bind to the PBS rods<sup>7</sup>, it differs in function from  $L_R^{8.9}$  in that it is not necessary for PBS stability or function. The homology match with  $L_R^{8.9}$  is good (52%) but it corresponds to a small region near the N-terminus of the FNR protein (69 out of 402 residues). The FNR binding affinity to PBS rods suggests that it may share the characteristics necessary to bind to PC at the same point as  $L_R^{8.9}$ ; however, the structure of FNR cannot be used as a template for three-dimensional structure predictions of  $L_R^{8.9}$ . The FNR section homologous to  $L_R^{8.9}$  represents only 17% of the total FNR protein sequence, and it lies in a region of nested  $\beta$ -strands and loop segments; the probability of maintaining the same structure in the absence of the other neighboring residues is low. The secondary structure predictions for  $L_R^{8.9}$  discussed in the next section follow more closely the  $L_C^{7.8}$  structure profile than that of FNR, even though the  $L_R^{8.9}$  sequence homology with  $L_C^{7.8}$  over this region is slightly lower. Despite the differences in structure, both FNR and  $L_C^{7.8}$  do possess loop regions within the segments homologous with  $L_R^{8.9}$  which could be the key binding sites to the PBS. Published studies on the similarity relationship between  $L_R^{8.9}$  and FNR have concluded that this region of shared homology is not necessary for the catalytic activity of FNR from a structural point of view.<sup>8</sup> The presence of this FNR fragment that is homologous to  $L_R^{8.9}$  seems to be a feature found only in cyanobacterial FNR, and its speculated purpose is to act as an alternative anchor for the FNR to modulate the photosynthetic electron transport cycle.<sup>7,8</sup> Once anchored to the PBS, the structure of this FNR segment may be modified, and its sequence homology to  $L_R^{8.9}$  could represent a shared binding affinity for this region.

The  $L_{RC}^{28.5}$  linker possesses the highest pI value of all three rod linkers as well as the highest number of hydrophobic residues (Table 3-2). These general properties approach those of  $L_C$  peptides rather than those of a rod linker. As a rod-core peptide,  $L_{RC}^{28.5}$  is expected to contain at least two distinct functional domains. In other words, one area to interact with the rod biliprotein PC is necessary while another area interacts with the core biliprotein APC. With this in mind, the  $L_{RC}^{28.5}$  sequence was divided in half and the pI value was calculated for each half. Dramatically different results were obtained; the N-terminal region had a calculated pI of 5.23 while that of the remainder of the peptide sequence was 10.48. The C-terminal region of  $L_{RC}^{28.5}$  is therefore positively charged at physiological pH, so it is more likely to be interacting with APC because this is a property typically found in core peptides. The N-terminal region is then expected to bind to the PBS rods. This assignment is supported by the fact that the strongest homology shared between  $L_{RC}^{28.5}$  and the rod linker  $L_R^{32.3}$  lies in its N-terminal region ( $L_{RC}^{28.5}$  residues 35-160, Figure 3-2). Interestingly, the first 180 amino acid residues of  $L_{RC}^{28.5}$  are better conserved (68% average conservation across species) than those in the rest of the protein (40% average conservation value). In *Synechococcus* sp. PCC 7002, the first 180 amino acid residues of  $L_{RC}^{28.5}$  (i.e., the N-terminal side) are sufficient to produce the characteristic PC/ $L_{RC}^{28.5}$  absorbance band.<sup>9</sup> Similar behavior has been published for a rod-core linker ( $L_{RC}^{31}$ ) from *M. laminosus*, where a 22 kDa proteolytic fragment of  $L_{RC}^{31}$  with the N-terminal sequence intact associated with the PC trimer and produced a characteristic red shift of 15 nm in its absorbance maximum.<sup>10</sup> The well-

conserved N-terminal regions of these two rod-core linkers are thus classified as the areas critically responsible in modulating the spectra of the PC chromophores.

Comparison of the amino acid sequences of the rod linkers in this study with each other revealed that there is no common homologous region among the three. In addition, there are no duplicate areas within each linker sequence. This implied not only that each linker type may have different modes of binding to the biliprotein but also that the linker/biliprotein interaction must be different with each of the three monomers. Among the three rod linkers in this study,  $L_R^{32.3}$  is the only one with the function of connecting two PC hexamer stacks. Despite the fact that the PC hexamer is identical on both top and bottom surfaces, the primary structure of  $L_R^{32.3}$  does not possess any obvious axis of symmetry, nor regions of homology between the N- and C-terminal segments. This presents the possibility of a single linker that can have two different binding interactions with a PC hexamer. This mechanistic duality particular to  $L_R^{32.3}$  type linkers was remarked early on by Yu and Glazer who proposed that the N-terminal end (~28 kDa) was necessary to stabilize the trimeric and hexameric complexes, while the remainder 4.5 kDa was necessary for the interaction between the hexameric complexes.<sup>11</sup> Comparing the  $L_R^{32.3}$  sequence with the other two rod linkers, we find two distinct areas of homology; one in the N-terminal region homologous with  $L_{RC}^{28.5}$  as already discussed, and a second one in the C-terminal region that is homologous with the  $L_R^{8.9}$  linker ( $L_R^{32.3}$  residues 253-288, Figure 3-2). These two regions are very different from each other. First, the amino acids in the N-terminal end (residues 1-180) are better conserved (79% average conservation value) across species than the C-terminal end (55%). Gläuser *et al.* compare the N-terminal sequences (1-190) of  $L_{RC}$  and  $L_R$  linkers associated with

PC or PEC and found six well-conserved regions of residues across species and linker type;<sup>12</sup> there was very little similarity found in the C-terminal ends. The distribution of electrostatically charged residues in each of the two  $L_R^{32.3}$  regions is very different; of all the charged residues in the region homologous with  $L_R^{8.9}$ , 89% are positively charged, while in the area homologous with  $L_{RC}^{28.5}$  there is a distribution of 38% positively charged residues and 62% negatively charged. There is no appreciable difference in the total percent content of hydrophobic, polar, or nonpolar residues in each of the segments, but there are three times the number of aromatic groups (16% versus 5.6%) in the segment homologous to  $L_{RC}^{28.5}$ . It is interesting to note that all of the conserved aromatic and negatively charged residues in both  $L_{RC}^{28.5}$  and  $L_R^{32.3}$  are located in the N-terminal region. This suggests a strong likelihood that this region evolutionarily possesses a more specific role than the C-terminal region, namely to modulate the spectral properties of the PC chromophores.

In joining two separate PC hexamers, each end of the  $L_R^{32.3}$  linker binds with a different mechanism: one which is similar to  $L_{RC}^{28.5}$  and the other similar to  $L_R^{8.9}$ . The  $L_R^{8.9}$  linker has not been recovered during protein isolation intact nor in association with PC. Its impact on the PC spectroscopy is unknown but is expected to be minimal, because it can associate indiscriminately with both PC and PEC in its role as a rod capping linker.<sup>5</sup> If the N-terminal region of  $L_R^{32.3}$  (homologous with  $L_{RC}^{28.5}$ ) interacts with the biliprotein chromophores to produce the observed spectral changes, then the remaining one-third of the  $L_R^{32.3}$  amino acid sequence is not involved in the PC spectroscopy but bound to a neighboring hexamer. Assuming that the  $L_R^{32.3}$  peptide has

two different modes of binding interactions, one which affects the PC chromophores and a second which does not, then a possible role arises for these remaining residues:

insertion in the next PC hexamer central cavity to make a stabilizing backbone for the rod structure. It is possible that the  $L_R^{32.3}$  linker binds at a specific site to orient the two PC hexamers with each other for optimal transfer of energy, however, the lower number of strongly conserved amino acids in the C-terminal region suggests that this is a less precise interaction than that of the N-terminal region.

#### *Discussion of secondary structure predictions*

Unfortunately, none of the homologous proteins identified for the rod linkers has an available crystal structure to provide a starting point for three-dimensional modeling. Each linker sequence in this study shared the highest primary structure homology with other linkers of the same respective type. These homologous regions have been evolutionarily conserved across species and are therefore considered important to the function of the protein. As a result, special interest is given to the predicted structure within the homologous regions identified in the primary structure analysis. The focus of this section is to evaluate the secondary structure predictions for each linker peptide and search for similarity in the distribution of specific structural elements ( $\alpha$ -helix,  $\beta$ -strand). The general idea behind these predictive algorithms is to evaluate the location and properties of each amino acid residue within a sequence and identify the most energetically favorable and stable structural conformation for that combination of amino acids. The evaluations are based on folding pattern statistics gathered from the SWISS-PROT protein database and, since most globular proteins have an uneven distribution of



secondary structure types (32%  $\alpha$ -helix, 21%  $\beta$ -strand, 47% coil),<sup>13</sup> the  $\alpha$ -helix remains by far the most well defined structural motif. All of the linkers in this study were submitted to the JPred consensus method for secondary structure prediction described in detail in Chapter 2.

The  $L_C^{7.8}$  linker sequence was submitted to the JPred algorithm to measure its prediction accuracy against a known structure, and the results are presented in Figure 3-4A. Although  $L_C^{7.8}$  was co-crystallized with APC, the prediction program is able to simulate its structure with an overall accuracy of 85%. This value is much higher than the reported average 72.9% accuracy<sup>14</sup> and implies that this family of proteins should be reasonably well-described by this prediction method. Of the six individual predictive algorithms comprising Jpred, the PHD method was the most accurate (83.6%), so its prediction reliability per amino acid residue is included as a graph in Figure 3-4A. It is in the “edge” regions of each structural prediction motif that the greatest differences between the secondary structure prediction and the crystallographic results are found (i.e., the algorithm is poor in defining where an  $\alpha$ -helix or  $\beta$ -sheet begins and ends). Otherwise, every expected structural element (three  $\beta$ -strands and two  $\alpha$ -helices) is predicted at more or less the correct position.

The  $L_C^{7.8}$  area (V31-V52) of shared homology with  $L_R^{8.9}$  includes a long  $\alpha$ -helix segment and the beginning of flanking  $\beta$ -strands; a similar pattern is found in the  $L_R^{8.9}$  prediction (Figure 3-4B). Although the consensus prediction is incorrect for  $L_C^{7.8}$  residues 31-34 (replacing the  $\alpha$ -helix and  $\beta$ -strand assignments with random loop predictions) this area is nearly 100% conserved between  $L_C^{7.8}$  and  $L_R^{8.9}$  and therefore both should possess an identical structure despite the prediction for shorter  $\alpha$ -helix and  $\beta$ -

strand segments in both linkers. In addition to the area of shared homology with  $L_C^{7,8}$ , the  $L_R^{8,9}$  secondary structure prediction also includes a  $\beta$ -strand segment (R14-Q21) near its N-terminus. There is no threefold symmetry found in the pattern of structural element positions as expected from the lack of repeating regions in the amino acid sequence.

The secondary structure predictions for both  $L_R^{32,3}$  (Figure 3-4C) and  $L_{RC}^{28,5}$  (Figure 3-4D) display a large number of  $\alpha$ -helices and few  $\beta$ -strands. An issue that arises from the primary structure analysis for  $L_R^{32,3}$  is the possibility of two distinct interaction binding sites since it has two areas of homology: one with each of the other two rod linkers. Its amino acid sequence does not contain repeating sequence segments, and therefore it is not surprising that there is no axis of symmetry in the pattern of predicted secondary structure elements. In fact the two regions are very different in predicted structure (Figure 3-4C) suggesting a different method of interaction for each region. The  $L_R^{32,3}$  C-terminal region (residues 253–288) that shares homology with  $L_R^{8,9}$ , produces the same pattern of already observed structural elements for this region: a long  $\alpha$ -helix flanked by short  $\beta$ -strands. The N-terminal area (residues 27-152), homologous with  $L_{RC}^{28,5}$ , has a consensus prediction of 7  $\alpha$ -helices and one very small  $\beta$ -strand starting at residue 85. The PHD prediction differs from the consensus prediction in this N-terminal region and indicates the presence of two additional  $\beta$ -strands. The first small  $\beta$ -strand (residues 82-88) roughly agrees in position with the consensus prediction and has the highest reliability prediction score (~40%) of all three. The consensus prediction shows only four other short  $\beta$ -strands predicted throughout the remainder of the protein.

Despite the similarities in the primary structures of  $L_R^{32,3}$  and  $L_{RC}^{28,5}$ , the homologous regions are not identical in predicted secondary structure. Starting at the

beginning of the homologous section in  $L_{RC}^{28.5}$  in Figure 3-4D, one finds two separate  $\alpha$ -helices instead of the three predicted in  $L_R^{32.3}$ . Interestingly, a  $\beta$ -strand located at  $L_{RC}^{28.5}$  residues W124-T130 has a PHD prediction reliability maximum score of 70% but is predicted to be an  $\alpha$ -helix (maximum reliability = 40%) in the homologous  $L_R^{32.3}$  segment (Figure 3-4C). It is a region that is well-conserved in each linker peptide across different species (Tables 3-3B and 3-3C) but which has very low homology between  $L_{RC}^{28.5}$  and  $L_R^{32.3}$  (Figure 3-2). This difference in structure may identify an area where one of the key distinctions between the linker/PC interaction mechanisms of these two linkers exists. The rest of the  $L_{RC}^{28.5}$  homologous region contains two  $\alpha$ -helices at similar locations to the  $L_R^{32.3}$  prediction.

General changes in the electrostatic environment of the central PC cavity due to linker association, as well as specific interactions between the linker and the PC chromophores, contribute to the observed spectral differences between PC trimers and PC/linker complexes. The distribution of electrostatically charged amino acids in the entire homologous region shared by  $L_R^{32.3}$  and  $L_{RC}^{28.5}$  is subtly different. While both possess a nearly equal number of charged amino acids,  $L_R^{32.3}$  has a significantly larger content of negatively charged residues (21 versus 16 in  $L_{RC}^{28.5}$ ). Of particular interest is the string of consecutive positively charged arginines (R159-R161) found in  $L_{RC}^{28.5}$  which is conserved ~100% across rod-core linkers of different species (Table 3-3C) and is structurally predicted as a random loop thus likely to be more important to the protein function rather than its structure. Both glutamate and arginine can serve to make strong binding charge-charge interactions with the PC near the chromophores, but arginine has

been calculated to have a strong effect on the chromophore spectra by interacting with the propionate side chains of the bilin.<sup>15,16</sup> This carboxylate-arginine interaction is expected to shift the spectral maximum of the covalently bound PC chromophore to a lower energy. L<sub>RC</sub><sup>28.5</sup> is the only one of the three rod linkers in this study to possess a consecutive string of arginines in its amino acid sequence, and the absorbance band maximum of the PC/L<sub>RC</sub><sup>28.5</sup> complex is found at a lower energy than that of the PC/L<sub>R</sub><sup>32.3</sup> (discussed in Chapter 4).

### **Tertiary structure analysis**

The availability of the three-dimensional crystal structure of the APC/L<sub>C</sub><sup>7.8</sup> complex provides insight into the biliprotein/linker interaction. The crystal structure clearly shows that the L<sub>C</sub><sup>7.8</sup> linker is physically interacting with only two of the three available APC monomers and thereby disrupting the C3 symmetry of the APC trimer (Figure 3-5, adapted from reference 1). This means that the protein environment of each APC chromophore is altered to a different extent by the linker peptide. The surface area of the linker is 45.3% buried within the APC trimer, and it does not fully occupy the entire central cavity of the biliprotein. The goal of this section is to identify sites in PC and the rod linker peptides that may be analogous to interaction sites in the APC/L<sub>C</sub><sup>7.8</sup> structure and to propose a model of biliprotein/linker interaction. The following discussion is divided into three parts: (1) examination of the contact regions between APC and L<sub>C</sub><sup>7.8</sup>, (2) comparison of the APC and PC tertiary structures, and (3) comparison of the amino acid sequences of L<sub>C</sub><sup>7.8</sup> and the rod linkers (L<sub>R</sub><sup>32.3</sup> and L<sub>RC</sub><sup>28.5</sup>) associated with PC trimers.

### *Examination of the of APC/L<sub>C</sub><sup>7.8</sup> interface region*

The L<sub>C</sub><sup>7.8</sup> linker is located between two monomers of the trimeric APC complex and binds “via multiple charged, polar, and hydrophobic contacts to the protein chains”.<sup>1</sup> There is no clear clustering of charged or polar residues at the linker/biliprotein interface, and there are two structurally distinct binding sites. The first APC monomer interacts with the long  $\alpha$ -helix region of L<sub>C</sub><sup>7.8</sup> (residues 33-46), while the second APC monomer is approached by a smaller loop segment from L<sub>C</sub><sup>7.8</sup> (residues 13-19) (Figure 3-5). For clarity in the following discussion, these two types of APC/linker interactions are labeled as ‘APC/L<sub>C</sub><sup>7.8</sup> ( $\alpha$ -helix)’ and ‘APC/L<sub>C</sub><sup>7.8</sup> (loop)’. Association with L<sub>C</sub><sup>7.8</sup> has two consequences on the overall structure of the APC trimer. First, looking at the APC trimer disk toroidal surface, a slight counter-clockwise rotation of each monomer is noted. Although this relocation does bring the central  $\beta$ 81 chromophores closer together, it does not significantly alter the center-to-center distance between neighboring  $\alpha$ 80 and  $\beta$ 81. This effect of linker association enhances the electronic interaction of the central  $\beta$ 81 chromophores without disrupting the  $\alpha$ 80/ $\beta$ 81 relationship. A rough calculation made by drawing a straight line across each chromophore to represent its transition dipole moment, and then calculating the angle between  $\alpha$ 80/ $\beta$ 81 pairs, showed a difference of  $\sim 3^\circ$  wider for the  $\alpha$ 80/ $\beta$ 81 pairs located at a linker binding site. According to Förster theory<sup>17</sup>, faster energy transfer occurs over shorter distances and between collinear transition dipoles; therefore, the slight changes in  $\beta$ 81 orientation induced by the linker should be beneficial to the flow of energy in the APC trimer. The second overall structural effect of this counter-clockwise rotation of APC induced by L<sub>C</sub><sup>7.8</sup> association

is best observed along the side of the cylindrical APC trimer disk. On binding of the linker, the entire complex collapses slightly towards the pseudo-C3 axis and the APC ring becomes flattened.  $L_C^{7.8}$  does not penetrate the APC to emerge at the opposite side but Reuter *et al.* describe a variability of  $\sim 3 \text{ \AA}$  in its central insertion position<sup>1</sup>, and the flattening of the APC trimer is more pronounced when  $L_C^{7.8}$  is inserted deeper within the trimer structure. This action may provide better contact for the trimer-to-trimer interactions that form a hexamer stack.

There are specific structural changes in the side chain orientations of APC residues located at the interface region with the  $L_C^{7.8}$  linker. The schematic representations of Figures 3-6A to 3-6C compare the structure of each APC monomer at the interface surface with the linker peptide (APC and  $L_C^{7.8}$  are artificially separated in the schematic drawing by  $\sim 5 \text{ \AA}$ ). The APC  $\alpha$ -helix (residues 105-120) enclosing the  $\beta 81$  chromophore binding pocket is in direct contact with the  $L_C^{7.8}$  linker (Figure 3-5) and is nearly 100% conserved in APC from different species, although there is no obvious concentration of a specific residue type (electrostatically charged, aromatic, polar, nonpolar). Comparing this  $\alpha$ -helix segment from APC against other phycobiliproteins there are few amino acid replacements: the exact identity of D105 and E106 seems to be interchangeable in PC, PE and PEC thereby conserving the negative charge though not necessarily the specific residue; aliphatic V108 is frequently replaced by a cysteine in other biliproteins while K113 is replaced by an arginine in PC and some types of PEC; N117 is replaced by a variety of aliphatic or polar residues and lastly S118 is replaced by an alanine in other biliproteins. Of the five charged residues in the APC  $\alpha$ -helix segment, three are unperturbed by linker association (D105, E106, K113); their side

chains remain in the same positions in both the linker free monomer and the linker-coupled monomers. The other two charged residues (R107, E114) which are ~100% conserved across species and phycobiliproteins, directly face the central cavity and exhibit obvious changes in side-chain orientation (Figures 3-6A to 3-6C). Compared to the linker-free APC monomer, the positively charged side chain of R107 is twisted slightly counterclockwise in APC/L<sub>C</sub><sup>7.8</sup> ( $\alpha$ -helix), and it is distinctly bent towards APC (instead of towards the central cavity) in APC/L<sub>C</sub><sup>7.8</sup> (loop). This effect may be induced by the same mechanism: the side chain of a positively charged residue from the L<sub>C</sub><sup>7.8</sup> linker (K45 in  $\alpha$ -helix and R20 in loop) places itself near the oxygen group at the base of the APC R107, and the increased positively charged environment forces the R107 side chains to repel from this area and towards neighboring APC negatively charged residues. In contrast, E114 appears to be affected by the linker in different ways. In APC/L<sub>C</sub><sup>7.8</sup> ( $\alpha$ -helix), the E114 negatively-charged side chain is moved ~2 Å closer to the  $\beta$ 81 chromophore and is twisted (~ 20°) clockwise (using the carboxylate carbon as the center for the rotating oxygens); this E114 re-orientation could be a result of attraction to the nearby positively charged L<sub>C</sub><sup>7.8</sup> K52 residue (~3.0 Å away). In APC/L<sub>C</sub><sup>7.8</sup> (loop), the E114 side chain is bent back towards the APC and, because there is no obvious positively charged opponent in the nearby L<sub>C</sub><sup>7.8</sup> loop segment, we propose that this E114 conformational change is due to steric interference from the nearby L<sub>C</sub><sup>7.8</sup> Y26 residue. The position of this charged side chain may impact the spectral characteristics of the  $\beta$ 81 chromophore since E114 (along with K113) is located close to the propionate groups of  $\beta$ 81.

The APC surface interfacing with  $L_C^{7.8}$  has only one aromatic residue (Y116) and, like the charged D105 and E106 residues extending towards the cavity, it is ~100% conserved in phycobiliproteins. Y116 is located in the  $\beta 81$  chromophore binding pocket and is perturbed by the presence of the linker. Compared to the linker-free APC monomer, Y116 in both APC/ $L_C^{7.8}$  monomers is moved ~1 Å closer to  $\beta 81$  pyrrole ring C and in APC/ $L_C^{7.8}$  ( $\alpha$ -helix) adopts a spatial orientation that is more planar to this pyrrole. There is no obvious reason for this structural change, because there is no immediate polar or steric influence introduced by the linker; however  $L_C^{7.8}$  directly perturbs the orientation of a second aromatic residue within the  $\beta 81$  binding pocket (Y87), and this may have an indirect effect on Y116.  $L_C^{7.8}$  residue F37, extending from the  $\alpha$ -helix segment disrupts the coplanar arrangement of the Y87 aromatic side chain and the  $\beta 81$  pyrrole ring D by inserting itself between them and causing these two rings to move in opposite directions from each other. The change in the chromophore conformation may induce a concomitant change in the Y116 residue located 5 to 6 Å away from ring D.

There is a number of polar residues scattered throughout the APC linker-interface region (APC: N110, T115, N117, S118) which roughly correlate in position with those found in the  $L_C^{7.8}$  linker segments ( $L_C^{7.8}$  ( $\alpha$ -helix): N35, Q40, Q41, Q44; (loop): Q13, T14, T18, Q19). In both APC/linker interactions, the orientation of the N110 side chains is affected, while in APC/ $L_C^{7.8}$  (loop) the S118 side chain is also perturbed. There are no linker residues within H-bonding distance ( $< 4$  Å) of these two side chains; therefore the observed structural changes must be produced indirectly through modifications in the spatial orientation of other nearby residues.



In addition to structural changes in the APC surface closest to the linker binding site, the structural conformation of the individual  $\beta 81$  chromophores in each APC monomer is also altered by the presence of the  $L_C^{7.8}$  linker peptide. Figure 3-7A presents a schematic of the *in situ* structures of the three APC  $\beta 81$  chromophores; these have been drawn on top of each other, and the overlap has been optimized at the central methine bridge. The position of the propionate side chains of the  $\beta 81$  chromophore is noticeably variable, and the relative orientation of the pyrrole rings is also altered by linker association (Figure 3-7A). A face view of the pyrrole rings (top Figure 3-7A) shows that the B and C rings are nearly identical, while rings A and D differ in position. Looking along an axis perpendicular to the pyrrole plane we see that ring D is the most perturbed by the linker presence (bottom Figure 3-7A). The APC/ $L_C^{7.8}$  crystal structure shows that the displacement of this  $\beta 81$  pyrrole ring is due to an aromatic residue extending from the  $L_C^{7.8}$  linker (F37) that disrupts the coplanar arrangement of ring D with a nearby APC (Y87) residue.

#### *Comparison of APC and PC tertiary structures*

The tertiary structure of PC isolated from *Fremyella diplosiphon* was resolved at 1.66 Å and shows nine  $\alpha$ -helices connected by short loop segments in each subunit.<sup>18</sup> Although the PC structure from *Synechococcus* sp. PCC 7002 was the first to be resolved<sup>19</sup>, it is no longer available through the Brookhaven protein structure databank, and the *F. diplosiphon* structure is considered to be representative of the PC family. Note that while the overall homology score between the *Synechococcus* sp. PCC 7002 and *F. diplosiphon* PC monomer sequences is ~ 72%, it is nearly 100% conserved near the

chromophore binding pockets and in the  $\alpha$ -helices comprising the central cavity. These homology results are the same when comparing the PC sequences from *F. diplosiphon* and *M. laminosus* (structure also not available). Unlike APC, PC possesses three phycocyanobilin chromophores per monomer unit. The  $\alpha$ 84 and  $\beta$ 84 chromophores are attached to cysteine residues located in  $\alpha$ -helices close to the central cavity while the  $\beta$ 155 chromophore is attached to a cysteine located in one of the loop segments between helices on the periphery of the protein. Analogous to the APC  $\beta$ 81 chromophore,  $\beta$ 84 is physically located closest to the central cavity of the PC trimer aggregate.

The backbone structure of APC and PC is very similar; the largest variation occurs in the  $\beta$ -subunit of PC where 10 additional amino acids are placed in the loop segment enclosing the  $\beta$ 155 chromophore (Figure 3-8). The structure of the internal cavity region is extraordinarily well-conserved among these two biliproteins despite being produced by different species. The linker-interface region identified in the APC/L<sub>C</sub><sup>7,8</sup> crystal structure is compared with the analogous area in PC (residues 108-123). The sequence homology between APC and PC in this  $\alpha$ -helix segment is 78% (Figure 3-8); all of the charged and aromatic residues are conserved, but two polar residues in APC (N117, S118) are replaced by aliphatic ones (L120, A121) in PC. Like in APC, the critical side chains that extend towards the PC central cavity belong to D109, R110, N113 and E117 (labeled in Figure 3-8); as shown in the APC/L<sub>C</sub><sup>7,8</sup> structure, the position of these side chains is modulated by the presence of the linker peptide; therefore the slight orientation differences observed between PC and APC are probably not indicative of inherently different linker-association behavior. Potentially more significant structural differences are the loss of the polar residues near the bottom end of the PC  $\alpha$ -

helix (noted above), and the lack of an aromatic residue in a position analogous to APC's Y87 (whose orientation was dramatically altered by an extending aromatic residue from LC<sup>7,8</sup>). Overall, seven of the total thirteen aromatic residues found in the APC  $\beta$ -subunit are directly conserved in PC; however, the area around the APC  $\beta$ 81 D-ring is endowed with a wealth of tyrosines that is not duplicated in PC  $\beta$ 84 environment. This could be one of the key structural differences leading to the different spectral properties displayed by PC and APC despite containing chemically identical chromophores. It would be interesting to see if genetic deletion or alteration of either these aromatic residues, especially APC Y87, or of the polar groups N117 and S118, would produce an APC mutant that has an absorbance band shape that is more symmetric (or "PC-like") than the native peak/shoulder profile.

A number of charged amino acids located in the PC chromophore binding pockets have been identified as important modulators of the chromophore spectral properties. Arginine (R86) and aspartate (D87) residues form salt bridges with the propionate side chains of the  $\alpha$ 84 and  $\beta$ 84 chromophores.<sup>16,20,21</sup> The residue R80, near the tail end of  $\beta$ 84 (~3 Å away) is also suspected of tuning the spectra of this chromophore.<sup>12</sup> In addition, aspartate residues providing structural stability to the chromophore have been identified as a common feature of bilin attachment: a single aspartate anchors a phycocyanobilin through a hydrogen bond between its carboxylate group and both nitrogen atoms of pyrrole rings B and C.<sup>22</sup> All of these key amino acid residues are also conserved in the APC chromophore binding pockets and, despite these strong influential structural similarities among chromophore binding environments, PC and APC display very different spectral characteristics.

It is the subtle differences in either chromophore conformation or environment that contribute to the distinct spectroscopic properties of each PC chromophore.<sup>15,16,23,24</sup> Figure 3-7B compares the chromophore conformation of linker-free APC  $\beta$ 81 and PC  $\beta$ 84 superimposed and centered at the central methine bridge; the top part shows a view of the tetrapyrrole plane while the bottom part of the figure shows an orthogonal perspective. From the top view of Figure 3-7B, one can see that there is good agreement between central rings B and C and that larger variations occur at the terminal rings A and D. The orthogonal view of the tetrapyrrole plane shows that there is slight twist of the central methine bridge in the APC chromophore that lifts ring C from the plane (bottom Figure 3-7B). There are slight variations throughout the entire conformation of these two chromophores and these variations are most severe at ring D. The APC/LC<sup>7.8</sup> ( $\alpha$ -helix) complex shows that this ring can be influenced by linker association. This suggests that, along with the labile propionate side chains, ring D is the most likely interaction site for protein modulation of the chromophore spectral properties.

#### *Comparison of linker characteristics*

The structure of the LC<sup>7.8</sup> linker contains three  $\beta$ -strands (residues L3-L9, Y26-P32 and K49-L55) and two  $\alpha$ -helices (L22-T25 and Y33-M46). As discussed in the previous section, the linker interacts with the APC  $\beta$ 81 chromophores in two distinct manners. In APC/LC<sup>7.8</sup> ( $\alpha$ -helix), the linker induces a noticeable change in the  $\beta$ 81 chromophore conformation, while in APC/LC<sup>7.8</sup> (loop) it conceals the chromophore pocket from the external solvent with a loop segment that contains a number of positively charged residues. Within this LC<sup>7.8</sup> loop segment, the closest arginine/chromophore

distance is found between the R20 side chain nitrogens and the  $\beta$ 81 oxygen at the end of pyrrole ring D ( $\sim 4.2$  Å and  $\sim 4.7$  Å). Although slightly outside van der Waals contact distance, the combined effect of R20 along with R17 (located  $\sim 5.3$  Å and  $\sim 6.6$  Å from  $\beta$ 81 ring C propionate groups) may attract the chromophore further towards the central cavity without altering its conformation.

Examination of the primary and predicted secondary structures of the rod linker peptides reveals possible sites for chromophore interactions analogous to those of the  $L_C^{7.8}$  linker. In the  $\alpha$ -helix segment of  $L_C^{7.8}$  there are three aromatic residues (Y33, W36 and F37), and it is F37 which extends to interfere with the chromophore planar arrangement. Although this area is homologous with  $L_R^{8.9}$  (Figure 3-1), the only aromatic residue that is conserved is Y33 (Y48 in  $L_R^{8.9}$ ) located at the beginning of the  $\alpha$ -helix oriented opposite the APC biliprotein and thus physically unlikely to interfere with the chromophore. This type of interaction would then be expected to be absent in  $L_R^{8.9}$ . The second type of linker/chromophore interaction is a good candidate in  $L_R^{8.9}$ , because this linker possesses three well-conserved arginine residues (R24, R37 and R38) in a predicted loop region that could interact in a combined manner similar to the active  $L_C^{7.8}$  loop.

Both  $L_R^{32.3}$  and  $L_{RC}^{28.5}$  contain a number of aromatic residues found in regions structurally predicted as  $\alpha$ -helix (Figures 3-4C and 3-4D). These residues could very well orient themselves to form a network of extended "aromaticity" that reinforces the binding interaction, as found in the PC hexamer structure. However, we focus our attention on choosing specific interactions between the linker and the PC chromophores that could produce the observed spectral changes.  $L_R^{32.3}$  has four phenylalanines (F61,

F71, F99, F168) and five tryptophans (Y39, Y81, Y120, Y132, Y138), while L<sub>RC</sub><sup>28.5</sup> has nine phenylalanines (F41, F53, F54, F57, F79, F89, F93, F102, F136), two tryptophans (Y49, Y146) and one histidine (H55) in structurally helical regions. Of these aromatic residues, the ones homologous to both L<sub>R</sub><sup>32.3</sup> and L<sub>RC</sub><sup>28.5</sup> (L<sub>R</sub><sup>32.3</sup> F71, Y39, Y81, Y138 and L<sub>RC</sub><sup>28.5</sup> F79, Y49, F89, Y146) can be discounted as interaction sites, because each linker has very different effects on the spectroscopy of the PC chromophores and therefore cannot be interacting through the exact same residue. Secondly, residues that are located at the edges of the  $\alpha$ -helix regions (L<sub>R</sub><sup>32.3</sup> Y132, F61 and L<sub>RC</sub><sup>28.5</sup> F93, F102, F136) can be discounted since these areas are not reliably predicted; also the L<sub>C</sub><sup>7.8</sup> F37 residue is located in the middle of its long  $\alpha$ -helix segment. This leaves three viable candidates for aromatic interaction between the L<sub>R</sub><sup>32.3</sup> linker and a PC chromophore (F99, Y138, F168; Figure 3-4C bold-face type) of which Y138 and F168 are ~100% conserved across species (Table 3-3B). L<sub>RC</sub><sup>28.5</sup> retains five viable aromatic residues (F41, F53, F54, F57, H55) which are clustered together in the first  $\alpha$ -helix segment (Figure 3-4D bold-face type). Only one of these, F53, is 100% conserved across species (Table 3-3C).

Both large rod linkers also contain loop regions with multiple arginine residues that could resemble the structural arrangement found in the L<sub>C</sub><sup>7.8</sup> linker. In L<sub>R</sub><sup>32.3</sup> the loop sequence (A245-R259) borders on the homologous region with L<sub>R</sub><sup>8.9</sup> and contains five arginine residues (R250, R253, R255-256, R259) that are clustered together and are highlighted in bold type in Figure 3-4C. Although all of these are outside of the area homologous with L<sub>C</sub><sup>7.8</sup> (Figure 3-3), the sequential pair R255-256 are so well-conserved

across species (Table 3-3B) that they are likely to be involved in some function. In  $L_{RC}^{28.5}$  there is a long loop segment after the last  $\alpha$ -helix that contains a number of arginines (R159-161, R166, 179, R185, R200-201, R203, R207 and R222) with two short cluster areas (159-166 and 200-207) that are likely to interact with a chromophore. The arginines in these two areas are highlighted in Figure 3-4D and except for R201, R207 and R222 they are all well-conserved (Table 3-3C). Additionally, the second  $L_{RC}^{28.5}$  cluster (200-207) is included in the region homologous to  $L_C^{7.8}$  (Figure 3-3) and there is direct conservation of R203 and R207. In the primary structure discussion above, it was noted that experiments with short  $L_{RC}$  sequences produced the same characteristic shifts in the biliprotein spectra as the whole linkers.<sup>7,8</sup> Because this observation excludes about half of the arginine residues in the  $L_{RC}^{28.5}$  loop segment, it indicates that this type of interaction between the linker and the biliprotein chromophores is not likely to be the major mechanism for  $L_{RC}^{28.5}$ .

### Concluding remarks

The APC/ $L_C^{7.8}$  crystal structure has shown that the linker penetrates the central APC cavity to interact with each of the three monomers in distinct manners. To bind at a position near the  $\beta 81$  chromophore pocket, and thus effectively modulate its spectra, the linker must penetrate at least  $\sim 12$  Å within the APC trimer.  $L_C^{7.8}$  association induces a counter-clockwise rotation of each APC monomer that results in bringing the central  $\beta 81$  chromophores closer together and could be beneficial to energy transfer. The orientation

of specific APC amino acid side chains at the surface interfacing with the  $L_C^{7.8}$  linker, as well as the conformation of the  $\beta 81$  chromophore, are also affected by linker association.

Each rod linker sequence in this study was examined for sequence homology with other proteins, and evolutionarily conserved regions were identified. Coupled with secondary structure predictions, these conserved regions provided a wealth of candidates for functional sites. The analysis showed that both  $L_{RC}^{28.5}$  and  $L_R^{32.3}$  possess separate regions of homology with other linker types as well as an N-terminal region of shared homology with each other. This N-terminal region (~2/3 of the whole sequence) is responsible for modulating the spectral properties of the PC trimer<sup>10,11</sup>, yet the entire  $L_R^{32.3}$  linker sequence is protected from trypsin digestion when the linker joins two PC hexamers<sup>11</sup>. Because there is also no symmetry or repeating segments of residues in the  $L_R^{32.3}$  sequence, these observations support the idea of two distinct biliprotein-binding mechanisms for a single linker: one that produces the observed spectral changes upon linker-biliprotein association and a second that could provide structural support without influencing the chromophore spectra. On a more microscopic level, specific residues in PC and the rod linkers were selected for biliprotein/linker binding interactions analogous to those found in the APC/ $L_C^{7.8}$  structure. Table 3-4 identifies the critical residues that may impact the biliprotein's spectral characteristics.

One general issue of biliprotein/linker association has not yet been considered: the physical bulk of the associated linkers in the biliprotein central cavity. The rod linkers  $L_R^{32.3}$  and  $L_{RC}^{28.5}$  are significantly larger than the  $L_C^{7.8}$  linker. It is unknown whether the linkers interact with each other through the PC hexamer's central cavity. Such an arrangement would provide optimal structural support to the rod assembly and could thus



enhance the transfer of energy; however, to minimize direct influence upon the central  $\beta 84$  chromophores, the non-modulating linker segment should not penetrate further than  $\sim 12$  Å. A study of crystallized B-phycoerythrin hexamers of *Porphyridium sordidum* in the presence and absence of a  $\gamma$ -subunit occupying the central space of the hexamer showed nearly identical physical positions for the amino acids directed towards the central area; only a decrease in their mobility was detected.<sup>20,25</sup> Although the  $\gamma$ -subunit was not resolved, it is approximately the same size as the  $L_R$ <sup>32,3</sup> linker peptide so one can expect that the physical bulk of the peptide will not greatly stress the structure of the central cavity. In addition, the PC  $\beta 84$  chromophore binding pocket seems to be strategically designed to foster specific interactions that can only be made through deeper penetration within the protein. Although  $\beta 84$  is exposed to the central cavity, its propionate side chains are pointed “downwards” towards the next PC trimer. Electrostatically charged residues are scattered throughout the protein, but there are concentrations near each chromophore site and at  $\beta 84$ , they occupy the side furthest from the solvent exposed surface. If a linker were to approach with either similar or oppositely charged residues, it would interfere with these neighboring interactions and effectively modulate the chromophore spectra. To accomplish this specific interaction, it would have to penetrate the PC cavity more fully.

### Summary observations

- There is no universal binding site among all three *Synechococcus* sp. PCC 7002 rod linkers.
- Except for L<sub>R</sub><sup>8.9</sup>, the rod linker peptides have high calculated pI values and are expected to exhibit hydrophobic behavior at pH 7.0.
- L<sub>R</sub><sup>32.3</sup> has a region homologous with L<sub>RC</sub><sup>28.5</sup> near its N-terminus and a region homologous with L<sub>R</sub><sup>8.9</sup> near its C-terminus.
- L<sub>R</sub><sup>32.3</sup> and L<sub>RC</sub><sup>28.5</sup> share a homologous sequence region near their N-terminus which is also shared with a small REP domain of the core anchor peptide L<sub>CM</sub>.
- Both L<sub>R</sub><sup>32.3</sup> and L<sub>RC</sub><sup>28.5</sup> possess two domains; one modulates the spectral characteristics of the PC chromophores and the other may help to form a skeletal backbone for the phycobilisome rod.
- In each rod linker sequence, conserved aromatic residues in predicted  $\alpha$ -helical segments as well as clusters of positively charged residues within loops segments were identified for possible linker active sites analogous to those found in L<sub>C</sub><sup>7.8</sup>.
- Comparative analysis of PC and the rod linkers to the APC/L<sub>C</sub><sup>7.8</sup> structure identified a number of analogous and conserved residues that could impact the spectral properties of the biliprotein complexes; these are summarized in Table 3-4. Genetic alteration of these sites could provide a great deal of insight for the biliprotein/linker interaction.

**References for Chapter 3**

1. Reuter, W., Wiegand, G., Huber, R., Than, M.E. (1999) *Proc. Nat. Acad. Sci. USA*, 96:1363-1368.
2. Glauser, M., Stirewalt, V.L., Bryant, D.A., Sidler, W., Zuber, H. (1992) *Eur. J. Biochem.*, 205:927-973.
3. Füglistaller, P., Suter, F., and Zuber, H. (1985) *Biol. Chem. Hoppe-Seyler*, 366:993-1001.
4. Eberlein, M. and Kufer, W. (1990) *Gene*, 94:133-136.
5. deLorimier, R., Bryant, D.A., Stevens Jr., S.E. (1990) *Biochim. Biophys. Acta*, 1019:29-41.
6. personal communication with W. Schluchter and D.A. Bryant (1996).
7. Schluchter, W.M., Bryant, D.A. (1992) *Biochem.* 31:3092-3102.
8. Fillat, M.F., Flores, E., Gómez-Moreno, C. (1993) *Plant Mol. Biol.* 22:725-729.
9. unpublished results, personal communication with D.A. Bryant, Pennsylvania State University (1997).
10. Gottschalk, L., Fischer, R., Lottspeich, F., Scheer, H. (1991) *Photochem. Photobiol.* 54:283-288.
11. Yu, M.-H., Glazer, A.N. (1982) *J. Biol. Chem.* 257:3429-3433.
12. Glauser, M., Stirewalt, V., Bryant, D., Sidler, W., Zuber, H., (1992) *Eur. J. Biochem.* 205:927-937.
13. Salamov, A.A., Solovyev, V.V. (1995) *J. Mol. Biol.* 247:11-15.
14. Cuff, J.A., Clamp, M.E., Siddiqui, A.S., Finlay, M., Barton, G.J. (1998) *Bioinformatics* 14:892-893.

15. Scharnagl, C., Schneider, S. (1989) *J. Photochem. Photobiol. B: Biol.*, 3:603-614.
16. Scharnagl, C., Schneider, S. (1991) *J. Photochem. Photobiol. B: Biol.*, 8:129-157.
17. Förster, T. (1967) In Comprehensive Biochemistry, eds. Florkin, M. and Stotz, E.H., Elsevier, Amsterdam, Vol 22, pp 93-137.
18. Duerring, K.M., Schmidt, G.B., Huber, R. (1991) *J. Mol. Biol.* 217:577-592.
19. Schirmer, T., Bode, W., Huber, R. (1987) *J. Mol. Biol.* 196:677-695.
20. Mimuro, M., Rübelen, R., Füglistaller, P., Zuber, H. (1986) *Biochim. Biophys. Acta* 851:447-456.
21. Betz, M. (1997) *Biol. Chem.* 378:167-176.
22. Kikuchi, H., Sugimoto, T., Mimuro, M. (1997) *Chem. Phys. Letters* 274:460-465.
23. Debreczeny, M.P., Sauer, K., Zhou, J., Bryant, D.A. (1993) *J. Phys. Chem.* 97:9852-9862; Debreczeny, M.P. (1994) Ph. D. Thesis, LBL-35672, University of California, Berkeley, CA, 72-102.
24. MacColl R., Guard-Friar, D. (1987) Phycobiliproteins, CRC Press, Boca Raton, FL.
25. Ficner, R., Lobeck, K., Schmidt, G., Huber, R. (1992) *J. Mol. Biol.* 228:935-950.

Table 3-1. Amino acid sequences of *Synechococcus* sp. PCC 7002 rod linker peptides

Linker	SWISS-PROT database AC number	Number of Residues	Sequence				
L <sub>R</sub> <sup>8.9</sup>	P31966	80	MLSQFANGTE MNQEMQRILR	AASRVFTYEV LGGKIVSIKP	QGLRQTEETD YTGATASDEE	NQEYAFRRSG	SVFINVPYAR
Reference: De Lorimier R., Bryant D.A., Stevens S.E. Jr. (1990) <i>Biochim. Biophys. Acta</i> 1019:29-41.							
L <sub>R</sub> <sup>32.3</sup>	Q05237	290	MPVTVAASRL SERLTAAESL HLLGRAPYDE RSYVVQTGHR VVGPSGVNEG YPRIRRSSRV	GTAAFDQSPV FTNGFISVRD AEVIEHLDRY TVGFTRMFSL WAFRSAADDY FFVPVSRLSQ	ELRANYSRDD FVRAVAQSEL QNEGFADIN QRGYANSRA HPGQSLGGST KLQEIQRMG	AQTVIRAVYR YKEKFLYNNF SYIDSAEYTE QIAGNASRLA GLSADDQVVR RVASISPAGQ	QVLGNDYVMS QTRVIELNFK NFGDNIVPYI QELARNTTSA VEVAALSTPR
Reference: De Lorimier R., Guglielmi G., Bryant D.A., Stevens S.E. Jr. (1990) <i>Arch. Microbiol.</i> 153:541-549.							
L <sub>R</sub> <sup>28.5</sup>	Q05238	247	TIPLLQYAPS QIFFHAFKCD RFVEQCIQRV GYDTVYQRR RFRTPDVVPQ	SONTRVAGYT RQQLLESQLR LGRDPFSEQE RSLASREQGE AGDPALFLNM	VGGDEQPFVF NGQITVRDFI KIAWSIVICT IPFNIKSPRY ARSAQIPKVN	TTDNVISDSD RGLLLSETFI KGLAAFVDQL DAYRSQJGF VRVSAADISL	FDVLINAAYR DSFYNKNSNY LNTDEYMENF PQVWQNAVR AAVPYRN
Reference: Bryant D.A., De Lorimier R., Guglielmi G., Stevens S.E. Jr. (1990) <i>Arch. Microbiol.</i> 153:550-560.							

Table 3-2. Statistics of amino acid composition of the linker peptides

Linker peptide		LC <sup>7.8</sup>	LR <sup>8.9</sup>	LR <sup>32.3</sup>	LRC <sup>28.5</sup>
Calculated molecular weight (Da)		7,741	9,055	32,273	28,380
Calculated isoelectric point <pI>		10.44	5.42	6.93	8.39
Amino acid composition (%)					
Aliphatic side chains					
Ala	A	4.5	8.8	10.3	7.7
Ile	I	7.5	5.0	4.1	6.1
Leu	L	7.5	5.0	6.6	7.3
Val	V	6.0	6.3	9.3	7.7
Aromatic side chains					
His	H	0	0	1.4	0.4
Phe	F	4.5	5.0	4.8	6.9
Trp	W	1.5	0	0.3	0.8
Tyr	Y	3.0	5.0	4.8	4.9
Nonpolar side chains					
Cys	C	1.5	0	0	1.2
Gly	G	9.0	7.5	6.9	4.5
Met	M	1.5	3.8	1.4	0.8
Pro	P	3.0	2.5	3.5	5.3
Polar side chains					
Asn	N	4.5	5.0	4.8	5.7
Gln	Q	10.5	7.5	5.5	7.7
Ser	S	1.5	7.5	9.3	6.9
Thr	T	10.5	7.5	4.8	4.5
Electrostatically charged side chains					
Asp	D	0	2.5	5.2	6.5
Glu	E	6.0	10.0	5.9	4.1
Arg	R	9.0	8.8	9.7	8.9
Lys	K	9.0	2.5	1.4	2.4
Total %	Nonpolar*	49.5	48.9	52.0	53.2
	Polar	27.0	27.5	24.4	24.8
	(-) Charged	6.0	12.5	11.1	10.6
	(+) Charged	18.0	11.3	11.1	11.3
	Aromatic	9.0	10.0	11.3	13.0
	Hydrophobic**	27.0	25.1	26.2	28.8

\* includes nonpolar, aliphatic and aromatic groups (except for His)

\*\* includes I, L, M, V, F

**Table 3-3A.** Fast-A3 results of SWISS-PROT database search for sequence homology (>25%) with L<sub>R</sub><sup>8.9</sup>

**Bold-face type residues** represent direct homology matches to the query sequence

**■** represents residues involved in chromophore interactions discussed in the text

*Code key to abbreviations used in Tables 3A-3D*

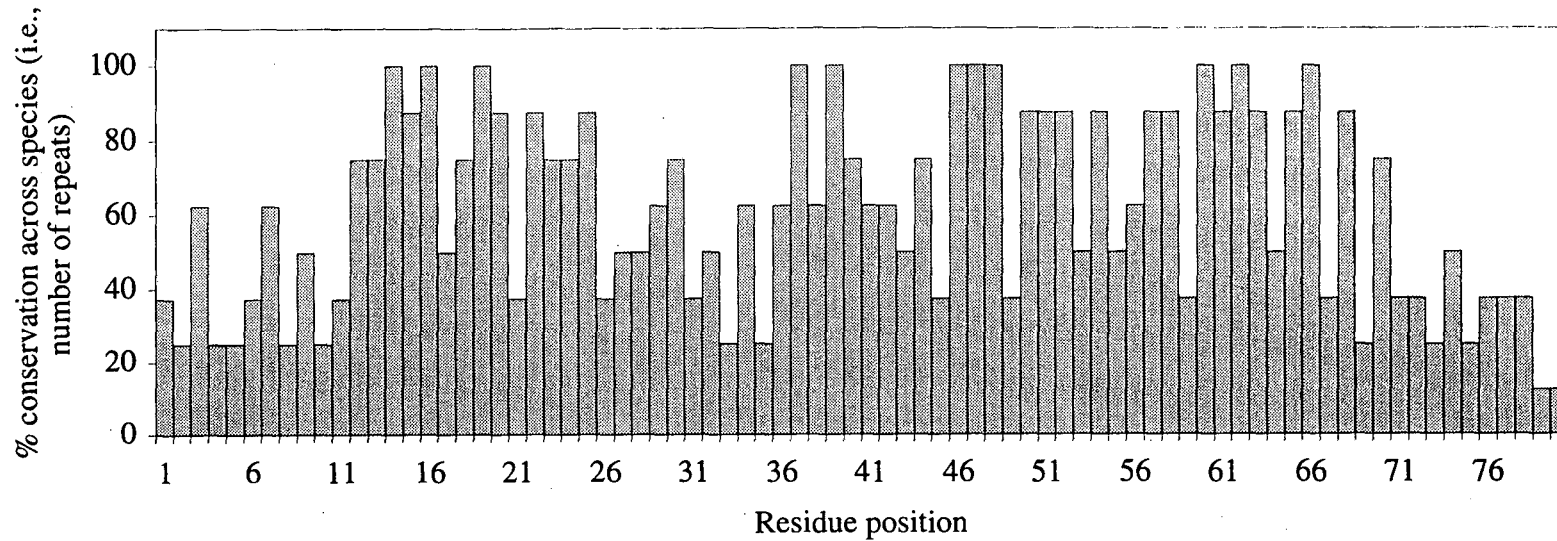
<u>Code</u>	<u>Gene name</u>	<u>Protein expressed</u>	<u>Code</u>	<u>Organism</u>
APCE	APCE	core-membrane linker L <sub>CM</sub> (APC)	AGLNE	Aglaothamnium neglectum
FENR	PETH	Ferredoxin-NADP <sup>+</sup> reductase enzyme	ANASO	Anabaena sp PCC 7119
PHEG	MPEC	Phycocerythrin-associated rod linker	ANASP	Anabaena sp. PCC 7120
PYC1	APCC	Allophycocyanin-associated core linker	ANAVA	Anabaena variabilis
PYG	CPCG	Rod-core linker (CPC and APC)	CYACA	Cyanidium caldarium; Galdieria sulphuraria
PYG1	CPCG1	“ “ “	FREDI	Fremyella diplosiphon; Calothrix sp. PCC 7601
PYG2	CPCG2	“ “ “	MASLA	Mastigocladus laminosus; Fischerella sp.
PYG3	CPCG3	“ “ “	PORPU	Porphyra purpurea
PYG4	CPCG4	“ “ “	PSAE9	Pseudanabaena sp. PCC 7409
PYR1	CPCC	C-Phycocyanin-associated rod linker	SPIPL	Spirulina platensis
PYR2	PECC	Phycocerythrin-associated rod linker	SYNEL	Synechococcus elongatus
PYR3	CPCI2	C-Phycocyanin-associated rod linker	SYNP2	Synechococcus sp. PCC 7002; Agmenellum quadruplicatum
PYR4	CPCI3	“ “ “	SYNP6	Synechococcus sp. PCC 6301; Anacystis nidulans
PYR5	CPCH2	“ “ “	SYNPY	Synechococcus sp. WH8020
PYR6	CPCH3	“ “ “	SYNY3	Synechocystis sp. PCC 6803
PYS1	CPCD	“ “ “	SYNY4	Synechocystis sp. PCC 6714
PYS2	CPCD3	“ “ “	SY9314	Synechocystis sp. PCC 9314

*E.g. PYS1\_SYNP2 is a C-Phycocyanin-associated rod linker from Synechococcus sp. PCC 7002*

Matching sequence	Score (%)	Query sequence (rod linker Lr <sup>8,9</sup> from <i>Synechococcus</i> sp. PCC 7002)
PYS1_SYN2	100	MLSQFANGTEAASRVFTYEVQGLRQTEETDQEQYAFRRSGSVFINVPYARMNQEMQRIILRLGGKIVSIKPYTGATASDEE
PYS1_SYNEL	62.2	MLSQFANGTEAASRVFTYEVQGLRQTEETDQEQYAFRRSGSVFINVPYARMNQEMQRIILRLGGKIVSIKPYTGATASDEE
PYS1_ANASP	56.6	MFGQTAGSAAALSIYFRYEVVGLRQNEETDQEQYAFRRSGSVFTIIVPYNRMNDEMRRISRLGGTIVNIRPYQADSNEQN-
PYS1_MASLA	55.3	TTLGAGSVSSASRVFRYEVVGLRQSSSETDKKNIYINRSGSVFTIIVPYNRMNDEMRRISRLGGTIVNIRPYQADSNEQN-
PYS1_SYN3	54.4	TTLGIDSVSSASRVFRYEVVGLRQNEENDKKNIYINRSGSVFTIIVPYNRMNDEMRRISRLGGTIVNIRPYQADSNEQN-
PYS1_SYN6	53.8	SSLVGYNTQAANRVFVYEVVGLRQTDANENSADHDIRRSGSVFIKVPYARMSEEMQRIHRLGGKIVKIEPLTRAAG---
PYS1_FREDI	45.0	SALVGARSGAVASRVFVYEVVGLRQSDQTDLNRYPVRSQTVQFVYARMNEELRRINRLGARIVSIKPLT-AEAE---
PYS1_SYN6	26.0	SVLTTRSSSSGSDNRVYEVVGLRQNEQTDNRYQIRNSSTIEIQVPSYRMNEEDRRITRLGGRIVNRIRPAGENPTEDAS
PYS2_FREDI	48.6	PWVARRPFGQTASRIFRVEIAAISKPGEP-----SVRRSNRSL-VPYEQLNNTLQQVNRSGGRVSVAPASL-----
FENR_ANAVA	56.2	-----MVIQSRSFQVEVGLHQNQVETNQNNYPIRSGSVFTIIPFSRFEELQRINRLGGKIVNIQPLNLQINEN---
FENR_ANASO	55.0	GAFEGAANVESGSRVYEVVGMKQNEETDQTNYPINKSGSVFIRVYPNRMNDEMQRITRMGGKIVSITPVVAS---
FENR_SYN3	50	GAFDGAANVESGSRVYEVVGMKQNEETDQTNYPINKSGSVFIRVYPNRMNDEMQRITRLGGKIVSIQTVSALQQLNGK
FENR_SYN2	47.5	YVATSSRQSDACNRLFVYEVVGLRQSSSTMTDGLDYPINKSGSVFTIIVPLKRMNQEMRRITRMGGKIVSIKPLEGDSPLPHT
PYR1_FREDI	39.4	STANSTGNQSYANRIFIYEVVGLRQDGRNENS--LVKRSQTTFTIIVPYNRMNDEMQRITKLGGKIVSIRPAEDAAQIVSE
PYR1_ANASP	33.8	ARTRLGVDASAGGKVYRIEVTGYRA--KTFNNISKFRRSNQVFL-VPYEKLSQYQRIHQQGGVIASTPV-----
PYR1_MASLA	35.2	PKNLGNVAVGEGDRVYRIEVTGIRSPGYP-----SVRRSSTVFI-VPYERLSDKIQQVHKQGGKIVSITSA---
PYR1_SYNEL	29.6	PRKTLGNVAVGENDRVYRIEVTGIVT-----SPGYPVRRSSYAIIVPYERLSEKIQQIHKLGGKIVSITSA---
PYR1_SYN2	26.9	PRACLGGSFGESGRVYRIEVAIGIQPGYP-----GVRSSSTAFI-VPYEQLSAKMQQLQRTGARII SVNPA---
PYR1_SYN3	27.9	-----SLCGSTGLSADDQVVRVEVAALSTPRYPRIARRSRVFF-VPVSRLSQKLQEIQRMGGRVASISPAQQ---
PYR1_SY9413	32.1	-----ALGGTVFPFQAASKLFRVEITAIASAPGYPKVRRSNKAVI-VPFEQLNQTLQQINRLGGKIVASITPASLS---
PYR2_MASLA	29.2	-----GPHRLYRIEIAAGPNLPRYFRVRRVNDVIVPYERLSEEMORINRQGGRISSITPV-----
PYR2_FREDI	25.8	ASTSSARTLVTSPVMGDARMFIVEAIAAGTLNTNVAARRSRQVY-TVPYDRLSATYQEIHKRGGKIVKITPAS---
PYR3_FREDI	28.8	-----AIKASSIVIGSGTEKRFKILLVQGSKFDSPRRISTTEYIVPASKMTPQIQRIINRITSGKIVSITTEIV---
PYC1_SYN2	21.5	-----RTFGRSSQGSTPRLYRIEVTGIVSLPRYPKVRSNKEFI-VPYEQLSSTLQQINKLGGKIVASITFAQ---
		-----MRMFKITACVPSQSRIRKTRQRELQNTYFTKLVPYDWNWFEREQRIIMKMGKIVKVQLATGKPGTNTG



**Comparison of PYS1 sequences ( $L_R^9$  linkers only) for evolutionarily conserved residues**



**For PYS1 sequences only:**

Sequence length = 80 residues

Average overall sequence homology score = 50.5%

Number of residues in the consensus sequence with greater than 80% conservation across species = 34 (or 42.5% of 80)

Average % conservation across species for central amino acids (residues 12-70) = 70.8%

Table 3-3B. Fast-A3 results of SWISS-PROT database search for sequence homology (>25%) with L<sub>R</sub><sup>32.3</sup>

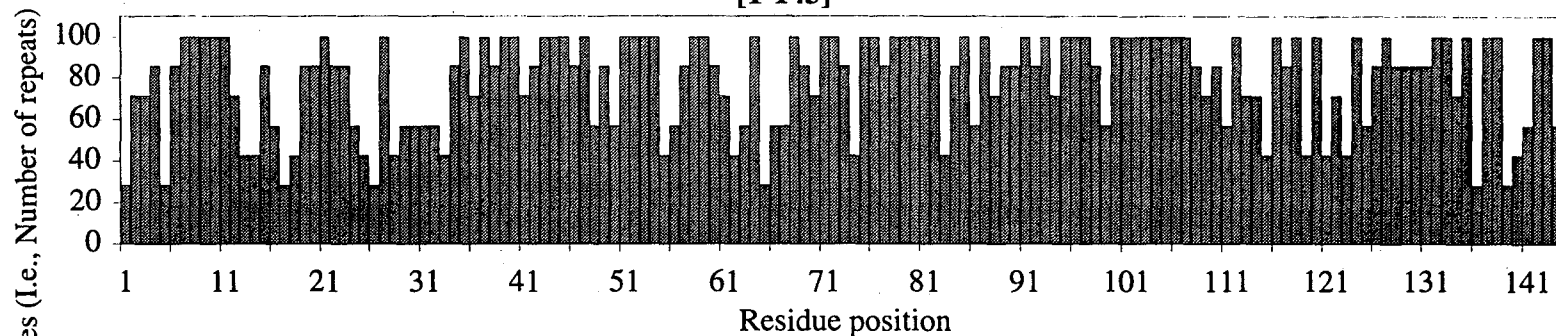
Matching sequence	Score (%)	Query sequence (rod linker L <sub>R</sub> <sup>32.3</sup> from <i>Synechococcus</i> sp. PCC 7002 residues 1-80):
		MPVTVAASRLGTAAFDQSPVELRANYSRDDAQTIVIRAVYRQVLGNDYVMSERLTAAESLFTNGFISVRDFVRAVAQSEL
PYR1_SYNP2	100	MPVTVAASRLGTAAFDQSPVELRANYSRDDAQTIVIRAVYRQVLGNDYVMSERLTAAESLFTNGFISVRDFVRAVAQSEL
PYR1_SYNY3	63.0	-AITTAASRLGVAPYNE <span style="font-variant: small-caps;">SPVELR</span> PDFSLDDAKM <span style="font-variant: small-caps;">VIRAVYRQVLGNDY</span> IMD <span style="font-variant: small-caps;">SERLKGAESLLTNGSISVREFV</span> RTVAKSEL
PYR1_SYNEL	62.7	-AITAAASRLGTS <span style="font-variant: small-caps;">AFDAPPVELRAN</span> WSEEDLE <span style="font-variant: small-caps;">TVIRAVYRQVLGNDYVMA</span> SERLVS <span style="font-variant: small-caps;">AESLLRNGKITVREFV</span> RAVAKSEL
PYR1_SY9413	58.3	MAITSAASRLGTS <span style="font-variant: small-caps;">AFDASPVELRPN</span> F <span style="font-variant: small-caps;">SQNDAKAVISAVYRQLGNDY</span> IM <span style="font-variant: small-caps;">ASERLTSAESLLCNGTITVKDFV</span> SLVAKSEL
PYR1_MASLA	57.8	-AITAAASRLGTE <span style="font-variant: small-caps;">PFSNAa</span> i <span style="font-variant: small-caps;">ELRSDASREEVEAVINAVYRHV</span> LGNDYIM <span style="font-variant: small-caps;">ASERLVS</span> AESLLRDG <span style="font-variant: small-caps;">NLTVREFV</span> RSVAKSEL
PYR1_ANASP	56.8	-AITTAASRLGTE <span style="font-variant: small-caps;">PFS</span> DAP <span style="font-variant: small-caps;">VELRPKASREEVESVIRAVYRHV</span> LGNDYIL <span style="font-variant: small-caps;">ASERLVS</span> AESLLRDG <span style="font-variant: small-caps;">NLTVREFV</span> RSVAKSEL
PYR1_FREDI	42.5	-MPFGPASRLGVSLFD <span style="font-variant: small-caps;">ETPVEWV</span> PGRS <span style="font-variant: small-caps;">QEEAETIRAIYRQVLGNA</span> YVME <span style="font-variant: small-caps;">SERLAVPESQ</span> FKRGEL <span style="font-variant: small-caps;">SVREFV</span> RAVAKSEL
PYR2_SYNY3	49.7	MTSLVSAQRLGIVAVDEap <span style="font-variant: small-caps;">LELRSRSTEE</span> EVDAVILAVYRQVLGNDH <span style="font-variant: small-caps;">LMSQERLTS</span> AESLLRGREISVRDFVRAVALSEV
PYR2_ANASP	41.6	--SSSVAERLAIRDAIGNK <span style="font-variant: small-caps;">VELRQNW</span> SEDDLQ <span style="font-variant: small-caps;">VFR</span> AAYEQIFGRQGIYASQ <span style="font-variant: small-caps;">KFTSAE</span> ALLRNGKISVRQFVEILAKSEF
PYR2_MASLA	41.5	--STSV <span style="font-variant: small-caps;">AERLAIK</span> DEV <span style="font-variant: small-caps;">DKKIELRPN</span> WSEDE <span style="font-variant: small-caps;">LQIVFKTAYEQV</span> FGROGLYASQ <span style="font-variant: small-caps;">RFATAE</span> ALLRNGKISV <span style="font-variant: small-caps;">KQFIE</span> LLAKSEF
PYR2_FREDI	38.4	-----MASQTILELWPSS <span style="font-variant: small-caps;">SLEEVQTI</span> RAVYKQVLG <span style="font-variant: small-caps;">NPHVM</span> ESERL <span style="font-variant: small-caps;">VTAESQL</span> CDRSITVREFVRSVAKSDF
PYR3_FREDI	61.8	-PITSAASRLGTTAYQ <span style="font-variant: small-caps;">TNP</span> IELRPN <span style="font-variant: small-caps;">WTAEDA</span> KI <span style="font-variant: small-caps;">VIQAVYRQVLGNDY</span> LM <span style="font-variant: small-caps;">QSERLTS</span> LESLLTNGKLSVRDFVRAVAKSEL
PYR4_FREDI	52.1	MPITTAASRLGTSaf <span style="font-variant: small-caps;">nAAP</span> IELRSNT <span style="font-variant: small-caps;">NKAEIAQVIA</span> AIYRQVLGNDYV <span style="font-variant: small-caps;">LQ</span> SERL <span style="font-variant: small-caps;">KGL</span> ESLLTNGNITVQ <span style="font-variant: small-caps;">EFV</span> ROLAKSNL
PYR5_PSEA9	57.7	MTSSAAAI <span style="font-variant: small-caps;">RLGF</span> PEPfn <span style="font-variant: small-caps;">ASPVELR</span> TN <span style="font-variant: small-caps;">WSDSDVQAVISATYRQV</span> FGNEH <span style="font-variant: small-caps;">LMLSERLTS</span> AESLLASGNISVREF-----
PYR5_FREDI	48.3	MTSS <span style="font-variant: small-caps;">TAARQLG</span> PEPfs <span style="font-variant: small-caps;">TAPTELRA</span> --SSDVI <span style="font-variant: small-caps;">HAAYRQV</span> F-QVFGNDH <span style="font-variant: small-caps;">VMQSERLTS</span> AESLLQ <span style="font-variant: small-caps;">QGNISV</span> RDV <span style="font-variant: small-caps;">RLLA</span> QSEL
PYR6_FREDI	46.9	MAPL <span style="font-variant: small-caps;">TEASRLG</span> V <span style="font-variant: small-caps;">RPF</span> d <span style="font-variant: small-caps;">SDKVELR</span> FVK <span style="font-variant: small-caps;">TAE</span> EVRS <span style="font-variant: small-caps;">VIWSAYRQVLGNEH</span> L <span style="font-variant: small-caps;">FESERLSS</span> AESLLQ <span style="font-variant: small-caps;">QAQISV</span> RDV <span style="font-variant: small-caps;">VRAIA</span> QSEL
PYS1_SYNP6	56.9	MAITVASSRLGTA <span style="font-variant: small-caps;">PFSNa</span> PVELR <span style="font-variant: small-caps;">PDGDRDQVQAVIRAVYRQVLGNDY</span> IM <span style="font-variant: small-caps;">KSERLTA</span> AESLLV <span style="font-variant: small-caps;">NGSISV</span> RDV <span style="font-variant: small-caps;">VRAV</span> AKSEL
PYS1_SYNEL	38.6	-----
PYS1_FREDI	30.9	-----
PYS1_ANASP	29.6	-----
PYS1_MASLA	28.4	-----
PYS1_SYNP2	25.8	-----
PYS2_FREDI	27.7	-----
FENR_SYNY3	36.1	-----
FENR_ANAVA	25.7	-----
FENR_SYNP2	25.4	-----
PHEG_SYNPY	33.1	-MGIGIGPRL <span style="font-variant: small-caps;">LSECPFAVTFDRY</span> SPDSSAAL <span style="font-variant: small-caps;">ERVI</span> VAA <span style="font-variant: small-caps;">YRQVLG</span> NLPPTDN <span style="font-variant: small-caps;">QRETS</span> LEVRL <span style="font-variant: small-caps;">MNGEITV</span> RDV <span style="font-variant: small-caps;">VNG</span> LAKSDF
APCE_SYNP6	32.0	-FKLTQACAGSS <span style="font-variant: small-caps;">SIRSKSVGN</span> PSIRQ <span style="font-variant: small-caps;">TESTTQAVIRA</span> AYRQV <span style="font-variant: small-caps;">FGRD</span> -LYEG <span style="font-variant: small-caps;">QRLTVPEIK</span> LEN <span style="font-variant: small-caps;">GEITVREFV</span> RQIAKSET
APCE_ANASP	30.5	RGQS <span style="font-variant: small-caps;">VEVGVT</span> TRRK <span style="font-variant: small-caps;">PARIYRL</span> TNGIG <span style="font-variant: small-caps;">QAEKQLVINAIYRQV</span> dsGQ <span style="font-variant: small-caps;">VPDYR</span> RTE <span style="font-variant: small-caps;">LDSKLR</span> N <span style="font-variant: small-caps;">GEISVREFV</span> REIASSEI
APCE_CYACA	27.9	NIILKLNQGVSKRRE <span style="font-variant: small-caps;">QTVIFARH</span> INNS <span style="font-variant: small-caps;">QSSLEQIIKAAYRQV</span> F <span style="font-variant: small-caps;">ERdy</span> TIGRE-FYSI <span style="font-variant: small-caps;">ETLFYAGSL</span> SVKE <span style="font-variant: small-caps;">FIEHLG</span> QSEL
APCE_SYNY4	25.8	EIAYRSN <span style="font-variant: small-caps;">QGVYRQRQ</span> TKV <span style="font-variant: small-caps;">FKLTSNYDKVAVKNAIRA</span> HRQV <span style="font-variant: small-caps;">FERdy</span> IINSE-FTALES <span style="font-variant: small-caps;">KLSNGEIN</span> VKE <span style="font-variant: small-caps;">FIEGL</span> GTSEL
APCE_SYNP2	23.1	QPGSLGAKV <span style="font-variant: small-caps;">FRLN</span> NEL <span style="font-variant: small-caps;">PSGKT</span> T <span style="font-variant: small-caps;">NVSF</span> SESAT <span style="font-variant: small-caps;">QK</span> VI <span style="font-variant: small-caps;">EAYRQV</span> FGR-MVYAG <span style="font-variant: small-caps;">QRQKVAEIK</span> LEN <span style="font-variant: small-caps;">GEITLREF</span> IRALAKSDV
PYG1_SYNP2	22.5	SQ <span style="font-variant: small-caps;">NTRVAGY</span> TVGGDE <span style="font-variant: small-caps;">QPFV</span> TT <span style="font-variant: small-caps;">DNVIS</span> SDSD <span style="font-variant: small-caps;">FDV</span> LINA <span style="font-variant: small-caps;">YRQIF</span> F <span style="font-variant: small-caps;">HAFK</span> CDR <span style="font-variant: small-caps;">QQLL</span> --ES <span style="font-variant: small-caps;">QL</span> NG <span style="font-variant: small-caps;">QITV</span> RD <span style="font-variant: small-caps;">FIRGL</span> LLSET
PYG2_SYNEL	26.0	SQ <span style="font-variant: small-caps;">NQRVAGY</span> EV <span style="font-variant: small-caps;">PNEET</span> P <span style="font-variant: small-caps;">WRY</span> SL <span style="font-variant: small-caps;">EDA</span> VD <span style="font-variant: small-caps;">QSD</span> ID <span style="font-variant: small-caps;">ELIWA</span> AYRQV <span style="font-variant: small-caps;">FSEHV</span> VL <span style="font-variant: small-caps;">KSTR</span> OP <span style="font-variant: small-caps;">HLES</span> QLAN <span style="font-variant: small-caps;">RAISV</span> RD <span style="font-variant: small-caps;">FIRGL</span> AKSET

Matching Sequence	Query sequence (rod linker L <sub>R</sub> <sup>32,3</sup> from <i>Synechococcus</i> sp. PCC 7002 residues 81-160)
	YKEKFLYNNFQTRVIELNFKHLLGRAPYDEAEVIEHLDRYQNEGFEADINSYIDSAEYTENFGDNIVPYIRSYVVQTGHR
PYR1_SYN2	YKEKFLYNNFQTRVIELNFKHLLGRAPYDEAEVIEHLDRYQNEGFEADINSYIDSAEYTENFGDNIVPYIRSYVVQTGHR
PYR1_SYNY3	YKKKFLYNNFQTRVIELNYKHLGRAPFSEDEVIFHLDLYENQGFADIDSYSIDSEYQENFGENIVPYR-FNNQVQDR
PYR1_SYNEL	YKEKFLYGNFQTRVIELNYKHLGRAPYDESEVIFHLDLYENEGFDADIDSYSIDSEYTNNSFGDWVVPYRGFNTQPGQK
PYR1_SY9413	YKKKFFYNNFQTRVIELNYKHLGRAPYDESEVIEHLDRYQNEGYDADIDSYSINSQENDNFGDYTPYRGFSTQRGQK
PYR1_MASLA	YKKKFFYNSFQTRFIELNYKHLGRAPYDESEVIFHLDLYQNKGYDAEIDSYSIDSEYQNNFGDNIVPYRGFETQPGQK
PYR1_ANASP	YKTKFFYNNFQTRVIELHCKHLLGRAPYSEAEVIEHLDRYETQGYDADVDYSIDSEYLETTFGESIVPYRDIEYHVQGG
PYR1_FREDI	YKECFFYKNSQVRLIELNYKHLGRAPYDQSEIADHVDIYAARGYDADIDAYIYSSEYENAFGNSIVPYRGFQSIQGMK
PYR2_SYNY3	YRQKFFYSNSQVRFIELNYKHLGRAPYDESEIAYHVDIYTSQGYEAEINSYIDSEYQONFGDSIVPYRQYQTTVQGGK
PYR2_ANASP	YKECFFYNNNSQVRFIELNYKHLGRAPYDQSEIAFHVDLYAAAGYDAEIESYIYSPEYDNAFGNFVVPYRGFQSIQGMK
PYR2_MASLA	YRNRYFQSCAPYRFVIELNFLHLLGRAPDQREVSEHIVRTVAEGYDAEIDSYSIDSEYEAAFGENVVPYRGRSSEANSK
PYR2_FREDI	YKDNFFHAVGAQRGIELLNHLLGRAPLNQQEVQNHKLQAEQFDALIDTLTDSAEYTEVFGADIVPYDRTKDSYAGMN
PYR3_FREDI	YKTKFLYPHFQTRVIELLNHLLGRAPYDESEVIEHLDRDQNGFDADIDSYSIDSAEYDTYFGDSIVPYRDLVTVGQR
PYR4_FREDI	YREKFFHNNAHNRFIELLNHLLGRAPYDQAEVAHAATYHSHGYDADINSYIDSAEYTESFGDNVVPYRQGFATIRAQK
PYR5_PSEA9	YKKKFFYNSFQTRFIELNYKHLGRAPYDESEIYHLDLYQNKGYDAEIDSYSIDSEYQSNFGDNVVPYRQGFATIRAQK
PYR5_FREDI	YRQKFFYSTPQVRFIELNYKHLGRAPYDESEISYHVNLYTEKGYEAEINSYIDSAEYQESFGERIVPHYRQGFATIRAQK
PYR6_FREDI	YRSRFFTSCARYRAIELNFRHLLGRAPPLDLEEMRSHSTILDTOGFEAEIDSYSIDGDEYQSTFGENIVPYIRGYKTEALQS
PYS1_SYN2	YKSKFFSNNFHSRVTELLNFHLLGRAPYDESEIYHLDLYQTKGYEADIDSYSIDSAEYQTNFADNIVPYRQGFNNQLGQK
PYS1_FREDI	-----
PYS1_ANASP	-----
PYS1_MASLA	-----
PYS1_SYN2	-----
PYS2_FREDI	-----
FENR_SYNY3	-----
FENR_ANAVA	-----
FENR_SYN2	-----
PHEG_SYNPY	YSRHFIERVSOIRSVELRFMHLGRPLKDESELINNFIREKGFESHIDSLMNSLEYEEHFGEDIVPFORHWNSPCGST
APCE_SYN2	YAQQFYEPFINSRALELAFRHLGRGPGSREEVQYFALISKGLPLLVDALVDSKEYEYFGEIVPYLRTLGEAAQEC
APCE_ANASP	YRKEFFEPYPNTKVIELGTHFLGRAPKQGEIRLYNQILRSQGLKSFVKSLSINSQYIEIFGDSIVPYRRFPPTLPAGTF
APCE_CYACA	FKTLVIKSNNSYRLVELALKRLLGRAPYKDEEIAWSIKIATNGWGFVDALLDSEYQSNFGENIVPYQRRRYKDRPFN
APCE_SYNY4	YRQVVAQTNSYRLVDSFKRFLGRATYKSEQIAWSIVIATQGLHGFIDALVDSEYRQNFQGDIVPQRRRFKDRPFN
APCE_SYN2	FYQLVVTNNYRLVEMSLKRLGRSPYNEEKIAWSIQIASKGWGFVDALIDSEYEQAFQDNIVPYQRKRL--TTDR
PYG1_SYN2	FIDSFYKNSNYRFVEQCIQRVLGRDPFSEQEKIAWSIVICTKGLAAFVDQLLNTDEYMNFGYDTPYQRRRSLASREQ
PYG2_SYNEL	FYRLVVSVMNYRLVDITLKRLLGRSSYNKDEQIAWSIVIGTKGFSGFVDALIDSEYTKNFGENIVPYQRKRMEGRPHN

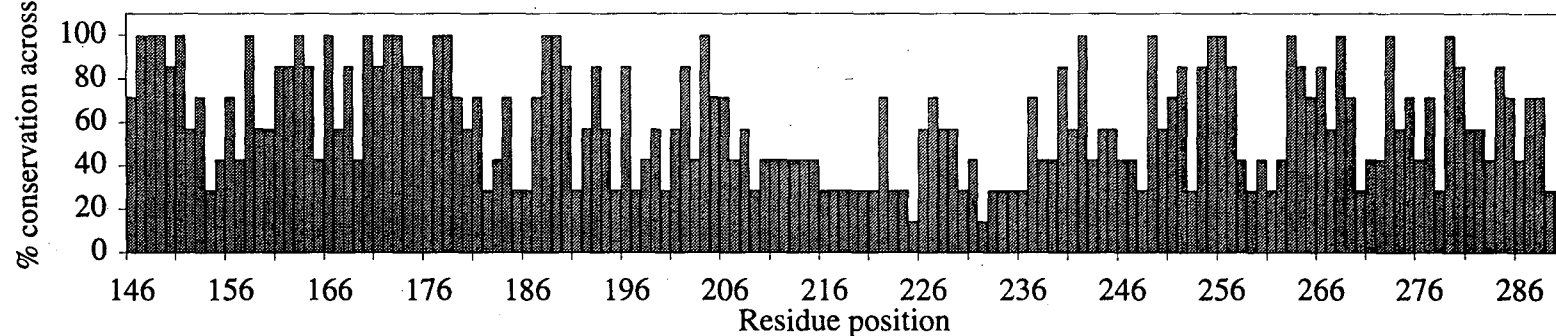
Matching Sequence	Query sequence (rod linker L <sub>R</sub> <sup>32,3</sup> from <i>Synechococcus</i> sp. PCC 7002 residues 161-240)
	TVGFTRM <del>SL</del> QRGYANS <del>DRAQ</del> IAGNAS <del>RLAQ</del> ELARNTTSAVVGPSGVNEGWA <del>FRSA</del> ADDYHPGQSLGGSTGLSADDQVVR
PYR1_SYN2	TVGFTRM <del>SL</del> QRGYANS <del>DRAQ</del> IAGNAS <del>RLAQ</del> ELARNTTSAVVGPSGVNEGWA <del>FRSA</del> ADDYHPGQSLGGSTGLSADDQVVR
PYR1_SYN3	TVGFTRM <del>RL</del> YRGYANS <del>DRS</del> QLERSS <del>RRLA</del> TELQNTVSAIVGPSG <del>SNAG</del> WAYRPSRAGNTPAKALGGTVPPGQASKLFR
PYR1_SYNEL	TVGFNRI <del>RL</del> YRGYANS <del>DRAQ</del> AEGSMS <del>RRLA</del> DLATN <del>RANT</del> VVPPSN <del>SDTA</del> FAYYTPSADVP <del>PRAC</del> LGGSFGE <del>SG</del> --RVYR
PYR1_SY9413	TAGFPRM <del>QL</del> YRGYAS <del>SDT</del> SGMSG <del>KKTR</del> LAE <del>LGR</del> NSV <del>SSV</del> VAPGGT <del>GAG</del> YL--PSGKGVTPKTGF <del>GGS</del> AMFGPGHRLYR
PYR1_MASLA	TVGFNRM <del>RL</del> YRGYANS <del>DRAQ</del> IEG <del>TKP</del> R <del>LARE</del> LATNKASSIVGPSG <del>SNPA</del> WGYRPSVD-ITP <del>PKTL</del> GNAVG--ENDRVYR
PYR1_ANASP	SVGLPRL <del>SL</del> YRGYAT <del>SDRT</del> SLGG <del>NN</del> SRLAAN <del>LA</del> AKSVAP <del>VV</del> APSGTNY---VASTKGEARC <del>GP</del> WVARRPFGQTASRIFR
PYR1_FREDI	TVGFNRI <del>CE</del> LYRGR <del>NSD</del> NAQMGR <del>TNS</del> R <del>LRT</del> KVSL <del>NL</del> PNGL <del>LP</del> PTSA-GTNEVSA <del>APT</del> LI-----SSATKGD <del>NRM</del> FV
PYR2_SYN3	TAGFPRF <del>QL</del> YRGYANT <del>DRAQ</del> NKSKG-QL <del>TW</del> DLAK <del>NL</del> VSP <del>II</del> Y-----PAD---AGSLT <del>GV</del> STG-NRGGNTYR
PYR2_ANASP	TVGFNRI <del>EL</del> YGRANS <del>DNAQ</del> FGGKSAR <del>LRS</del> KIS <del>MNL</del> ANTIVPPTSP <del>IA</del> ASTSSARTL <del>VTS</del> PMG-----DAR <del>MF</del> I
PYR2_MASLA	QVGFNRI <del>AL</del> DRGPAQ <del>IDS</del> AV---KSAQL <del>VY</del> AVAT <del>NS</del> ANA <del>IK</del> ASS <del>TV</del> IGSGTEK <del>RFK</del> ILVQ <del>GSK</del> FDS <del>PRRI</del> STTEYIVP
PYR2_FREDI	TRSFNLM <del>RD</del> L <del>gm</del> KVAIS <del>DNAQ</del> -GRQ <del>SKT</del> VNALASAS <del>REST</del> K <del>QP</del> FSYVSV <del>TQ</del> IPVKLP <del>QQQ</del> YTGHNVPAMSDYV <del>PF</del> RP
PYR3_FREDI	TVGFTRS <del>RL</del> YRGYANS <del>DRS</del> QLAGSS <del>RRLA</del> SDLAT <del>NSA</del> TAI <del>IAP</del> SGGTQ <del>GW</del> SYLPSKQGTAP <del>SRT</del> FGRSSQ <del>G</del> STP-RLYR
PYR4_FREDI	TVGFNRI <del>QL</del> YRGYAT <del>SDRS</del> ATS-QRS <del>RLD</del> QEVARN <del>AAS</del> PVLIGS-----TGSALVGT <del>C</del> GGSRDQ <del>L</del> fVQ
PYR5_PSEA9	TAGFNRI <del>RL</del> YRGYANS <del>DRAQ</del> VEG <del>TKS</del> R <del>ia</del> GNLAS <del>NK</del> ASTIVGPSG <del>TND</del> SWGFRASAD-VAP <del>KKN</del> LGNAVG-EGDRVYR
PYR5_FREDI	TVGFNRM <del>QL</del> YRGYANS <del>DRS</del> QGNK <del>SA</del> WLTQ <del>DLA</del> LNLAS <del>NI</del> QTP---NFG-----KGLTGVVAG <del>d</del> QLYR
PYR6_FREDI	MVQFTHT <del>QL</del> VRGASS <del>S</del> skGDL <del>SG</del> KAP <del>KL</del> NALVIQSTPTAVISPASA--GATFSTPPTGART <del>RL</del> GV <del>DA</del> SAG---GK <del>VY</del> R
PYS1_SYN6	TVGFTRI <del>QL</del> YRGYAT <del>SDRS</del> QIPGASARLANELARNSASTVIAPAGSNNGFAYRASVKGKTPSTAFQGSQAFGSG-RLYR
PYS1_SYNEL	-----MFGQTASGSAALsGARVFR
PYS1_FREDI	-----RRSSSGSDNRVFEVVEG
PYS1_ANASP	-----MFGQTTLGAGSVSSASRVFR
PYS1_MASLA	-----DSVSSASRVFRFEVVG
PYS1_SYN2	-----NGTEAASRVFTYEVOG
PYS2_FREDI	-----MVYQSRSFQVEVSGL
FENR_SYN3	-----MYS <del>PG</del> YVATSSRQSDAGNRLFV
FENR_ANAVA	-----EGAANVESGSRVFEVVEVGM
FENR_SYN2	-----NSTGNQSYANRLFIEYV
PHEG_SYNPY	TSSFIKTALFRKGFASSDNVIYK-----
APCE_SYN6	RNWGAQIKLLNYSARFQKTPQFITL <del>FAG</del> YK <del>NL</del> PLDQHPYGGQNDPLEIQFGAIFPKETLQTKAAFFGKD-----
APCE_ANASP	PNTDILYKNLTKQKFFILMP <del>SY</del> KNTKLLSV-----
APCE_CYACA	LVT <del>PRY</del> GNYWRDKLESERYIEGD <del>IK</del> NFLELAKSIEIK <del>TV</del> TFTPVSTANIKIPD <del>T</del> TRNTTPTGIPISVNPSANFPVR----
APCE_SYN2	PFSFT <del>PRY</del> GADYRDRAGIVRPG <del>MS</del> NW <del>NN</del> SANQNYDGVAILGVLLAISAGMTFLFVLNWLGISSSF-----
PYG1_SYN2	GEIPFNIKSPRYDAYRSQ <del>L</del> GF <del>PQ</del> VVWQNAVRRFRTPD <del>RV</del> PQAGDPALFLNMARSAQIPKVNVRVSAADISLAAVP----
PYG2_SYNEL	LVT <del>PRY</del> GEDFQEKAGTVQTDWRFTLDK <del>FY</del> SRKSQEKQLREGDPRK <del>FAD</del> LAASVGNQGN <del>YA</del> QRISAFDIDYLN <del>AV</del> PN----

Matching Sequence	Query sequence (rod linker L <sub>R</sub> <sup>32.3</sup> from <i>Synechococcus</i> sp. PCC 7002 residues 241-290)
	VEVAALSTPRYPRI <b>RRSSRVFFV</b> PVSRLS <b>SQKLQEI</b> QRMGGRVASISPAGQ
PYR1_SYNP2	<b>VEVAALSTPRYPRI</b> RRSSRVFFV <b>PVSRLS</b> SQKLQEI <b>QRMGGRVASI</b> SPAGQ
PYR1_SYNY3	<b>VEITAI</b> SAPGYPKV <b>RRSNKAVI</b> VPFEQLN <b>QTLQ</b> QINRL <b>GGKVASIT</b> PASL
PYR1_SYNEL	IEVAGIR <b>QPGYP</b> GV <b>RRSSTA</b> FLV <b>PYEQL</b> SAKM <b>QQLQ</b> RTGARIISVNPA--
PYR1_SY9413	IEIAGPNL <b>PRYP</b> RV <b>RRVNQ</b> DVIV <b>PYERL</b> SEEM <b>QRINR</b> Q <b>GGRIS</b> SITPV--
PYR1_MASLA	IEVTGVR <b>TPGYP</b> SV <b>RRSSYA</b> IIV <b>PYERL</b> SEKI <b>Q</b> IHK <b>LGGKIV</b> SITSA--
PYR1_ANASP	<b>VEIAA</b> ISKPGF <b>PSVRR</b> SNR <b>SL</b> LVP <b>YEQL</b> NN <b>T</b> LQ <b>VNR</b> S <b>GGRV</b> SVAPASL
PYR1_FREDI	IEALAGGLN <b>T</b> NAV <b>RRSRQ</b> VY <b>T</b> VS <b>YERL</b> SATY <b>Q</b> EI <b>HKR</b> GGKIVKISQV--
PYR2_SYNY3	IRTTQAAS <b>PNSP</b> RI <b>RS</b> IS <b>RVV</b> VP <b>FDQ</b> LS <b>N</b> L <b>LQ</b> LN <b>RQ</b> GRK <b>VI</b> SIALS--
PYR2_ANASP	<b>VEA</b> IAGTLN <b>T</b> NAV <b>RRSRQ</b> VY <b>T</b> VP <b>YDR</b> LSATY <b>Q</b> EI <b>HKR</b> GGKIVK <b>IT</b> PAS-
PYR2_MASLA	---ASK <b>MTPQ</b> I <b>QRIN</b> ET <b>SGKIV</b> SIT <b>EIV</b> ----
PYR2_FREDI	-----
PYR3_FREDI	IEVTGISL <b>PRYP</b> KV <b>RRSN</b> KEF <b>IV</b> VP <b>YEQL</b> S <b>ST</b> L <b>Q</b> QIN <b>L</b> GGK <b>VASIT</b> FAQ-
PYR4_FREDI	V <b>Q</b> Q <b>Q</b> GV <b>T</b> ART <b>S</b> NR <b>IR</b> RS <b>N</b> ATY <b>F</b> VS <b>Y</b> E <b>K</b> LN <b>AT</b> L <b>Q</b> RV <b>H</b> A <b>Q</b> GG <b>RIV</b> SIT <b>PA</b> --
PYR5_PSEA9	LEV <b>T</b> GIR <b>S</b> PG <b>Y</b> PSV <b>RRS</b> TV <b>F</b> IV <b>PYER</b> LS <b>DKI</b> Q <b>V</b> H <b>K</b> Q <b>GGKIV</b> SV <b>T</b> SA--
PYR5_FREDI	VR <b>VI</b> Q <b>ADR</b> GR <b>T</b> T <b>Q</b> IR <b>RS</b> IO <b>EY</b> LV <b>S</b> Y <b>D</b> Q <b>L</b> S <b>P</b> T <b>L</b> Q <b>R</b> LN <b>Q</b> R <b>G</b> SR <b>V</b> NI <b>S</b> PA--
PYR6_FREDI	IEVTGYRA <b>K</b> TF <b>N</b> NI <b>RRS</b> N <b>Q</b> V <b>L</b> VP <b>YE</b> KL <b>S</b> Q <b>EY</b> Q <b>R</b> I <b>H</b> Q <b>Q</b> GG <b>V</b> IASIT <b>P</b> V--
PYS1_SYNP6	<b>VEVA</b> AIS <b>Q</b> PAI-Q <b>VRR</b> INKRSI <b>H</b> RR <b>T</b> I <b>Q</b> LF <b>P</b> TS <b>S</b> TS <b>Q</b> W <b>Q</b> I <b>A</b> SV <b>T</b> PL----
PYS1_SYNEL	Y <b>EV</b> VGL <b>e</b> T <b>D</b> r <b>f</b> P- <b>IR</b> SG <b>S</b> T <b>F</b> i <b>VP</b> Y <b>NR</b> M <b>N</b> E <b>E</b> M <b>Q</b> R <b>I</b> T <b>R</b> M <b>G</b> G <b>KIV</b> SIT <b>P</b> V <b>V</b> A
PYS1_FREDI	LR <b>Q</b> NE <b>Q</b> T <b>D</b> NN <b>R</b> Y <b>Q</b> I <b>R</b> NS <b>S</b> T <b>i</b> Q <b>V</b> P <b>Y</b> SR <b>M</b> NE <b>E</b> DR <b>R</b> I <b>T</b> RL <b>G</b> GR <b>I</b> V <b>N</b> IR <b>P</b> AG <b>E</b>
PYS1_ANASP	Y <b>EV</b> VGL <b>t</b> DK <b>N</b> K <b>Y</b> NI <b>R</b> NS <b>G</b> SV <b>F</b> i <b>VP</b> Y <b>SR</b> M <b>N</b> E <b>E</b> Y <b>Q</b> R <b>I</b> T <b>RL</b> GG <b>KIV</b> K <b>I</b> E <b>Q</b> LV <b>S</b>
PYS1_MASLA	MR <b>Q</b> NE <b>E</b> ND <b>K</b> N <b>K</b> Y <b>N</b> I <b>R</b> NS <b>G</b> SV <b>i</b> VP <b>Y</b> NR <b>M</b> SE <b>E</b> M <b>Q</b> R <b>I</b> H <b>RL</b> GG <b>KIV</b> K <b>I</b> E <b>P</b> L <b>T</b> R
PYS1_SYNP2	LR <b>Q</b> TE <b>E</b> T <b>D</b> N <b>Q</b> E <b>Y</b> A <b>F</b> IR <b>SG</b> SV <b>F</b> i <b>VP</b> Y <b>AR</b> M <b>N</b> Q <b>E</b> M <b>Q</b> R <b>I</b> L <b>RL</b> GG <b>KIV</b> S <b>I</b> K <b>P</b> Y <b>T</b> G
PYS2_FREDI	H <b>Q</b> NE <b>V</b> T <b>N</b> Q <b>NN</b> Y <b>P</b> - <b>IR</b> SG <b>S</b> V <b>F</b> i <b>i</b> P <b>F</b> SR <b>F</b> NE <b>E</b> L <b>Q</b> R <b>I</b> N <b>RL</b> GG <b>KIV</b> N <b>I</b> Q <b>P</b> L <b>N</b> L
FENR_SYNY3	Y <b>EV</b> IG <b>L</b> S <b>Q</b> St <b>Y</b> P- <b>IR</b> SG <b>S</b> T <b>F</b> i <b>VP</b> L <b>K</b> R <b>M</b> N <b>Q</b> E <b>M</b> RR <b>I</b> T <b>R</b> M <b>G</b> G <b>KIV</b> S <b>I</b> K <b>P</b> L <b>E</b> G
FENR_ANAVA	R <b>Q</b> NE <b>E</b> T <b>D</b> Q <b>T</b> N <b>Y</b> P- <b>IR</b> K <b>S</b> G <b>S</b> V <b>F</b> i <b>VP</b> Y <b>NR</b> M <b>N</b> Q <b>E</b> M <b>Q</b> R <b>I</b> T <b>RL</b> GG <b>KIV</b> S <b>I</b> Q <b>T</b> V <b>S</b> A
FENR_SYNP2	V <b>G</b> L <b>G</b> G <b>D</b> GR <b>N</b> EN <b>S</b> L <b>V</b> R <b>K</b> SG <b>T</b> T <b>F</b> i <b>VP</b> Y <b>AR</b> M <b>N</b> Q <b>E</b> M <b>Q</b> R <b>I</b> T <b>KL</b> GG <b>KIV</b> S <b>I</b> R <b>P</b> A <b>E</b> D
PHEG_SYNPY	-----
APCE_SYNP6	-----
APCE_ANASP	-----
APCE_CYACA	-----
APCE_SYNY4	-----
APCE_SYNP2	-----
PYG1_SYNP2	-----
PYG2_SYNEL	-----

**Comparison of PYR1 sequences ( $L_R^{33}$  linkers only) for evolutionarily conserved residues [1-145]**



**[146-290]**



**For PYR1 sequences only:**

Sequence length = 290 residues

Average overall sequence homology score = 56.9%

Number of residues in the consensus sequence with greater than 80% conservation across species = 164 (or 56.6% of 290)

Average % conservation across species for first 180 amino acids (N-terminus) = 78.5%

Average % conservation across species for remainder (C-terminus) = 54.7%

Table 3-3C. Fast-A3 results of SWISS-PROT database search for sequence homology (>25%) with LRC<sup>28.5</sup>

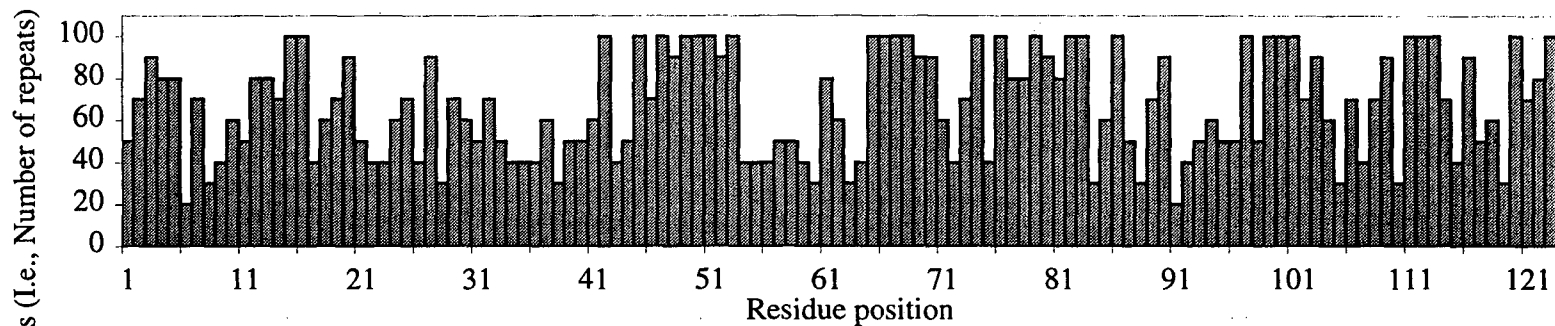
Matching sequence	Score (%)	Query sequence (rod-core linker LRC <sup>28.5</sup> from <i>Synechococcus sp.</i> PCC 7002 residues 1-80)
		TIPLLQYAPSSQNTRVAGYTVGGDEQPFVFTTDNVISSDFDVLINAAYRQIFFHAFKCDRQQLLESQLRNGQITVRDFI
PYG1_SYNP2	100	TIPLLQYAPSSQNTRVAGYTVGGDEQPFVFTTDNVISSDFDVLINAAYRQIFFHAFKCDRQQLLESQLRNGQITVRDFI
PYG1_ANASP	43.3	SIPLLQYAPSSQNQRVEGYEVPNEDPTIYRLAAAIDDADVDAIWAAYRQIFSeiIKSNRQSFLESQLRNRAINVRDFI
PYG1_SYNEL	40.1	SLPLLAVKPRTLQDQRVVSYEVPEDDPVIYRLTDATDAAAVDALIWAAYRQIFshLILAsrQPFLLESQLRNRAISVRDFI
PYG1_MASLA	39.3	PIPLLSPFLTTQNRVEGYEVPNEDPTIYRLTDTSSDTEIDAIWAAYRQIFSEHLILerQPYLESQLRNRAINVRDFI
PYG_SYNY3	67.0	-----MRVEGYEIGSEEEKPVVFTTENILSSDMDNLIEAAYRQIFFHAFKWDREKVLLESQLRNGQITVRDFV
PYG_CYACA	53.5	AIPLLYYPYSSQNHRVNSFEIqNDEQPRIFSTDSLPTSSEMDEIIWAAYRQIFhQILKFNDRDYLESQLRNFQITVREFI
PYG_PORPU	52.9	SIPLLNYSLSTQNQRVDGYEVSgEEQPRAYNTDNLPSAVEMDEVIWAAYRQIFhQILSSTSAPYLESQLRNFQIKVKDFI
PYG_CYACA	52.1	SIPLLYYPLSTQNQRVESFEIsNEEQSKIYTTDTLPASSEMDELIIWAAYRQIFhQILKFTRQRFLESQLRNFQIKVREFI
PYG_AGLNE	50.9	SIPILNYSLSTQNQRVNGFelpGDELPKIYTTDNLPTSIEMDEIIWAAYRQIFhQMLSSCMDRFLESQLRNFQIKVKDFI
PYG_SYNY3	48.5	TLPLIAYAPVSONQRVTNYEVSgDEHARIFTTTEGTLSPSAMDNLI AAYRQVNEciQSNRQIALESQFKNQITVRDFI
PYG2_SYNEL	46.2	TIPLLSYAPSSQNQRVAGYEVNNEETPWRYSLLEDAVDQSDIDELIWAAYRQVSevLKSTRQPHLESQLANRAISVRDFI
PYG2_MASLA	44.1	AIPLLQYKPPSSQNQRVPGYEVNEDTPRIYRLEDAPSYSEIQELIWAAYRQIFheILKFHRQINLESQKRNRTITVRDFI
PYG2_ANASP	39.1	SIPLLEYKPPSSQNQRVPGYEVNEDTPRIYRIEDAAYDSELKELIWAAYRQVSeiLKFFRQGNLESQKRNRAISVRDFV
PYG3_MASLA	41.3	ALPLLEYKLSQNHRVKSFGVaneDTPYIYRLEDVSSFTDIQNIWAAYRQVheILRFNRQKHLESQKSGLITVRDFI
PYG3_ANASP	39.0	ALPLLEYKPTTQNQRVQSFGTaneDTPYIYRLENANSPSEIEELIWAAYRQVNEqIKFNROIGLETQKNRSITVKDFI
PYG4_SYNEL	42.1	DLPLLAYKPTTQNRVASFGIaneDTPYIYRLEDAAASYSEIREVIWACYRQVshalaFNROITLESQLVNRVITVRDFI
PYG4_ANASP	40.9	ALPLLQYKPPSSQNHRVTSFGAaneDTPYIYRIEDVSSYTDIQNIWASRYRQVheILKFNRQKTLESQVKNNGSISVRDFI
APCE_SYNP6	29.2	-----TPTKIFKLTQACAGSSSIRSKSVGNPSIRQTESSTQAVIRAAYRQVGrlyEGQRLTVPEIKLENGETVREFV
APCE_PORPU	27.1	DKLRKRESTDLQGLRLPQYTSQAGVSTPRFVMKSSLSADEKNTVVKACYRQIFEraKTYDLssNLESQVKNNGQISIKEFI
APCE_CYAPA	26.7	DLVRKRTSTSLQGLKLPQIYANAVQKPRFQMKSTLSTTEKETVIKAVYRQIFERDVRrakNYDLESKVKNGQLSIKEFV
APCE_FREDI	26.2	GTVTDTRTEPDIQFRINQGVSKQREQTKVFKQVANISdaaVQTLISAAYRQIFeryIAKNEFSALESKLSNGEITVKEFI
APCE_SYNP2	23.3	QLGNPRARATQPGSLGAKVFRLLNELPSGKTTNVSESATQKVI EAAYRQVGrvyAGORQKVAEIKLENGETLREFI
PYR1_SYNP2	22.2	-----MPVTVAASRLGTAAFDQSPVELRANYSRDDAQTIVIRAAYRQVVLGNDYVMSSERLteSLFTNGFISVRDFV
PYR5_PSEA9	27.1	-----MTSSAAAIRLGFEPFVNASPVELRTNWSDDVQAVISATYRQVleHMLSERLTSAESLLASGNISVREF-
PYR6_FREDI	27.7	-----TEASRLGVRPFADSDKVELRFVKTAEVRSVIWSAYRQVleHLFESERLSSAESLLQQAQISVRDFV
PYS1_SYNP6	26.4	-----VASSRLGTAPFSNAAPVELRPDGRDQVQAVIRAAYRQVVLGNDYkSERLTAESLLVNGSISVRDFV

Matching Sequence	Query sequence (rod-core linker Lrc <sup>28.5</sup> from <i>Synechococcus</i> sp. PCC 7002 residues 81-165)
	RGLLLSETFIDSFYNKNSNYRFVEQCIQORVLGRDPFSEQEKIAWSIVICTKGLAAFVDQLLNTDEYMFNGYDTPVYQRRSLAS
PYG1_SYNP2	RGLLSETFIDSFYNKNSNYRFVEQCIQORVLGRDPFSEQEKIAWSIVICTKGLAAFVDQLLNTDEYMFNGYDTPVYQRRSLAS
PYG1_ANASP	RGLGKSEVYRTQVADLNSNYRLVDITLKRFLGRAAYNQDEEIAWSIVIGSQGLHGFIDALLDSDEYRENFQGGDIVPYQRRH----
PYG1_SYNEL	RGLGKSEVYREQVAAVNSNYRLVDISFKRFLGRPTYGQQEQIAWSIILATRGLEGFIDALVDSDEYQQNFQADIVPYQRRH-----
PYG1_MASLA	RGLGKSEVYRQEVAAQTNNSNYRLVDISFKRFLGRATYQKSEQIAWSIVIATQGLHGFIDALVDSEYRQNFQGGDIVPFQRRHFK--
PYG_SYNY3	RGLLLSNTFRNSFYEKNSNYRFVEHCVQKILGRDVYSEREKIAWSIVVATKGYQGLIDLLNSDEYLNNGYDTPVYQRRHNLPG
PYG_CYACA	RGLLLSERFRILNVDVNNNYRFSEMCVQORVLGRDIYNDKEKIAWSIVLCTRGIKAFVDSLNSKEYEENFGDNTVPYQRRHVFIFQ
PYG_PORPU	KGLILSEFRKLNVDVNNNYRFVEICVQRI LGRDVYNEREKIAWSIVIASKGLESFINMLIESDEYEENFGDSIVPYQRRHIIAQ
PYG_CYACA	RGLAISDSFRKLNVDVNNNYRFVELCVQRI LGRDVYNEKEKIAWSIVICNKGVI GFIDCLINSQEYLENFGDNIVPYQRRHIIIFQ
PYG_AGLNE	KGLVLSAFRNLDVNNNYRFVEMCIQORVLGRDIYNEREKIAFAIIIASQGIETFVDFLLNSDEYIENFGDNTVPYQRRHIIAQ
PYG_SYNY3	RGLALSDFSRRRNFEVNNNYRFVQMCIQORLLGRDVYSEEEKIAWSIVIATKGLPGFINELLNSQEYLENFGYDTPVYQRRHILPQ
PYG2_SYNEL	RGLAKSETFRRLVVEVETNSNYRLVEIALKRLGRAPYNKQEEIAWSIRIATDQWQKFDVTLVDSDEYTONFGDNTVPYQRRH----
PYG2_MASLA	RGLAKSEAFQRLVVEVETNSNYRIVEISLKRILGRAPYNREEEIAWSIKIATDQFGGFVDALVDSEYQINFGDNTVPYQRRHFK--
PYG2_ANASP	RGLAKSEAFKTLVVIKSNNSNYRLVELALKRLLGRAPYNKDEEIAWSIKIATNGWDGFVDALLDSEYQSNFGENIVPYQRRH-----
PYG3_MASLA	RGLAKSEAFYRLVVSNNNYRLVDVVLKRLGRSAYNKEEEIAWSIVIGTKGFDGFDVDAVDSDEYTOAFGDNTVPYQRRH-LVD
PYG3_ANASP	RGLAKSERFYQLVVTPNNNYRLVEMSLKRLGRSPYNEEEKIAWSIQIASKGWGGFVDALIDSTEYEQAFGDNTVPYQRRHKLTTD
PYG4_SYNEL	RGLAKSERFYNTVVAVNDNYRLVDVCLRRFLGRSAYNEEEKIAWSIKIGTLGFYGFVDALLDSEYTNAFGDFTPVYQRRHNM---
PYG4_ANASP	RGLAKSEAFYRLVVSNNNYRLVDITLKRLLGRSSYNKDEQIAWSIVIGTKGFSGFVDALIDSEYTKNFGENIVPYQRRHNM---
APCE_SYNP6	RQIAKSETFRKLYWNNLYVVKAVEYIHRRLGRPTTGRAEINAYFDISAKKGFYALVDAILDSPEYIAAFGEDTVPYERHYPKGP
APCE_PORPU	RSLGTSNIYRKQFYEPFVNSRALELAFRHF LGRGPGSSLEEFQKYFAILSATGLSGLVNAI LNSAEYADYFGEETVPYFRKLGEEP
APCE_CYAPA	RALGKSKLYAQQFYEPFINSRALELAFRHF LGRGPGSREEVQYEFALISKGGLPLLVDALVDSKEYEYFGEETVPYFRKLGEEA
APCE_FREDI	EGLGYSNLYIKEFYTPYPTKVIELGTHKFLGRAPLDQVEIRKYNQILATQGI RAFIGALVSSAEYAEVFGEDTVPYFRAYPTLPA
APCE_SYNP2	RALAKSDVFRNTYWSLYVTKAVEYIHRRLGRPTYGRQEINSYFDTCAKKGFYALVDAIIDSKEYEAFGEDTVPYFRAYLTPGG
PYR1_SYNP2	RAVAQSELYKEKFLYNNFQTRVIELNFKHLLGRAPYDEAEVIEHLDRYQNEGFADINSYIDSAEYTENFGDNTVPYFRAYSVVQT
PYR5_PSEA9	-----
PYR6_FREDI	RAIAQSELYRQKFFYSNSQVRFIELNYKHLGRAPYDESE-IAVHVDIYTGGEAEINSYIDSVEYQQNFQGGDIVPYFRAYGYQTTV
PYS1_SYNP6	RAVAKSELYKTKFFYNNFQTRVIELHCKHLLGRAPYSEAEVIEHLDRYETQGYDADVDSYIDSSEYLETFGESIVPYFRAYDIEYHV

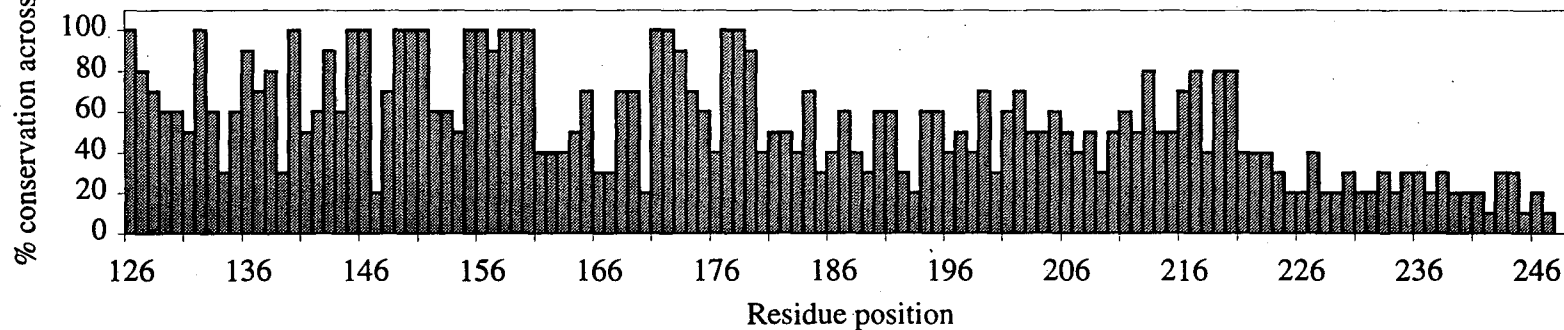


Matching Sequence	Query sequence (rod-core linker LRC <sup>28.5</sup> from <i>Synechococcus</i> sp. PCC 7002 residues 166-247)
	EQQEIPFNIKSPRYDAYYRSQLGFPQVWQNAVRRFTPDVRVQAGDPALFLNMARSAQIPKVNVRVSAADISLAAVPYRN
PYG1_SYNP2	EQQEIPFNIKSPRYDAYYRSQLGFPQVWQNAVRRFTPDVRVQAGDPALFLNMARSAQIPKVNVRVSAADISLAAVPYRN
PYG1_ANASP	--YKDRPFNLVNPYINAYWRDRQTLNALGGRSFYSARTSGTLTKDDIRRAIPANFMALAGKILTPERNYQRTI-----
PYG1_SYNEL	-----PFNLVNPYSDYWRneISLSGRSYYQARYYASGpkQIVRGAIIPANFLSMARSIVVPTLDTqVARATSSLVKVPNTA
PYG1_MASLA	-----DRPFNLVNPYADYWRNRLLEQFLGGQSFYRVVWAgervGAGAIIPSTFLSMAASIAPSGISYQRTADSARTFISTVKL
PYG_SYNY3	KEAGELPFNIKSPRYDAYYHRQLGFPQIVWQNEVRRFIPQEKKLTAGNPMNFLGMARSIN-PAANtkVSAQNIINiaSVPRR-
PYG_CYACA	SKGEVPPFNLKTPYGSDFKERLGMPOYAWQGAIRTFRPOEIKPRPGDPVLYLDMRELLFSRFR-----
PYG_PORPU	SKGEMPPFNLKTPYGADFKKFKGMPQFIWQGPVRFQFPQEQRPKAGDPALFLGMVNDLATV-----
PYG_CYACA	VNGETPPFNLKTPYGFIEFDKLVKPKFIWQGAIRTFRPOEQRPRAGDPVLFNLMVYDLNipKFNRYPGLVNR-----
PYG_AGLNE	SKGEIPFNLKTPFLDYNFLYKKNMPQLLWSGVPVRFQFPQEQPKAGDPALFLNMVLDVSPSYLISS-----
PYG_SYNY3	ISGELPF-ARMPYGADHREKLEAIGYFRNQAPLTYRWEWQKQPYPAGVYLAGKVVLVYGGALVSLGIIV-----
PYG2_SYNEL	--YKDRPFNLVTPYSDYWRDKLENSRYKW-----GDIRNFLEMARSVKVTPVQFkvSTANVQIPDTTRRD
PYG2_MASLA	-----DRPFNLVTPYGNYSWRDKLENERIYI-----AGDIKNFLDLAKSIEIKTVSYqvSTANISIPDTTRNT
PYG2_ANASP	--YKDRPFNLVTPYGNYSWRDKLESERYIEGDIKNFLELAKSIEIKTVTFTPVSTANIKIPDTTRNTTPTGIP-----
PYG3_MASLA	-----PHNLVTPYGEDFOETAGTVKTDWRFTLQNFkFQERQLREGDPRKYTDMA-AAIAPKGNynIRAADLd1NLVPSRT
PYG3_ANASP	-----PFSF-TPYGADYRDRAGIVRPGRMSNWNNSANQNYDGVAILGVLLAISAGMTFLFVLLNWLGISS-----
PYG4_SYNEL	--EGR-PFNLVTPYGFYRDKVGTTKTDWRFTLEKFKFQERRLPEGDPRKYRDLA-AAIAPKTRyqLRAADI1LAKVPRRR
PYG4_ANASP	--EGR-PHNLVTPYGEDFOEKAGTVQTDWRFTLDKFKsQOEKOLREGDPRKFAADLAASvnOGNYAQRISAFDI1LNAVPNRS
APCE_SYNP6	LALRSVRGLESLE-----
APCE_PORPU	QECRNWGPQIDLLNYSAPFRK---VPQFITLFSYKQSLPDQHPYgnDP---LSIQFGAIFPKENKDPKRQAIFFGKDTRRI
APCE_CYAPA	QECRNWGAQIKLLNYSARFQKT---PQFITLFAGYKNPLPDQHPqGNDP---LEIQFGAIFPKETLQTKAAFFGKDTRRILI
APCE_FREDI	ANFPNTEKLYNQLTQNDLLVVPSPFKTVQPRLLTAGTSSSGKNGFTDLGRSSTSAQGQLGETAN-----
APCE_SYNP2	YSLRQTRPGALREDVGKVKVEKTARFIELGTSSTKNLPVTDVdarLKQGVNIQRQOTKAFKLTDTF-----
PYR1_SYNP2	GHRTVGFTRMFSLQRYANSRRAQIAGNASRLAQLAENPTSAVVGPSGVNEGWAFRSAADYHPGQSLGGSTGLSADD---
PYR5_PSEA9	-----
PYR6_FREDI	GQKTAGFPRFFOLYGYANTDIAQN-----
PYS1_SYNP6	GQQSVGLPRLFSLYGYATSDTSLG-----

Comparison of PYG1/PYG sequences ( $L_{RC}$  linkers only) for conserved residues [1-124]



[125-247]



**For PYG1 and PYG sequences only:**

Sequence length = 247 residues

Average overall sequence homology score = 49.7%

Number of residues in the consensus sequence with greater than 80% conservation across species = 107 (or 43.3% of 257)

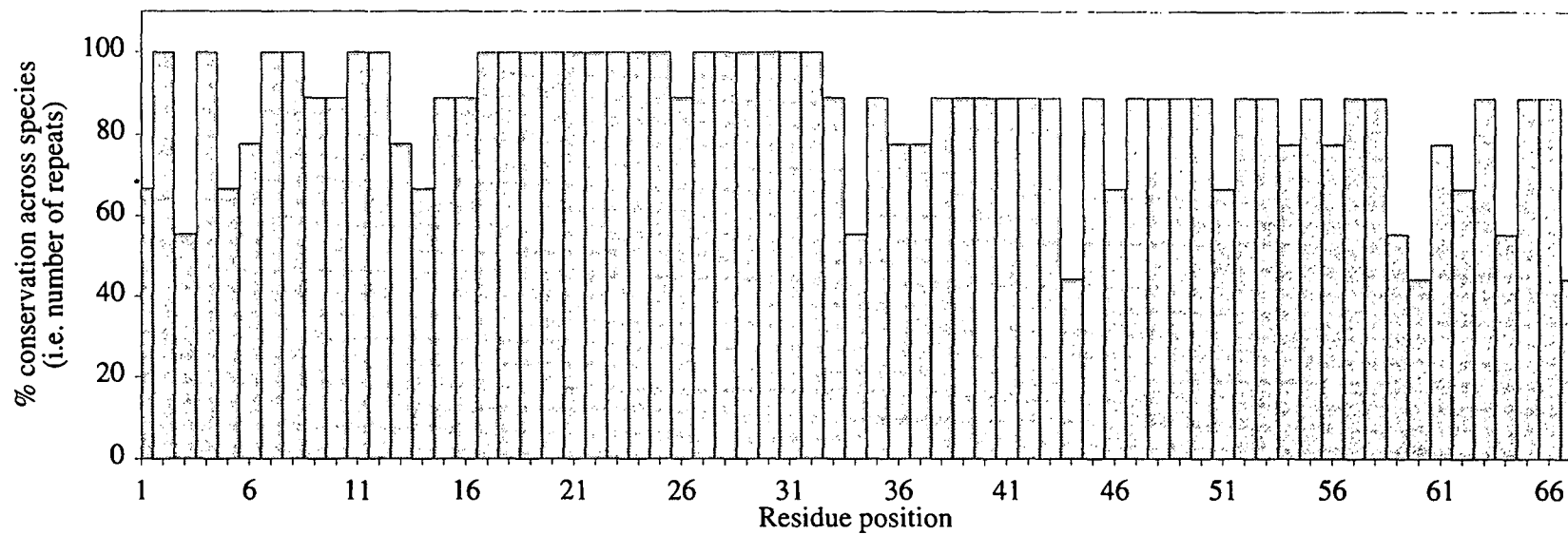
Average % conservation across species for first 180 amino acids (N-terminus) = 68.2%

Average % conservation across species for remainder (C-terminus) = 40.4%

Table 3-3D. Fast-A3 results of SWISS-PROT database search for sequence homology (>20%) with L<sub>C</sub><sup>7.8</sup>

Matching sequence	Score (%)	Query sequence (core linker L <sub>C</sub> <sup>7.8</sup> from <i>Mastigocladus laminosus</i> ):
		GRLFKITACVPSQTRITQRELQNTYFTKLVPEYENWREQQRIQKMGGKIVKVELATGKQGINTGLA
PYC1_MASLA	100	GRLFKITACVPSQTRITQRELQNTYFTKLVPEYENWREQQRIQKMGGKIVKVELATGKQGINTGLA
PYC1_SYNEL	92.5	MRMFKITACVPSQTRITQRELQNTYFTKLVPEYENWREQQRIQKMGGKIVKVELFTGKPGVNTGLA
PYC1_FREDI	91.0	ARLFKVTACVPSQTRITQRELQNTYFTKLVPEYENWREQQRIQKMGGKIVKVELATGKQGTNTGLL
PYC1_SPIPL	88.6	MRVFKVTACVPSQTRITQRELQNTYFTKLVPEYDN-----
PYC1_SYNY4	86.6	MRMFRITACVPSQTRITQRELQNTYFTKLVPEYDNWREQQRIQKMGGKIVKVELATGRPGTNAGLA
PYC1_SYNP2	86.6	MRMFKITACVPSQTRITQRELQNTYFTKLVPEYDNWREQQRIQKMGGKIVKVELATGKPGTNTGLT
PYC1_SYNY3	85.1	MRMFRITACVPSQTRITQRELQNTYFTKLVPEYDNSREQQRIQKMGGKIVKVELATGRPGTNAGLA
PYC1_ANASP	85.1	SRLEFKITALVPSLSRITQRELQNTYFTKLVPEYENWREQQRIQKAGGKIIVKVELATGKQGTNAGLQ
PYC1_SYNP6	74.6	MRMFRITACLPSPSKIRITQRELQNTFFTKLVPEYDAWREQQRIQKLGKKIIVKVELATGRPNTNTGLL
PYR1_SY9413	30.4	HRLYRIEIAAGPNLPIYPRVIRVNQDV---IVPYERLSEEMQRINRQGGRISSITPV-----
PYR1_MASLA	26.8	GENDRVYRIEVTGVSPGYPSVRRSSYAIIVPYERLSEKIQQIHKLGKKIVSITSA-----
PYR1_ANASP	26.8	GEGDRVYRLEVTGIPSPGYPSVRRSSTVFIVPYERLSDKIQQVHKQGGKIVSVTSA-----
PYR1_FREDI	25.0	GKVYRIEVTGYRAKTFNNISKFRRSNQVFLVPYEKLSQEYQRIHQGGVIASITPV-----
PYR1_SYNEL	23.2	GESGRVYRIEVAGIQPGYPGVRRSSTAFVLPYEQLSAKMQLQRTGARIISVNPA-----
PYR2_MASLA	24.6	MGDARMFIVEAIAAGTLNTNVAVRRSRQVYTVPYDRLSATYQEIHKRGGKIVKITPAS-----
PYR2_ANASP	23.2	KGDNRMFVIEAIAAGGLNTNVAVRRSRQVYTVSYERLSATYQEIHKRGGKIVKISQV-----
PYR3_FREDI	21.1	GSTPRLYRIEVTGISLPRYPKVRRSNKEFIVPYEQLSSTLQQINKLGGKIVASITFAQ-----
PYS1_MASLA	27.9	FRFEVVGMRQNEENDKNKYNIIRRSQSVYITVPYNRMSEEMQRIRHLGGKIVKIEPLTRAAG-----
PYS1_ANASP	27.9	FRYEVLGLRQSSSETDKNKYNIIRNSQSVFITVPYSRMNEEYQRITRLGGKIVKIEQLVSAEA-----
PYS1_SYNEL	27.1	YEVVGLRQNEETDRMEFPIIRSGSTFIT--VPYNRMNEEMQRITRMGGKIVSITPVVAS-----
PYS1_SYNP2	23.1	FTYEVQGLRQTEETDNQEYAFRRSGSVFINVPYARMNQEMQRILRLGGKIVSIKPYTGATASDEE--
PYS1_SYNP6	21.0	FVYEVEGLHQSDQTDLNRYPVRQSGTVQFKVPYARMNEELRRINRLGARIVSIKPLTAEAE-----
FENR_SYNP2	26.9	FIYEVLGLGGDGRNENSLVIRKSGSTFIT--VPYARMNQEMQRITKLGGKIVSIRPAEDAAQIVSEGG
FENR_ANAVA	22.4	FVYEVLGMRQNEETDQTNYPPIRKSQSVFIRVPYNRMNQEMQRITRLGGKIVSIQTVSALQQLNGKTT
FENR_ANASO	22.4	FVYEVLGMRQNEETDQTNYPPIRKSQSVFIRVPYNRMNQEMQRITRLGGKIVTIQTVSALQQLNGRRT
FENR_SYNY3	22.4	YEVIGLSQSTMTDGLDYPIIRSGSTFIT--VPLKRMNQEMRRITRMGGKIVSIKPLEGDSPLPHTEG

### Comparison of PYC1 sequences ( $L_C$ linkers only) for evolutionarily conserved residues



#### For PYC1 sequences only:

Sequence length = 67 residues

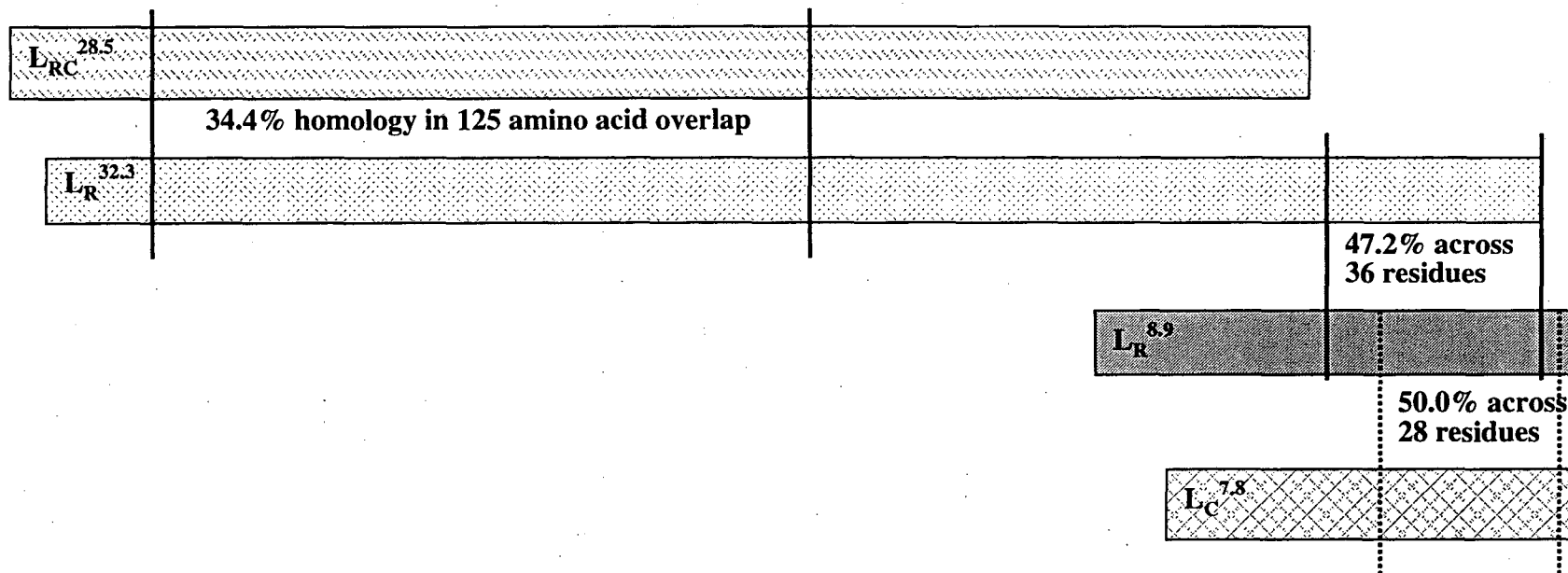
Average overall sequence homology score = 86.3%

Number of residues in the consensus sequence with greater than 80% conservation across species = 54 (or 80.6% of 67)

**Table 3-4.** Summary of amino acid residues that could be affected as a result of the linker-biliprotein interaction

Residue	Reasoning	Analogous sites		
<b>LC<sup>7.8</sup></b>		<b>LR<sup>8.9</sup></b>	<b>LR<sup>32.3</sup></b>	<b>LRC<sup>28.5</sup></b>
F37  R15 R17 R20	Aromatic residue disturbs coplanar arrangement of APC $\beta$ 81 chromophore with APC Y87 and changes its conformation  Located in loop close to the APC $\beta$ 81 side chains	Y48  R24 R37 R38	Y138 F168 R255 R256	F53  R159-161 R200 R203
<b>APC</b>		<b>PC</b>		
Y87  E114	Coplanar arrangement with $\beta$ 81 ring D may extend the "aromaticity" of the chromophore and result in lower energy band  Located near $\beta$ 81 side chains and side chain orientation is affected by linker association	Missing  E115		

**Figure 3-1.** Schematic diagram of homologous regions in the amino acid sequences of *Synechococcus* sp. PCC 7002 rod linker peptides and  $L_C^{7.8}$  (*M. laminosus*). Sequence details shown in Figures 3-2 and 3-3.



**Figure 3-2.** Schematic of amino acid sequence homology between the *Synechococcus* sp. PCC 7002 rod linker peptides

Only regions of homology are shown; for complete sequences refer to Table 3-1.

- (|) represents a direct match
- (:) represents a related amino acid
- ( ) means that there is no match

**34.4% identity in 125 amino acid overlap**

	10	20	30	40	50	60	70	80	90
<b>LRC<sup>28.5</sup></b>	YAPS	SONTRVAGYT	VGGDEQPFVF	TTDNVISDSD	FDVLINAAYR	QIFFHAF--K	CDRQQLLESQLR	NGQITVRDFI	RGLLLSETFI
<b>LR<sup>32.3</sup></b>	MPVTVAASRL	GTAAFDQSPV	ELRANYSRDD	AQTVIRAVYR	QVLGNDYVMS	SERLTAAESL	FTNGFISVRDFV	RAVAQSELYK	
		10	20	30	40	50	60	70	82

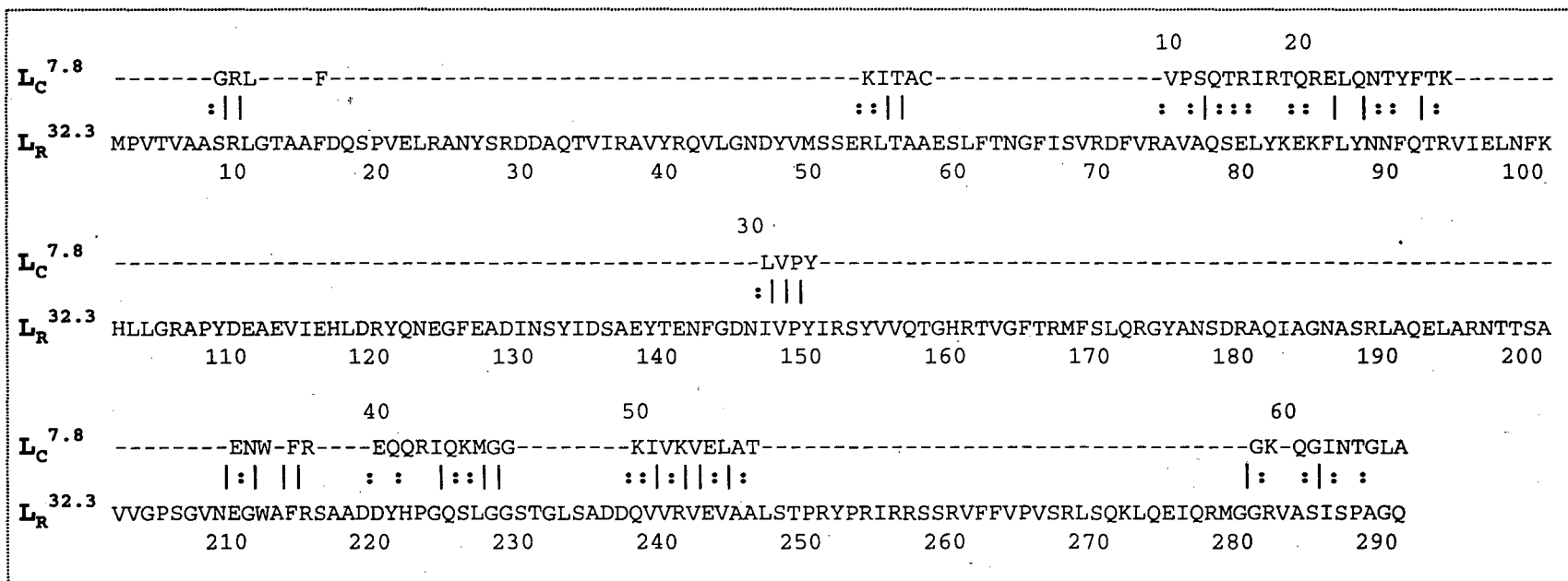
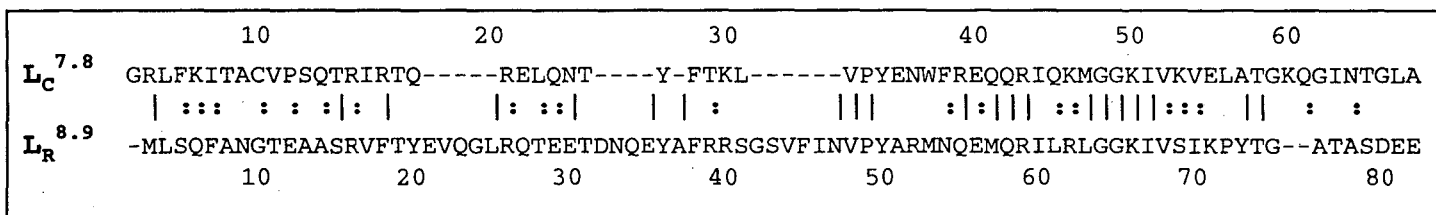
	100	110	120	130	140	150	160	170	177
<b>LRC<sup>28.5</sup></b>	DSFYNKNSNY	RFVEQCIQRV	LGRDPFSEQE	KIAWSIVICT	KGLAAFVDQL	LNTDEYEMENF	GYDTVYPYQRF	RSLASREQGE	IPFNIKS
<b>LR<sup>32.3</sup></b>	EKFLYNNFQT	RVIELNFKHL	LGRAPYDEAE	VIEHLDRYQN	EGFEADINSY	IDSAEYTENF	GDNIVPYIRS	YVVQTGHRTV	GFTRMFS
	92	102	112	122	132	142	152	162	169

**47.2% identity in 36 amino acid overlap**

	231	250	260	270	280	290
<b>LR<sup>32.3</sup></b>	..GLSADDQVVR	VEVAALSTPR	YPRIRRSSRV	FF-VPVSRLSQ	KLQEIQRMGG	RVASISPAGQ
<b>LR<sup>8.9</sup></b>	AASRVFTYEV	QGLRQTEETD	NQEYAFRRSGSV	FINVPYARMNQ	EMQRILRLGG	KIVSIKPYTGATASDEE
	20	30	42	53	63	80

**Figure 3-3.** Comparison of amino acid sequences from *Synechococcus* sp. PCC 7002 rod linkers and L<sub>C</sub><sup>7.8</sup> (*M. laminosus*) prepared using BLAST database search and Multalin sequence alignment programs.

(|) represents a direct match  
 (: ) represents a related amino acid  
 ( ) means that there is no match





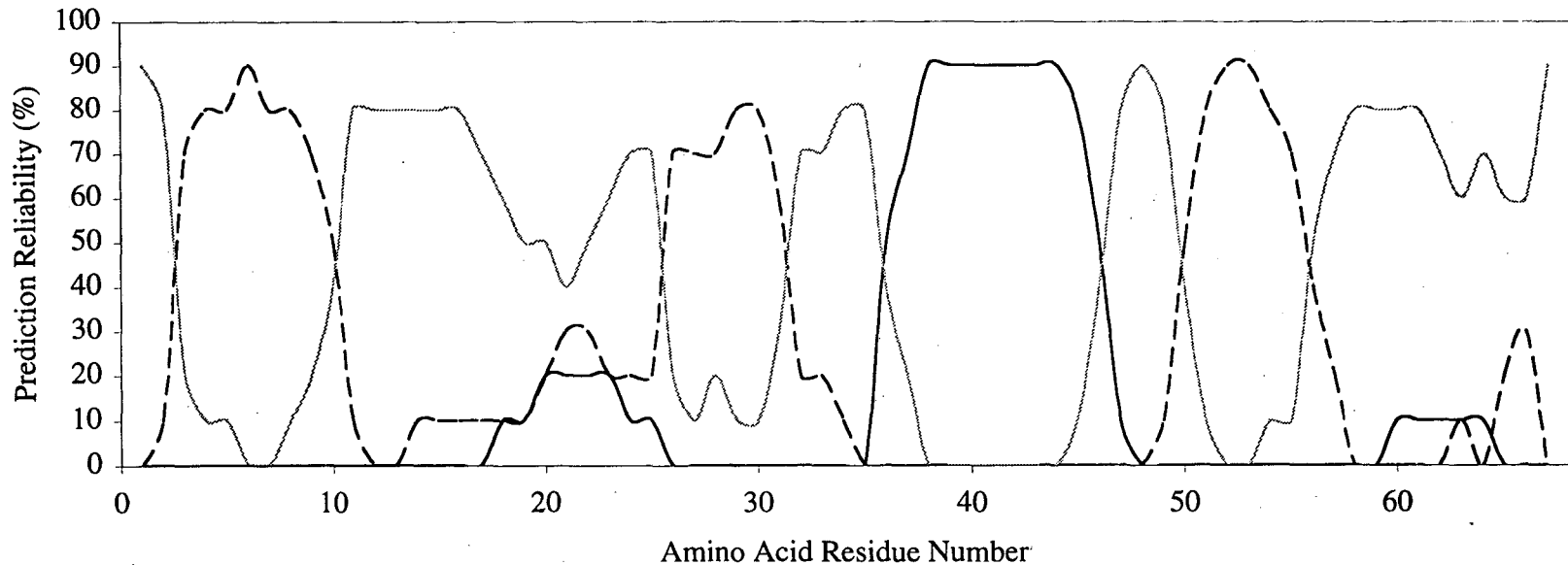


**Figure 3-4A.** Secondary structure predictions for the amino acid sequence of *M. laminosus* core linker L<sub>C</sub><sup>7.8</sup>

**Highlighted** residues indicate areas of linker/chromophore interaction

		Accuracy
Amino acid sequence:	GRLFKITACVPSQTRIR <sup>■</sup> TR <sup>■</sup> QRELQNTYFTKLV <sup>■</sup> PYENW <sup>■</sup> FR <sup>■</sup> EQQRIQKMGGKIVKVELATGKQGINTGLA	
Crystallography results:	-- <b>EEEEEEE</b> ----- <b>HHHHEEEEEEEHHHHHHHHHHHHHHH</b> ----- <b>EEEEEEE</b> -----	100%
Consensus prediction:	-- <b>EEEEEEE</b> ----- <b>HHHHHHHHHEEEEE</b> ----- <b>HHHHHHHHHHHH</b> ----- <b>EEEEEEE</b> -----	(85%)
DSC prediction:	--- <b>EEEE</b> ----- <b>EEHHHH</b> --- <b>EEEE</b> ----- <b>HHHHHHHHHHHH</b> --- <b>EEEEEEE</b> -----	74.6%
MULPRED prediction:	--- <b>EEEE</b> ----- <b>EHHHHH</b> ----- <b>EE</b> ----- <b>HHHHHHHHHH</b> --- <b>EEEE</b> -----	61.2%
NNSSP prediction:	- <b>EEEEEEEE</b> ----- <b>HHHHHHHHHHHH</b> ----- <b>HHHHHHHHHHHHHH</b> --- <b>EEEEEEE</b> ----- <b>EE</b>	73.1%
ZPRED prediction:	<b>HHHEEEEE</b> ----- <b>EEEE</b> ----- <b>EEEE</b> ----- <b>HHHHHHHH</b> --- <b>EEEEEH</b> ----- <b>EEEE</b>	61.1%
PREDATOR prediction:	--- <b>EEEE</b> ----- <b>HHHHHHHHHEEEEE</b> ----- <b>HHHHHHHHHHHH</b> --- <b>EEEE</b> -----	79.1%
PHD prediction:	-- <b>EEEEEEE</b> ----- <b>HHHHHHHHHEEEEE</b> ----- <b>HHHHHHHHHH</b> --- <b>EEEE</b> -----	83.6%

— (H) α-helix    - - - - (E) β-strand    - - - - (-) random loop



**Figure 3-4B.** Secondary structure predictions for the amino acid sequence of *Synechococcus* sp. PCC 7002 rod linker L<sub>R</sub><sup>8.9</sup>

(Highlighted residues indicate areas of possible linker/chromophore interaction)

Sequence: MLSQFANGTEAASRVFTYEVOGLRQTEETDNQEYAFRRSGSVFINVPYARMNQEMQRILRLGGKIVSIKPYTGATASDEE  
 Consensus: -----EEEEEEEE-----EEEE-----HHHHHHHHHH-----EEEE-----  
 PHD: -----EEEE-----EEEEEEEE-----EEE-----EEEE-----HHHHHHHHHH-----EEEE-----

— (H) α-helix    - - - (E) β-strand    ····· (-) random loop

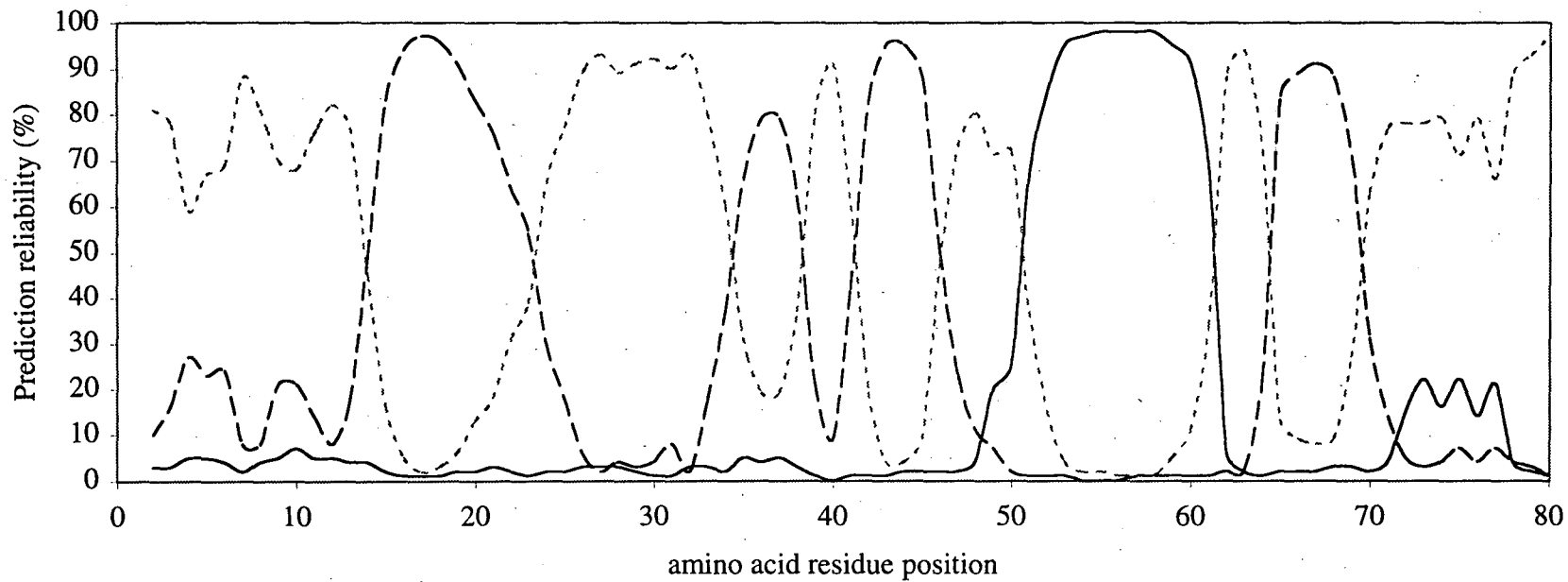


Figure 3-4C. Secondary structure predictions for the amino acid sequence of *Synechococcus* sp. PCC 7002 rod linker L<sub>R</sub><sup>32.3</sup>

**Residues 1-80**

Sequence:	MPVTVAASRLGTAAFDQSPVELRANYSRDDAQT VIRAVYRQVLGNDYVMSERLTAESLFTNGFISVRDFVRAVAQSEL
Consensus:	--- <b>HHHHHH</b> ----- <b>HHHHHHHHHHHHHHHH</b> ----- <b>HHHHHHHHHHHH</b> ----- <b>HHHHHHHHHHHH</b>
PHD:	-----HHHHHHHHHHHHHHHH-----HHHHHHHH-----EHHHHHHHHHHHHHH
Reliability:	9863322345632237999310247775799999999999971553121111147888633896222799997554343

**Residues 81-160**




Sequence:	YKEKFLYNNFQTRVIELNFKHLLGRAPYDEAEVIEHLDRYQNEGFEADINSYIDSAEYTENFGDNIVPYIRSYVVQTGHR
Consensus:	<b>HH-HEE</b> ----- <b>HHHHHHHHHHHH</b> ----- <b>HHHHHHHHHHHH</b> ----- <b>HHHHHH</b> ----- <b>HHHHHH</b> -----
PHD:	HE-EEEEEE---EEEEEEHHHHH-----HHHHHHHHHHH-----HHHHHH---HHHHHH-----EEEE-----E
Reliability:	30113443342144431213313889974234532221241356320132211333444413676202122223356621

**Residues 161-240**

Sequence:	TVGFTRMISLQRGYANS DRAQIAGNASRLAQELARNTTSAVVGPSGVNEGWAFRSAADDYHPGQSLGGSTGLSADDQVVR
Consensus:	--- <b>HHHHHHHH</b> ----- <b>HHHHHHHHH</b> ----- <b>EE</b> ----- <b>EEE</b>
PHD:	EE--EEEEEE-----HHHHHHHHHH-----EE-----EEE
Reliability:	21224356521233336633468984145788754246621157788877533457667888768888655677653599

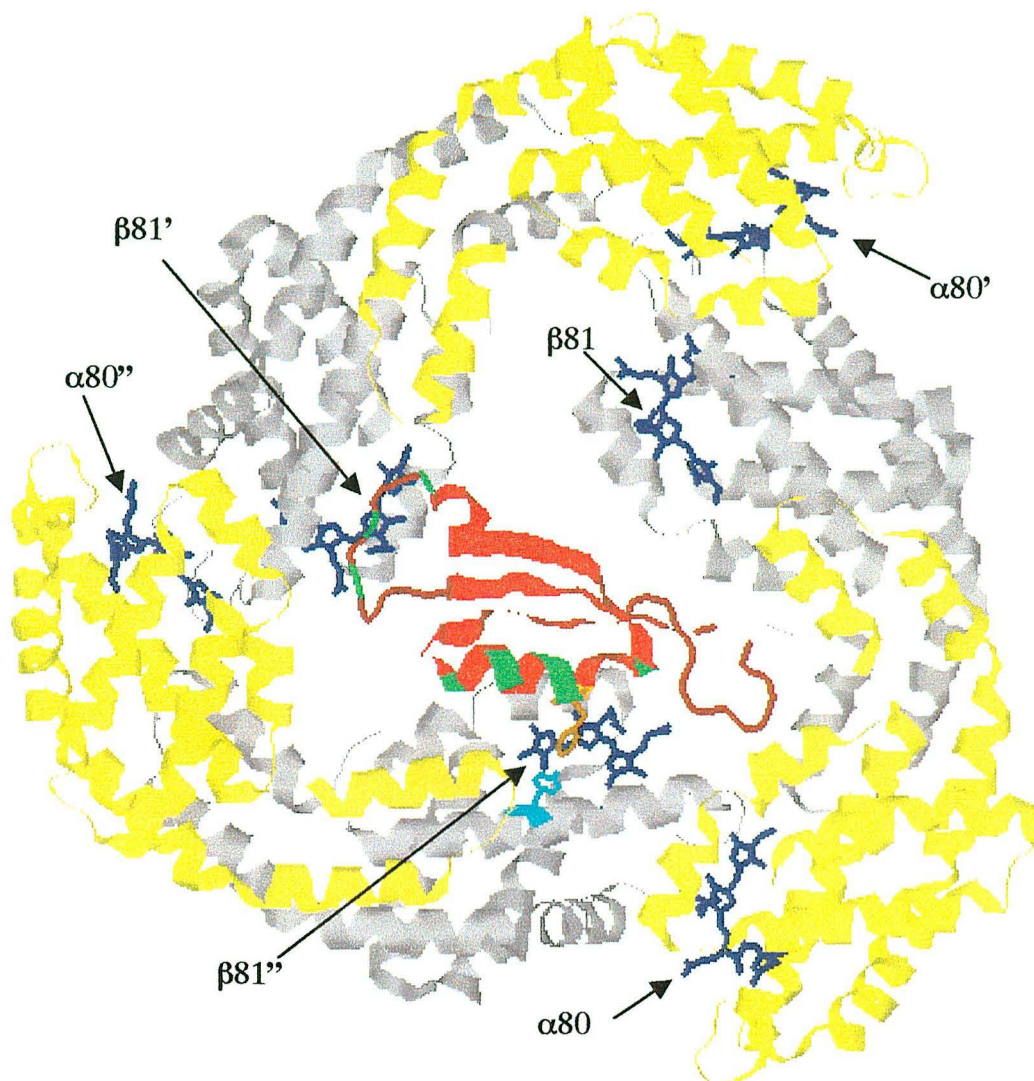
**Residues 241-290**

Sequence:	VEVAALSTPRYPRIRRSSVFFVPSRLSQKLQEIQRMGGRVASISPAGQ
Consensus:	<b>EEEE</b> ----- <b>EEEE</b> ----- <b>HHHHHHHHHH</b> ----- <b>EEEE</b> -----
PHD:	EEEE-----EE--EEEE-----HHHHHHHHH-----EEEEEE-----
Reliability:	99874157999641111225674256724899999851996588851679

	Region homologous with L <sub>RC</sub> <sup>28.5</sup>
	Region homologous with L <sub>R</sub> <sup>8.9</sup>
	<b>Highlighted</b> residues indicate areas of possible linker/chromophore interaction



**Figure 3-5.** Schematic representation of APC/L<sub>C</sub><sup>7.8</sup> isolated from *M. laminosus* (adapted from reference 1)



**Electrostatically charged residues in L<sub>C</sub><sup>7.8</sup> near APC**

(α-helix region near β81'')

Arg38\*      Glu34

Arg42      Glu39

Lys45\*

(loop region near β81')

Arg15

Arg17

Arg20

\* = residues < 4 Å from APC

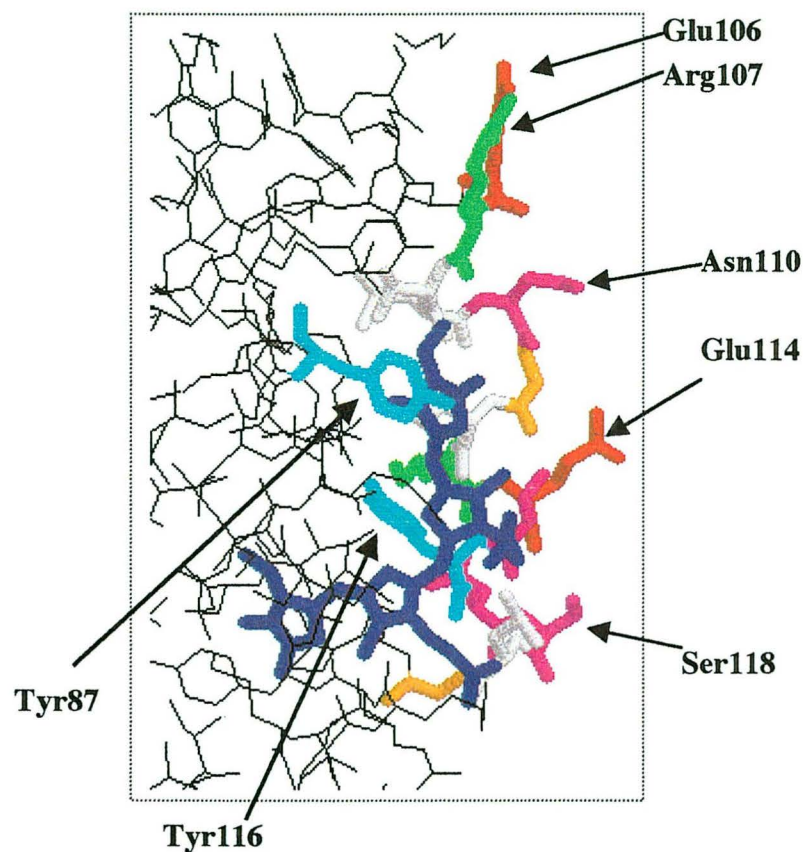
**Legend**

- APC α-subunit
- APC β-subunit
- APC chromophores
- L<sub>C</sub><sup>7.8</sup> linker
- L<sub>C</sub><sup>7.8</sup> (Phe37)
- APC (Tyr87)

**Figure 3-6[A-C].** Close-up view of the APC  $\beta$ -81 chromophore at the linker binding region. (Adapted from reference 1)

The APC  $\alpha$ -helix that extends into the central cavity (residues 116-120) as well as the  $\beta$ 81 chromophore are portrayed in stick format and specific amino acids are identified and coded according to the legend key below. The surrounding protein is shown in black wireframe format.

**[A] linker-free  $\beta$ 81**

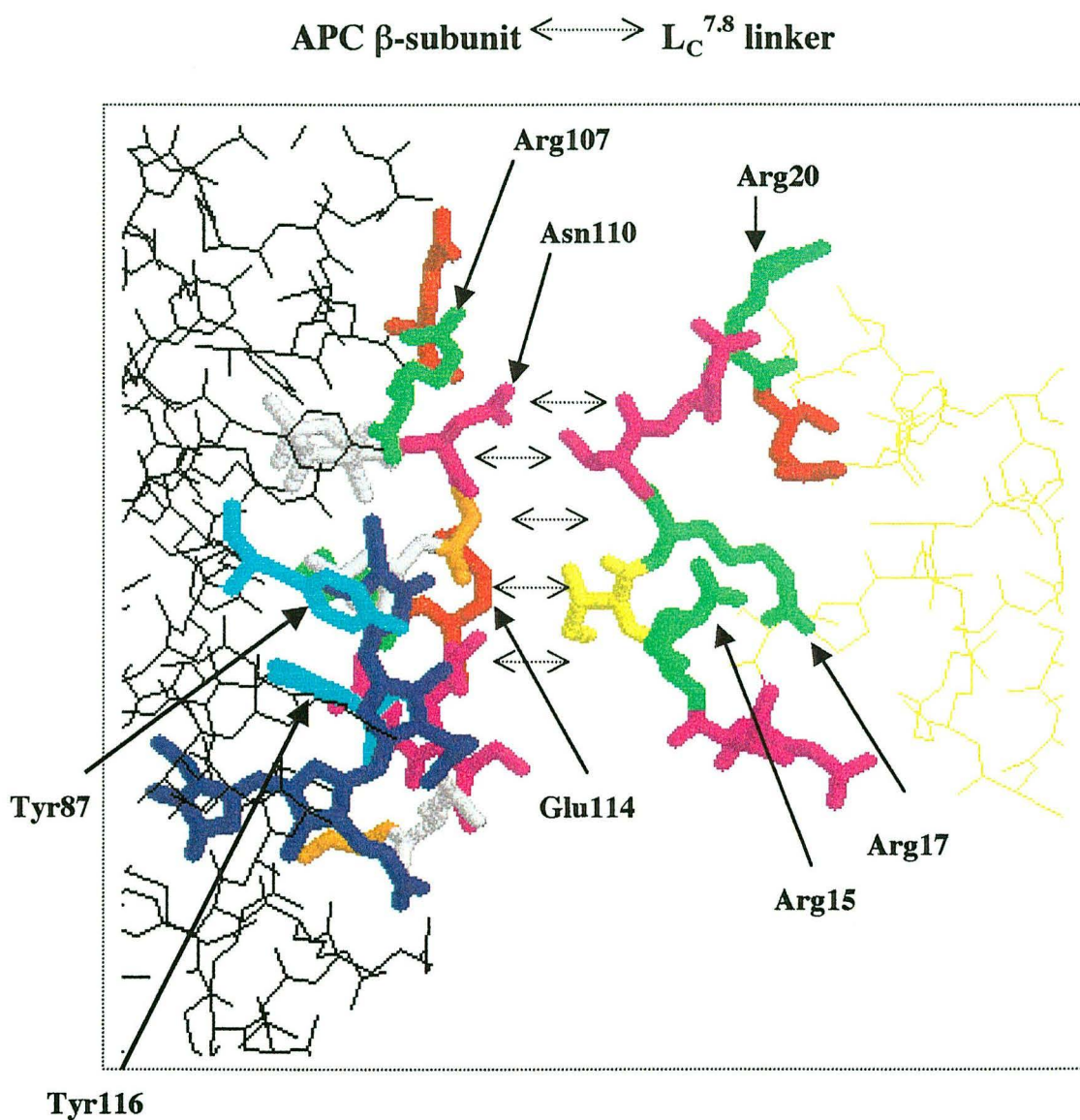


**Legend for Figures 3-6 and 3-8**

- APC  $\beta$ \_subunit
- APC  $\beta$ 81 chromophore
- $L_C^{7,8}$  linker
- (+) charged residues
- (-) charged residues
- aromatic residues
- polar residues
- nonpolar residues

### 3-6[B]. APC $\beta$ 81 at the $L_C^{7.8}$ loop segment binding site

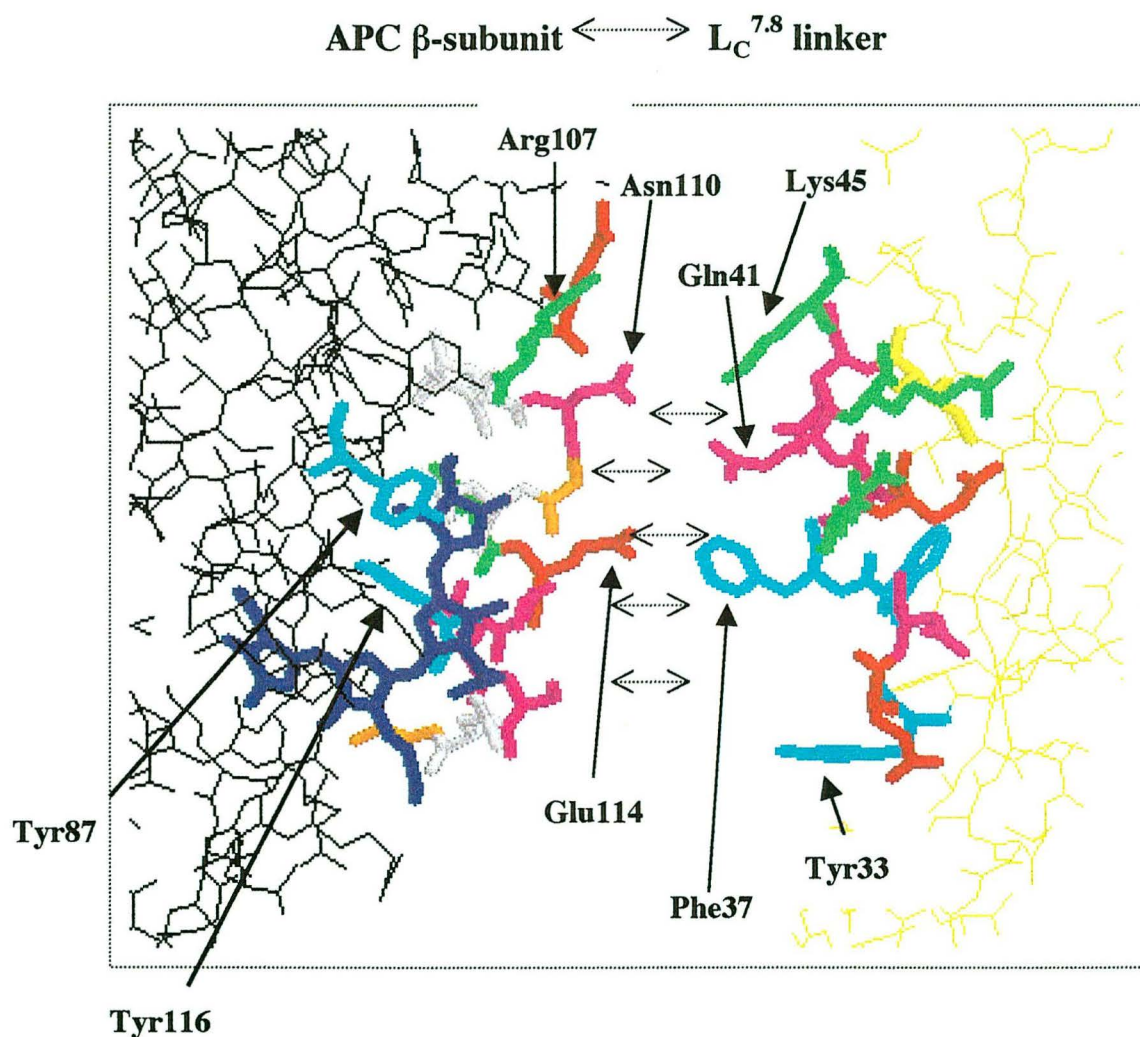
The APC  $\alpha$ -helix that extends into the central cavity (residues 116-120), the  $\beta$ 81 chromophore, and the  $L_C^{7.8}$  loop segment (residues 13-21) are portrayed in stick format. The surrounding protein is shown in wireframe format. APC and  $L_C^{7.8}$  are artificially separated by  $\sim 5$  Å. Specific amino acids are identified and coded according to the legend key in figure [A].



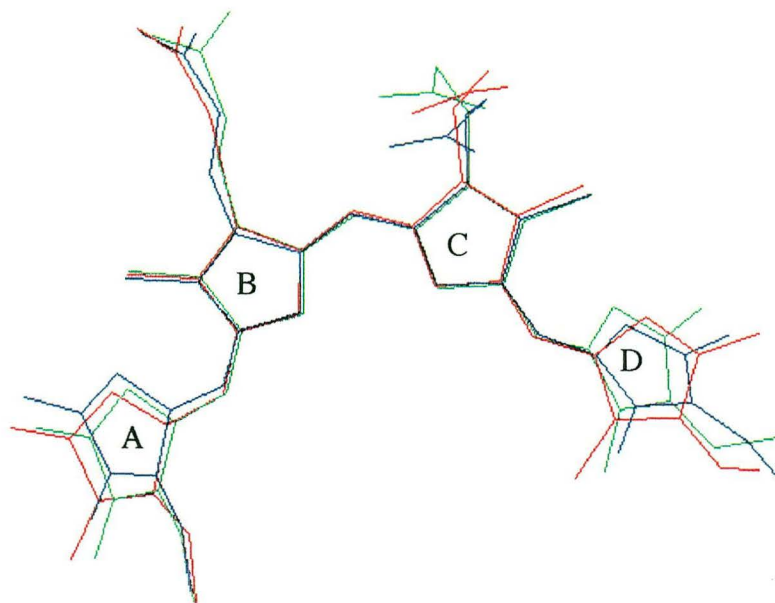


### 3-6[C]. APC $\beta$ 81 at the $L_C^{7.8}$ $\alpha$ -helix binding site

The APC  $\alpha$ -helix that extends into the central cavity (residues 116-120), the  $\beta$ 81 chromophore, and the  $L_C^{7.8}$   $\alpha$ -helix segment (residues 33-45) are portrayed in stick format. The surrounding protein is shown in wireframe format. APC and  $L_C^{7.8}$  are artificially separated by  $\sim 5$  Å. Specific amino acids are identified and coded according to the legend key in figure [A].

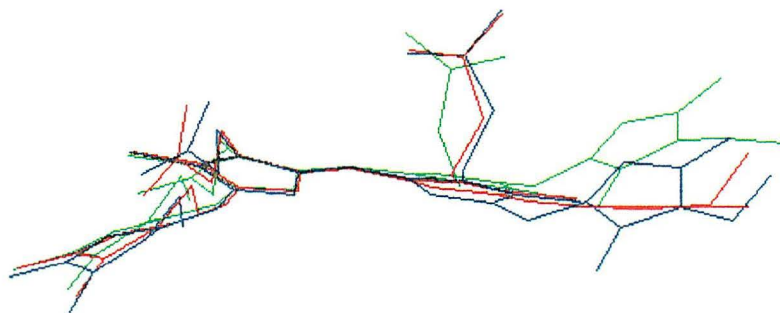


**Figure 3-7A.** Phycocyanobilin ( $\beta 81$ ) chromophores from *M. laminosus* APC/ $L_C^{7.8}$  crystal structure (reference 1)

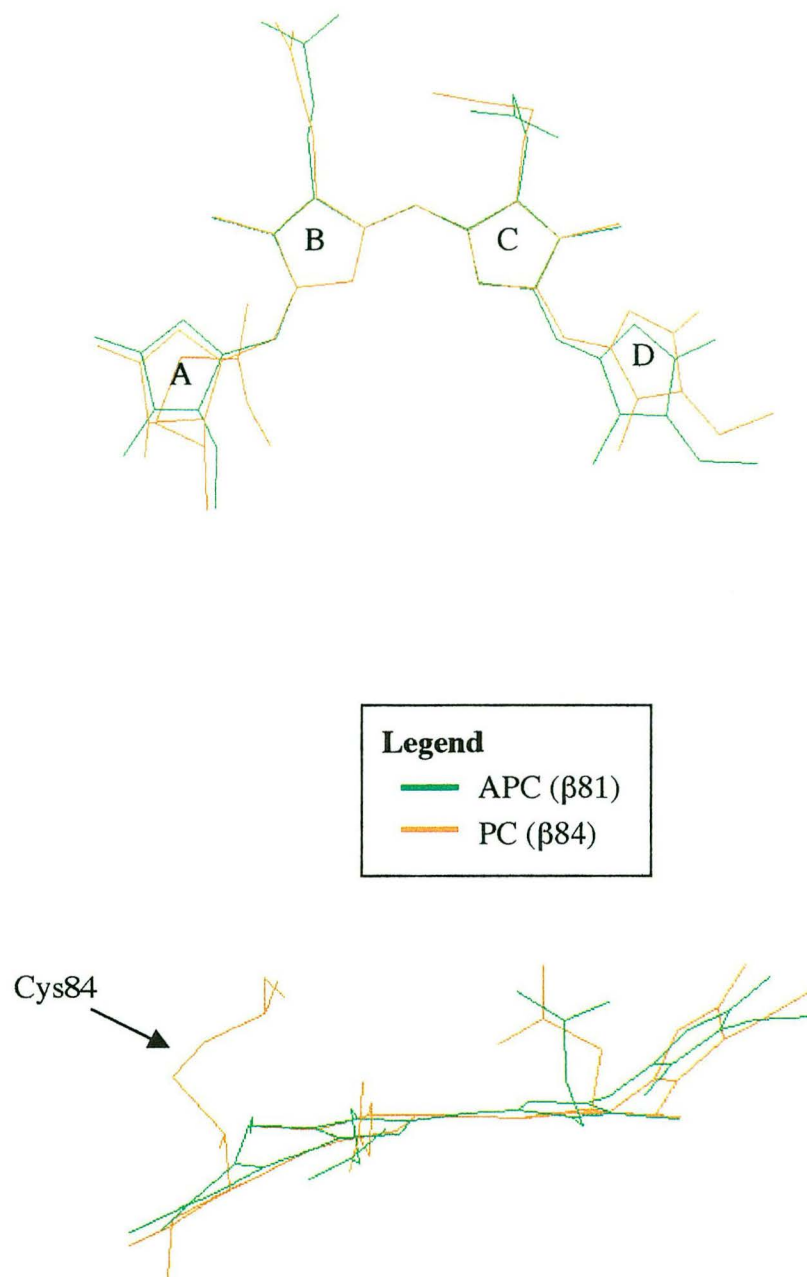


**Legend**

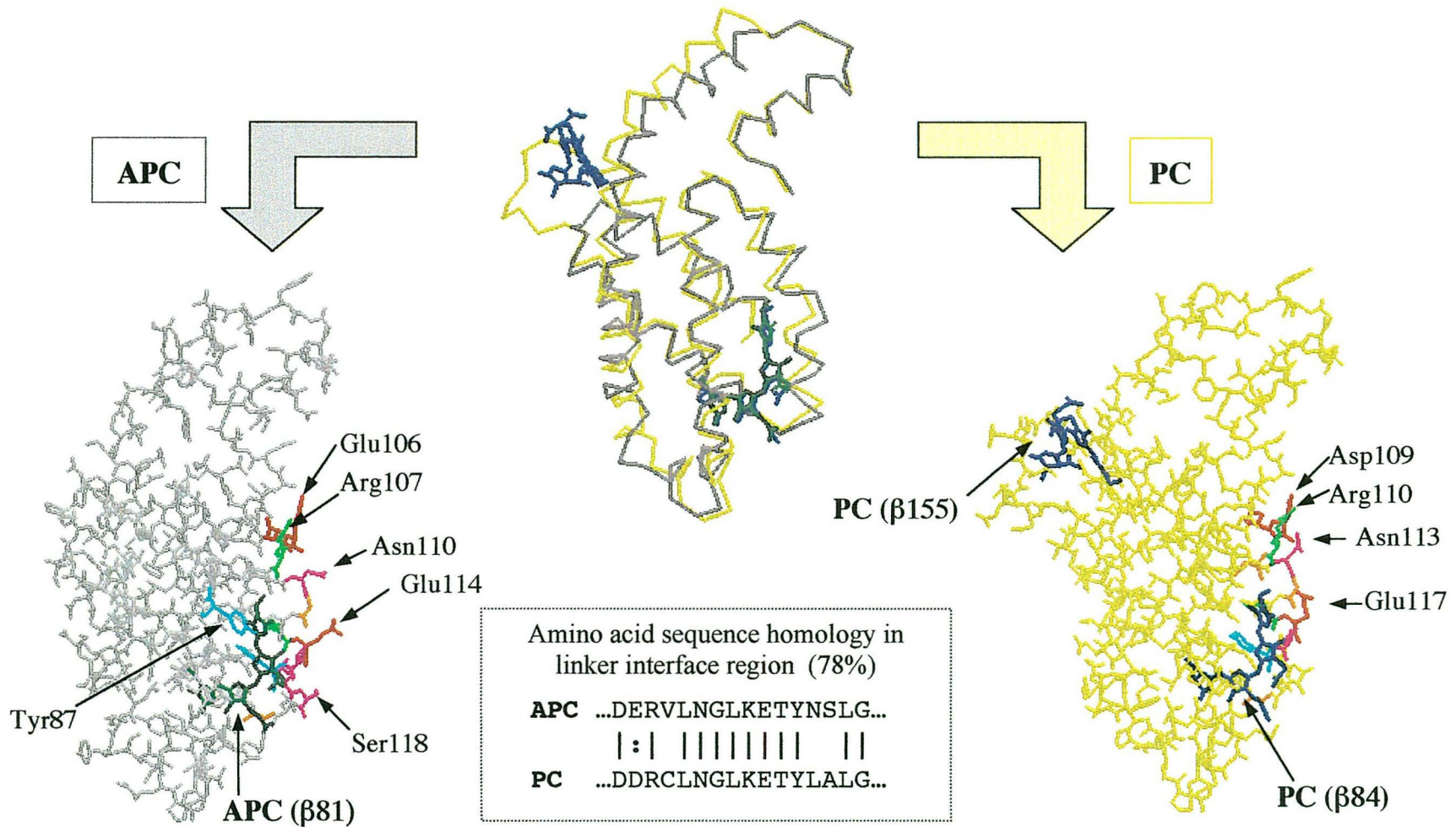
- $\beta 81$  (linker-free)
- $\beta 81'$  (near  $L_C^{7.8}$  loop region)
- $\beta 81''$  (near  $L_C^{7.8}$   $\alpha$ -helix region)



**Figure 3-7B.** Comparison of phycocyanobilin chromophores from *M. laminosus* APC/LC<sup>7,8</sup> (linker free  $\beta$ 81) and *F. diplosiphon* PC ( $\beta$ 84)



**Figure 3-8.** Comparison of the  $\beta$ -subunits from APC/L<sub>C</sub><sup>7,8</sup> (*M. laminosus*-linker free monomer) and PC (*F. diplosiphon*)



## Chapter 4. Steady State Spectroscopy

### Introduction

The phycobilisome (PBS) is a macromolecular structure composed of chromophore-containing proteins and 'colorless' linker peptides, that acts as a light-harvesting antenna for the cyanobacterial photosynthetic system.<sup>1</sup> A link between light absorbed by the PBS proteins and photosynthetic activity was first observed in the 1940's.<sup>2-5</sup> Since then, various spectroscopic techniques have been used to explore the spectral characteristics and energy transfer dynamics of this efficient system. This chapter focuses on the spectroscopy of C-phycoerythrin (PC) complexes isolated from the PBS rods of the cyanobacterium *Synechococcus* sp. PCC 7002. There are four isolated PC complexes examined: PC monomer, PC trimer, PC/L<sub>R</sub><sup>32.3</sup> and PC/L<sub>RC</sub><sup>28.5</sup>, where the PC/linker complexes are composed of one PC trimer and one linker polypeptide.

Although it is the spectral properties of the PC chromophores that are actually being measured, the surrounding protein matrix is of paramount importance. It not only functions as a structural support for the chromophores but also modulates their spectral properties to guide the absorbed energy through a well-ordered scaffold. This study explores the effect of associated linker peptides on the spectral characteristics of the biliprotein chromophores using a variety of techniques. Absorption and fluorescence measurements of each PC complex show that the PC chromophores are sensitive to the presence of the linker. To attempt to pinpoint the direct interaction of each chromophore with the linker peptides, spectroscopic measurements were carried out on each isolated PC complex over a range of temperature. Additionally, fluorescence excitation anisotropy and circular dichroism (CD) measurements performed at room temperature

examine the strength of coupling interactions among the chromophores in each PC complex. Lastly, the Kennard-Stepanov relation between absorbance and fluorescence bands is introduced and explored as a method to compare the inhomogeneity present in each PC complex.

## Results

### *Experimental note concerning low temperature measurements*

The temperatures reported throughout this report were gathered from a thermocouple inserted directly above the sample cuvette (refer to Figure 2-2). The liquid nitrogen cryostat used possessed its own thermocouple detector within the heating control unit located ~3 cm above the sample compartment. At each selected temperature, ~20 minutes was allowed for equilibration of these two thermocouples to the same value.

However, at the lowest temperature there is an irreconcilable difference, the instrument showed that liquid nitrogen temperature was achieved (77 K) while the measured temperature above the sample was 88 K. Despite allowing 3+ hours for equilibration, the values from these two thermocouples did not come to a consensus and this remains a limitation of the instrument design. We therefore regard the spectra reported here at 88 K to be equivalent to published "77 K" results where similar equipment was used.

### *Absorbance*

Figures 4-1A to 4-1D show the absorbance of each PC complex studied in a range of temperatures from 298 K to 88 K taken at approximately 20-degree intervals. To compare the absorption bands of the four PC complexes with each other at 298 K and 88 K, they are plotted together in Figures 4-2A to 4-2D. The maximal absorption

wavelengths at these two temperatures are noted in Table 4-1. As the PC aggregates become larger, there is a shift in the maximal absorbance position towards lower energy, but the magnitude of this energy shift differs. Aggregation of PC monomers into trimers results in a red-shift of ~6 nm while association of PC trimer with  $L_R^{32.3}$  brings very little change to the PC trimer absorbance maximum position. The largest shift from the PC trimer maximum ( $\Delta_{MAX}$  (absorbance) = 7.6 nm) occurs in the  $PC/L_{RC}^{28.5}$  absorbance, and it is accompanied by a change in the band shape that mimics the absorbance of the core biliprotein allophycocyanin (APC).

The first observed effect upon cooling from room temperature is a narrower absorption band. Secondly, a gradual shift in the maximal absorbance wavelength position towards lower energy is detected with decreasing temperature. These maximal absorbance positions are plotted versus temperature on the inset graphs in each of the Figures (4-1A to 4-1D). Except for  $PC/L_{RC}^{28.5}$ , the trend of these positions as the temperature drops is linear and has a steep rising slope until ~180 K, where the curve rises at a gentler slope as the temperature is decreased. The trend of maximal absorbance positions for  $PC/L_{RC}^{28.5}$  rises slowly to a maximum at ~180 K and then decreases, within a total range of 1.6 nm over the entire temperature region (inset Figure 4-1D). Unlike the other PC aggregates, which show a difference in their maximal absorption position of ~10 nm when compared at 298 K and 88 K, the  $PC/L_{RC}^{28.5}$  complex has a very small shift (1 nm, Table 4-1).

At the lowest temperatures, distinct features in the absorption spectra are resolved. In the PC monomer, two distinct peaks are resolved from the single 298 K band with absorption maxima at 630 nm and 604 nm, and at 88K there is an additional

shoulder feature at ~577 nm (Figure 4-1A). The PC trimer absorbance has several features that emerge at lower temperatures (Figure 4-1B). Once again, two predominant band maxima (636 nm and 608 nm) become resolved as the temperature drops, and at 88 K these peaks are accompanied by two pronounced shoulders (~579 nm and ~629 nm). The PC/L<sub>R</sub><sup>32.3</sup> 88 K absorbance is very similar to that of the PC trimer, with very slight increases in intensity at the peak near 607 nm and the shoulder feature at ~629 nm (Figure 4-2D). In contrast, the 88 K absorbance spectrum of the PC/L<sub>RC</sub><sup>28.5</sup> complex is very different from that of the PC trimer. There are two band maxima (605 nm and 638 nm) whose positions are slightly different from those found in the PC trimer, but the long-wavelength region (620-660 nm) has a very different profile (Figure 4-2D). The shoulder feature previously observed at ~629 nm in both the PC trimer and PC/L<sub>R</sub><sup>32.3</sup> 88 K absorbance is absent while a novel one appears at ~646 nm. These changes are even more apparent in the first derivative of the absorption spectra (Figure 4-3). In comparison to the PC trimer, the first derivative of the 88 K absorbance of the PC/L<sub>R</sub><sup>32.3</sup> aggregate shows slight changes in the intensity of the maxima/minima and a new small minimum at ~653 nm. The PC/L<sub>RC</sub><sup>28.5</sup> aggregate has an entirely different first derivative shape and the minimum at ~653 nm is the most intense feature.

#### *Circular dichroism (CD) measurements*

To test the influence of the protein environment on the optical activity of the PC chromophores, we measured the CD of each PC complex at room temperature. An excitonically coupled pair of chromophores is easier to detect by CD than from an absorbance spectrum, because the CD is dominated by the strength of the coupled



oscillator. In other words, slight changes in the distance and geometric orientation of the two chromophores can greatly impact the magnitude of the CD bands. Figure 4-4 shows the CD traces of all of the PC complexes. The PC monomer possesses a broad single band similar to its absorbance profile, while all of the other PC complexes display a two-component band (i.e., a peak feature with a less intense shoulder on its high energy side) in the visible wavelength region. The approximate maximal positions of these CD bands are recorded in Table 4-2. Comparing the CD traces of the two PC/linker complexes to that of the PC trimer, one can see that the magnitude of the shoulder feature is decreased in both PC/linker complexes and that, additionally, the peak band of PC/LRC<sup>28.5</sup> is shifted to lower energy where a negative band is also observed (Figure 4-4).

#### *Excitation anisotropy measurements*

Polarized fluorescence excitation measurements were gathered for each PC complex to examine the effect of the different protein environments on the emission characteristics of the PC chromophores. The trimeric complexes show low overall anisotropy values in comparison to the PC monomer (Figure 4-5). Across all excitation wavelengths, the PC monomer has markedly higher anisotropy values which increase at long wavelengths (> 600 nm) to reach near maximal anisotropy (0.4) at ~655 nm. This rise in anisotropy at the red wavelength region is also observed in the traces of the other PC complexes, although at much lower magnitudes. The anisotropy trace for the PC monomer reveals a smaller rise in the central wavelength region (~580 nm). This feature is much smaller in the other complexes but can be seen when the y-scale is amplified. Attempts were made to quantify the magnitude of the long wavelength rise in the

anisotropy trace of each PC complex by drawing a horizontal line through the 450-550 nm region and then subtracting it from the anisotropy values at 630 nm ( $\Delta_{630}$  values summarized in Figure 4-5). As confirmed from visual inspection, the largest anisotropy amplitude  $\Delta_{630}$  is observed in the PC monomer. The PC trimer and PC/linker complexes show significant differences in  $\Delta_{630}$  where the PC/L<sub>R</sub><sup>32.3</sup> complex rises twice as much as the other two:  $\Delta_{\text{anisotropy}}_{630}$  (PC/L<sub>R</sub><sup>32.3</sup>) = 0.07 versus 0.03 for the PC trimer. Generally, the two linkers appear to have opposite polarization effects on the PC trimer, showing anisotropy traces that are either higher (PC/L<sub>R</sub><sup>32.3</sup>) or lower (PC/L<sub>RC</sub><sup>28.5</sup>) than that of the PC trimer at all wavelengths.

### *Fluorescence*

Each PC complex was excited at 585 nm and the fluorescence spectrum was measured at a series of temperatures. Other excitation wavelengths were used ( $\lambda_{\text{EX}} = 630$  nm and 660 nm) to test the heterogeneity of the absorption band, however, all of the information included in the data tables was obtained at  $\lambda_{\text{EX}} = 585$  nm. Figures 4-6A to 4-6D present the fluorescence spectrum of each PC complex taken over a series of temperatures (298 K to 88 K at 20-degree intervals). As the temperature drops, the emission bands become narrower, increase in intensity and the maximum position shifts towards shorter wavelengths. Table 4-3 contains the fluorescence maximum positions for each PC complex at 298 K and 88 K. While there are small variations in the maximal fluorescence position of the PC monomer, PC trimer and PC/L<sub>R</sub><sup>32.3</sup> complexes, PC/L<sub>RC</sub><sup>28.5</sup> has an emission maximum that is red-shifted ~6 nm from the others (Table 4-3). The full-width at 50% of the maximum intensity (FWHM) of the fluorescence band

of each PC complex is shown in Table 4-4. At both 298 K and 88 K, the PC monomer possesses the widest band while the PC/LRC<sup>28.5</sup> has the narrowest one. At 88 K, all of the bands become narrower by ~15 nm. The absorbance and fluorescence maximal positions are combined to calculate the Stokes shift of each PC complex, and these are presented in Table 4-5. At 298 K, the PC monomer has the largest Stokes shift (28 nm), and PC/LRC<sup>28.5</sup> has the smallest one (15 nm). At 88 K, the Stokes shift is reduced similarly for all of the PC complexes except for PC/LRC<sup>28.5</sup> where it appears to retain the value observed at 298 K. However, PC/LRC<sup>28.5</sup> is the only complex to possess a shoulder feature on the long wavelength side of its absorption maximum, which may be the emission source.

The fluorescence maximal positions do not change so much as those of the absorption bands, and their behavior as a function of temperature is different for every aggregate. Like in the absorbance measurements, the fluorescence maximum position at each temperature is plotted in inset figures for each complex. The trend of maximal fluorescence position as a function of decreasing temperature can be described with a single line in all complexes except the PC monomer, which has two trendlines. From 298 K to 178 K, the maximum fluorescence position of the PC monomer decreases more rapidly than in the lower temperature region (178 K to 88 K) (inset Figure 4-6A). Once again there is differing behavior for the two PC/linker complexes; PC/LR<sup>32.3</sup> (inset Figure 4-6C) behaves similarly to the trimer (inset Figure 4-6B), while PC/LRC<sup>28.5</sup> maintains the same fluorescence maximum position throughout the entire temperature range (inset Figure 4-6D).

Besides the shift in fluorescence maximum as a function of temperature, a second effect is observed in all of the PC complexes at low temperatures: new features emerge in the long-wavelength region of each spectrum. At 88 K there is a number of small features that appear in the long-wavelength tail of the fluorescence band and the approximate peak positions of these are also included in Table 4-3. A single feature at ~710 nm in the PC monomer 88 K fluorescence spectrum is replaced by two features in the trimeric complexes (Figures 4-7C and 4-7D). Compared to the PC trimer 88 K emission, the emission band of the PC/L<sub>R</sub><sup>32.3</sup> complex shows a slightly higher intensity at ~682 nm as well as a shoulder feature at ~649 nm (Figure 4-7D). The fluorescence of the PC/L<sub>RC</sub><sup>28.5</sup> complex displays the least amount of variation as the temperature drops; the FWHM does become narrower (Table 4-4), but the band maximum position remains constant (Figure 4-6D). When normalized to its maximum value, the long-wavelength tail appears at a lower intensity than for the other trimeric complexes at 88 K. If the spectra in Figure 4-7D are normalized at 740 nm instead of their respective maximum positions then the long-wavelength tail of PC/L<sub>RC</sub><sup>28.5</sup> approaches the intensity level of the PC trimer, and the intensity at its maximum position is increased by 16% relative to the trimer.

#### *Fluorescence measurements at different excitation wavelengths*

The fluorescence emission of each PC complex was tested at 298 K with three different excitations:  $\lambda_{\text{EX}} = 585$  nm, 630 nm, and 660 nm. The PC monomer emission is sensitive to long-wavelength excitation ( $\lambda_{\text{EX}} = 660$  nm), which produces a ~6 nm shift of the entire emission band (Figure 4-8A). In contrast, the fluorescence emission of the PC

trimer, and that of the PC trimer/linker complexes (not shown), does not vary with excitation wavelength (Figure 4-8B).

#### *Excitation measurements at different fluorescence emission wavelengths*

The excitation spectrum of each PC complex was gathered at various emission wavelengths ( $\lambda_{EM}$ ) chosen according to the features observed in the 88 K fluorescence spectra. At 298 K, all of the PC complexes produce excitation traces that are independent of emission wavelength (Figures 4-9A to 4-9D at 298 K). It is at 88 K that differences are observed. The PC monomer has an excitation band that is slightly narrower on the long-wavelength side when measured at  $\lambda_{EM} = 710$  nm rather than the maximum emission wavelength,  $\lambda_{EM} = 648$  nm (Figure 4-9A at 88 K). Although to different extents, this is also observed in the PC/linker complexes: the excitation band gathered at  $\lambda_{EM} = 714$  nm is narrower on the long-wavelength side. The excitation spectra of the PC/linker complexes displayed more variation than the PC monomer when measured at different emission wavelengths. In addition to the slightly narrower excitation band produced at  $\lambda_{EM} = 714$  nm discussed above, the PC/LR<sup>32.3</sup> complex possesses a small long-wavelength (~660 nm) component that is clearly visible at  $\lambda_{EM} = 682$  nm (Figure 4-9C at 88K). This small band may also be present in the PC/LRC<sup>28.5</sup> excitation scan taken at  $\lambda_{EM} = 682$  nm, but it is much smaller (Figure 4-9D at 88K). The variations in the excitation band of PC/LRC<sup>28.5</sup> taken at different emission wavelengths are interesting. The excitation scans taken at  $\lambda_{EM} = 682$  nm and  $\lambda_{EM} = 714$  nm emission wavelengths are identical while that taken at  $\lambda_{EM} = 729$  nm is significantly broader on the long-wavelength edge (Figure 4-9D at 88 K). It is difficult to determine whether the excitation scan

measured at the emission maximum ( $\lambda_{EM} = 655$  nm) agrees with that from  $\lambda_{EM} = 729$  nm because the lamp excitation profile interferes in this region. However, in PC/L<sub>R</sub><sup>32.3</sup> the excitation scan at  $\lambda_{EM} = 729$  nm (not shown) follows the same profile as that taken at the maximum emission ( $\lambda_{EM} = 652$  nm).

## Discussion

### *PC Monomer*

The PC monomer contains three phycocyanobilin chromophores covalently bound to cysteine residues  $\alpha 84$ ,  $\beta 84$ , and  $\beta 155$ . The shortest center-to-center interchromophore distance is found between the  $\beta$  chromophores (34 Å apart); and the third chromophore,  $\alpha 84$ , is located closest to  $\beta 155$  (50 Å). The 77 K absorbance of the PC monomer isolated from *Synechococcus* sp. PCC 7002 has been previously measured by Debreczeny<sup>6</sup> and our results concur: two distinct peaks are resolved from the single 298 K band with absorption maxima at 630 nm and 604 nm (Figure 4-1A). Although the three PC chromophores are chemically identical, their spectroscopic properties are sensitive to changes in their structural conformation and surrounding protein environment. The absorbance and fluorescence spectra of the PC monomer have been deconvoluted into three individual component bands (representing each of the chromophores) by different research groups.<sup>6,7,9</sup> A survey of the published literature<sup>6-11</sup> provides a range of maximal positions and relative amplitudes for these spectral bands, and these parameters are summarized in Table 4-6. In each case, the resolved spectral bands were assigned to the three resident PC chromophores assuming that the subunits in the PC monomer do not interact with each other spectroscopically (i.e., they are independent of each other). The

Appendix 4-1 section at the end of this chapter, simulates the absorbance spectra of the PC monomer using the spectra from either the isolated subunits, or that of the individual chromophores (deconvoluted by M. P. Debreczeny)<sup>6</sup>. It concludes that the PC monomer absorbance can be adequately described as a linear combination of the individual components at both 298 K and 88 K. If a substantial coupling interaction existed between the subunits, then this would not be possible. CD measurements and polarized excitation anisotropy traces support this assumption. The CD spectrum for the PC monomer (Figure 4-4) shows a broad single band (450-650 nm region) similar to the isotropic absorbance measurement at 298 K (Figure 4-1A). Although this CD band is somewhat wider than the absorbance band, it does not possess individual components of opposite signs or an observed band splitting that would be indicative of strong excitonic coupling among the chromophores. Mimuro *et al.*<sup>8</sup> found that in excitation anisotropy measurements of PC complexes isolated from *M. laminosus*, the highest polarization occurred in the isolated  $\alpha$ -subunit with its single chromophore, followed by the monomer and then the  $\beta$ -subunit at lower anisotropy. The authors concluded that although some depolarization occurred in the  $\beta$ -subunit due to energy transfer between  $\beta 155$  and  $\beta 84$  (hence lower anisotropy than the  $\alpha$ -subunit), there was minimal interaction between the subunits themselves since addition of the two subunit anisotropy traces adequately reproduced that of the PC monomer.

The protein binding pockets of both  $\alpha 84$  and  $\beta 84$  are located in the central cavity area of the PC trimer, and in the PC hexamer arrangement, a pair of  $\beta 84$  chromophores have the closest center-to-center distance (26 Å).<sup>11</sup> Spectroscopically, the  $\alpha 84$  and  $\beta 84$  chromophores possess nearly overlapping absorption and fluorescence spectra

(determined for the PC monomer and subunits)<sup>6-9</sup> but the  $\beta$ 84 chromophore has been assigned to the lowest energy band based on studies that chemically alter the spectral properties of the  $\alpha$ 84 chromophore;<sup>7,12</sup> as such, the  $\beta$ 84 chromophore is the final emitter in a chain of 'downhill' energy transfer. The fluorescence excitation anisotropy traces presented in Figure 4-5 show that there is a rise in the anisotropy of each PC complex as excitation approaches the long-wavelength region ( $> 600$  nm) indicating a conservation of polarization in the fluorescence emission. The chromophore fluoresces rather than transfer its energy 'uphill' and so the polarized emission in this region is enhanced. This rise is particularly pronounced in the PC monomer where both subunits are fluorescing at these wavelengths.

The fact that the PC monomer fluorescence is not independent of excitation wavelength (Figure 4-8A), showing a bigger Stokes shift for excitation on the red-edge of absorption ( $\lambda_{\text{EX}} = 660$  nm), demonstrates that this is a poor system of energy transfer. If the chromophores were efficiently transferring their energy to one final emitter, then the emission band would be independent of excitation wavelength. In other words, if the two subunits were energetically linked, then the absorbed energy would always be directed to the lowest lying chromophore (or excitonic state) and the same fluorescence band would be observed regardless of excitation wavelength. Instead the PC monomer absorbance, as a combination of individual components, possesses site heterogeneity so that fluorescence from both PC subunits is observed according to excitation wavelength. There is some energy transfer between the  $\beta$ -subunit chromophores (otherwise the anisotropy would be highly polarized throughout the entire wavelength region) but the PC monomer is an inefficient system overall. Along with the steep rise in excitation



anisotropy at long-wavelengths ( $\lambda > 600$  nm) already discussed, the excitation-dependent fluorescence band of the PC monomer indicates a strong absorption sensitivity in the long-wavelength region. Although both  $\alpha 84$  and  $\beta 84$  have similar spectral bands in this long-wavelength region and therefore are both contributing to the observed sensitivity, those of the  $\beta 84$  chromophore are lowest in energy.<sup>6,7</sup> This means that although the PC monomer is overall an inefficient system of energy transfer, its  $\beta$ -subunit is channeling its energy towards the  $\beta 84$  chromophore since there is no  $\lambda_{EX}$  sensitivity in the emission at short wavelengths ( $< 620$  nm) where  $\beta 155$  absorbs maximally, and the excitation anisotropy in this region is low compared to the long-wavelength components.

#### *PC trimer*

Formation of the PC trimer from the individual PC monomer units, induces changes in the immediate protein environment surrounding  $\alpha 84$  and  $\beta 84$ . The  $\alpha 84$  chromophore binding pocket is now completely covered by the adjacent PC monomer protein, while  $\beta 84$  remains partially exposed to the central cavity.<sup>11</sup> Because the PC monomer structure is unknown, we cannot directly compare the  $\alpha 84$  and  $\beta 84$  chromophore conformations in PC monomers and trimers; however, formation of the PC trimer will have the greatest impact on the chromophores that are close to the junction site between the monomers. PC trimer formation furnishes new interactions between the chromophores and amino acid residues from neighboring monomers that can impact the spectral characteristics of the chromophores. For example, the propionate side chain of the  $\alpha 84$  chromophore forms a salt bridge to residue R57 of the neighboring monomer.

Perhaps the most important change introduced by PC trimer formation is an interchromophore distance between neighboring  $\alpha 84$  and  $\beta 84$  chromophores (21 Å) that is in the range of weak coupling. The distance between the outermost pyrrole rings of these two chromophores is  $\sim 13$  Å, and the relative orientation factor is favorable for energy transfer on a picosecond timescale. These favorable electronic interaction conditions as well as changes in the immediate protein environment of  $\alpha 84$  and  $\beta 84$  contribute to the differences observed in the PC monomer and PC trimer spectra. At room temperature, the absorbance band maximum of the PC trimer is red-shifted  $\sim 7$  nm relative to that of the PC monomer (Figure 4-2A) and, although the fluorescence maximum remains at about the same position when both are excited at  $\lambda_{EX} = 585$  nm (Figure 4-7A), the PC monomer fluorescence band red-shifts  $\sim 6$  nm when  $\lambda_{EX} = 660$  nm (Figure 4-8A) due to site heterogeneity. Efficient energy transfer pathways between adjacent monomers are formed in the PC trimer aggregate, and these affect the spectral properties of the chromophores to enhance the flow of energy towards a terminal emitter. The fluorescence band of the PC trimer (Figure 4-7A) is narrower than that of the monomer by  $\sim 5$  nm (Table 4-5) and polarized excitation measurements show low anisotropy values for the PC trimer (Figure 4-5), indicating efficient energy transfer to the final emitter rather than direct fluorescence from the individual subunits as seen in the PC monomer.

The short interchromophore distance (21 Å) between neighboring  $\alpha 84$  and  $\beta 84$  chromophores and favorable geometry provide an opportunity for excitonic coupling. Numerous studies have used CD measurements to identify coupling interactions in phycobiliproteins.<sup>8,13-24</sup> If a pair of chromophores were strongly coupled then their CD

spectrum would possess positive and negative bands. However, phycobiliproteins are more complex than a simple chromophore dimer system; although PC has chemically identical chromophores, these are placed in different environments at slightly different conformations and resulting positive and negative bands from the individual chromophores in the trimer aggregate would be difficult to resolve. However, both the absorbance (Figure 4-2A) and CD spectra (Figure 4-4) of the PC trimer is significantly different from that of the PC monomer. Compared to a typical absorption measurement, CD is much more sensitive to coupling interactions between chromophores and even small exciton band splittings can be detected. The CD of the PC trimer is different from that of the PC monomer in that it contains two components, a peak at 635 nm and a shoulder at ~593 nm, instead of a single broad band with maximum at 605 nm (Table 4-2, Figure 4-4). Clearly, PC trimer formation provides new electronic interactions between the chromophores or changes in their structural conformations that are affecting the optical activity of the PC chromophores.

The possibility of excitonic coupling between neighboring  $\alpha 84$  and  $\beta 84$  chromophores in PC and APC trimers has been the topic of much speculation. Experimentally, both PC and APC trimers have been shown to have much faster fluorescence anisotropy decay times than their monomer counterparts and in some cases the experimental resolution has revealed femtosecond components<sup>25-31</sup> theoretically expected from coherent energy transfer.<sup>32-35</sup> In the case of APC trimers, the initial polarization anisotropy at  $t = 0$  has been observed in a range (0.58 to 0.7)<sup>25,26</sup> that is impossibly high for uncoupled relaxation processes where the theoretical limit is expected to be 0.4.<sup>36,37</sup> Although the  $\alpha 84$  and  $\beta 84$  chromophores in APC and PC are very

similar in geometry and distance, the highest initial anisotropy observed in PC trimers has not been as high ( $r(0)_{PC} = 0.4$ ).<sup>27,29</sup> In steady-state measurements, the absorbance and CD spectra of APC and PC trimers are very different. In particular, the CD spectra of APC<sup>16,17,38</sup> is more promising (it has positive and negative components)<sup>16</sup> for predicting strong excitonic interactions between the two chromophores in comparison to that measured for PC<sup>8</sup>. However, the PC trimer CD spectrum does show two positive components (Figure 4-4) in contrast to the single peak found in the absorbance (Figure 4-2A) indicating that there is either weak coupling among the chromophores<sup>39,40</sup> or a change in the chromophore conformation induced by the trimer formation.<sup>38</sup> To evaluate excitonic coupling in the PC trimer, Sauer and Scheer<sup>41</sup> used the PC trimer crystal structure to calculate the expected band splitting of the chromophores. The largest excitonic band splitting is expected from the most strongly coupled pair, i.e., the pair with the shortest center-to-center distance,  $\alpha 84 \leftrightarrow \beta 84$ . This splitting is calculated to be 112  $\text{cm}^{-1}$  and may not be easily resolved in the absorbance or CD spectra.

The low temperature absorbance measurements in this study provide some insight to the possible  $\alpha 84$ – $\beta 84$  coupling interaction. The PC trimer absorbance and fluorescence spectra contain a number of features that are absent in those of the PC monomer. The PC trimer absorbance band has two predominant band maxima, similar to the PC monomer, that arise as the temperature is lowered and at 88 K these peaks are accompanied by two pronounced shoulders (Figure 4-1B). The biggest differences between the PC monomer and trimer absorbance spectra occur at long wavelengths where the  $\alpha 84$  and  $\beta 84$  bands are strongest. The PC monomer absorbance peak at 630 nm is replaced in the PC trimer spectrum by a sharp shoulder at ~629 nm and a well-resolved

peak at 636 nm (Figure 4-2C) . The separation of this peak and shoulder ( $175\text{ cm}^{-1}$ ) is larger than the calculated value of  $112\text{ cm}^{-1}$ , but is reasonable considering that (1) the appropriate maximal value of the shoulder feature is difficult to estimate by inspection and, (2) the calculation is based on the PC trimer crystal structure at 298 K which probably changes slightly at low temperatures and could produce a larger splitting.

The individual absorption bands, resolved from the PC monomer, which are assigned to  $\alpha 84$  and  $\beta 84$  are very similar in shape and maximal wavelength position. Appendix 4-1 uses these individual bands in linear combinations to satisfactorily describe the PC monomer. Aggregation of PC monomers into the PC trimer unit could unequally perturb each of the chromophore absorption bands and this would be accentuated at low temperatures. The long-wavelength peak/shoulder feature observed in the PC trimer 88 K absorbance could be due to either changes in the individual chromophore absorption bands (from changes in conformation or in the immediate protein environment) or to newly formed electronic coupling interactions between the chromophores. Unfortunately, the magnitude of the influence of environmental changes versus excitonic coupling upon the chromophore spectra is difficult to separate. If the observed peak/shoulder absorbance feature is interpreted as a separation of the  $\alpha 84$  and  $\beta 84$  bands due to changes in their individual protein environments then there would be minimal coupling between the chromophores and one should be able to reproduce the spectrum combining separate components as for the PC monomer. The simulation exercise detailed in Appendix 4-1 shows that a linear combination of the individual chromophore spectra is sufficient to reproduce the 298 K PC trimer absorbance, but at 88 K it is impossible to attain an adequate fit of the spectra using the low temperature absorbance

bands of the individual subunits. From Appendix Figure A6 it is clear that to reproduce the long-wavelength component of the 88 K PC trimer spectra, the amplitude of the second  $\beta$ -subunit peak (i.e.,  $\beta$ 84) must increase (leaving the first peak,  $\beta$ 155, roughly unchanged) while that of the  $\alpha$ -subunit decreases. However, even if we were to alter the  $\beta$ -subunit spectra as suggested, a suitable spectral fit is still not possible because the linewidth of the individual subunit absorbance spectra is too wide to reproduce the sharpness of the observed PC trimer peak/shoulder profile. Clearly, the chromophore-protein environment of the PC monomer and trimer are very different and one cannot be used to model the other. There are theoretical reasons for expecting excitonic components to be narrower than the non-interacting species from which they arrive,<sup>33,35,42</sup> so we used Gaussian functions with narrower linewidths to simulate the long-wavelength peak/shoulder shape in the 88 K PC trimer spectra. The separation of the Gaussian band maxima is even larger ( $275\text{ cm}^{-1}$ ) than the calculated  $112\text{ cm}^{-1}$  splitting, however, the rest of the spectrum is not included in the simulation; if trailing features were added to the short-wavelength side of these Gaussian waves (to appear more like the PC chromophores) their maximal positions would shift closer together. It is a complex situation and an accurate description of the absorbance bands of the PC trimer chromophores obviously lies somewhere between these narrow Gaussian waves and the broader bands of the individual subunits. The simulations showed that slight changes in amplitude, maximum wavelength position or width of the individual components can have a significant impact on the resulting spectral band. However, drastic changes in the individual chromophore absorbance spectra would be necessary to reproduce the experimentally observed 88 K PC trimer spectra. This means that linear combinations of

individual absorption bands were unable to satisfactorily explain the complex 88 K PC trimer absorbance; therefore, excitonic coupling interactions between the chromophores must be considered to impart narrower linewidths.

Like in the absorbance measurements, the low temperature fluorescence spectrum of the PC trimer is also red-shifted in comparison to that of the PC monomer (Figure 4-7C). The fluorescence band of the PC trimer has a narrower FWHM (Table 4-5) at all temperatures, and the long-wavelength tail resolves three small peak features suggested in the PC monomer 88 K fluorescence (Table 4-4, Figure 4-7C). Unlike the PC monomer, the fluorescence emission scans of the PC trimer taken at different excitation wavelengths are nearly identical (Figure 4-8B), reflecting an efficient system of energy transfer to the final emitter. In agreement, excitation scans of the PC trimer taken at different emission wavelengths at both 298 K and 88 K are identical, which is consistent with a single source of emission (Figure 4-9B). Because the PC trimer emission is independent of excitation and its excitation anisotropy is much lower than that of the PC monomer at all wavelengths, we conclude that a significant amount of energy transfer between the chromophores exists in the PC trimer. The complexity of the 88 K PC trimer absorbance, along with the multi-component CD trace and the fast relaxation components reported in the literature, suggests that neighboring  $\alpha 84$ - $\beta 84$  pairs are weakly coupled.

#### *PC/linker complexes*

The two PC/linker complexes in this study show distinct variations from the spectral properties of the PC trimer. The PC/L<sub>R</sub><sup>32.3</sup> complex typically shows minor deviations from the behavior of the PC trimer, while the PC/L<sub>RC</sub><sup>28.5</sup> aggregate is

characteristically different in absorbance, fluorescence, and polarization properties. This difference in behavior is suggestive of specific differences in the chromophore-linker interaction mechanism of each aggregate. There is only one available structure of a biliprotein/linker complex (APC/L<sub>C</sub><sup>7.8</sup>), and it reveals that the linker disrupts the C3-symmetry of the APC trimer by binding to only two of the three available monomers.<sup>43</sup> Linker association induces subtle changes in the APC trimer conformation which serve to bring the β84 chromophores slightly closer together without affecting the α84–β84 distance. The L<sub>C</sub><sup>7.8</sup> linker interacts with each APC monomer differently; in one case there is a specific change of the β84 conformation due to a phenylalanine residue extending from the linker towards the chromophore; in the second case, a loop segment containing multiple electrostatically charged residues enclose the β84 binding area but are not directly perturbing its structure. The spectroscopic effect of L<sub>C</sub><sup>7.8</sup> on the APC trimer absorbance is a small red-shift in the maximal position (2-3 nm) and a significant suppression of the short-wavelength shoulder.<sup>31,44</sup> The excitation anisotropy traces of both APC trimer and APC/L<sub>C</sub><sup>7.8</sup> are identical<sup>44</sup> showing that the linker does not impede the flow of excitation in the trimer.

The distinctive behavior of the two biliprotein/linker complexes in this study is best observed in the absorbance spectrum. Figure 4-2B compares the absorbance spectra of PC/L<sub>R</sub><sup>32.3</sup> and PC/L<sub>RC</sub><sup>28.5</sup> to that of the PC trimer and it is clear that each linker is perturbing the PC trimer chromophores in a different manner. The spectra are normalized to their respective maximum absorbance to demonstrate relative changes in wavelength position, but protein characterization studies done by Yu *et al.*<sup>45</sup> showed that PC/linker complexes, at the same protein concentration as the PC trimer, had an



increased molar extinction coefficient (~30% higher). Besides this increase, the absorbance profile of the PC/L<sub>R</sub><sup>32.3</sup> aggregate does not deviate greatly from that of the PC trimer, while the PC/L<sub>RC</sub><sup>28.5</sup> absorbance displays a different band shape as well as a large red-shift in the maximal absorbance position.

The spectral modulating effect of the L<sub>RC</sub> linker peptide at room temperature has been previously observed in other cyanobacterial biliprotein/linker complexes.<sup>45-49</sup> This change in the PC absorption band shape is expected to facilitate the transfer of energy across the junction of rod and core biliproteins and to guide the flow of energy in one direction: towards the APC core.<sup>50</sup> The PC/L<sub>RC</sub><sup>28.5</sup> absorbance mimics the shape of the APC trimer spectrum which is speculated to be caused by either exciton interactions between the neighboring  $\alpha$ 84 and  $\beta$ 84 chromophores<sup>51</sup> or changes in the chromophore structural conformation upon trimer aggregation<sup>52</sup> (the APC monomer absorbance has a more symmetrical shape). Although the backbone structures of PC and APC are very similar there are some key differences in the chromophore binding pockets that may lead to different spectral properties for these two biliproteins. In the APC  $\alpha$ -subunit chromophore binding region there are two deleted amino acid residues that lead to significant structural changes in the immediate chromophore environment.<sup>53</sup> The structure of the  $\beta$ -chromophore binding pocket is well-conserved between PC and APC except for two critical residues which are isoleucines in PC and are replaced by tyrosines in APC. One of the tyrosines (Y90) was suspected of electronically and sterically interacting with the  $\beta$ 84 chromophore, and this was confirmed by the APC/L<sub>C</sub><sup>7.8</sup> crystal structure. An aromatic residue (F37) extends from the L<sub>C</sub><sup>7.8</sup> linker and inserts itself between the APC (Y90) residue and the coplanar  $\beta$ 84 pyrrole ring D and forces them to

repel in opposite directions. Since the PC  $\beta$ -subunit does not possess an aromatic residue directly analogous to APC(Y90), the  $L_{RC}^{28.5}$  linker may have to interact with the  $\beta 84$  chromophore in a different manner to induce the observed “APC-like” absorbance spectrum. Genetic alterations of this APC (Y90) residue (or of  $L_C^{7.8}$  (F37)) would provide an essential insight in determining the impact of this chromophore conformation change on the APC spectral properties and point towards a mechanism for  $L_{RC}^{28.5}$  interaction with the PC chromophores. It is interesting to note that although the  $L_C^{7.8}$  linker perturbs the conformation of only one APC  $\beta 84$  chromophore, this action may disrupt the interaction between  $\alpha 84$  and  $\beta 84$ , thus indirectly affecting the  $\alpha 84$  spectral properties also. It could enhance the coupling of  $\alpha 84$ – $\beta 84$  by inducing a more favorable geometric factor, and the asymmetric absorbance band shape could be a reflection of this excitonic interaction. The CD spectra of both PC/linker complexes show a prominent decrease in the low energy shoulder band feature (Figure 4-4) that indicates a difference in the relative geometry or distance of the chromophores with each other due to the associated linker peptide. These slight changes in structure would impact coupling interactions between chromophores that contribute to the strength of the CD band. Since the magnitude of the decrease is roughly the same for both  $PC/L_R^{32.3}$  and  $PC/L_{RC}^{28.5}$ , we speculate that it is a general effect from the linker peptide occupying the PC trimer central cavity. Specific interactions between the linker residues and the chromophores are then needed to produce distinct perturbations in the PC trimer spectral properties.

At 88 K, the absorbance spectra of  $PC/L_R^{32.3}$  and  $PC/L_{RC}^{28.5}$  are very different from each other (Figure 4-2D). Since there is a lack of new features in the  $PC/L_R^{32.3}$  spectra, we assume that this linker does not seriously interfere with the established

$\alpha 84$ – $\beta 84$  relationship in the PC trimer. The presence of the linker in the central cavity should interrupt the symmetrical interaction of the three  $\beta 84$  chromophores and provide them with distinct environments. The observed slight red-shift in the 298 K PC/L<sub>R</sub><sup>32.3</sup> absorbance could be due to a linker-induced differentiation among the three  $\beta 84$  chromophores to place one of them at a slightly lower energy than the others and thus guide the flow of energy. The fluorescence spectrum of PC/L<sub>R</sub><sup>32.3</sup> shows noticeable differences from that of the PC trimer (Figures 4-7B and 4-7D), indicating that PC/L<sub>R</sub><sup>32.3</sup> is influencing the spectral properties of the terminal emitter ( $\beta 84$ ). Additionally, the polarized excitation anisotropy measurements show that PC/L<sub>R</sub><sup>32.3</sup> has a slightly higher anisotropy than the PC trimer at all wavelengths (Figure 4-5). The long-wavelength region (> 600 nm), where the anisotropy of all PC complexes rises, is interesting. As in the PC trimer, both PC/linker complexes exhibit a slight rise in anisotropy in this region, but it appears that the influence of L<sub>R</sub><sup>32.3</sup> is more pronounced than that of L<sub>RC</sub><sup>28.5</sup>.

In contrast to PC/L<sub>R</sub><sup>32.3</sup>, the 88 K absorbance spectrum of the PC/L<sub>RC</sub><sup>28.5</sup> complex is very different from that of the PC trimer. There are changes in the maximum positions of individually resolved peaks: the shoulder clearly resolved in the PC trimer at ~629 nm is suppressed, and a new band appears at long wavelength. These changes are even more apparent in the first derivative of the absorption spectra (Figure 4-3), which accentuates the small inflections flanking the absorbance maximum peak (near ~628 nm and ~648 nm). The ~628 nm inflection is probably the remains of the sharp shoulder observed in the PC trimer 88 K absorbance spectra; but the second band at ~648 nm is a new band created by the linker-chromophore interaction, because it does not coincide with any features in the PC trimer absorbance. Clearly, PC trimer association with the

$L_{RC}^{28.5}$  linker alters the immediate environment of  $\beta 84$  (exposed to the central cavity in the PC trimer) and disturbs the established interchromophore relationships. Since the  $L_{RC}^{28.5}$  and  $L_R^{32.3}$  amino acid sequences are generally similar (discussed in Chapter 3), we must assume that the effects of  $L_{RC}^{28.5}$  association are not arising from a general perturbation of the PC trimer structure but rather a specific interaction with the chromophores that is distinct from the  $L_R^{32.3}$  interaction. Because the  $PC/L_{RC}^{28.5}$  88 K absorbance is so different from that of the PC trimer, we propose that this linker interaction disrupts the coupling between the neighboring  $\alpha 84$  and  $\beta 84$  chromophores to create a unidirectional path of energy transfer to the APC core. The polarized excitation anisotropy of  $PC/L_{RC}^{28.5}$  is lower than that of the PC trimer at all wavelengths (Figure 4-5) implying a larger amount of disorder in the emission. If the  $L_{RC}^{28.5}$  linker interrupts the PC trimer C3-symmetry to preferentially interact with one of the  $\beta 84$  chromophores as the single final emitter to APC, then it would increase the number of energy transfer pathways resulting in lower anisotropy. It is unclear whether the  $L_{RC}^{28.5}$  linker is beneficial or detrimental to the  $\alpha 84$ – $\beta 84$  coupling interaction but it is definitely affecting their established relationship in the PC trimer. The maximal absorbance and fluorescence positions for each PC complex were recorded as the temperature was lowered. The trend for  $PC/L_{RC}^{28.5}$  is different from all the others and shows the least variation as the temperature is lowered (inset Figures 4-1D and 4-6D). This suggests that the PC chromophore/ $L_{RC}^{28.5}$  interaction is not only specific but also strong enough to resist the influence of the surrounding solvent environment.

Like the absorbance and excitation anisotropy measurements, the fluorescence measurements of the PC/linker complexes also display distinct behavior. Neither linker introduces dramatic changes in the PC trimer fluorescence band shape at 298 K (Figure 4-7B), but the band maximum is red-shifted by different amounts (Table 4-3). The PC/linker complexes have even narrower bands than the PC trimer (Table 4-4) indicating further quenching of the fluorescence through linker-chromophore interactions that enhance energy transfer pathways. At low temperatures, new features emerge in each PC/linker complex that were absent in the 88 K PC trimer fluorescence (Figure 4-7D) indicating that both linkers impact the emission properties of the terminal emitter to different extents. The PC/L<sub>R</sub><sup>32.3</sup> complex enhances the intensity of the observed trimer peak at ~680 nm and produces a new feature at ~649 nm, while PC/L<sub>RC</sub><sup>28.5</sup> has a suppressed long-wavelength tail and a small novel shoulder at ~641 nm. Excitation scans taken at different emission wavelengths show that, except for  $\lambda_{EM} = 682$  nm, there are no appreciable differences that could be attributed to different sources of emission for each feature observed in the 88 K fluorescence. At  $\lambda_{EM} = 682$  nm both PC/linker complexes have a small excitation shoulder at ~660 nm that is not present at the other emission wavelengths (Figures 4-9C and 4-9D). Although this could be a general effect of linker association, it is more likely due to a purification artifact, because its intensity differed among preparations. Although the protein isolation procedure includes multiple chromatography columns to separate APC and PC complexes, the PC/linker complexes elute in a region close to APC and would suffer from the largest amount of APC impurity. A small APC impurity that is not detectable at room temperature can be exposed when the excitation band becomes narrower at low temperatures.

In summary, the significant differences in the spectral properties of the two PC/linker complexes suggest that chromophore-linker interactions in each case are proceeding by different mechanisms. Because the fluorescence spectra of both PC/linker complexes is different than that of the PC trimer, we expect that the linkers are influencing the terminal emitter ( $\beta 84$ ). Since the absorbance and fluorescence spectra of the PC/LRC<sup>28.5</sup> complex are so remarkably different from that of the PC trimer, we propose that its interaction mechanism is also affecting the  $\alpha 84$ – $\beta 84$  relationship. The CD measurements of the PC/linker complexes support this proposal, because they display a decrease in the magnitude of the low energy band component relative to the PC trimer CD spectrum.

### **Kennard-Stepanov relation between absorption and fluorescence**

This section exploits the measured spectroscopic properties of the PC complexes to gain an understanding of the relaxation behavior of the system. In the early 1900's physicists sought to establish a link between the molecular properties of absorption and luminescence of light. Kirchoff's law postulated that good absorbers are good emitters, but it was not until the dual nature of matter and light was established that ideas began to solidify. E. H. Kennard was probably the first to predict a general relation between the shapes of absorption and emission bands.<sup>54,55</sup> In 1957, B. I. Stepanov revived interest in the relation, and capitalizing on recently published spectral measurements that suggested a mirror image symmetry between absorption and emission bands, he proposed to use his 'universal' relation as a means to produce the absorption spectrum of a molecule from the emission spectrum and vice versa.<sup>56</sup> Based on thermodynamic considerations of

blackbody radiation, and Einstein's coefficients for transitions between two quantum states, the Kennard-Stepanov relation can be expressed as follows:

$$F(\nu) = \ln \left[ \frac{c^2 I(\nu)}{8\pi h \nu^3 \sigma(\nu)} \right] = -\frac{h\nu}{k_B T} + D(T)$$

where  $I(\nu)$  and  $\sigma(\nu)$  are the experimentally observed emission and absorption spectra in a frequency scale ( $\nu$ ). The other factors are:  $h$ , Planck's constant,  $k_B$ , Boltzmann's constant and  $D(T)$ , a quantity that is independent of frequency. Stepanov stipulated two conditions to uphold the universal relation for the molecule under study: (1) complete vibrational thermal equilibrium of the excited state before emission, and (2) negligible "non-exciting" absorption (i.e., minimal transitions between two vibrational state levels). If a fluorescent system is homogeneous, then the vibrational distribution of the ground state is reflected in the excited electronic state resulting in a mirror symmetry between absorption and fluorescence spectra.

Different groups have formulated the Kennard-Stepanov (KS)  $F(\nu)$  relation to obtain radiative lifetime and thermodynamic information from organic molecules,<sup>57-60</sup> semiconductors<sup>61</sup> and chromophore-protein assemblies like rhodopsin,<sup>62</sup> phycobiliproteins<sup>63</sup> and other photosynthetic systems.<sup>64,65</sup> Typically,  $F(\nu)$  is plotted versus frequency,  $\nu$ ; then a straight line is drawn through the function so that the slope of the line ( $-h/k_B T$ ) and subsequently, the temperature  $T$ , are calculated. If the vibrational levels of the excited state relax by rapidly exchanging energy with their environment (before emission occurs) then the calculated KS temperature should reflect the ambient temperature of the experiment. In practice, the KS value is frequently higher than the experimentally ambient temperature. Van Metter and Knox<sup>66</sup> discuss the reasons that have been proposed for this elevated KS temperature: (1) inhomogeneous broadening

caused by a mixture of fluorescing species<sup>67,68</sup> and (2) a “warm fluorescence” situation where the calculated  $T$  is actually the temperature of the excited state during the fluorescence lifetime.<sup>69,70</sup> Differentiation between these two possibilities can be achieved by calculating  $F(\nu)$  and  $T$  using fluorescence spectra obtained at different excitation wavelengths. If the same  $T$  is obtained, then it excludes the “warm fluorescence” possibility, which assumes that higher excitation energies lead to higher effective temperatures.<sup>70</sup>

Recently, Knox and coworkers<sup>71,72</sup> have investigated  $F(\nu)$  in more detail by analyzing its first derivative:

$$T^*(\nu) = \frac{-h/k_B}{dF(\nu)/d\nu}$$

where curvature or non-linearity in  $F(\nu)$  will be more pronounced. Analysis of  $T^*(\nu)$  for a simple dye solution can reveal unusual properties of the excited state or the relaxation behavior of that chromophore, and can sensitively measure impurities.<sup>71</sup> Of further interest is the application of  $T^*(\nu)$  analysis to large molecular assemblies such as photosynthetic antenna systems, where it can be used to detect a pair of poorly coupled states or incomplete thermal relaxation of the excited state.<sup>62,71,73,74</sup> It is in this capacity that it is being used here. A light-harvesting antenna system with efficient energy transfer among the pigments should have a well-equilibrated excited-state manifold. Sawicki and Knox examine the effect of oscillator strength and energy transfer rates on  $T^*(\nu)$  calculations using simulated Gaussian functions.<sup>71</sup> They conclude that both energy transfer and oscillator strength contribute to  $T^*(\nu)$  features in distinct manners: the oscillator strength controls the position of the feature (the peak follows the weaker band)



while the kinetic transfer rate determines its intensity (more pronounced as the transfer rate is slower).

The isolated PC complexes in this study present a relatively small but organized system of chromophores that is ideal for this kind of analysis. If the transfer of energy among the chromophores is not fast enough to allow complete vibrational relaxation of the excited state before emission, then this will be detected in the  $T^*(\nu)$  plot as a broad band. In essence,  $T^*(\nu)$  becomes a probe of thermal relaxation and efficient energy transfer, as a function of protein aggregation in these complexes.

## Results

Because  $F(\nu)$  is obtained from the overlap of the absorbance and fluorescence bands, artifacts can be produced in the areas where one amplitude becomes small. For this reason, we arbitrarily restrain our observations to a spectral region where there is a normalized spectral intensity that is  $>10\%$ . The calculated  $F(\nu)$  for each PC complex at 298 K is presented in Figures 4-10A to 4-10D.  $F(\nu)$  is initially plotted over the entire spectral overlap area with the absorbance and fluorescence bands in the background, and then it is constrained to the  $>10\%$  normalized spectral intensity region. A linear fit of  $F(\nu)$  in this region is made and the KS temperature is resolved from the slope of this fit. Because there is obvious curvature in  $F(\nu)$  at low energy, this region was then excluded to obtain a better linear fit (as judged by  $R^2$  error parameter); the KS temperature obtained in each fit is reported in each figure as well as in Table 4-8A. None of the calculated KS temperatures match the experimentally ambient temperature (298 K), and the largest discrepancy is found in the PC monomer (330 K).

Although the  $F(\nu)$  trace of each PC complex shows some curvature, it is in the  $T^*(\nu)$  plots that inflections are well-resolved.  $T^*(\nu)$  is calculated at each energy position using a slope width of 7 points (3 nm span), and the results are presented in Figures 4-11A to 4-11D. In each case, the  $T^*(\nu)$  values steeply increase towards low frequency. Because the largest degree of curvature was observed in the PC monomer  $F(\nu)$  low-frequency region, it is not surprising that PC monomer also has the highest  $T^*(\nu)$  value (~1050 K). Towards higher frequency,  $T^*(\nu)$  approaches ambient temperature (represented by the straight black line at 298 K), except in the case of  $PC/L_{RC}^{28.5}$  where a broad feature appears with a maximum  $T^*(\nu) = 335$  K at  $15700\text{ cm}^{-1}$  (637 nm).

Calculations of  $F(\nu)$  and  $T^*(\nu)$  are carried out for each PC complex at a series of temperatures, and the resulting  $T^*(\nu)$  plots are shown in Figures 4-12A to 4-12D. At low temperature the  $T^*(\nu)$  rise at low energy is more pronounced and band features are better resolved; for comparison, the 88 K  $T^*(\nu)$  trace for each PC complex is plotted in Figure 4-13. Both the PC monomer and trimer appear to have a peak feature at  $\sim 15420\text{ cm}^{-1}$  ( $\sim 649$  nm) which is not readily apparent from the 298 K  $T^*(\nu)$  plot. Notably, the  $PC/L_{RC}^{28.5}$   $T^*(\nu)$  has a broad band with maximum at  $15576\text{ cm}^{-1}$  (642 nm), while  $PC/L_R^{32.3}$  has a peak at  $15480\text{ cm}^{-1}$  (646 nm) and, unlike the other PC complexes, is suppressed in the  $15420\text{ cm}^{-1}$  region.

## Discussion

Although Kennard provided the underlying formulation for the relation between absorbance and luminescence in 1918, it was not until the 1960's that it was rigidly tested for various pigments in solution at room temperature.<sup>57,59,60</sup> The predicted linear

dependence of  $F(\nu)$  on  $\nu$  was readily observed in the central spectral overlap region, but at the extreme regions on each side of the overlap, where one of the spectral bands was at its minimum, the linearity broke down. Scientists reasoned that these extreme regions could not be measured reliably, because they were subject to contamination from impurities or stray light, and excluded them when fitting the central linear area to gain the KS temperature. However, even such  $F(\nu)$  fits, constrained to the central (i.e., most linear) region, produced KS temperatures which were higher than the ambient temperature. Initially, these elevated KS temperatures were ascribed to slow vibrational relaxation of the excited molecule, but this idea was quickly discarded in lieu of more thermodynamically feasible suggestions such as solute-solvent interactions, inhomogeneous broadening or warm fluorescence (discussed in reference 66).

The transfer of excitation energy among the chromophores of the PC monomer and PC trimer occurs in picoseconds while the excited state lifetime is in the nanosecond range,<sup>6</sup> yet all of the PC complexes in this study produced KS temperatures that are higher than ambient (Table 4-8A). The highest KS temperature was found in the PC monomer (Figure 4-10A) which is not surprising since its emission band is dependent on excitation wavelength (Figure 4-8A) and therefore in direct violation of Stepanov's stipulation for complete thermal equilibrium. The fact that the fluorescence of the PC monomer changes at long wavelength excitation, and that its emission is significantly polarized (Figure 4-5), indicates multiple sources of emission; therefore the equilibration manifold of excited states expected for an antenna system is not present in the PC monomer. KS calculations performed with fluorescence bands obtained at different excitation wavelengths confirm that the high KS temperature obtained from the PC

monomer is due to inefficient energy transfer and not from incomplete or slow vibrational relaxation. Excitation at 660 nm shifts the fluorescence band to longer wavelength (smaller wavenumber) and, because this emission is highly polarized it is presumed to be direct emission from the long-wavelength absorbing chromophores. The resulting effect of  $\lambda_{\text{EX}} = 660$  nm on  $F(\nu)$  is to tilt the slope slightly more clockwise to produce a KS temperature closer to ambient.

The trimeric PC complexes possess KS temperatures that are closer to 298 K, although still elevated (Table 4-8A). Considering the possibility of warm fluorescence discussed by Van Metter and Knox<sup>66</sup> as explanation for the elevated temperatures, additional calculations of the PC trimer  $F(\nu)$  were made using fluorescence spectra gathered at different excitation wavelengths. If the KS temperature represented a Boltzmann-like population of vibrational states established at a higher temperature, then excitation at a different energy should systematically affect the KS temperature. Our calculations produced similar KS temperatures ( $\pm 3$  K) with no correlation to excitation wavelength. A second and more likely explanation for the observed elevated temperature is inhomogeneous broadening. Slight differences in chromophore conformation and protein environment have produced distinct spectral properties for each PC chromophore. It is not unreasonable to expect different levels of broadening for each chromophore arising from different solvent or protein interactions. The effect should be more intense at low experimental temperatures where the solvent re-orientation around the excited molecule is impaired.<sup>75</sup>  $F(\nu)$  calculations made at 88 K show that the discrepancy between the ambient and KS temperatures in each PC complex is larger at 88 K (Table 4-8B) than at 298 K. The effect is least pronounced in the PC trimer, suggesting that

introduction of an associated linker peptide contributes to the inhomogeneous broadening of the chromophore spectra. A previous study by Björn and Björn<sup>63</sup> applies the KS  $F(\nu)$  relation to individual phycobiliproteins and phycobilisomes with the focus of comparing the protein environment to that of frozen or viscous solutions. Although there is no effort to characterize the protein aggregation state, the KS temperatures for all of the phycocyanins (PC-1 and PC-2 from *P. luridum*) in the study are higher than ambient. Our PC trimer KS temperature (315 K) is well within the Björn and Björn average value for pooled PC-1 and PC-2 fractions ( $318 \pm 19$  K) and is closest to the KS value for PC-2 dissolved in water (314 K). Björn and Björn point out that although the calculated KS temperatures for the isolated phycobiliproteins are higher than ambient, they are not unusually high compared to those obtained from organic molecules in solution and dismiss their significance. However, their results do show variation in phycocyanin KS temperatures of  $\pm 19$  K for the different buffer environments used (water, 0.75 M phosphate buffer) suggesting that protein aggregation state is a significant factor. In our PC/linker complexes, the PC/L<sub>R</sub><sup>32.3</sup> complex gives a KS temperature that is similar to that of the PC trimer alone, while the KS temperature for PC/L<sub>RC</sub><sup>28.5</sup> is noticeably higher, suggesting a rise in the system inhomogeneity or incomplete dissipation of the excitation energy through energy transfer.

The  $T^*(\nu)$  analysis introduced by Knox and co-workers<sup>71,72,76</sup> is sensitive to non-linear behavior in  $F(\nu)$ , and reveals a number of interesting properties of the PC complexes. In the low energy region, where the absorbance overlap with the fluorescence is diminishing,  $T^*(\nu)$  rises to high values which are reproducible. Although it varies among the PC complexes, the maximum  $T^*(\nu)$  value reached in this rise can be

as high as 1500 K. The beginning of the rise occurs within an area of significant spectral overlap so it cannot be attributed entirely to the low values at the extremes as an artifact and, because it has been observed in multiple preparations, it is unlikely to be an impurity. Knox *et al.*<sup>76</sup> have reported similar behavior in chlorophyll *a* and conclude that it can be due to a mixture of spectroscopic components. Since this rising  $T^*(\nu)$  feature is highest in the PC monomer, it is perhaps indicative of a spectral species that is not thermally equilibrated until its energy is transferred. As a test, the PC monomer  $T^*(\nu)$  was calculated using the excitation spectrum instead of the absorbance, and the same elevated KS temperature was obtained; however, the steep rise at low energy was conspicuously absent. Unlike the excitation spectra, the absorbance of coupled chromophores should be independent of energy transfer among them, and this rise in  $T^*(\nu)$  may represent a spectral species that is absorbing but not achieving thermal equilibration through energy transfer.

Considering the acknowledged function of the linker peptides to facilitate energy transfer among biliproteins, it is interesting to note that the  $PC/L_R^{32.3}$  complex possesses a  $T^*(\nu)$  plot that at low frequencies rises to a higher value than that of the PC trimer (Figures 4-11B,C)--as if the presence of this linker decreases the thermal equilibration of the long-wavelength absorbing chromophores in the PC trimer. Since the  $T^*(\nu)$  plot of the  $PC/L_{RC}^{28.5}$  complex also rises at short frequencies (although not as high as for  $PC/L_R^{32.3}$ ), we propose that insertion of a linker peptide in the central biliprotein cavity creates a general structural stress on the PC trimer assembly that adds to the system inhomogeneity.

A second interesting phenomena observed in the  $T^*(\nu)$  plots is the appearance of peaks within a region of significant spectral overlap in the PC/linker complexes (Figures 4-11C,D), and that are particularly pronounced at low temperatures (Figure 4-13). Bellacchio and Sauer<sup>77</sup> observe similar features in bacteriochlorophyll *a* in different organic solvents and propose that these are indicators of spectral heterogeneity (i.e., components that are not “spectroscopically identical”) and/or incomplete excited-state relaxation of a fluorescing system. To extend this idea to a photosynthetic antenna system, Sawicki and Knox<sup>71</sup> attempt to reproduce such  $T^*(\nu)$  features simulating a system of coupled chromophores with various kinetic transfer rates and oscillator strengths. The simulations show that peaks in the central  $T^*(\nu)$  region are more pronounced at slow kinetic rates between two chromophores, but since they are still visible at fast rates of transfer, Sawicki and Knox propose that these are an indication of an inhomogeneity in the system despite fast energy transfer. At 298 K, the  $T^*(\nu)$  plot for the PC/L<sub>RC</sub><sup>28.5</sup> complex is the only one to display a peak feature instead of a flat curve in the central overlap region (Figure 4-11D). According to the simulations discussed above, such a band could indicate either (1) a slower rate of energy transfer compared to the PC trimer or PC/L<sub>R</sub><sup>32.3</sup> or (2) a higher degree of spectral inhomogeneity. In the absence of kinetic rate information for the PC/linker complexes it is difficult to conclusively choose between these two alternatives, however the low temperature spectroscopy described in this study indicates that the L<sub>RC</sub><sup>28.5</sup> linker is producing an uneven perturbation among the PC chromophores. In particular, it appears to affect the coupling relationship of at least one neighboring pair of  $\alpha 84$ – $\beta 84$  chromophores, and this asymmetrical interaction may prevent complete vibrational relaxation of one or more fluorophores leading to a high KS

temperature calculated from  $F(\nu)$ . The thermodynamic advantage to the phycobilisome antenna in arranging the protein/linker in this manner is a mystery. The presence of  $L_{RC}^{28.5}$  shifts both the PC trimer absorbance and fluorescence maxima towards red wavelengths. The obvious advantage in energy transfer for this action is to produce a better spectral overlap with the core biliprotein. However, it also causes a greater overlap of the absorbance and fluorescence of  $PC/L_{RC}^{28.5}$  and thus the ratio of  $I(\nu)/\sigma(\nu)$  increases, which ultimately gives a rise in  $T^*(\nu)$ , an indication that the excited state is not reaching full thermal equilibrium. Perhaps the advantages of enhanced energy transfer in one direction outweigh the disadvantage of an incomplete equilibration over the entire system, and the heat produced is easily distributed within the surrounding protein.

$F(\nu)$  and  $T^*(\nu)$  calculations were repeated for spectra taken at a series of low temperatures which impede the mobility of the surrounding environment for each chromophore. As a result of low temperature, the absorbance and fluorescence spectra move towards each other increasing the spectral overlap. The calculated KS temperature values are again higher than the experimental temperatures and the discrepancy grows at lower temperatures (Tables 4-7A,B). The low temperature spectral shifts also cause the  $T^*(\nu)$  rise found at low frequencies to begin at more intermediate frequencies as the experimental temperature decreases (Figures 4-12A to 4-12D) and to reveal features which were not readily apparent at 298 K. The dramatic variations in the 88 K  $T^*(\nu)$  plots of the PC/linker complexes (Figure 4-13) could represent at least two separate populations of excitation in the PC trimer. The first population, located at  $\sim 15480 \text{ cm}^{-1}$  (646 nm), is part of the rising  $T^*(\nu)$  of the PC trimer and is perturbed to a higher value by  $L_R^{32.3}$  association. The second population, located at  $\sim 15576 \text{ cm}^{-1}$  (642 nm), is relaxed



in the PC trimer and PC/L<sub>R</sub><sup>32.3</sup> complexes (low T\*(v) value) but is perturbed to a higher T\*(v) level in PC/L<sub>RC</sub><sup>28.5</sup>. Linker association inherently introduces a break in the symmetry of the PC trimer which can easily produce a mixture of nonidentical spectral components that are manifested as separate bands in the 88 K T\*(v) plot.

### Summary of observations

- The PC monomer fluorescence is composed of multiple components because (1) the fluorescence emission is not independent of excitation wavelength, (2) the anisotropy trace shows two distinct rising features demonstrating individual polarized sources of emission.
- Association of PC trimer with  $L_{RC}^{28.5}$  produces more dramatic changes in the PC trimer spectral properties than association with  $L_R^{32.3}$ .
- Fluorescence excitation anisotropy traces of each PC complex show an increase in intensity in the long-wavelength region (~630-650 nm).  $PC/L_{RC}^{28.5}$  and  $PC/L_R^{32.3}$  have opposing effects on the anisotropy of the PC trimer, and in the long-wavelength region ( $\lambda > 600$  nm) the  $PC/L_R^{32.3}$  anisotropy is enhanced more than that of the PC trimer or  $PC/L_{RC}^{28.5}$ .
- $PC/L_{RC}^{28.5}$  has the narrowest fluorescence emission band and the smallest Stokes shift (at 298 K) of all the complexes. Unlike the other complexes in this study, the maximal position of its emission band does not vary with temperature.
- The Kennard-Stepanov relation between absorbance and emission  $F(\nu)$  is used to probe the relaxation state of each PC complex. All calculated KS temperatures are higher than ambient conditions, but agree with published results. The  $T^*(\nu)$  function is explored at a range of temperatures (298 K to 88 K).

## References

1. MacColl, R., Guard-Friar, D. (1987) Phycobiliproteins, CRC Press, Boca Raton, FL.
2. Emerson, R., Lewis, C.M. (1942) *J. Gen. Physiol.* 25:579-595.
3. Arnold, W., Oppenheimer, J.R. (1950) *J. Gen. Physiol.* 33:423-435.
4. Duysens, L.N.M. (1951) *Nature* 168:548-550.
5. French, C.S., Young, V.K. (1952) *J. Gen. Physiol.* 35:873-890.
6. Debreczeny, M.P., Sauer, K., Zhou, J., Bryant, D.A. (1993) *J. Phys. Chem.* 97, 9852-9862; Debreczeny, M.P. (1994) Ph. D. Thesis, LBL-35672, University of California, Berkeley, CA.
7. Demidov, A.A., Mimuro, M. (1995) *Biophys. J.* 68:1500-1506.
8. Mimuro, M., Füglistaller, P., Rümbeili, R., Zuber, H. (1986) *Biochim. Biophys. Acta* 848:155-166.
9. Sauer, K., Scheer, H., Sauer, P. (1987) *Photochem. Photobiol.* 46:427-440.
10. Juszczak, L., Geacintov, N.E., Zilinskas, B.A., Breton, J. (1988) Photosynthetic Light-Harvesting Systems: Organization and Function, eds. Scheer, H., and Schneider, S., Walter de Gruyter, Berlin, pp 281-292.
11. Schirmer, T., Bode, W., Huber, R. (1987) *J. Mol. Biol.* 196:677-695.
12. Siebzehnriibl, S., Fischer, R., Scheer, H. (1987) *Z. Naturforsch.* 42:258-262.
13. Glazer, A.N., Fang, S., Brown, D.M. (1973) *J. Biol. Chem.* 248:5679-5685.
14. Brown, A.S., Foster, J.A., Voynow, P.V., Franzblau, C., Troxler, R.F. (1975) *Biochem.* 14: 3581-3588.
15. Frackowiak, D., Grabowski, J., Manikowski, H. (1976) *Photosynthetica* 10:204-207.

16. Csatorday, K., MacColl, R., Csizmadia, V., Grabowski, J., Bagyinka, C. (1984) *Biochem.* 23:6466-6470.
17. Canaani, O.D., Gantt, E. (1980) *Biochem.* 19: 2950-2956.
18. Langer, E., Lehner, H., Rüdiger, W., Zickendraht-Wendelstadt, B. (1980) *Z. Naturforsch.* 35c:367-375.
19. MacColl, R., Lam, I., Choi, C.Y., Kim, J. (1994) *J. Biol. Chem.* 269:25465-25469.
20. MacColl, R., Williams, E.C., Eisele, L.E., McNaughton, P. (1994) *Biochem.* 33:6418-6423.
21. MacColl, R., Eisele, L.E. (1996) *Biophys. Chem.* 61:161-167.
22. Edwards, M.R., MacColl, R., Eisele, L.E. (1996) *Biochim. Biophys. Acta* 1276:64-70.
23. MacColl, R., Eisele, L.E., Williams, E.C., Bowser, S.S. (1996) *J. Biol. Chem.* 271:17157-17160.
24. MacColl, R., Malak, H., Gryczynski, I., Eisele, L.E., Mizejewski, G.J., Franklin, E., Sheikh, H., Montellese, D., Hopkins, S., MacColl, L.C. (1998) *Biochem.* 37:417-423.
25. Edington, M.D., Riter, R.E., Beck, W.F. (1997) *J. Phys. Chem. B* 101:4473-4477.
26. Riter, R.E., Edington, M.D., Beck, W.F. (1997) *J. Phys. Chem. B* 101:2366-2371.
27. Gillbro, T., Sharkov, A.V., Kyukov, I.V., Khoroshilov, E.V., Kyukov, P.G., Fischer, R., Scheer, H. (1993) *Biochim. Biophys. Acta* 1140:321-326.
28. Debreczeny, M.P., Sauer, K., Zhou, J., Bryant, D.A. (1995) *J. Phys. Chem.* 99:8412-8419; Debreczeny, M.P. (1994) Ph. D. Thesis, LBL-35672, University of California, Berkeley, CA.

29. Debreczeny, M.P., Sauer, K., Zhou, J., Bryant, D.A. (1995) *J. Phys. Chem.* 99:8420-8431; Debreczeny, M.P. (1994) Ph. D. Thesis, LBL-35672, University of California, Berkeley, CA
30. Xie, X., Du, M., Mets, L., Fleming, G.R. (1992) Time-Resolved Laser Spectroscopy in Biochemistry III, ed. Lakowicz, J.R., SPIE, Bellingham, WA, Vol. 1640, pp 690-706.
31. Sharkov, A.V., Kyukov, I.V., Khoroshilov, E.V., Kryukov, P.G., Fischer, R., Scheer, H., Gillbro, T. (1992) *Chem. Phys. Lett.* 191:633-638.
32. Rahman, T.S., Knox, R.S., Kenkre, V.M. (1979) *Chem. Phys.* 44:197-211.
33. Wynne, K., Hochstrasser, R.M. (1993) *Chem. Phys.* 171:179-188.
34. Knox, R.S., Güllen, D. (1993) *Photochem. Photobiol.* 57:40-43.
35. van Amerongen, H., Struve, W.S. (1995) In Methods in Enzymology, ed. Sauer, K., Academic Press, New York, Vol 246, pp 259-283.
36. Demidov, A.A. (1994) *Biophys. J.* 67:2184-2190.
37. Demidov, A.A., Andrews, D.L. (1995) *Chem. Phys. Lett.* 235:327-333.
38. MacColl, R., Csatorday, K., Berns, D.S., Traeger, E. (1980) *Biochem.* 19:2817-2832.
39. John, W., Fischer, R., Siebzehrübl, S., Scheer, H. (1985) In Antenna and Reaction Centers of Photosynthetic Bacteria, ed. Michel-Beyerle, M.E., Springer, Berlin, pp17-25.
40. Scharnagl, C., Schneider, S. (1991) *J. Photochem. Photobiol. B:Photobiol.* 8:129-157.
41. Sauer, K., Scheer, H. (1988) *Biochim. Biophys. Acta* 936:157-170.
42. van Amerongen, H., Struve, W.S. (1995) *J. Luminesc.* 51:29-38.

43. Reuter, W., Wiegand, G., Huber, R., Than, M.E. (1999) *Proc. Natl. Acad. Sci. USA* 96:1363-1368.
44. Maxson, P. (1988) Ph. D. Thesis, LBL-26163, University of California, Berkeley, CA, 52-76.
45. Yu, M.H., Glazer, A.N. (1982) *J. Biol. Chem.* 257:3429-3433.
46. Gottschalk, L., Fischer, R., Lottspeich, F., Scheer, H. (1991) *Photochem. Photobiol.* 54:283-288.
47. Lundell, D.J., Williams, R.C., Glazer, A.N. (1981) *J. Biol. Chem.* 256:3580-3592.
48. Bryant, D.A., Stirewalt, V.A., Glauser, M., Frank, G., Sidler, W., Zuber, H. (1991) *Gene* 107:91-99.
49. Glazer, A.N. (1988) In Methods in Enzymology, eds. Packer, L. and Glazer, A.N., Academic Press, New York, Vol 167, pp 304-312.
50. Glauser, M., Sidler, W., Zuber, H. (1993) *Photochem. Photobiol.* 57:344-351.
51. MacColl, R., Csatorday, K., Berns, D.S., Traeger, E. (1981) *Arch. Biochem. Biophys.* 208:42-48.
52. Murakami, A., Mimuro, M., Ohki, K., Fujita, Y. (1981) *J. Biochem.* 89:79-86.
53. Brejc, K., Ficner, R., Huber, R., Steinbacher, S. (1995) *J. Mol. Biol.* 249:424-440.
54. Kennard, E.H. (1918) *Phys. Rev.* 11:29-38.
55. Kennard, E.H. (1926) *Phys. Rev.* 28:672-683.
56. Stepanov, B.I. (1957) *Dokl. Akad. Nauk. SSSR* 112:839-842 [English translation: *Soviet Phys. Doklady* 2:81-84].
57. Neporent, B.S., (1958) *Sov. Phys. Dokl.* 3:337-340.
58. Ross, R.T. (1975) *Photochem. Photobiol.* 21:401-406.

59. Strickler, S.J., Berg, R.A. (1962) *J. Chem. Phys.* 37:814-822.
60. Birks, J.B., Dyson, D.J. (1963) *Proc. Royal Soc.* 135-147.
61. Band, Y.B., Heller, D.F. (1988) *Phys. Rev. A* 38:1885-1895.
62. Ross, R.T. (1970) *Photochem. Photobiol.* 12:261-268.
63. Björn, L.O., Björn, G.S. (1986) *Photochem. Photobiol.* 44:535-542.
64. Knox, R.S., Van Metter, R.L. (1979) Chlorophyll organization and energy transfer in Photosynthesis, Ciba Foundation Symposium 61, Excerpta Medica, Amsterdam, pp 177-190.
65. Dau, H. (1996) *Photosynthesis Research* 48:139-45.
66. Van Metter, R.L., Knox, R.S. (1976) *Chem. Phys.* 12:333-340.
67. Mazurenko, Yu.T. (1972) *Opt. and Spectr.* 33:22-26.
68. Mazurenko, Yu.T. (1974) *Opt. and Spectr.* 36:283-286.
69. Ringler, A., Szalay, L. (1974) *Acta Phys. et Chem. Szeged.* 20:19.
70. Kozma, L., Szalay, L., Hevesi, J. (1964) *Acta Phys. et Chem. Szeged.* 10:67-77.
71. Sawicki, D.A., Knox, R.S. (1996) *Phys. Rev. A* 54:4837-4841.
72. Laible, P.D., Knox, R.S., Owens, T.G. (1998) *J. Phys. Chem. B* 102:1641-1648.
73. Ross, R.T. (1967) *J. Chem. Phys.* 46:4590-4593.
74. Dau, H., Sauer, K. (1996) *Biochim. Biophys. Acta* 1273:175-190.
75. Klochkov, V.P., Korotkov, S.M. (1968) *Opt. and Spectr.* 22:189-194.
76. Knox, R.S., Laible, P.D., Sawicki, D.A., Talbot, M.F.J. (1997) *J. Luminesc.* 72-74:580-581.
77. Bellacchio, E., Sauer, K. (1999) *J. Phys. Chem. B* 103:2279-2290.

**Table 4-1.** Summary of maximal absorbance positions for *Synechococcus* sp. PCC 7002 PC complexes at 298 K and 88 K

PC complex	Absorbance maxima (nm)				
	298K	88K			
PC monomer	619	630	604	577sh	
PC trimer	626	636	608	579sh	629sh
PC/L <sub>R</sub> <sup>32.3</sup>	627	636	607	579sh	628sh
PC/L <sub>RC</sub> <sup>28.5</sup>	637	638	605	--	627sh 646sh

“sh” denotes distinct shoulder feature (not an isolated peak maximum)

Experimental Notes: (1) Maximal absorbance 0.1. (2) PC monomer in 75% glycerol/5 mM phosphate pH 7. All other aggregates in 3 M sucrose (3) Baseline scans of the appropriate buffer were subtracted from each sample scan. (4) Observed fluctuation in absorbance maximal position is  $\pm 0.2$  nm.

**Table 4-2.** Summary of maximal CD band positions for *Synechococcus* sp. PCC 7002 PC complexes at 298 K

PC complex	CD maximum (nm)	
PC monomer	605	
PC trimer	593sh	635
PC/L <sub>R</sub> <sup>32.3</sup>	595sh	635
PC/L <sub>RC</sub> <sup>28.5</sup>	596sh	644

“sh” denotes distinct shoulder feature (not an isolated peak maximum)

Experimental Notes: (1) PC monomer in 1 M KSCN/5 mM phosphate pH 7. All other aggregates in 50 mM K phosphate pH 7. (2) CD scans of the appropriate buffer were subtracted from each sample scan. (3) Error in maximum position is  $\pm 2$  nm.



**Table 4-3.** Summary of maximal fluorescence emission positions for *Synechococcus* sp. PCC 7002 PC complexes at 298 K and 88 K

PC complex	Fluorescence maxima (nm)				
	298K	88K			
PC monomer	647	642	(710)		
PC trimer	646	643	(680)	(700)	(716)
PC/L <sub>R</sub> <sup>32.3</sup>	649	643	(649)	(682)	(700) (716)
PC/L <sub>RC</sub> <sup>28.5</sup>	652	652	(714) (729)		

“( ) ” denotes approximate maximum position of broad band

Experimental Notes: (1) Maximal absorbance 0.1. (2) PC monomer in 75% glycerol/5 mM phosphate pH 7. All other aggregates in 3 M sucrose. (3) Observed fluctuation in fluorescence maximal position is  $\pm 0.5$  nm. (4) Fluorescence excitation at 585 nm.

**Table 4-4.** Full width at half maximum (FWHM) of the fluorescence emission band for PC complexes at 298 K and 88 K

PC Complex	FWHM (nm)	
	298K	88K
PC monomer	39.5	23
PC trimer	34.5	19.5
PC/L <sub>R</sub> <sup>32.3</sup>	33.0	19.5
PC/L <sub>RC</sub> <sup>28.5</sup>	29.5	13

**Table 4-5.** Stokes shift of PC complexes at 298 K and 88 K

PC Complex	Stokes shift (nm)	
	298K	88K
PC monomer	28	12
PC trimer	21	7.5
PC/L <sub>R</sub> <sup>32.3</sup>	21	7.5
PC/L <sub>RC</sub> <sup>28.5</sup>	15	14.5

**Table 4-6.** Range of absorbance and fluorescence parameters available for the PC monomer chromophores at 298 K (gathered from references 6-11)

PC chromophore	Absorbance		Fluorescence	
	maximum (nm)	Relative intensity	maximum (nm)	Relative quantum yield
$\alpha$ 84	617-624	1	639-644	1
$\beta$ 84	622-628	0.60-0.69	644-648	0.58-0.83
$\beta$ 155	594-600	0.89-1.13	623-629	1.1-1.33

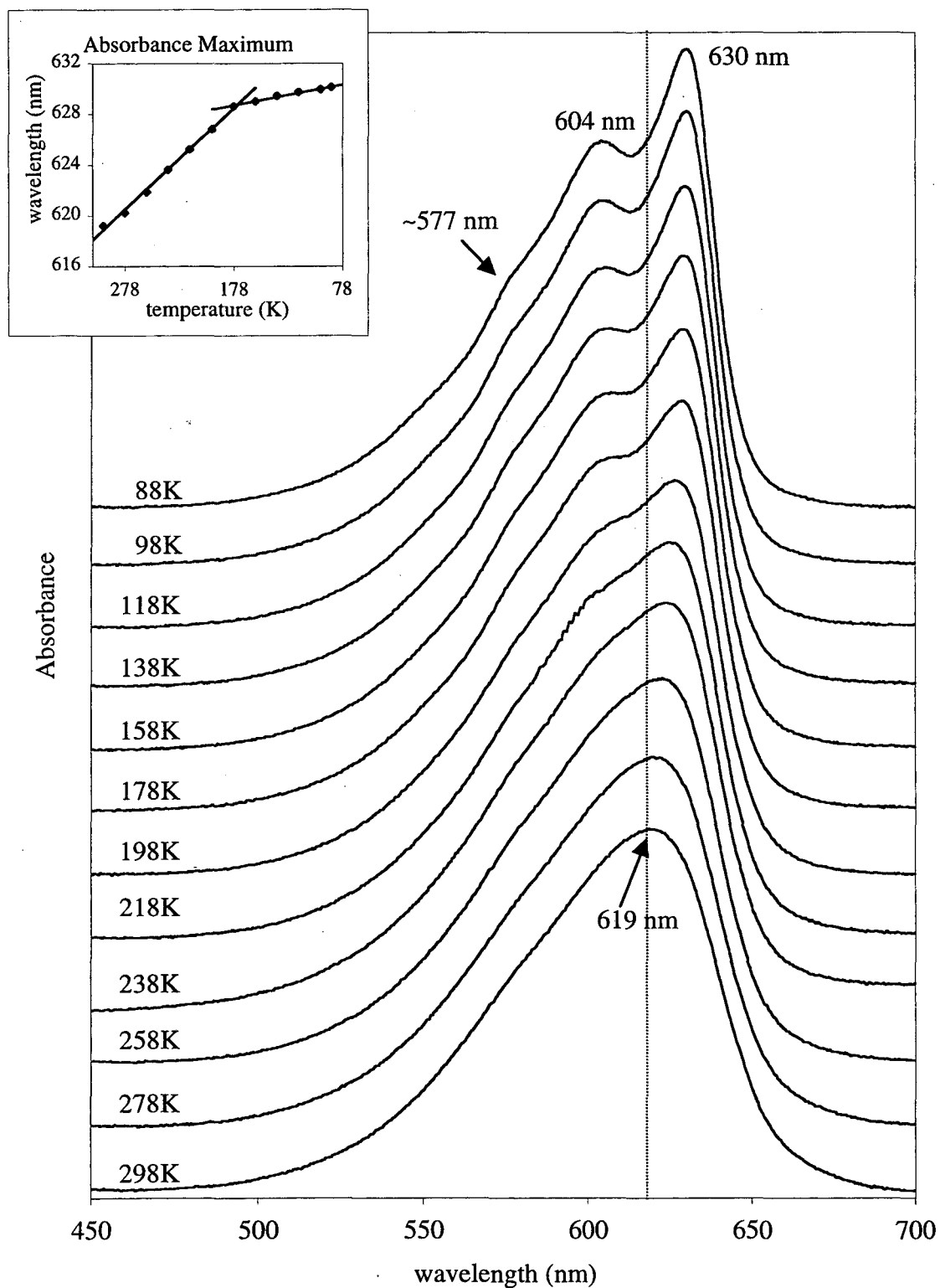
**Table 4-7A.** Kennard-Stepanov temperature values extracted from  $F(\nu)$ . The slope of the best linear fit (least-squares method) was taken over two frequency regions in each PC complex at 298 K (presented in Figures 4-10A to D)

PC Complex	Case 1: $F(\nu)$ fit is restricted to region of >10% normalized spectral intensity		Case 2: $F(\nu)$ fit is restricted to most linear section	
	Temperature (K)	$R^2$ value	Temperature (K)	$R^2$ value
PC monomer	363.21	0.9909	329.53	0.9987
PC trimer	323.39	0.9984	315.17	0.9993
$PC/L_R^{32.3}$	328.70	0.9968	313.66	0.9993
$PC/L_{RC}^{28.5}$	322.88	0.9999	322.23	0.9998

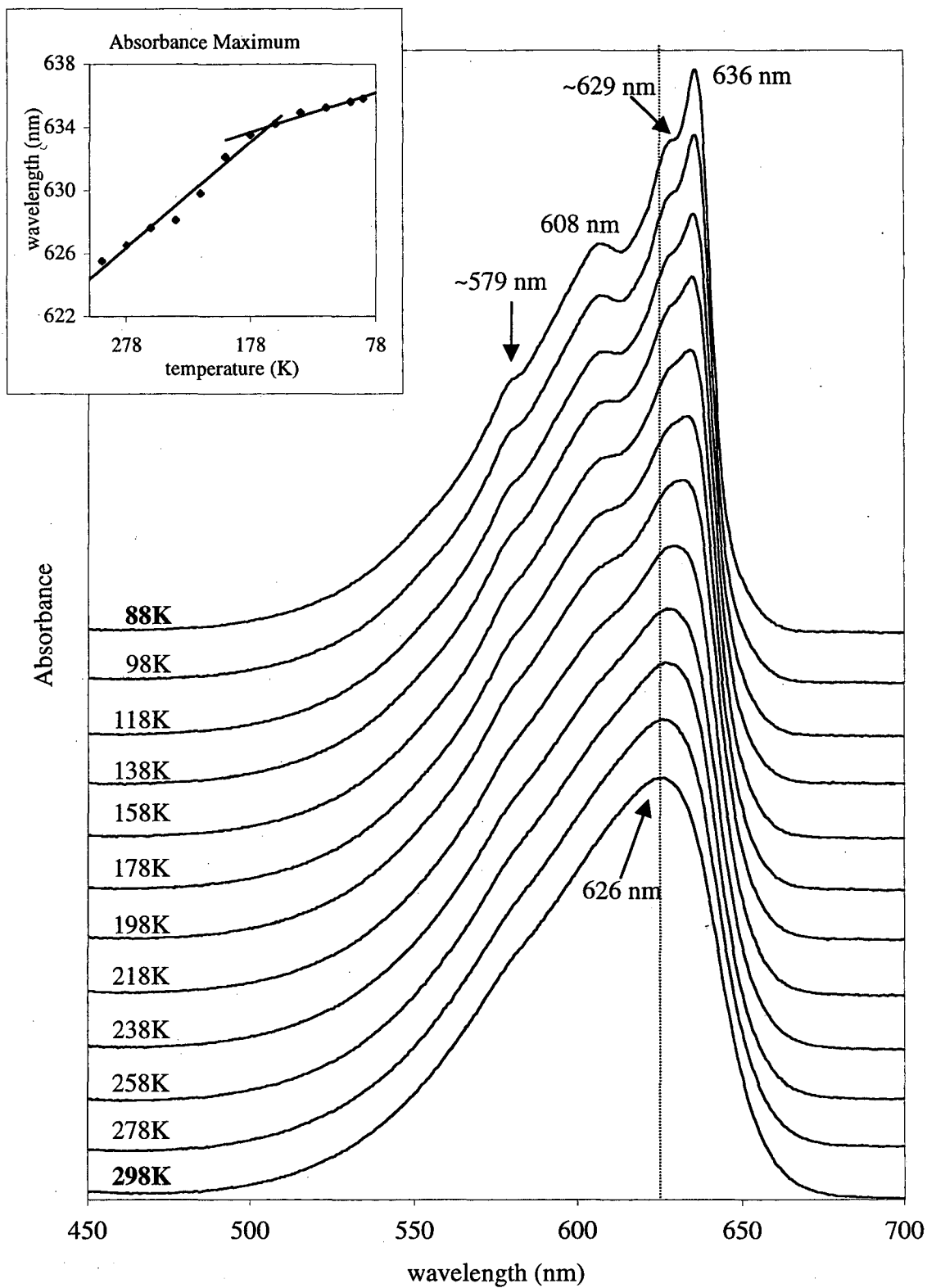
**Table 4-7B.** Kennard-Stepanov temperature values extracted from  $F(\nu)$  calculated at 88K (fit restricted to most linear section, i.e., Case 2 above)

PC Complex	F( $\nu$ ) fit at 88K	
	Temperature (K)	$R^2$ value
PC monomer	174.83	0.9921
PC trimer	106.62	0.9985
$PC/L_R^{32.3}$	110.64	0.9988
$PC/L_{RC}^{28.5}$	166.54	0.9932

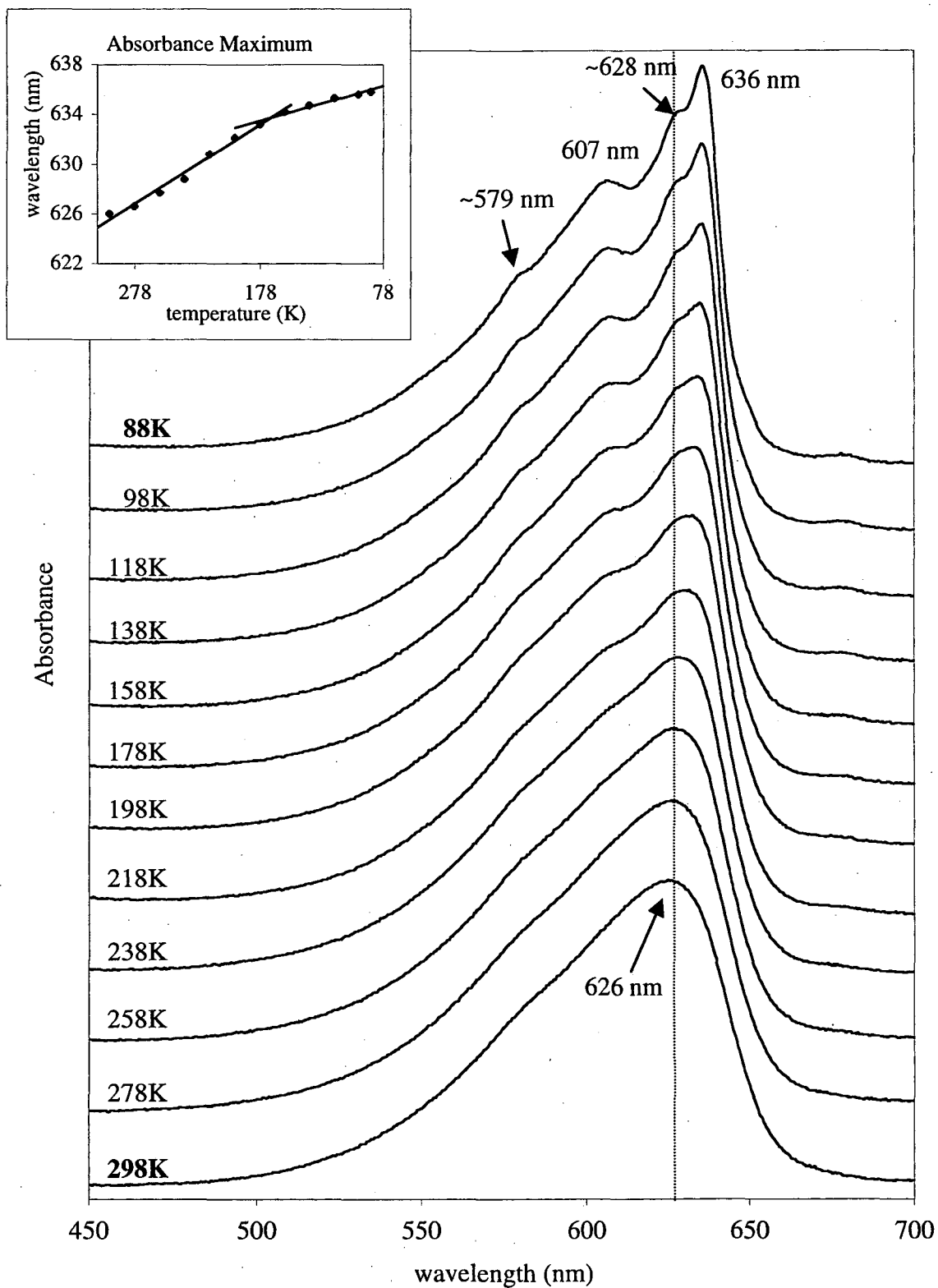
**Figure 4-1A.** Absorbance of PC monomer at different temperatures  
Maximum Absorbance 0.1 in 75% glycerol/5 mM phosphate pH 7.0 buffer



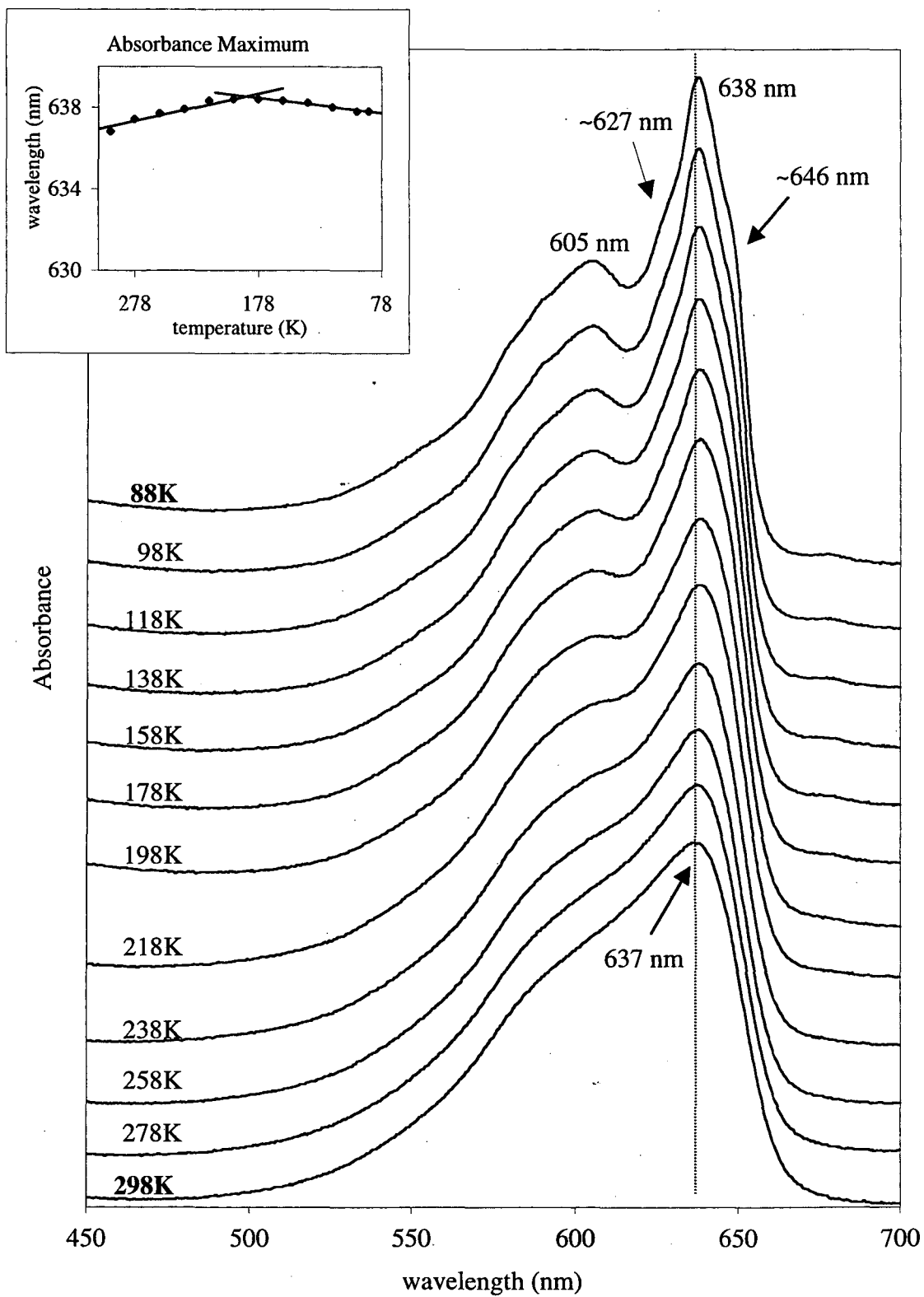
**Figure 4-1B.** Absorbance of PC trimer at different temperatures  
Maximum Absorbance 0.1 in 3 M sucrose



**Figure 4-1C.** Absorbance of  $PC/L_R^{32.3}$  at different temperatures  
Maximum Absorbance 0.1 in 3 M sucrose

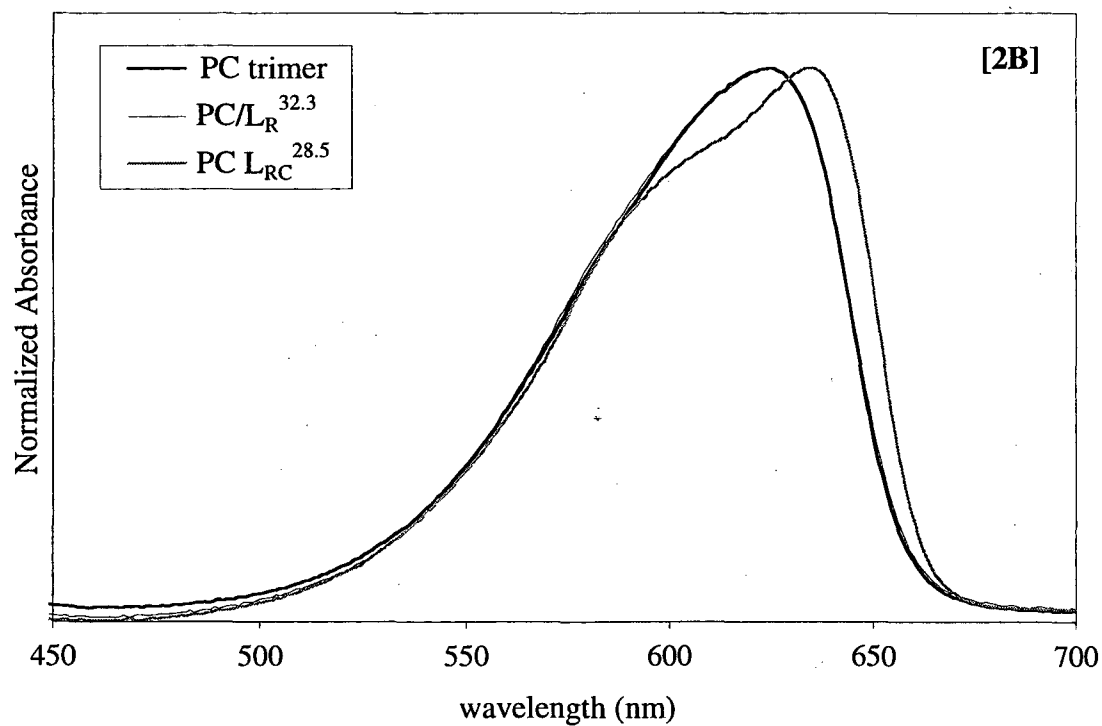
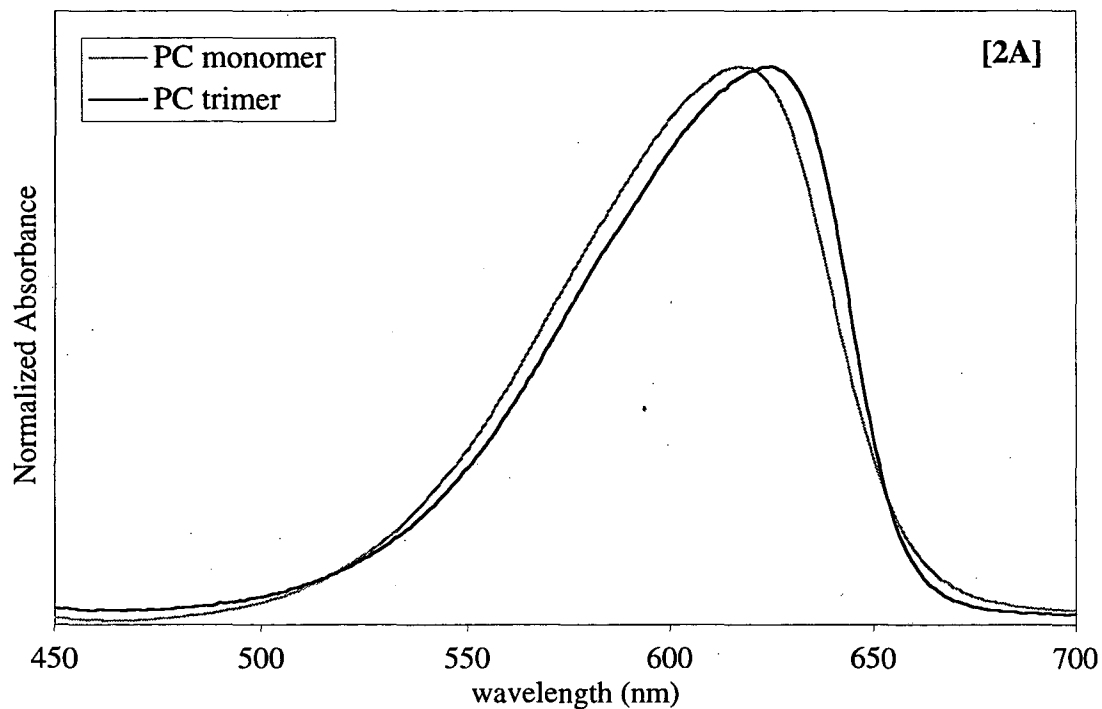


**Figure 4-1D.** Absorbance of  $PC/L_{RC}^{28.5}$  at different temperatures  
Maximum Absorbance 0.1 in 3 M sucrose



**Figure 4-2[A,B]. Absorbance of PC complexes at 298K**

PC monomer is in 1 M KSCN/5 mM phosphate pH 7 while all other PC complexes are in 50 mM phosphate pH 7 buffer. Maximum Absorbance 0.1.



**Figure 4-2[C,D]. Absorbance of PC complexes at 88K.**

PC monomer is in 75% glycerol/5 mM phosphate pH 7 while all other PC complexes are in 3M sucrose. Maximum Absorbance 0.1.

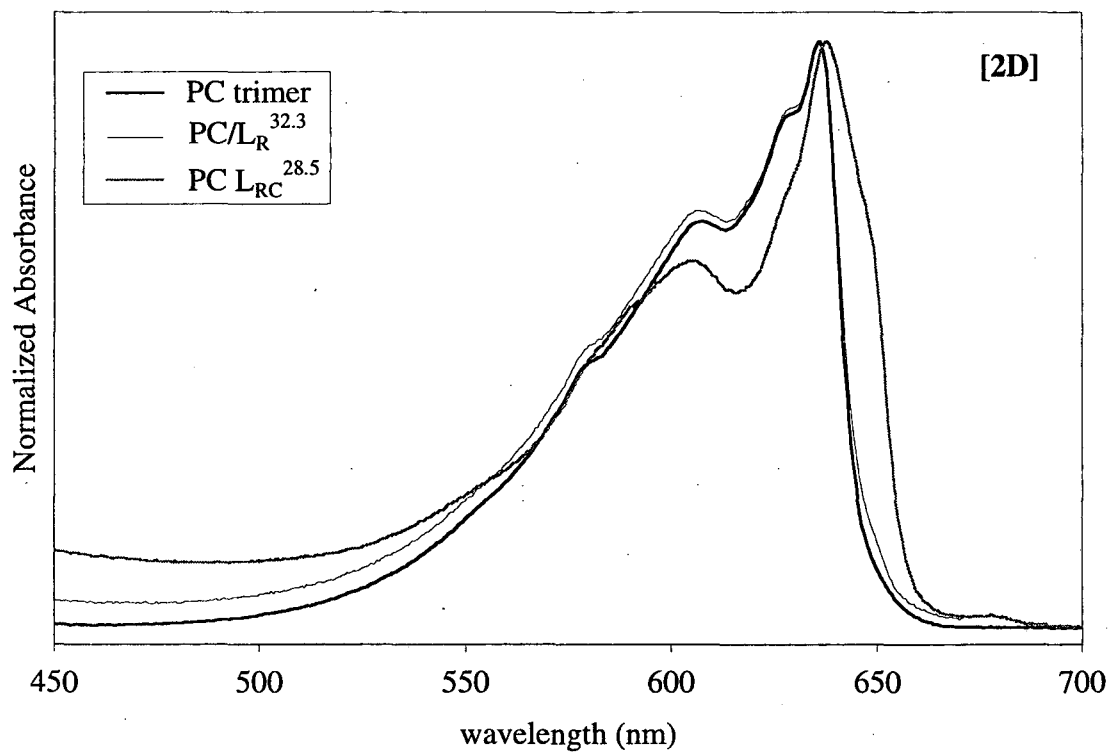
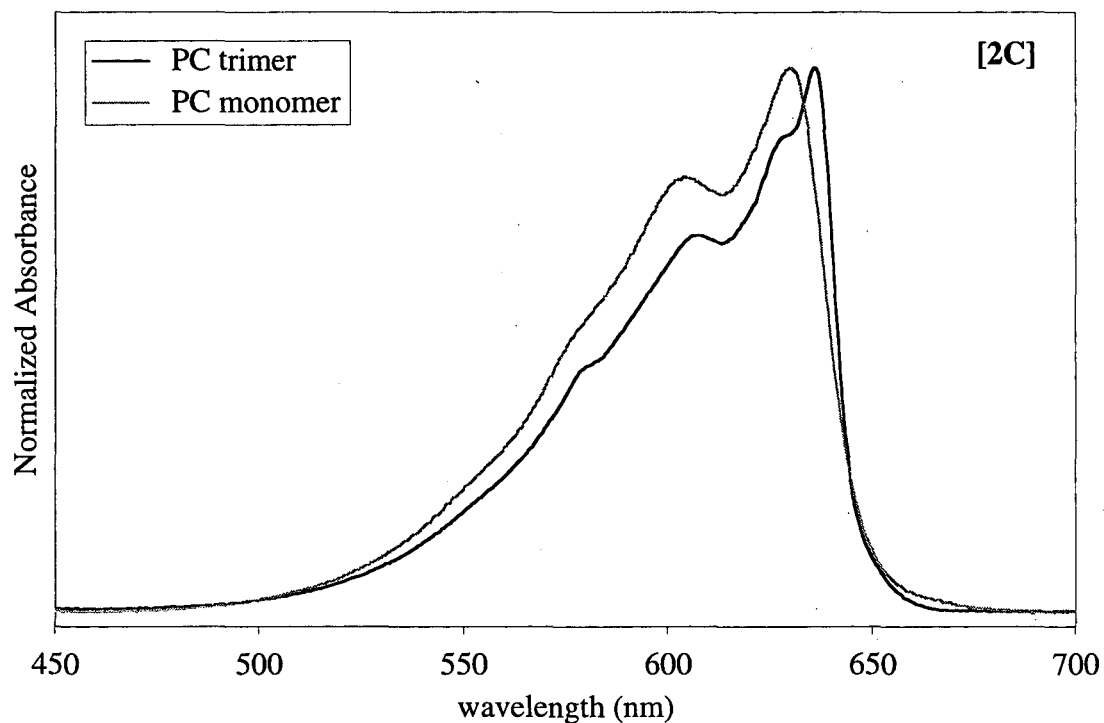




Figure 4-3. First derivative of absorbance of trimeric PC complexes at 88K

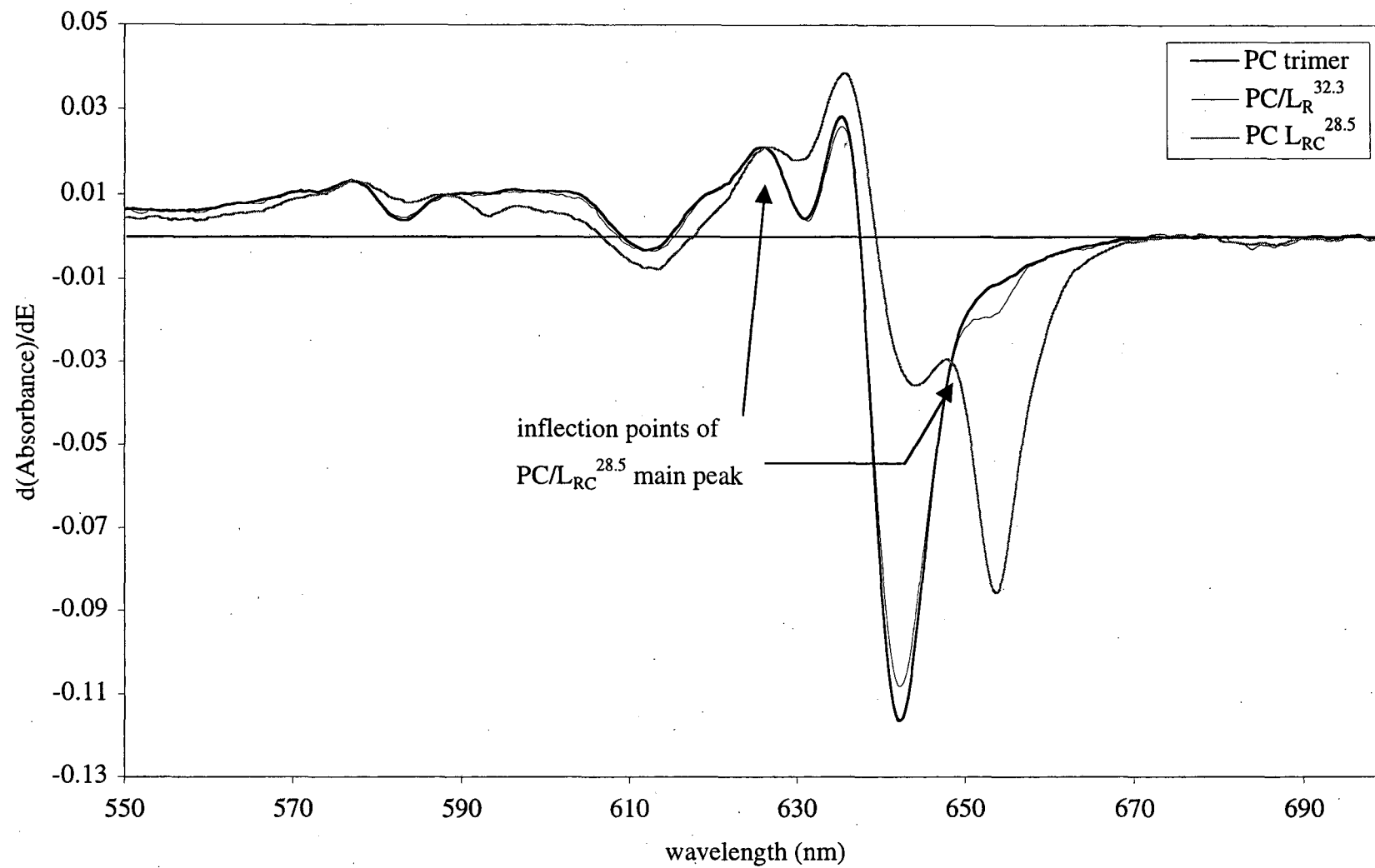
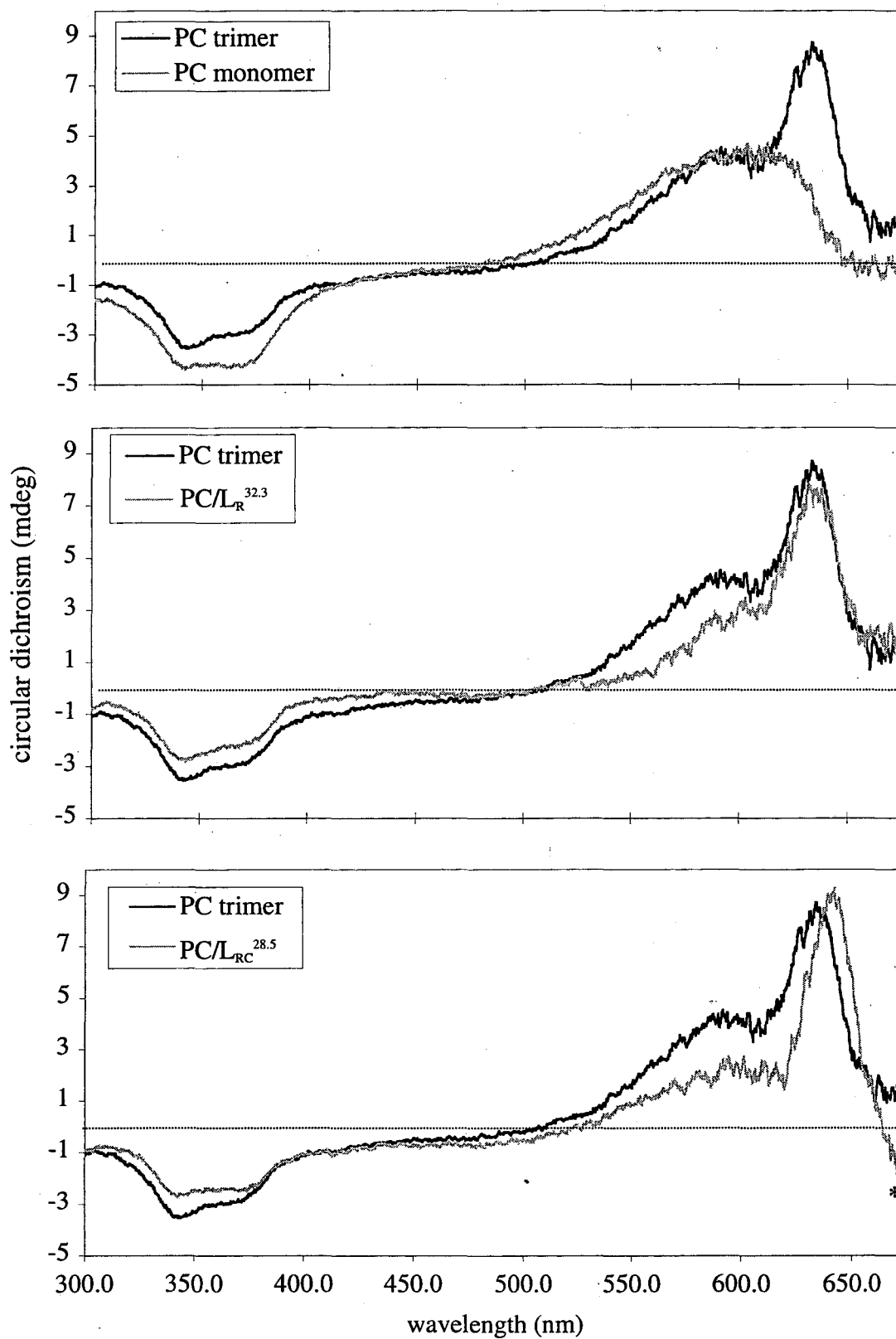
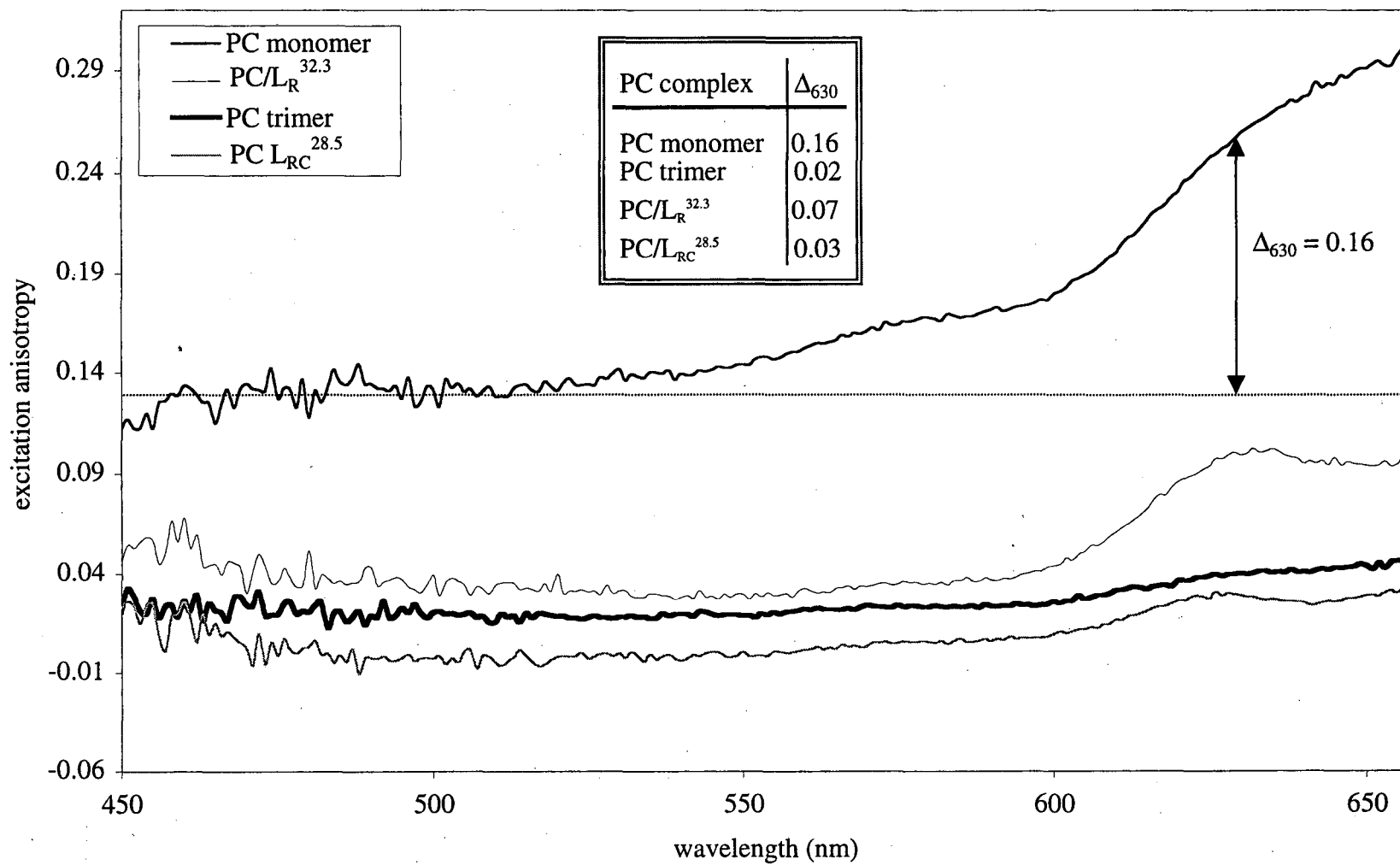


Figure 4-4. CD spectra of PC complexes at 298 K



**Figure 4-5.** Fluorescence excitation anisotropy of PC complexes ( $\lambda_{EM} = 660$  nm) at 298 K



**Figure 4-6A.** Fluorescence emission of PC monomer at different temperatures in 75% glycerol/5mM phosphate buffer (Excitation = 585nm)

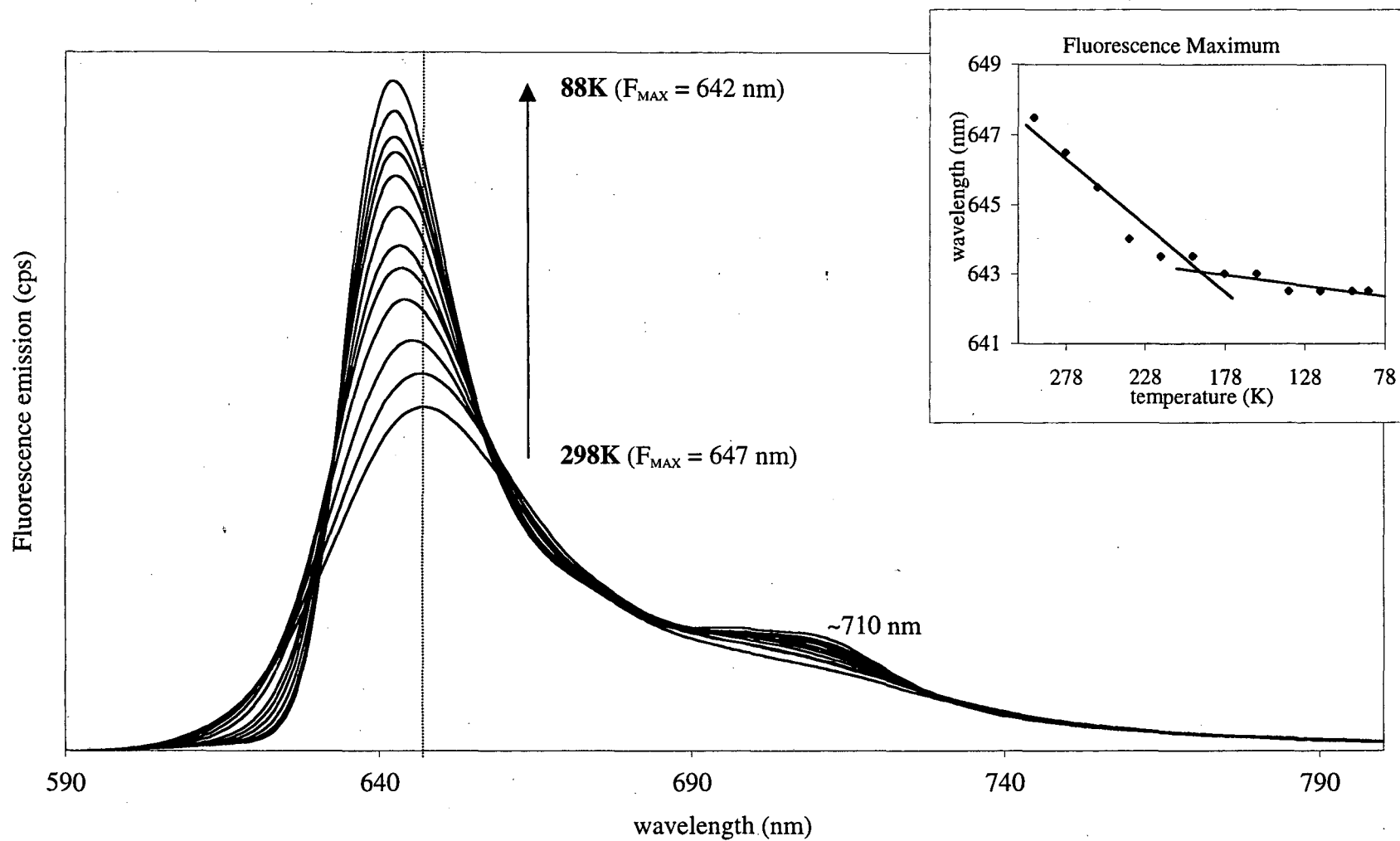


Figure 4-6B. Fluorescence emission of PC trimer at different temperatures in 3 M sucrose (Excitation = 585nm)

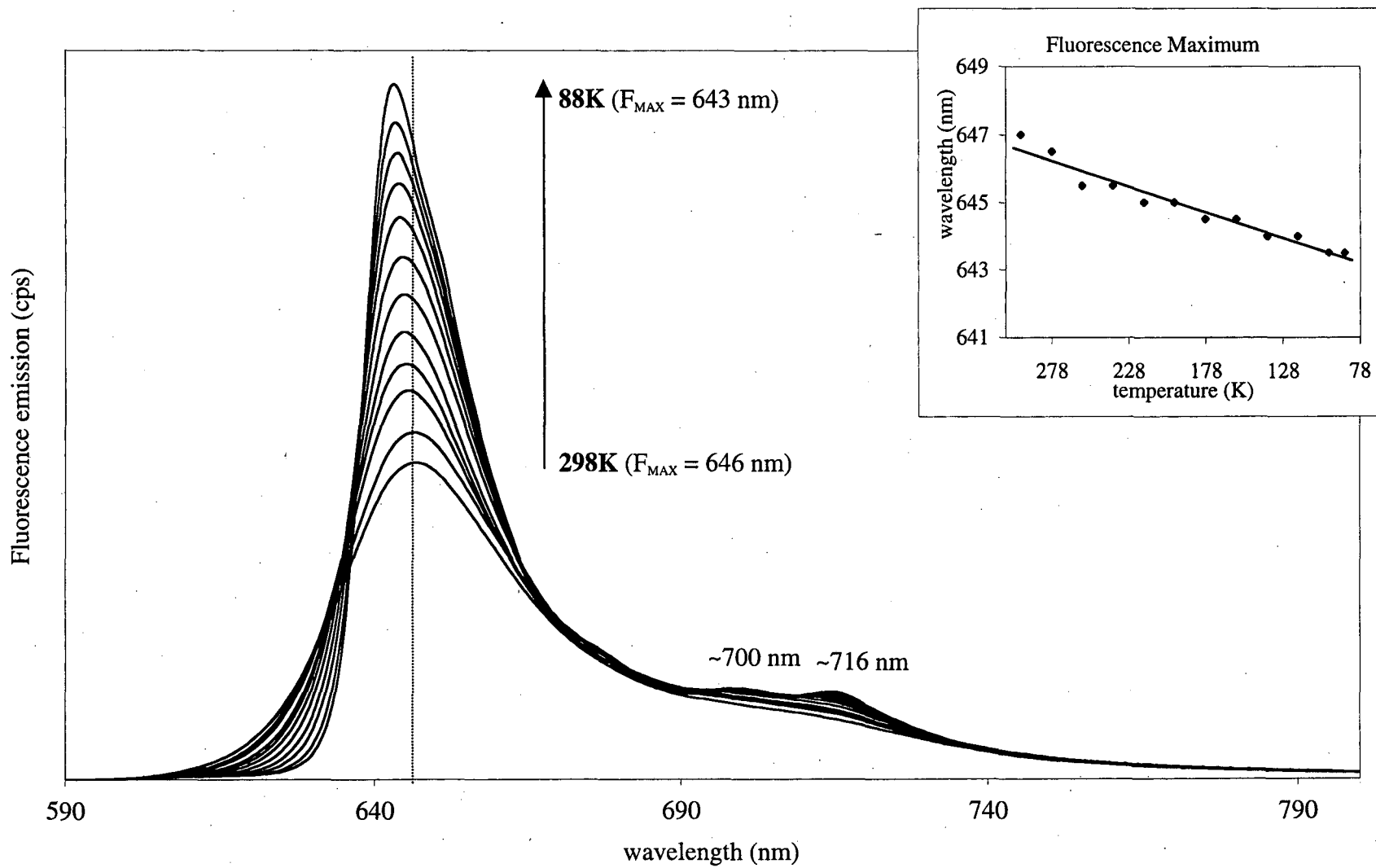


Figure 4-6C. Fluorescence emission of PC/L<sub>R</sub><sup>32.3</sup> at different temperatures in 3 M sucrose (Excitation = 585nm)

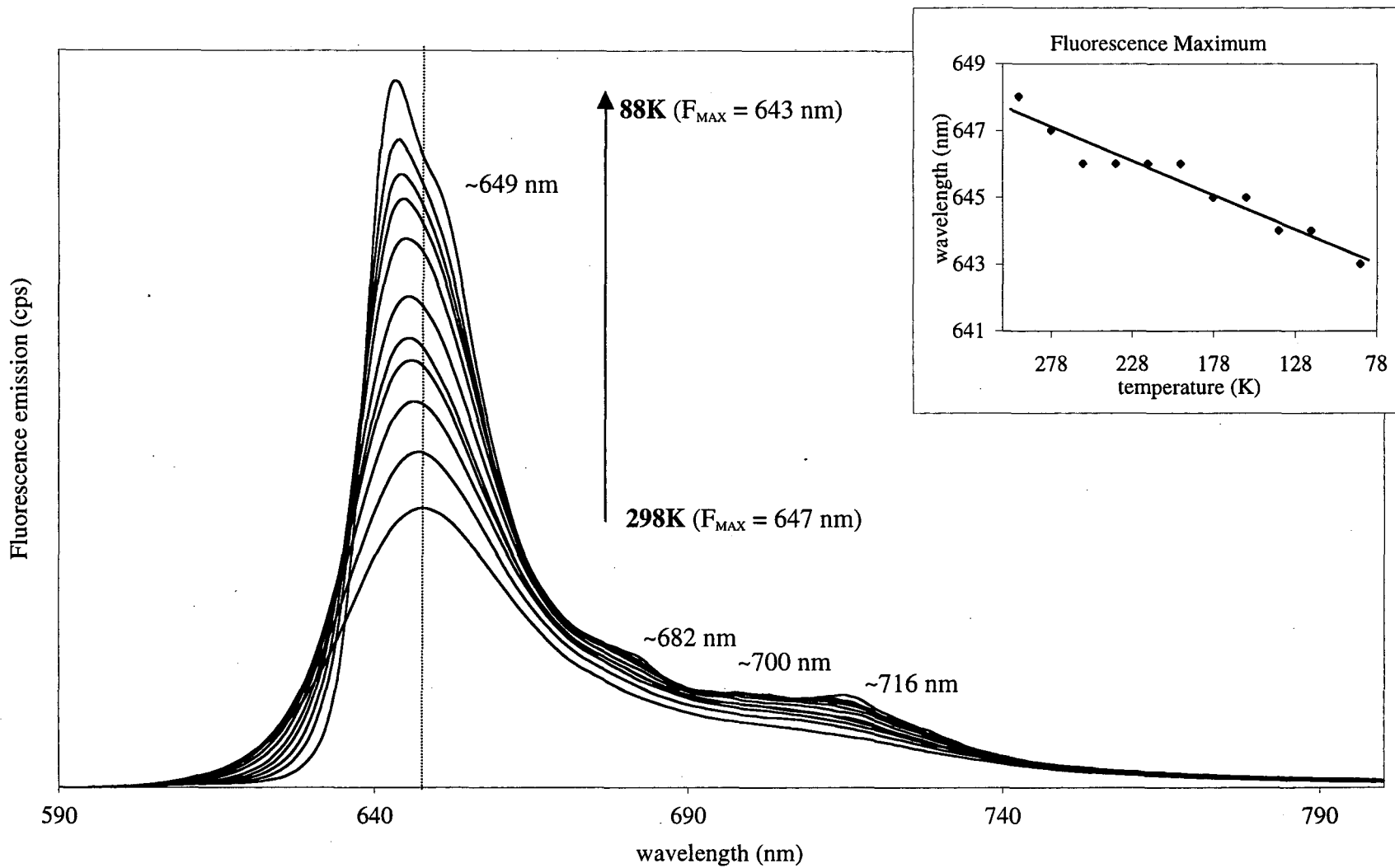


Figure 4-6D. Fluorescence emission of PC/L<sub>RC</sub><sup>28.5</sup> at different temperatures in 3 M sucrose (Excitation = 585nm)

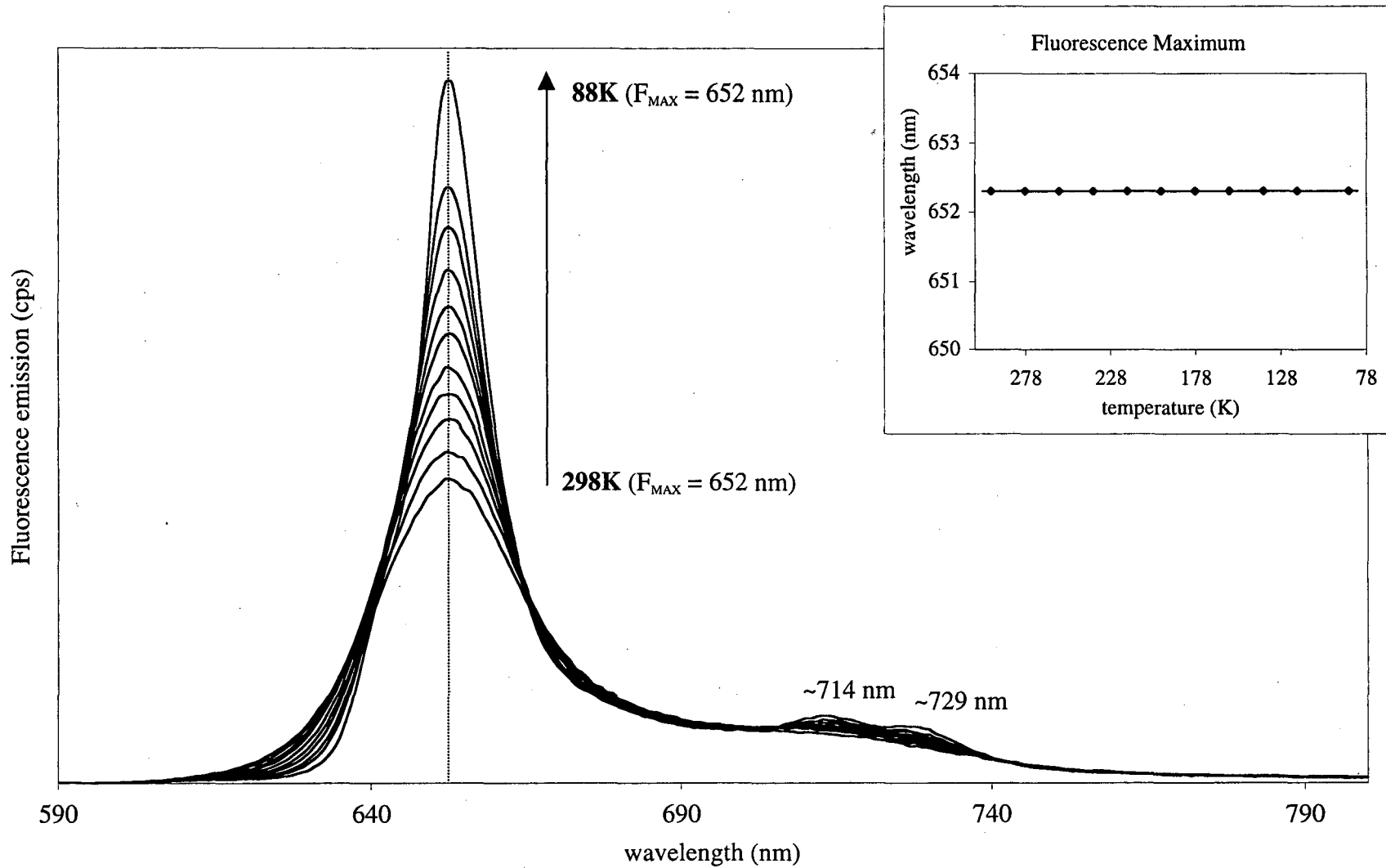
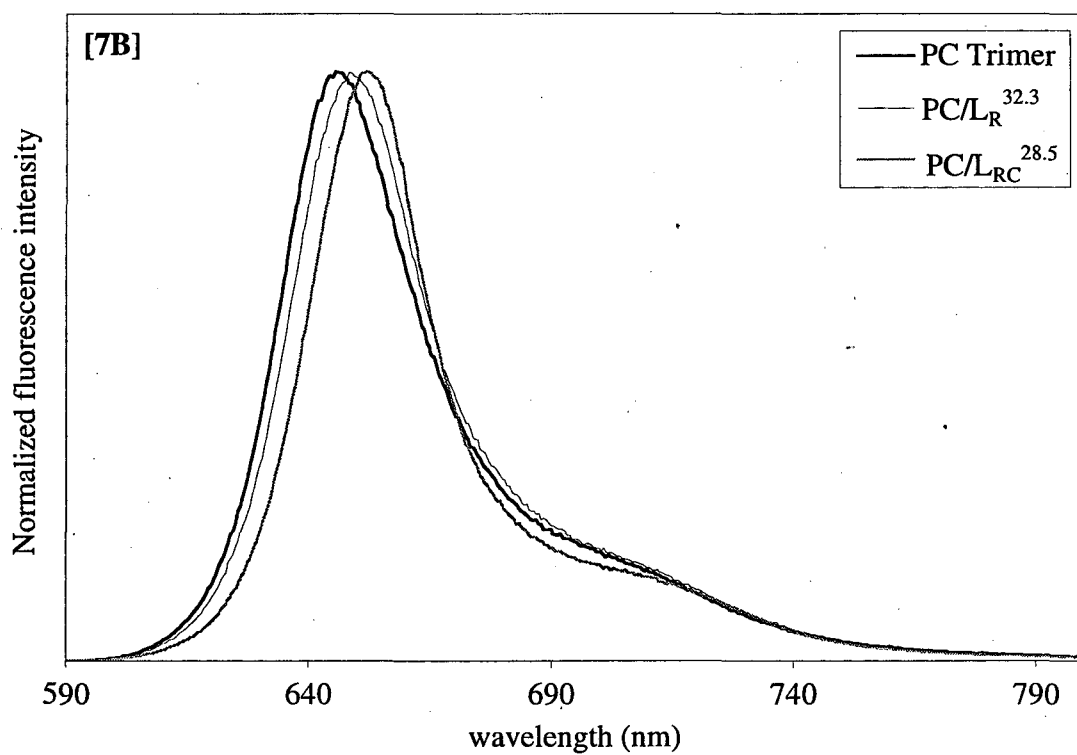
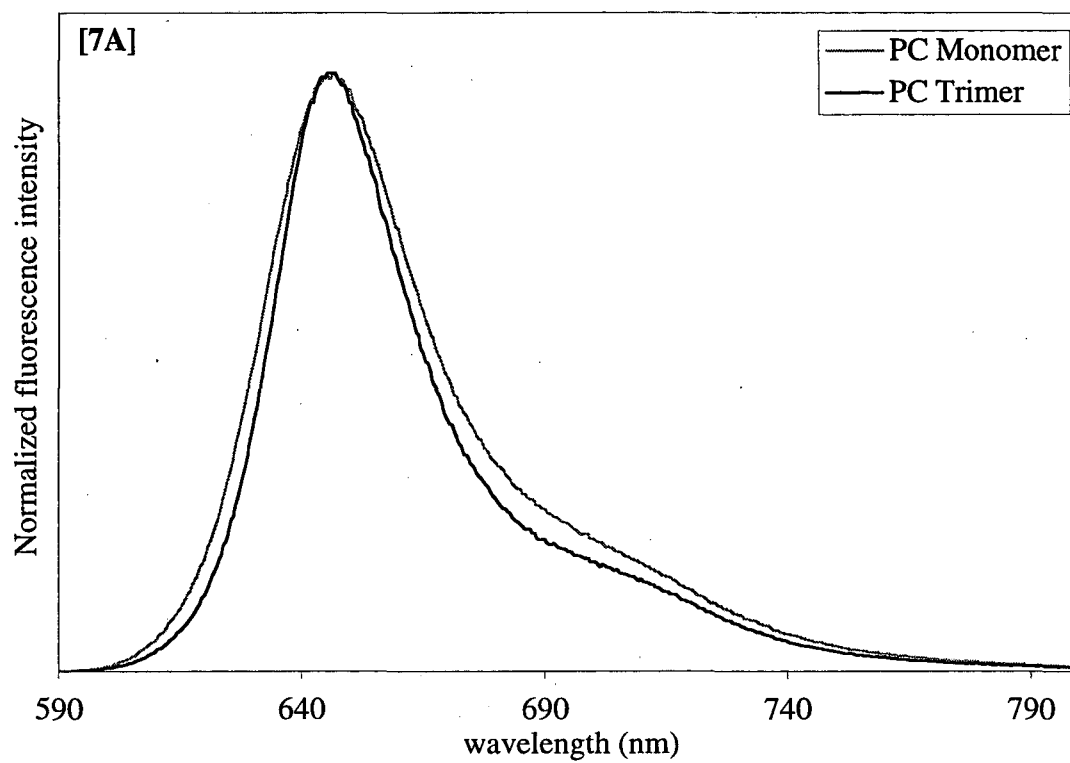
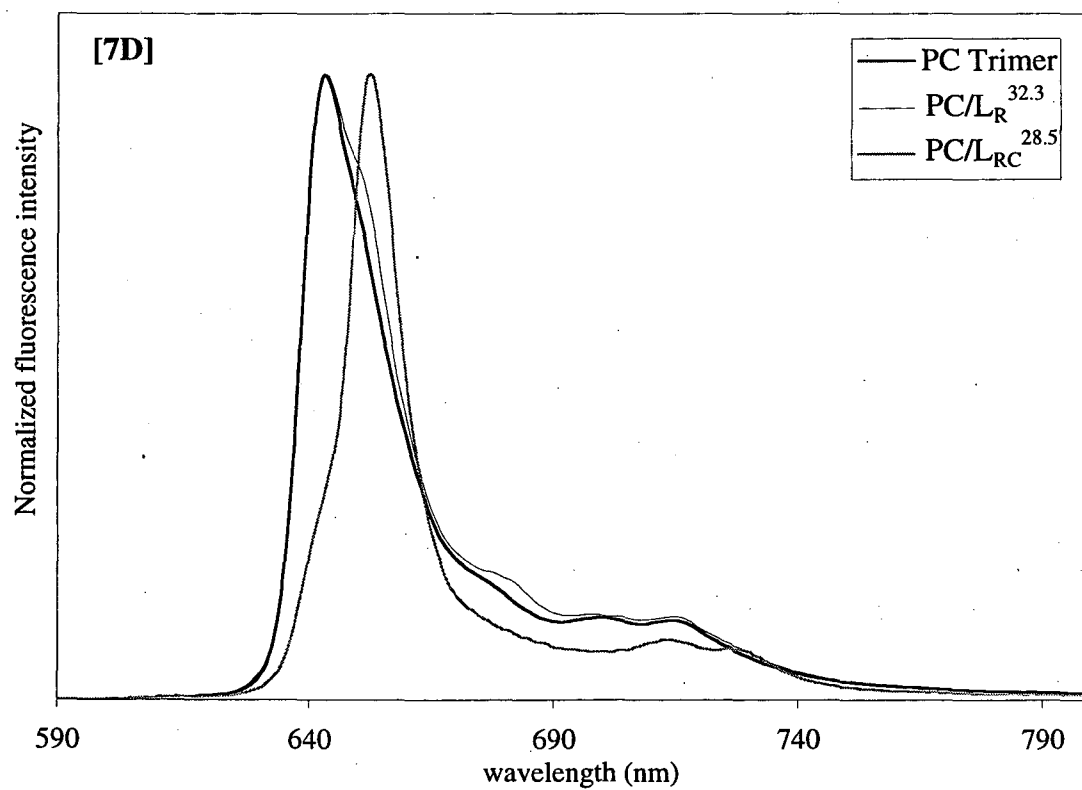
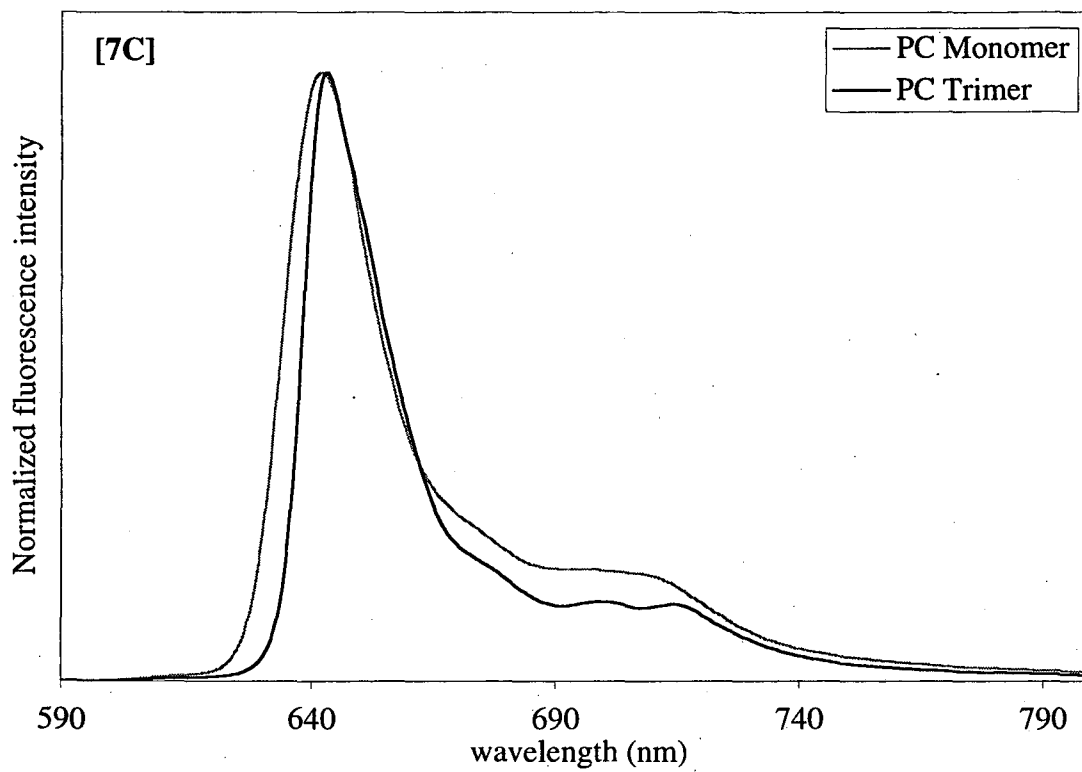


Figure 4-7[A,B]. Fluorescence emission of PC complexes at 298K  
(Excitation = 585nm)



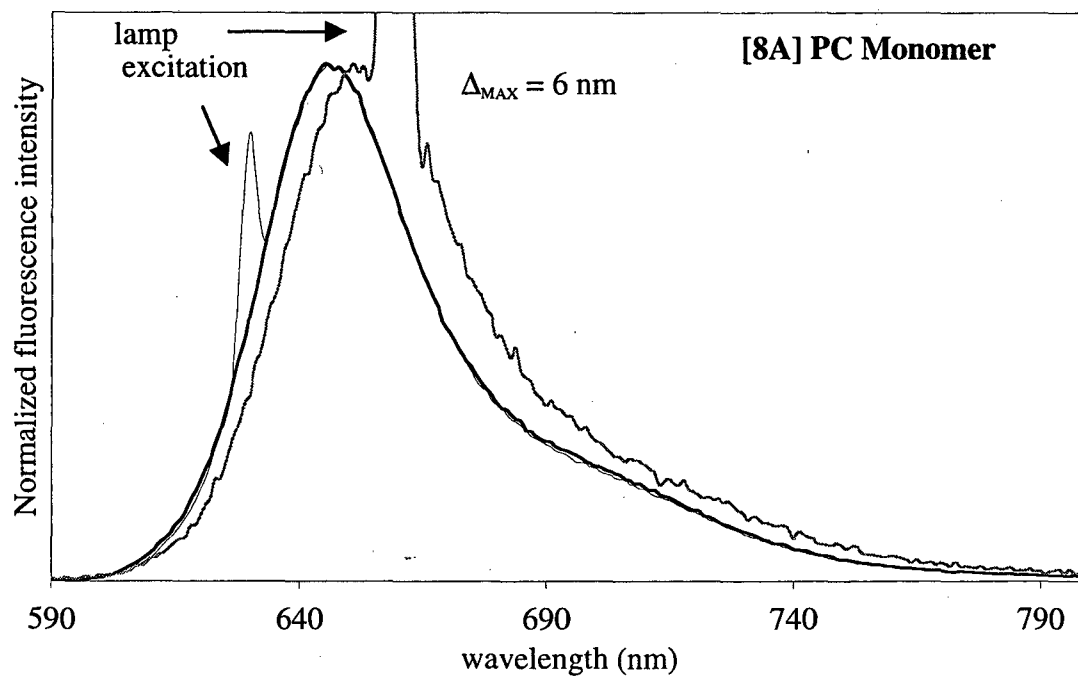


**Figure 4-7[C,D].** Fluorescence emission of PC complexes at **88K**  
(Excitation =585nm)

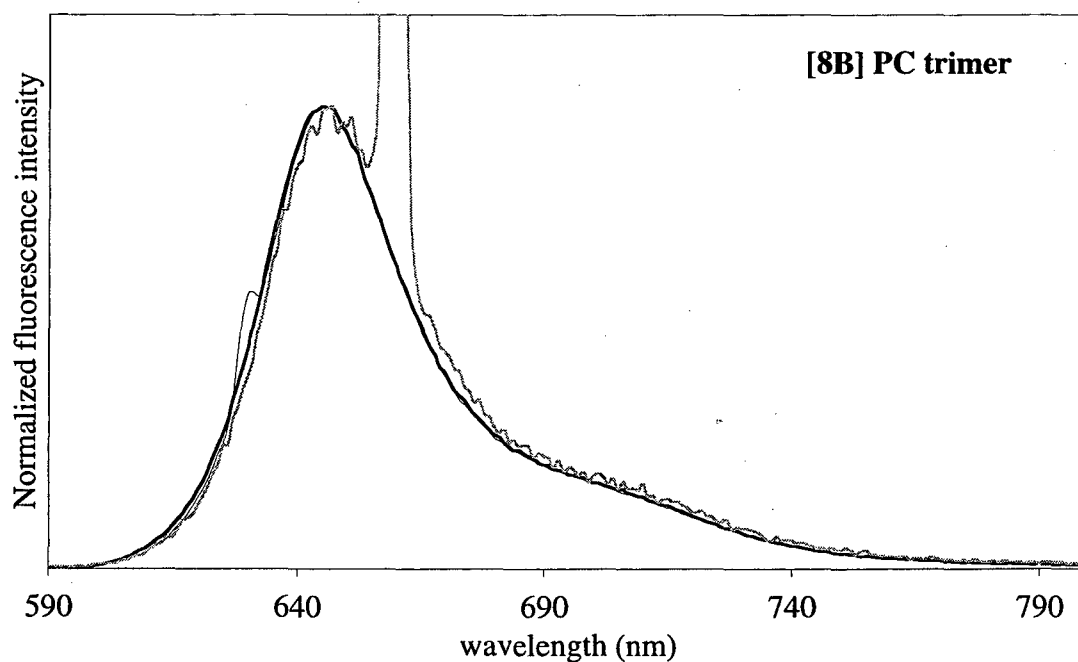


**Figure 4-8[A,B].** Fluorescence emission scans of PC monomer and PC trimer at different excitation wavelengths at **298K**

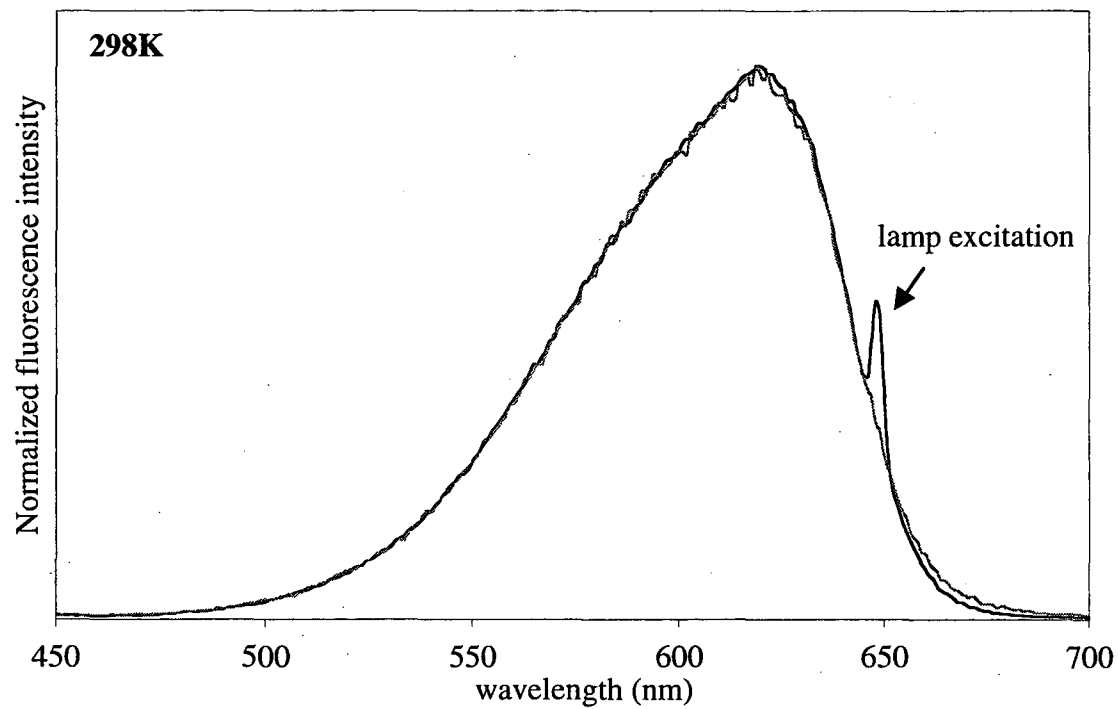
(Note: PC monomer is in 1M KSCN/5mM K phosphate pH 7 buffer while PC trimer in 50 mM K phosphate pH 7 buffer. Maximum absorbance 0.1)



— Excitation=585nm    — Excitation=630nm    — Excitation=660nm

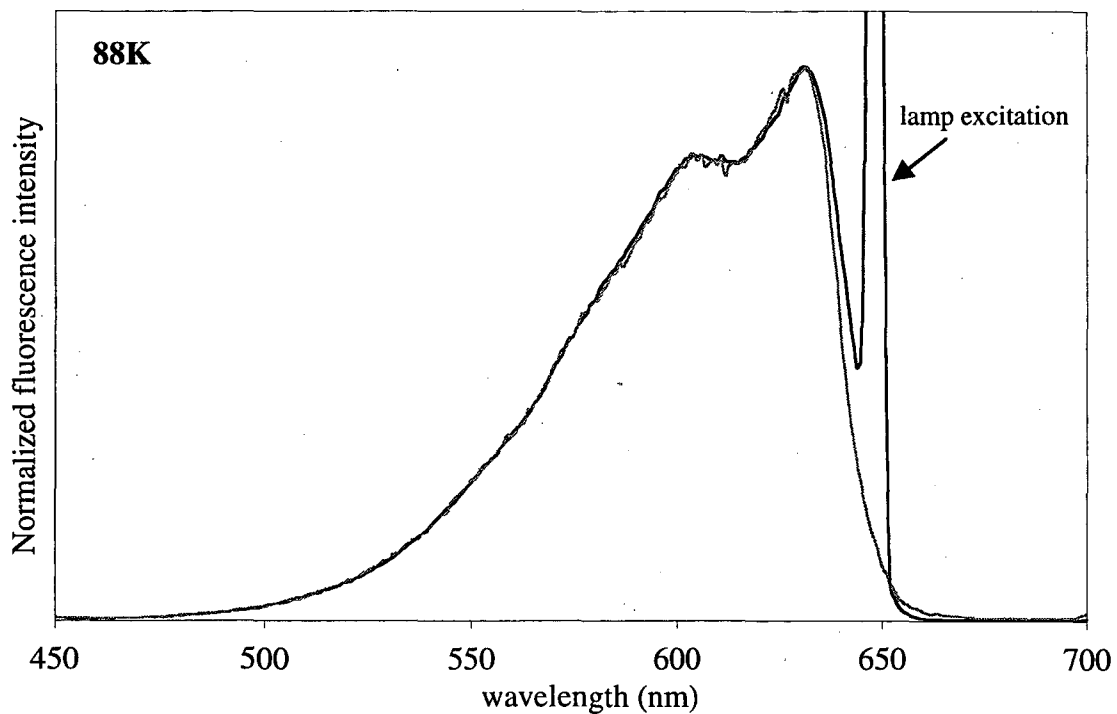


**Figure 4-9A.** Excitation scan of PC monomer at 298K and 88K at different emission wavelengths

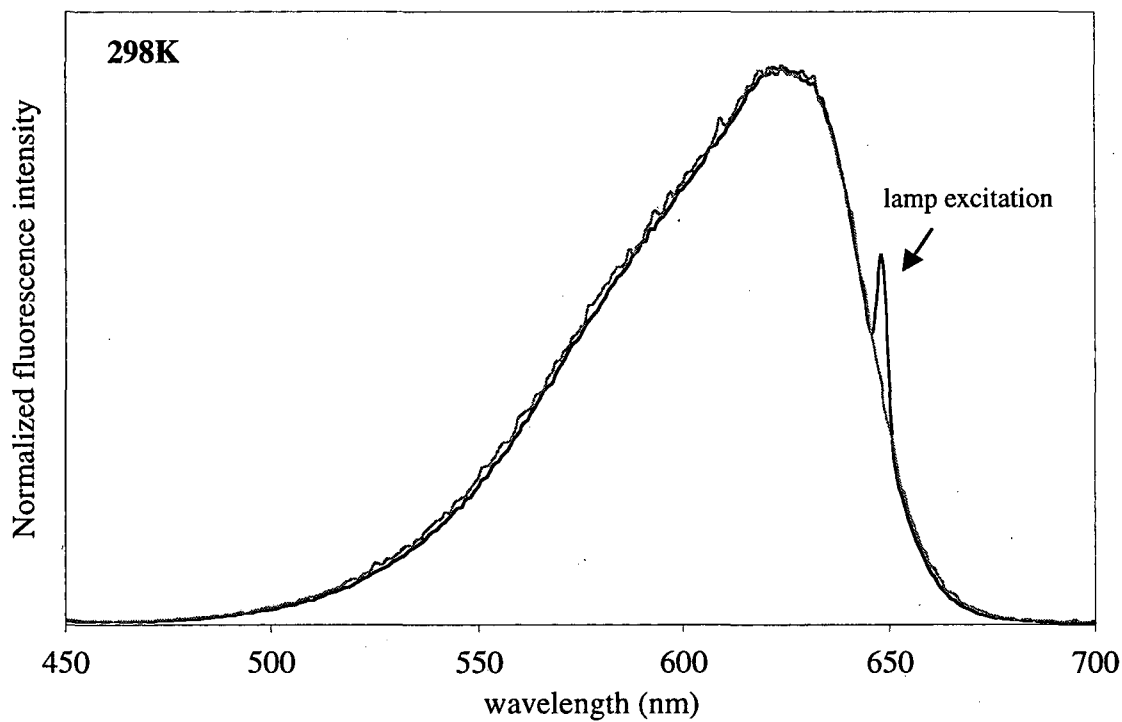


— Emission=648nm

— Emission=710nm



**Figure 4-9B.** Excitation scan of PC trimer at 298K and 88K at different emission wavelengths



— Emission = 648nm

— Emission = 710nm

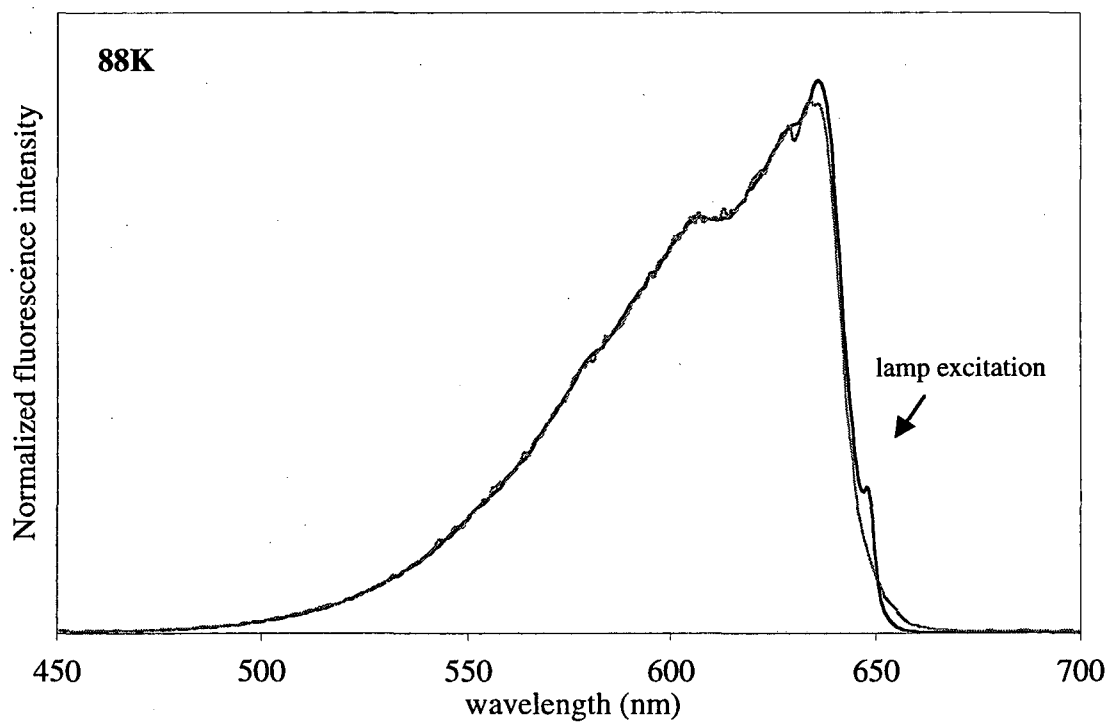
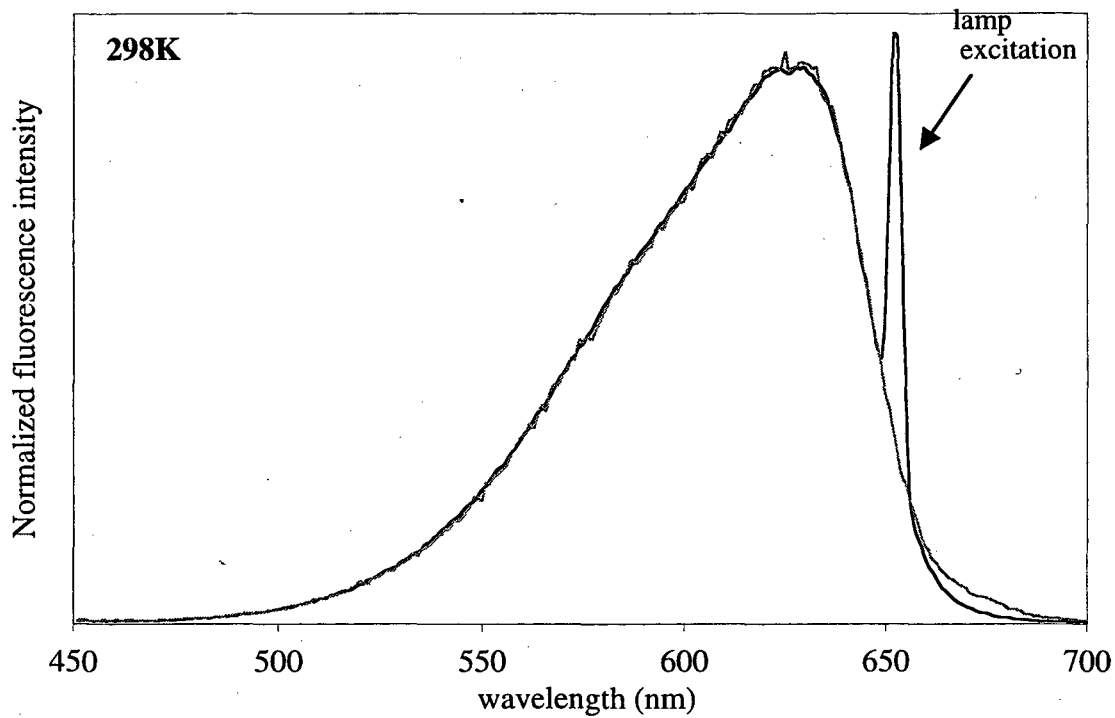


Figure 4-9C. Excitation scan of PC/L<sub>R</sub><sup>32.3</sup> at 298K and 88K at different emission wavelengths



— Emission=652nm    - - - Emission=682nm    - - - - - Emission=714nm

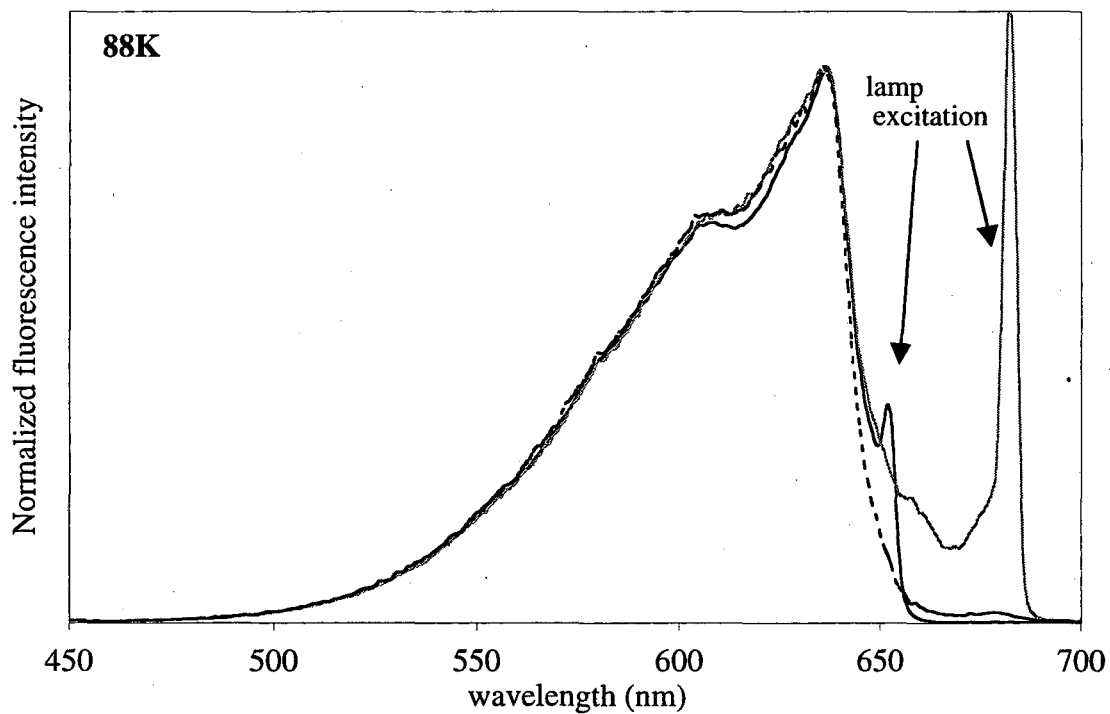
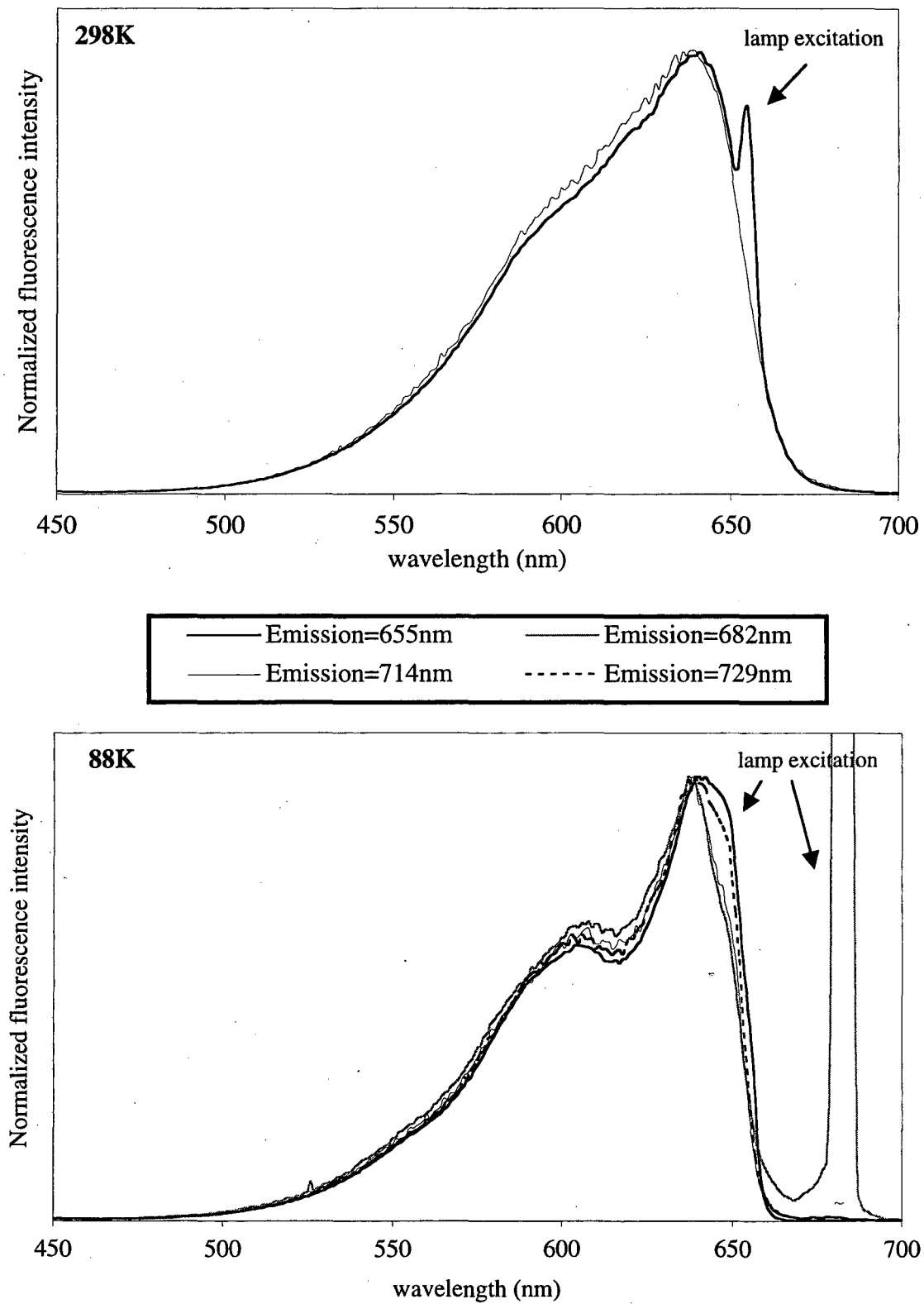
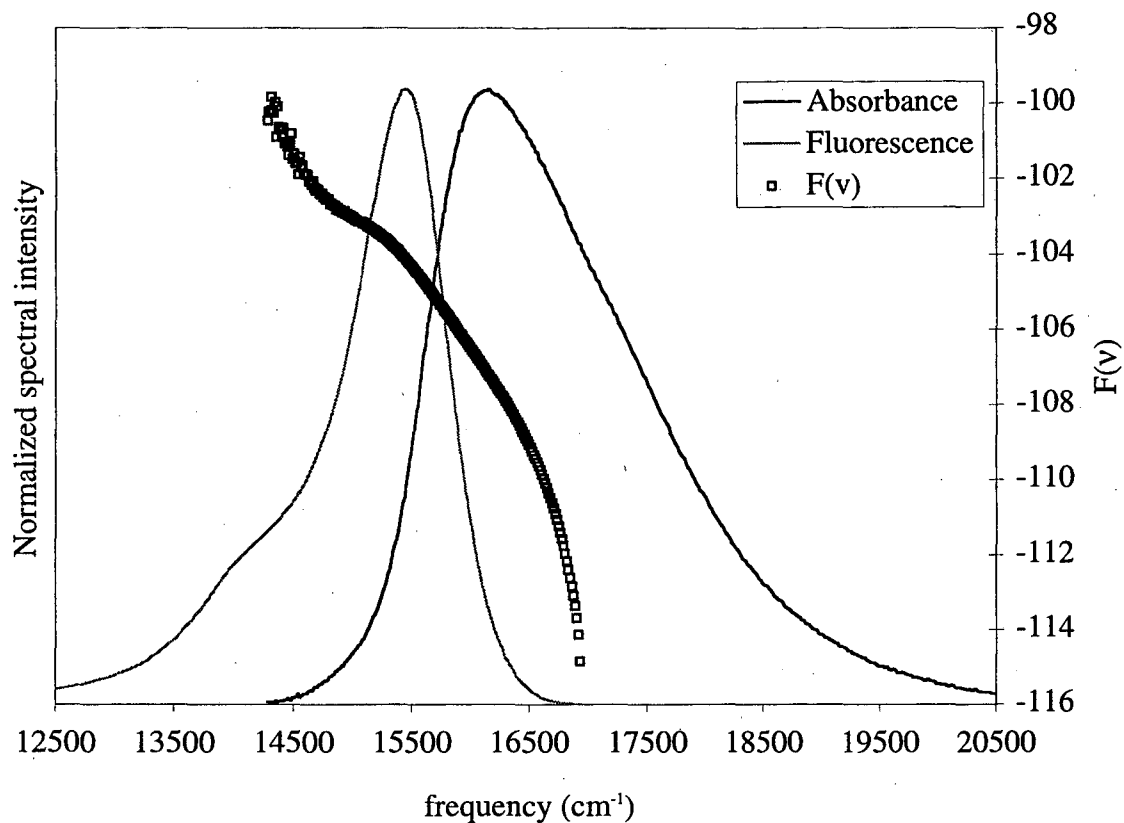


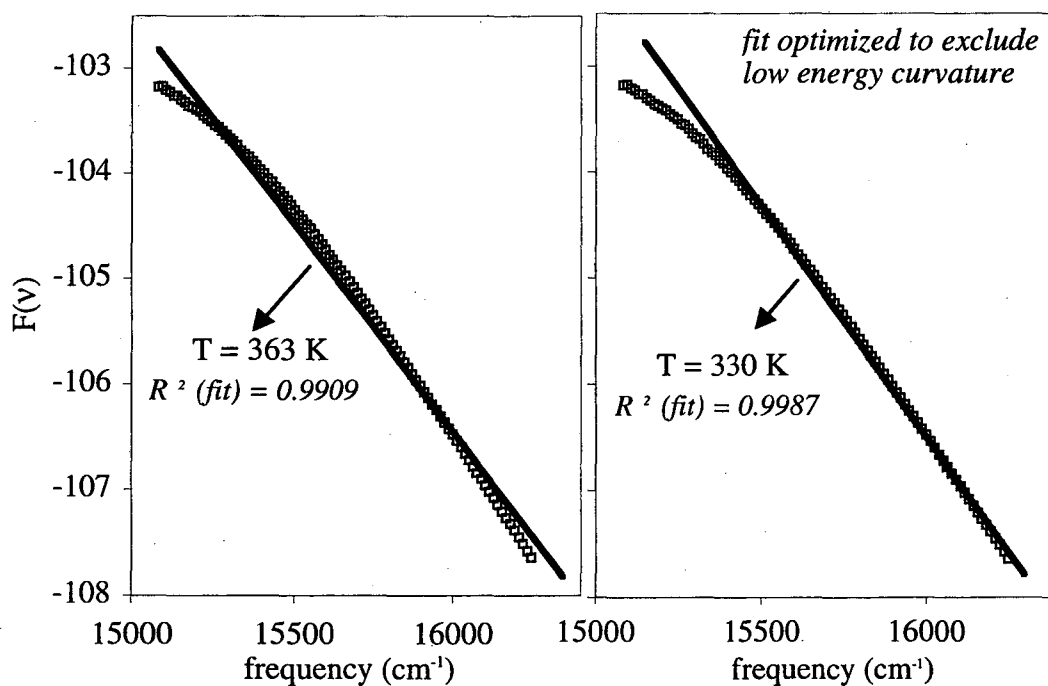
Figure 4-9D. Excitation scan of  $PC/L_{RC}^{28.5}$  at 298K and 88K at different emission wavelengths



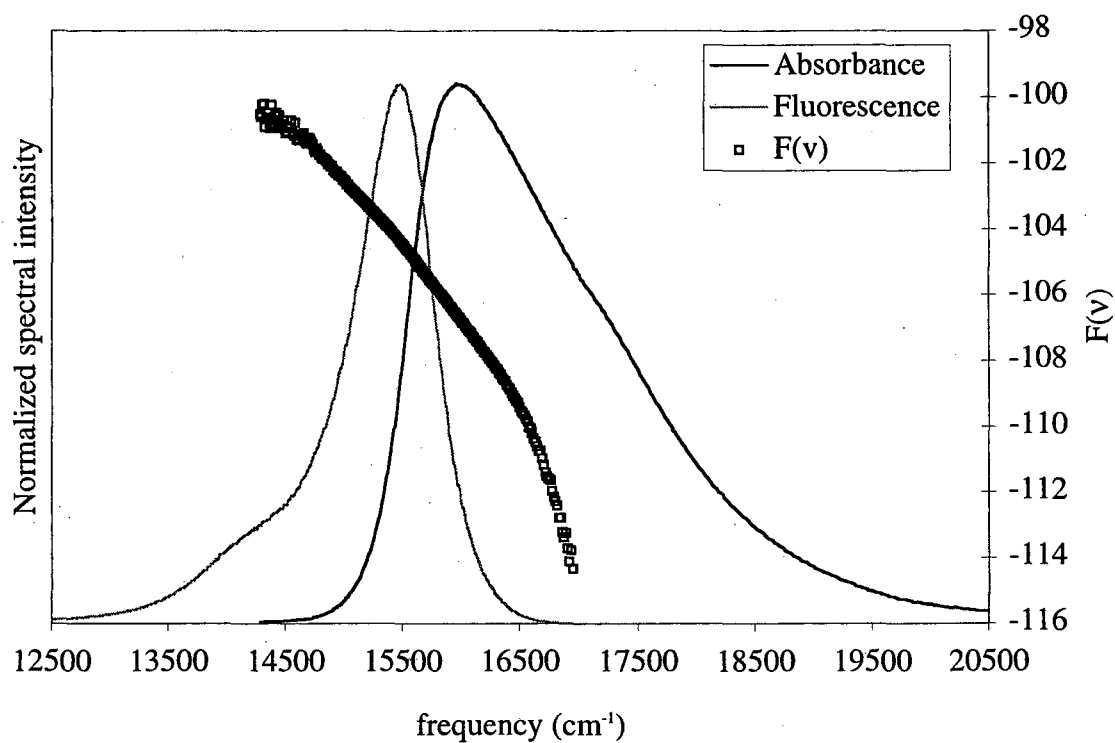
**Figure 4-10A.** Kennard-Stepanov  $F(\nu)$  calculation for **PC monomer** at 298K



Linear fits of  $F(\nu)$  in region of **>10%** normalized spectral intensity  
(Temperature calculated from the slope of the linear fit)



**Figure 4-10B.** Kennard-Stepanov  $F(\nu)$  calculation for **PC trimer** at 298K



Linear fits of  $F(\nu)$  in region of  $>10\%$  normalized spectral intensity  
(Temperature calculated from the slope of the linear fit)

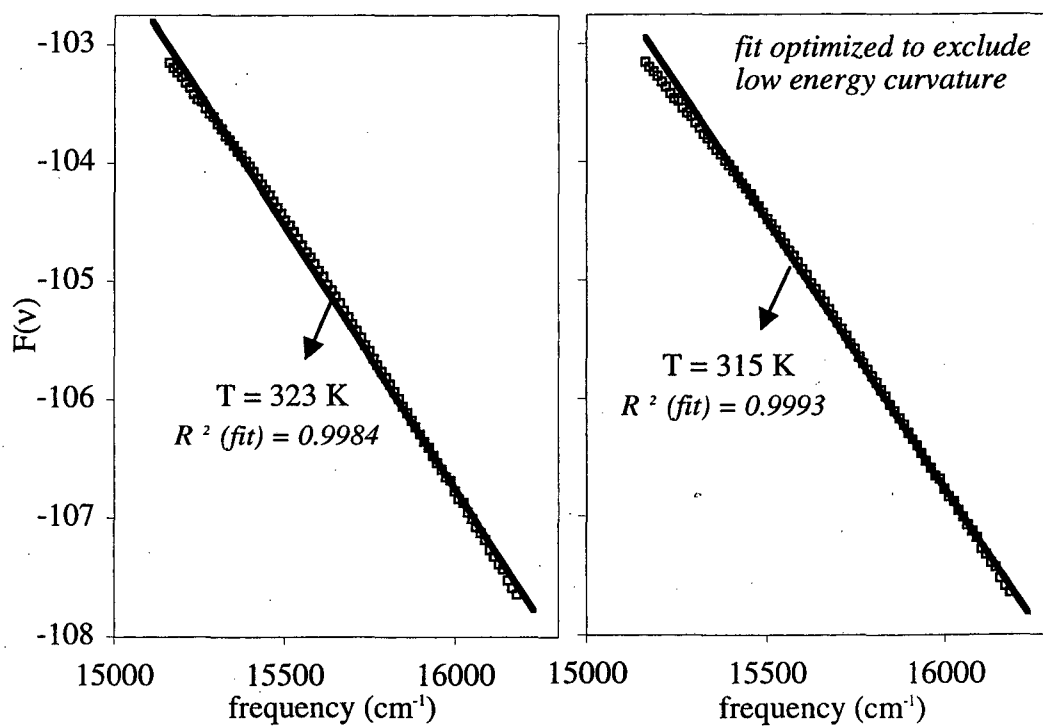
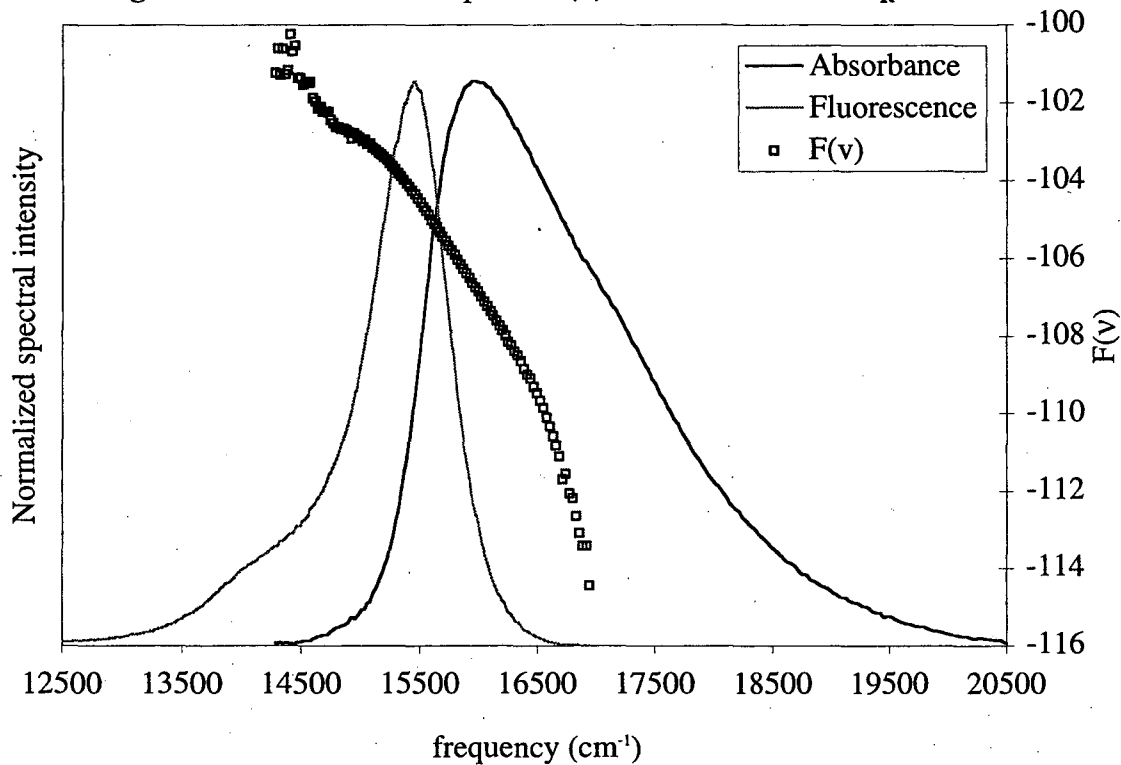




Figure 4-10C. Kennard-Stepanov  $F(\nu)$  calculation for  $\text{PC}/\text{L}_R^{32.3}$  at 298K



Linear fits of  $F(\nu)$  in region of  $>10\%$  normalized spectral intensity  
(Temperature calculated from the slope of the linear fit)

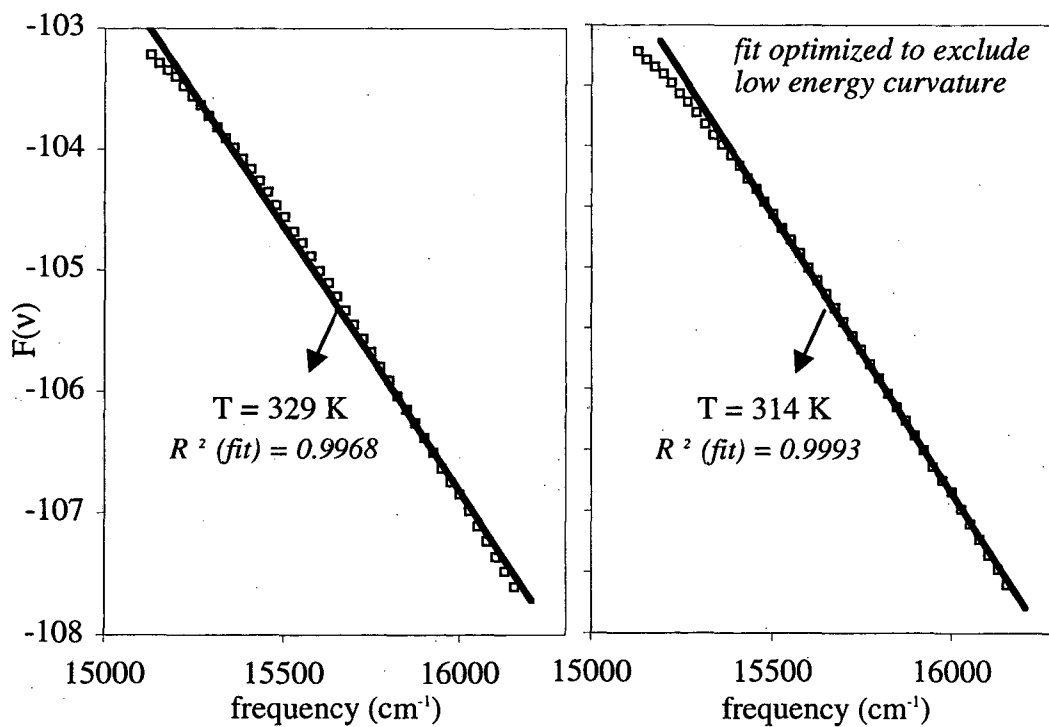
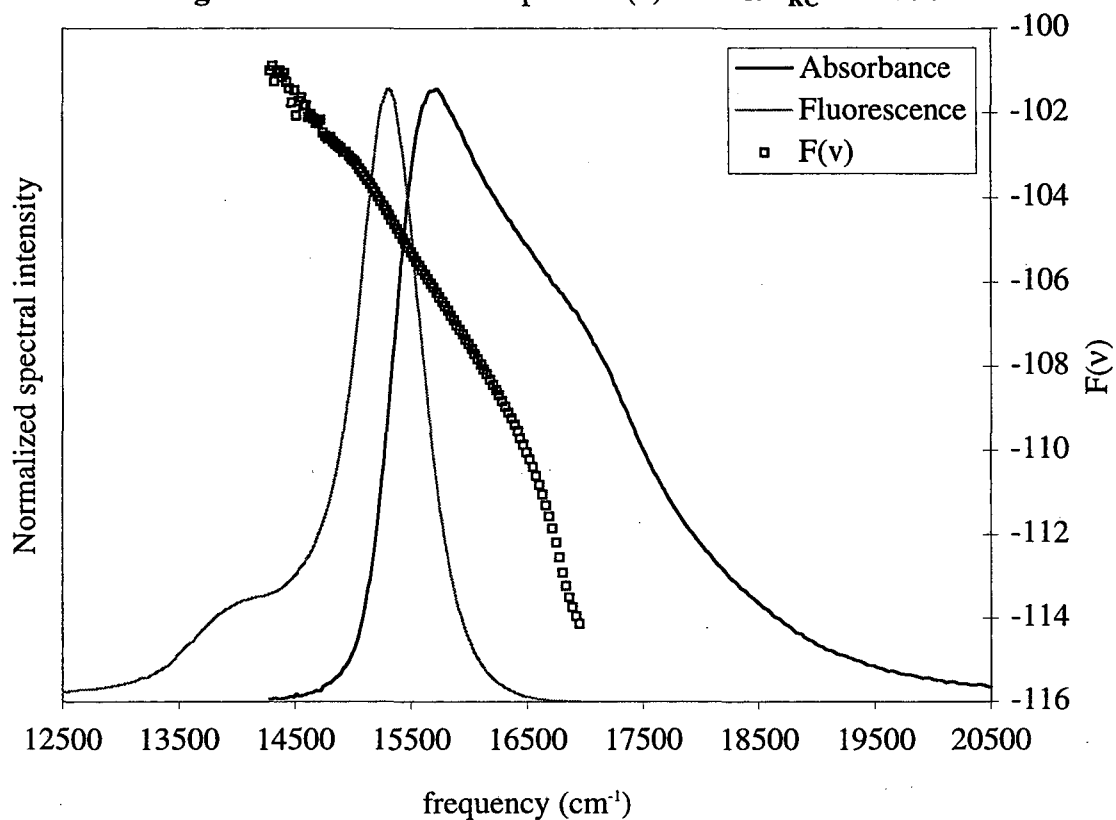
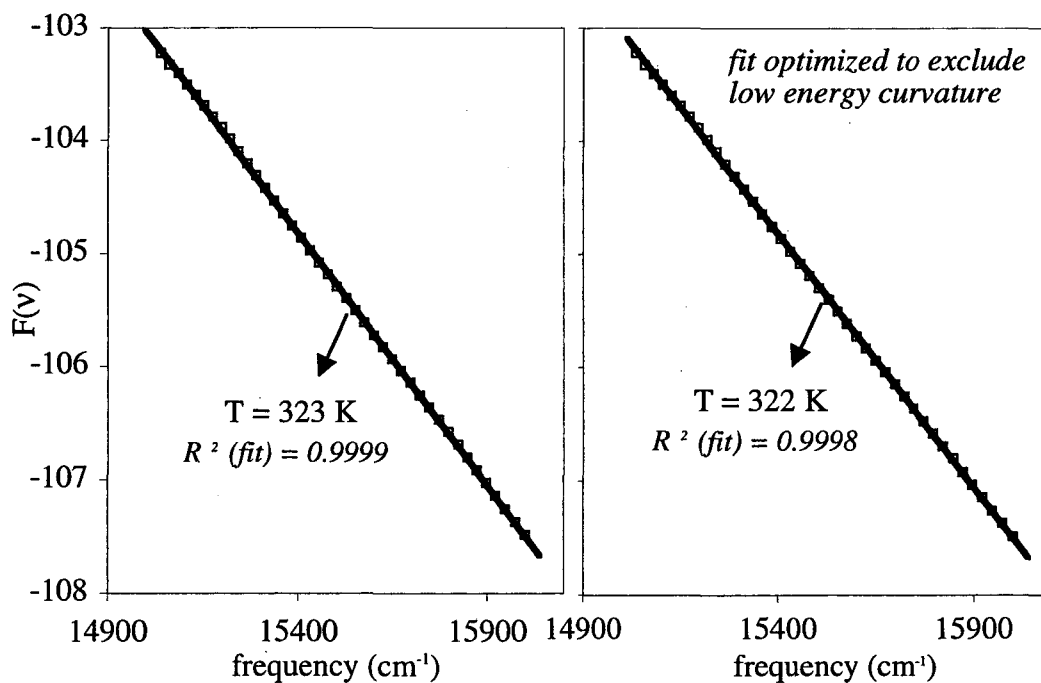


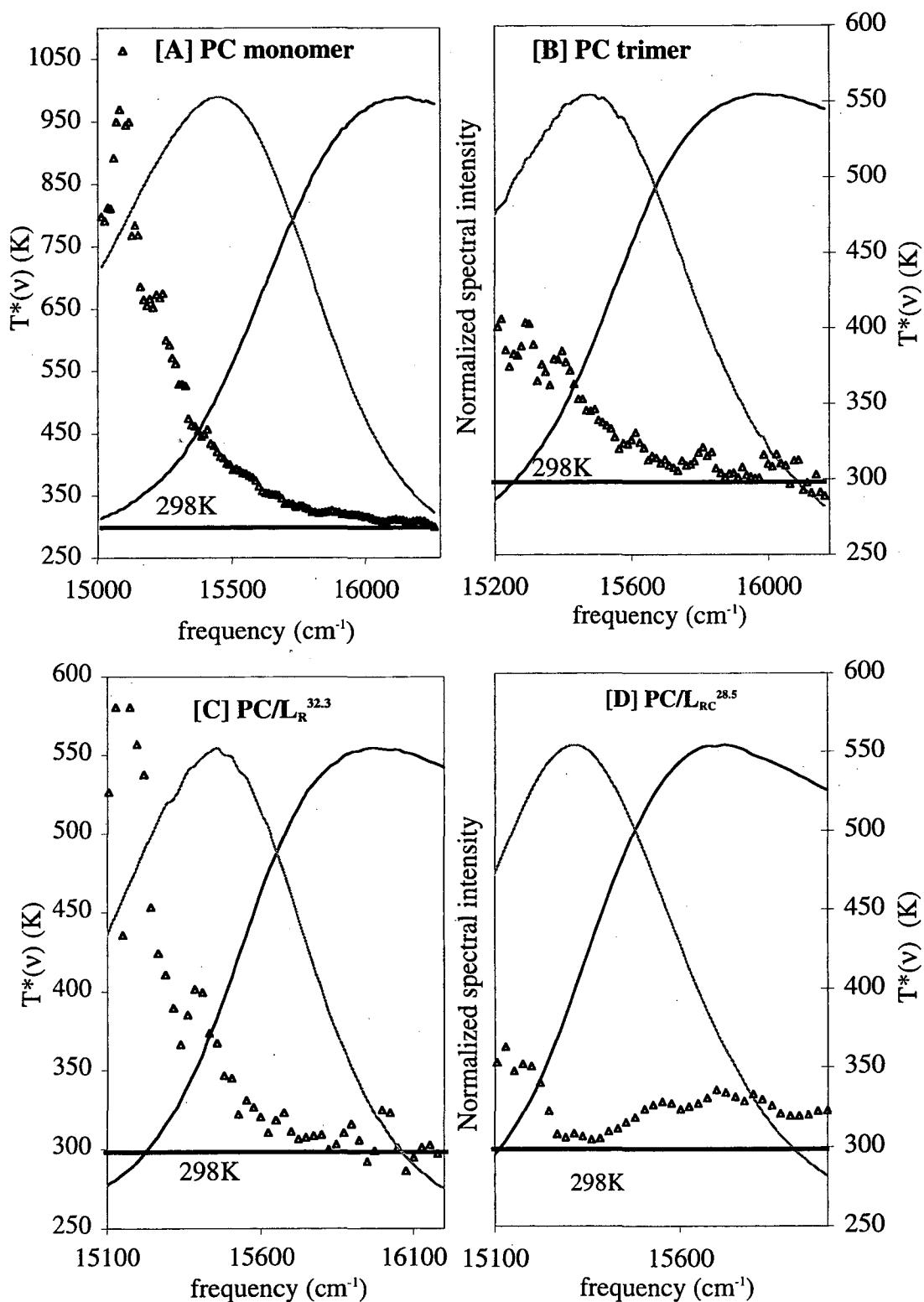
Figure 4-10D. Kennard-Stepanov  $F(\nu)$  for  $\text{PC}/L_{\text{RC}}^{28.5}$  at 298K



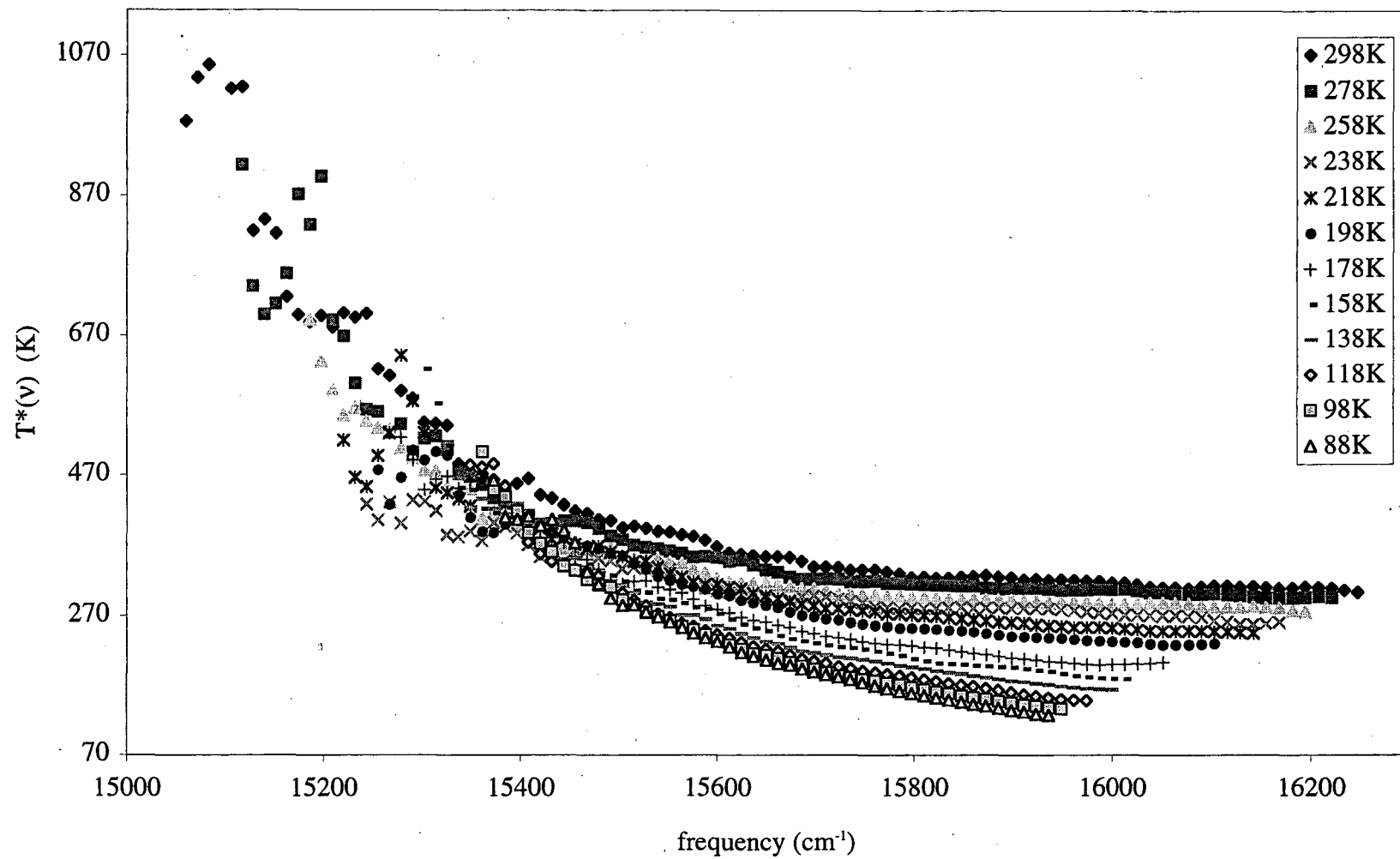
Linear fits of  $F(\nu)$  in region of  $>10\%$  normalized spectral intensity  
(Temperature calculated from the slope of the linear fit)



**Figure 4-11[A-D].** Calculated  $T^*(\nu)$  of PC complexes at 298K plotted against the absorbance and fluorescence spectra.  
Shown only in region of >10% normalized spectral intensity



**Figure 4-12A.** Calculated  $T^*(\nu)$  for PC monomer at different experimental temperatures  
(shown only in region of >10% normalized spectral intensity)



**Figure 4-12B.** Calculated  $T^*(\nu)$  for **PC trimer** at different experimental temperatures  
(shown only in region of >10% normalized spectral intensity)

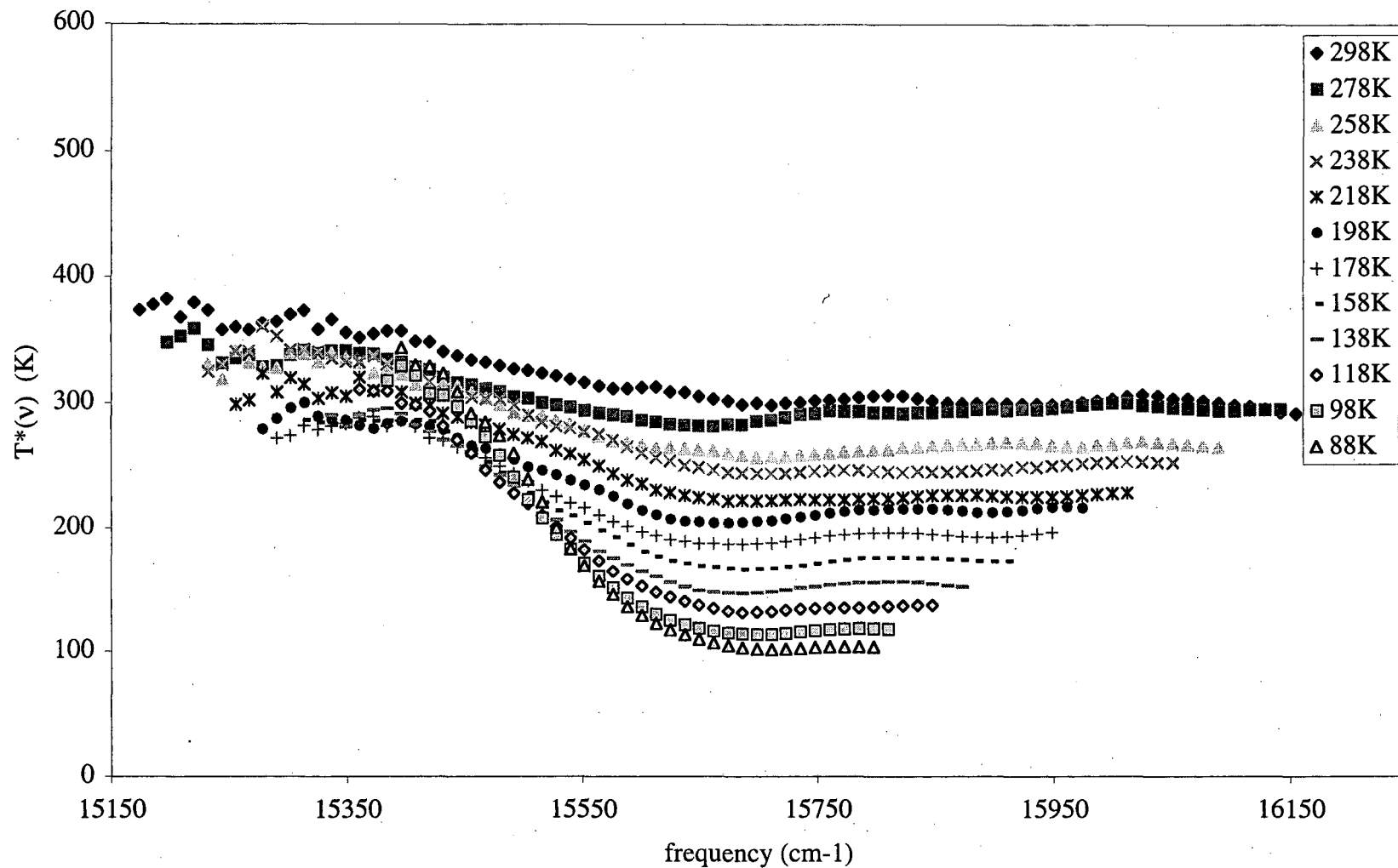


Figure 4-12C. Calculated  $T^*(\nu)$  for  $PC/L_R^{32.3}$  at different experimental temperatures  
 (shown only in areas of >10% normalized spectral intensity)

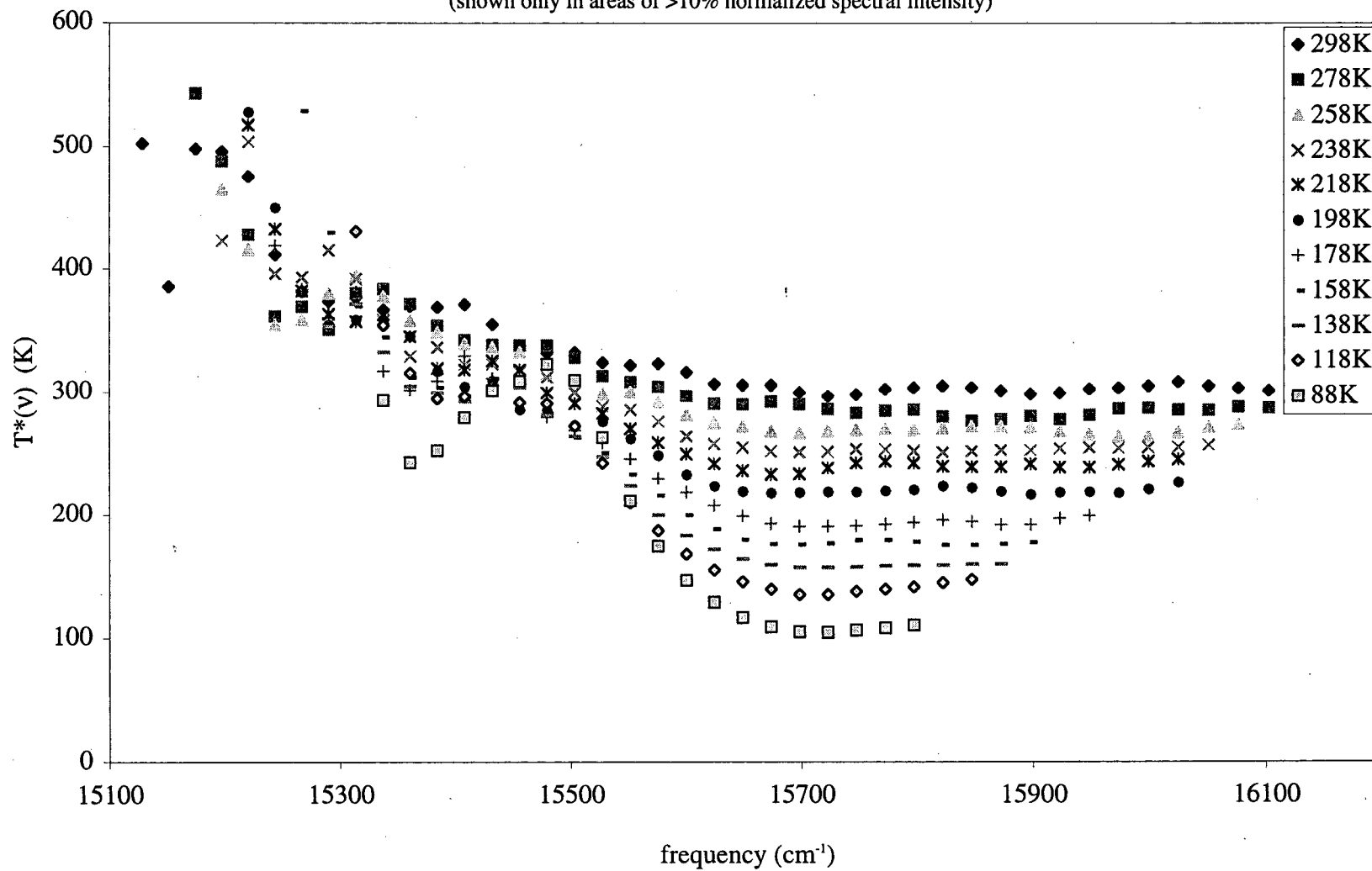
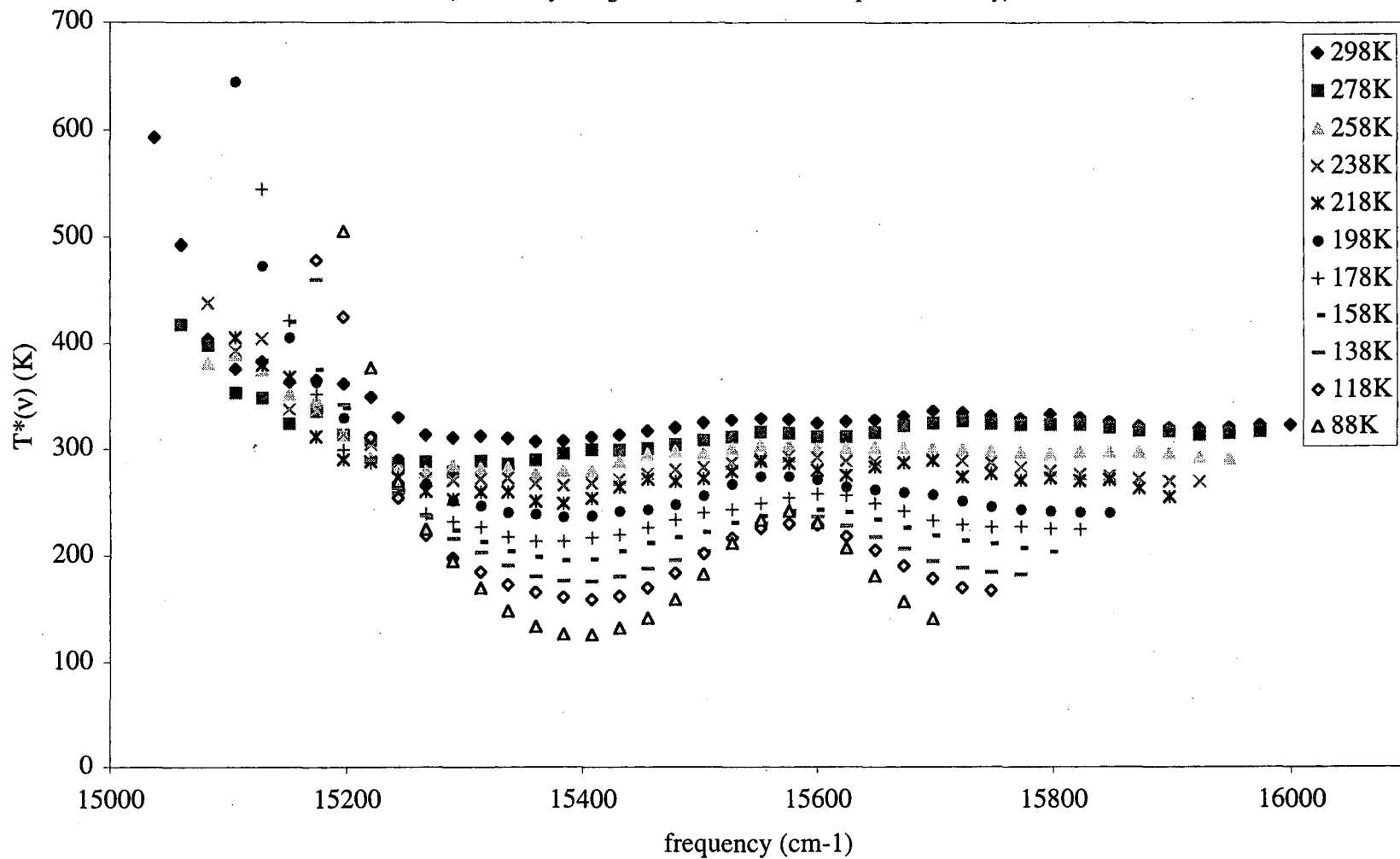
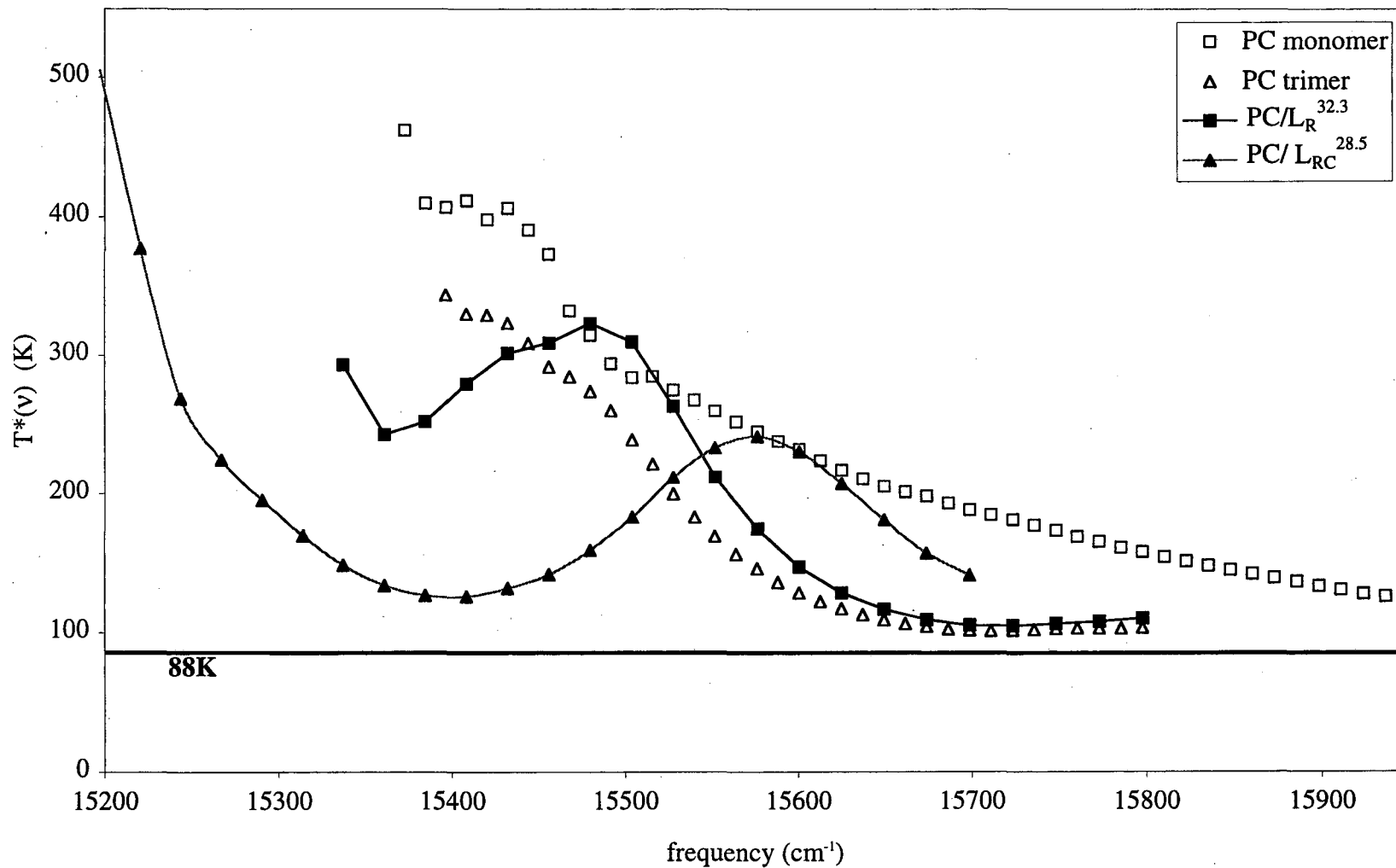


Figure 4-12D. Calculated  $T^*(\nu)$  for  $PC/L_{RC}^{28.5}$  at different experimental temperatures  
 (shown only in region of >10% normalized spectral intensity)



**Figure 4-13.** Calculated  $T^*(\nu)$  for PC complexes at 88K  
(shown only in region >10% normalized spectral intensity)





## Appendix 4-1. Simulations of PC absorbance spectra

The purpose of this appendix is to simulate the absorbance spectra of the PC monomer and trimer at 298 K and 88 K using the absorbance bands of the  $\alpha$ - and  $\beta$ -subunits or those of the individually resolved chromophores. The spectra of the subunits at 88 K and the resolved PC monomer chromophore bands were extracted from the work done by M.P. Debreczeny<sup>1</sup>. The analytical tool “Solver” from the Microsoft Excel<sup>®</sup> program was used to obtain the best fit for each case by minimizing the residual between the data and the fit.

The underlying concept being tested by these simulations is the magnitude of interaction between the chromophores. For a mixture of non-interacting dyes in solution, the absorbance spectrum is accurately represented as a sum of the individual absorbance bands. *Can the absorbance spectra of the PC monomer and trimer be treated as a linear combination of the individual components despite possible coupling between pairs of chromophores?* In the PC monomer the closest distance is found between the two  $\beta$ -chromophores (35 Å). This is beyond the accepted limit for coupling that significantly perturbs the absorption spectrum; however, in the PC trimer, chromophores from neighboring monomers achieve a more favorable interchromophore distance (21 Å).

### PC monomer simulations

The PC monomer is composed of two subunits ( $\alpha$  and  $\beta$ ) and contains three chromophores ( $\alpha$ 84,  $\beta$ 84, and  $\beta$ 155) which are chemically identical yet have different spectral properties. The spectroscopy of each chromophore is modulated by changes in its physical conformation or electrostatic interactions with charged amino acids placed in

or near the binding site. Figure A1 compares the PC monomer absorbance band measured at 298 K with a simulation that uses the absorbance bands of the individual  $\alpha$ - and  $\beta$ -subunits, as fitting curves. If both subunits are normalized to an equal amplitude [ $A_{\text{MAX}}$  for  $\alpha:\beta = 1:1$ ] then the combination results in a poor fit for the short wavelength side of the PC monomer absorbance curve—changing their relative amplitudes to [ $A_{\text{MAX}}$  for  $\alpha:\beta = 1:0.68$ ] results in a much better fit. Debreczeny<sup>1</sup>, as well as other groups<sup>2,3</sup> have found that the oscillator strengths of the chromophores in the PC monomer are not equal. The strongest absorbance intensity is found for  $\alpha 84$ , and relative to it, the  $\beta 155$  is at  $\sim 98\%$  and  $\beta 84$  at  $\sim 61\%$  (calculated from maximal absorbance measured at equal concentration). These values were used as initial parameters for  $A_{\text{MAX}}$  in simulating the PC monomer absorbance using the individually resolved chromophore spectra. Figure A2 shows that although a decent fit is obtained with these initial values, a better one can be gained if both the  $A_{\text{MAX}}$  amplitudes and the maximal wavelength positions for each curve are slightly adjusted (parameters summarized in Table A1). The fit was optimized by shifting the maximal positions of the  $\beta$ -chromophores 2 nm to longer wavelengths and then adjusting their  $A_{\text{MAX}}$  amplitudes to minimize the residual error between data and simulation. This optimized fit (Figure A2) slightly raises the  $\beta 84$  amplitude from the initial value by 2% while lowering the  $\beta 155$  amplitude by 13% (Table A1). Although the optimized  $\beta 155$  amplitude is lower than the lowest experimentally resolved relative intensity (Chapter 4, Table 4-7), all of the other parameters are well within the expected limits. This simulation results in a satisfactory representation of the PC monomer 298 K absorbance spectra.

At 88 K, the absorbance of the PC monomer shows two well-resolved peaks; the individual  $\beta$ -subunit absorbance band (measured by Debreczeny ) is also resolved into two peak features while the  $\alpha$ -subunit, possessing a single chromophore, has only one band maximum. Using these isolated subunit bands to model the 88 K PC monomer absorbance band we obtain an excellent fit with the following amplitude parameters [ $A_{MAX} \alpha::\beta(1):\beta(2) = 1:0.84:0.66^*$ ] (\* only the amplitudes for  $\alpha$  and  $\beta(1)$  are adjustable, the  $A_{MAX}$  for  $\beta(2)$  is linked to  $\beta(1)$  at a 0.79:1 ratio). Also, the 77 K  $\beta$ -subunit absorbance spectrum is simulated using two of the 77 K  $\alpha$ -subunit absorbance bands at different maximal wavelength positions to test whether all three chromophores have similar band shapes at 77 K. The result, shown in Figure A4, raises two interesting points. First, the overall  $\beta$ -subunit fit is poor; although the  $\alpha$ -subunit shape is a more realistic approximation than artificial Gaussian functions, the three chromophores do not possess identical band shapes. Secondly, the curve placed at the long-wavelength peak position, representing  $\beta_{84}$ , has a much higher amplitude than expected. The individually resolved chromophore bands at 298 K show that  $\beta_{84}$  has the lowest relative intensity, about 63% of  $\beta_{155}$ . The unexpectedly high  $\beta_{84}$   $A_{MAX}$  amplitude found in the 77 K  $\beta$ -subunit fit could be decreased if the individual fitting curve peaks were slightly wider resulting in an overall fit with lower amplitude for  $\beta_{84}$ . However if the band “tails” extending into the short wavelength region, are maintained at roughly the same intensity, then the value of the  $\beta_{84}$  amplitude cannot be lower than that of the short-wavelength component (i.e.,  $\beta_{155}$ ) because they are additive—it would result in a very poor fit of the short wavelength region. This implies that at low temperature there is either a large difference in the absorption band shape of the two  $\beta$ -chromophores (particularly in the extending

short-wavelength tails) or there is some coupling between the two chromophores which alters their relative amplitudes.

### **PC trimer**

The assumption of non-interacting chromophores to make linear combinations has thus far been adequate in describing the PC monomer, where few interactions are expected between the chromophores.<sup>3</sup> PC trimer formation induces changes in the immediate protein environment of chromophores near the neighboring junction of the monomers ( $\alpha 84$  and  $\beta 84$ ). It also provides opportunities for excitonic coupling between these two chromophores because their interchromophore distance is now  $\sim 20$  Å. Figure A5 presents the simulation of the 298 K PC trimer absorbance using the individual chromophore spectra from the PC monomer. The amplitudes and maximal wavelength positions were adjusted from the initial values, and these are summarized in Table A1. The maximal wavelength positions of  $\beta 84$  and  $\alpha 84$  were increased equally (3 nm) while  $\beta 155$  remained at the same initial position. This is consistent with the expected structural changes due to trimer formation, the chromophores closest to the monomer junction should be the most affected. The relative amplitude of  $\beta 84$  was increased from the initial value and is consistent with published values (Table 4-6), but the  $\beta 155$  amplitude is once again lower than expected although consistent with the 298 K PC monomer fit (Table A1). The overall shape of the fit simulates the PC trimer absorbance well except for a tail on the red-wavelength side ( $>650$  nm) that is not present in the trimer. Even so, a linear combination of the individual chromophore spectra appears to suitably reproduce the 298 K PC trimer absorbance with minor manipulation.

The 88K PC trimer spectrum is more complex than that of the monomer and contains both well-resolved peaks and sharp shoulder features. Examination of the PC trimer crystal structure<sup>5</sup> shows that all three chromophore types have different conformations, and therefore one expects distinct spectral properties. Figure A6 presents the simulation of the 88 K PC trimer absorbance using the spectra of the individual monomer subunits at 77 K. The shoulder/peak feature at ~635 nm is suspected to be a band splitting from an excitonically coupled energy level produced by  $\alpha 84$  and  $\beta 84$ . The individual chromophore spectra at 77 K were not available for this exercise, but using the subunit spectra we can see that this PC trimer feature will not be fit by simply adding the subunit bands (Figure A6). Clearly, the intensity of the longest-wavelength absorbing chromophore ( $\beta 84$ ) needs to increase, while  $\alpha 84$  decreases and  $\beta 155$  remains unchanged, to produce the peak/shoulder feature. This scenario is a possible consequence of excitonic coupling between neighboring  $\alpha 84$  and  $\beta 84$  chromophores. A number of linear combinations of the subunits at different amplitudes and wavelength positions were tested and did not produce a better fit of the PC trimer spectra. The biggest limitation in reproducing the PC trimer sharp shoulder/peak is the width of the fitting curve peaks. The PC trimer feature width at 75% absorbance is ~18 nm while the corresponding  $\alpha$ -subunit width is 17 nm. If all three chromophore bands possess the same shape as the  $\alpha$ -subunit, then their combined absorption would be too wide to fit the long-wavelength trimer feature. We used Gaussian waves of varying widths, amplitudes, and maximal wavelength positions to simulate the trimer shape. The result of this effort is included in Figure A6 and the parameters for the Gaussian waves are presented in Table A2. The optimal width of the Gaussian waves at 75% amplitude is 7 nm, significantly narrower

than the  $\alpha$ -subunit peak. The separation between these artificial Gaussian waves is 276  $\text{cm}^{-1}$ , much larger than the 112  $\text{cm}^{-1}$  splitting calculated by Sauer and Scheer<sup>5</sup> using the refined PC trimer crystal structure.

The PC trimer absorbance spectrum represents the influence of various factors. The PC trimer is a stable, native form of the protein, and molar quantities of a chaotropic salt (KSCN) are necessary to reduce it to monomers. Such strong bonding interaction in the trimer complex imposes a rigid structure that is optimal in aligning chromophores in a specific way to facilitate energy transfer. The combination of an imposed guiding structure and the availability of nearby chromophores for excitonic interaction contribute to the observed absorbance spectrum of the PC trimer. It is not a simple result of additive non-interacting spectral bands and, while the simulation succeeds in imitating the general profile of the overall 298 K absorbance band, it fails severely in capturing the details of the interaction between the spectral bands at 88 K.

## References

1. Debreczeny, M.P., Sauer, K., Zhou, J., Bryant, D.A. (1993) *J. Phys. Chem.* 97:9852-9862; Debreczeny, M.P. (1994) Ph.D. thesis, LBL-35672, University of California, Berkeley, CA.
2. Mimuro, M., Füglistaller, P., Rübelen, R., Zuber, H., (1986) *Biochim. Biophys. Acta* 848, 155-166.
3. Sauer, K., Scheer, H., Sauer, P. (1987) *Photochem. Photobiol.* 46, 427-440.
4. Duerring, K.M., Schmidt, G.B., Huber, R. (1991) *J. Mol. Biol.* 217:577-592.
5. Sauer, K., Scheer, H. (1988) *Biochim. Biophys. Acta* 936, 157-170.

**Table A1.** Parameters for individual chromophore spectra used to fit the PC monomer and PC trimer absorbance spectra (298 K)**PC Monomer 298K** [Figure A2]

chromophore	Relative Amplitude		Wavelength position of absorbance maximum (nm)	
	initial*	optimized	initial*	optimized
$\alpha$ 84	1	1	624	624
$\beta$ 84	0.61	0.63	627	629
$\beta$ 155	0.98	0.85	601	603

**PC Trimer 298K** [Figure A5]

chromophore	Relative Amplitude		Wavelength position of absorbance maximum (nm)	
	initial*	optimized	initial*	optimized
$\alpha$ 84	1	1	624	627
$\beta$ 84	0.61	0.68	627	630
$\beta$ 155	0.98	0.85	601	601

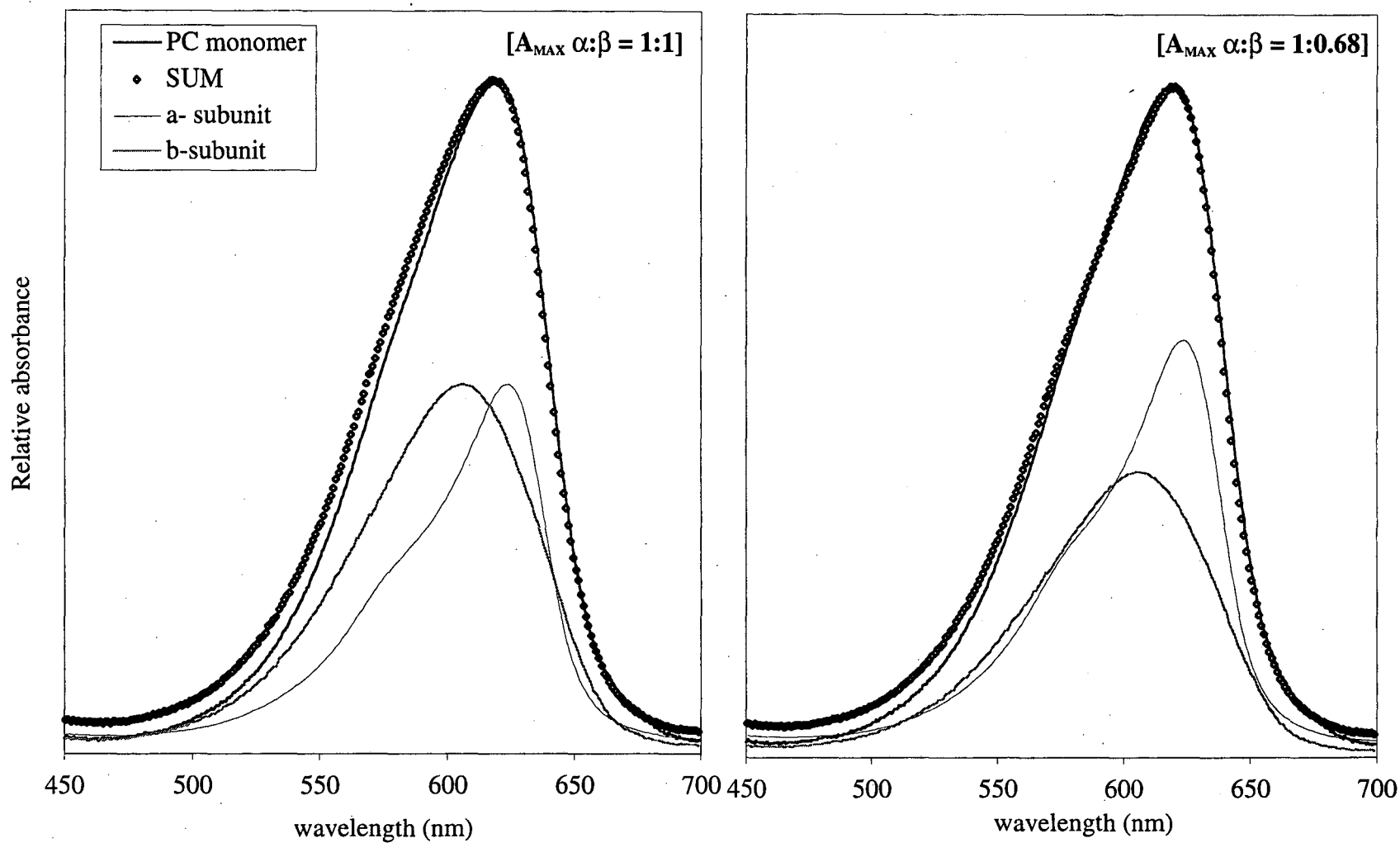
\*initial values and chromophore spectra obtained by M.P. Debreczeny (1994) Ph. D. Thesis, University of California, Berkeley, CA.

**Table A2.** Parameters for individual Gaussian waves used to fit the PC trimer absorbance spectra (88 K) in Figure A6

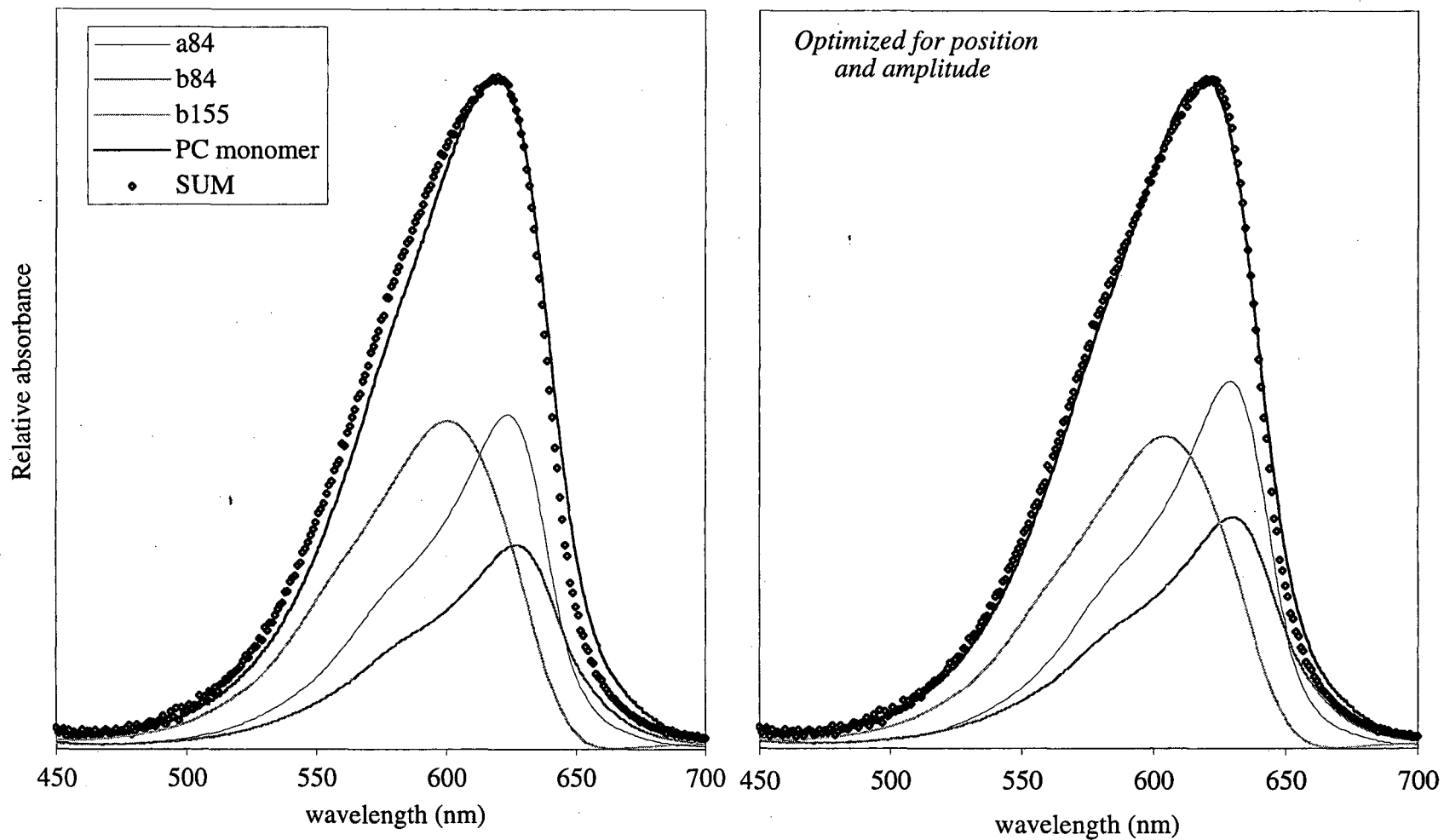
	Relative amplitude	Maximal position	Width at 75% amplitude
wave 1	1	637	7
wave 2	0.85	626	7



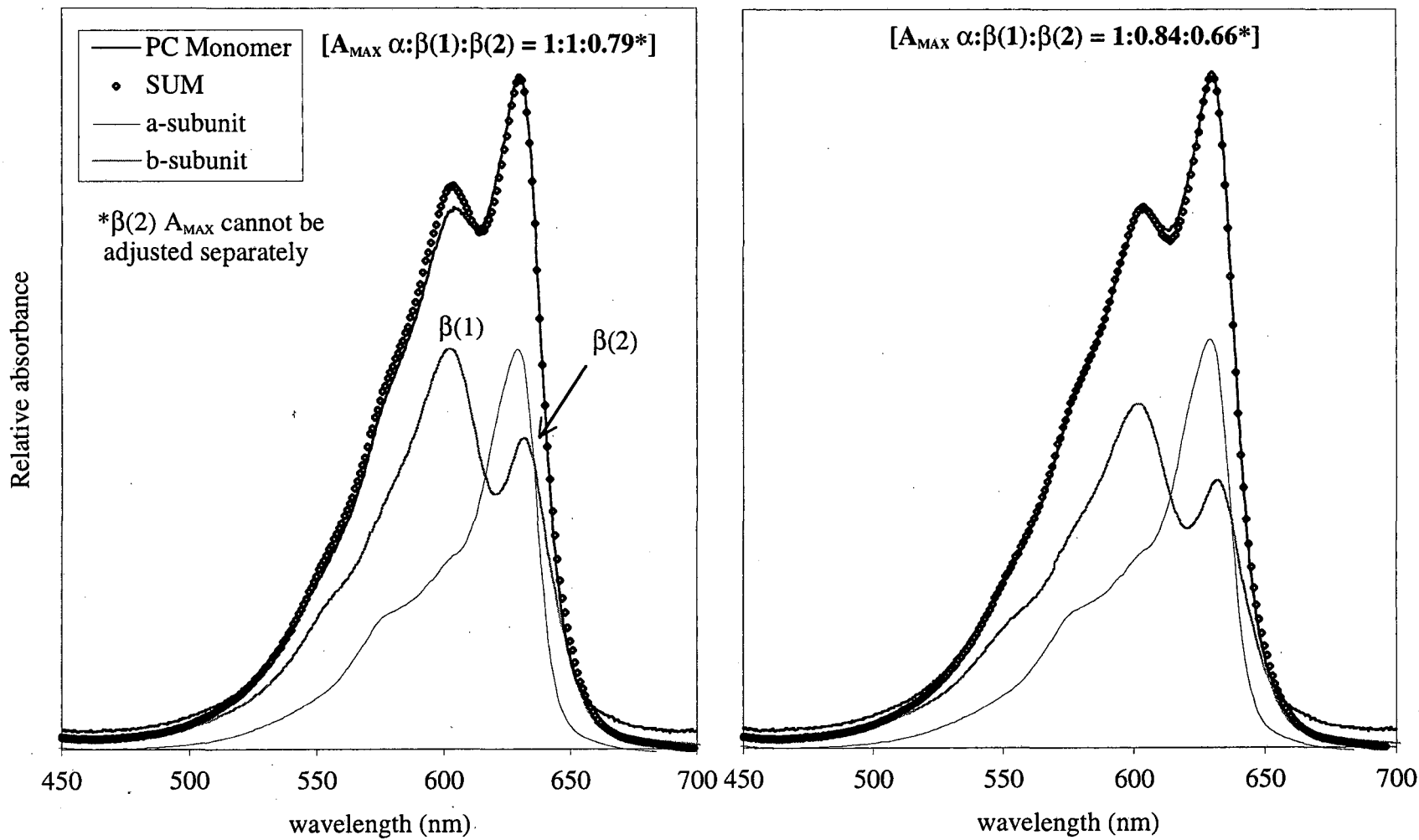
**Figure A1.** Absorbance of PC monomer at 298 K compared with the sum of isolated  $\alpha$ - and  $\beta$ -subunits at different amplitudes



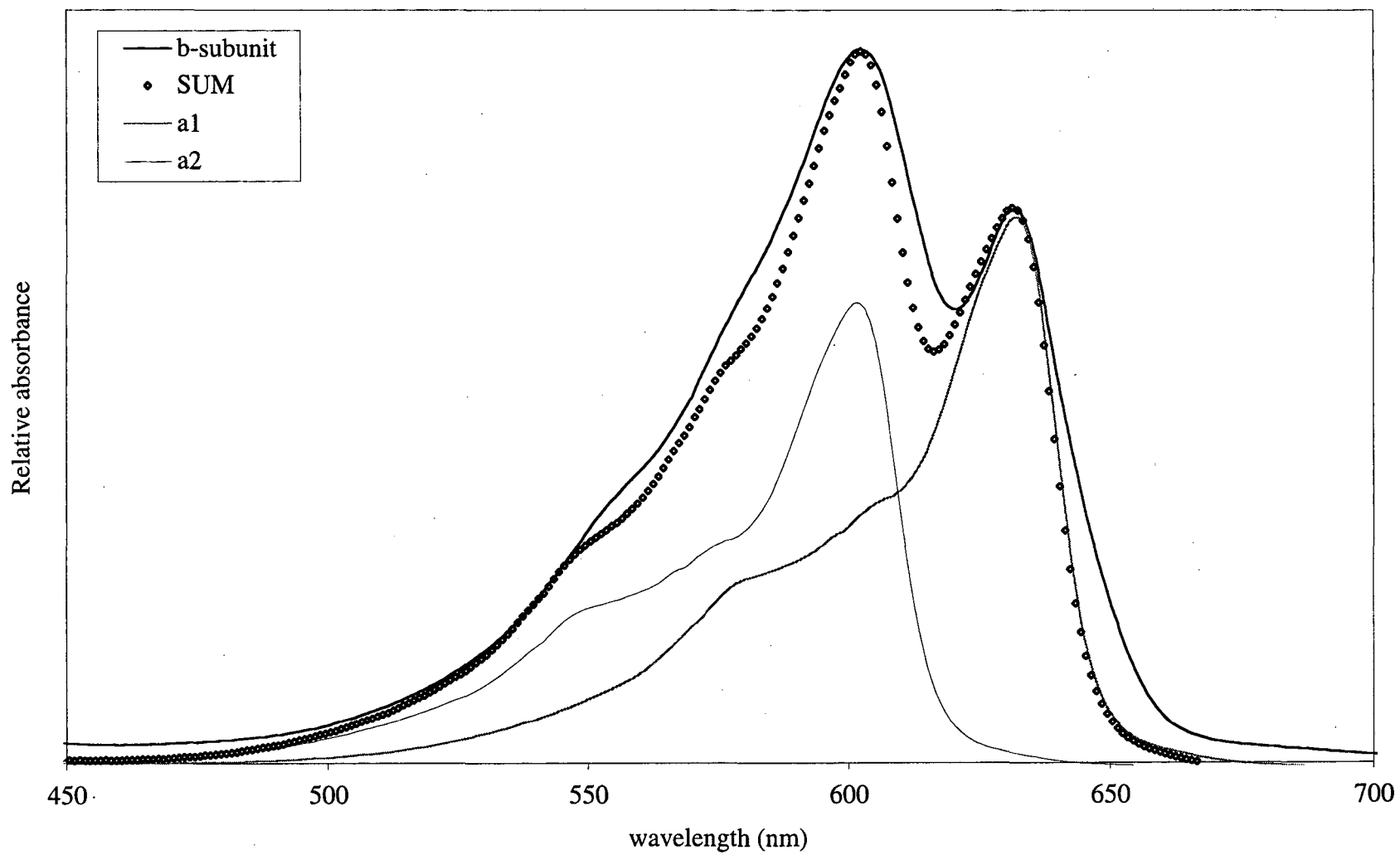
**Figure A2.** Absorbance of **PC monomer** at 298 K compared with the sum of individual ( $\alpha$ 84,  $\beta$ 84,  $\beta$ 155) chromophore spectra



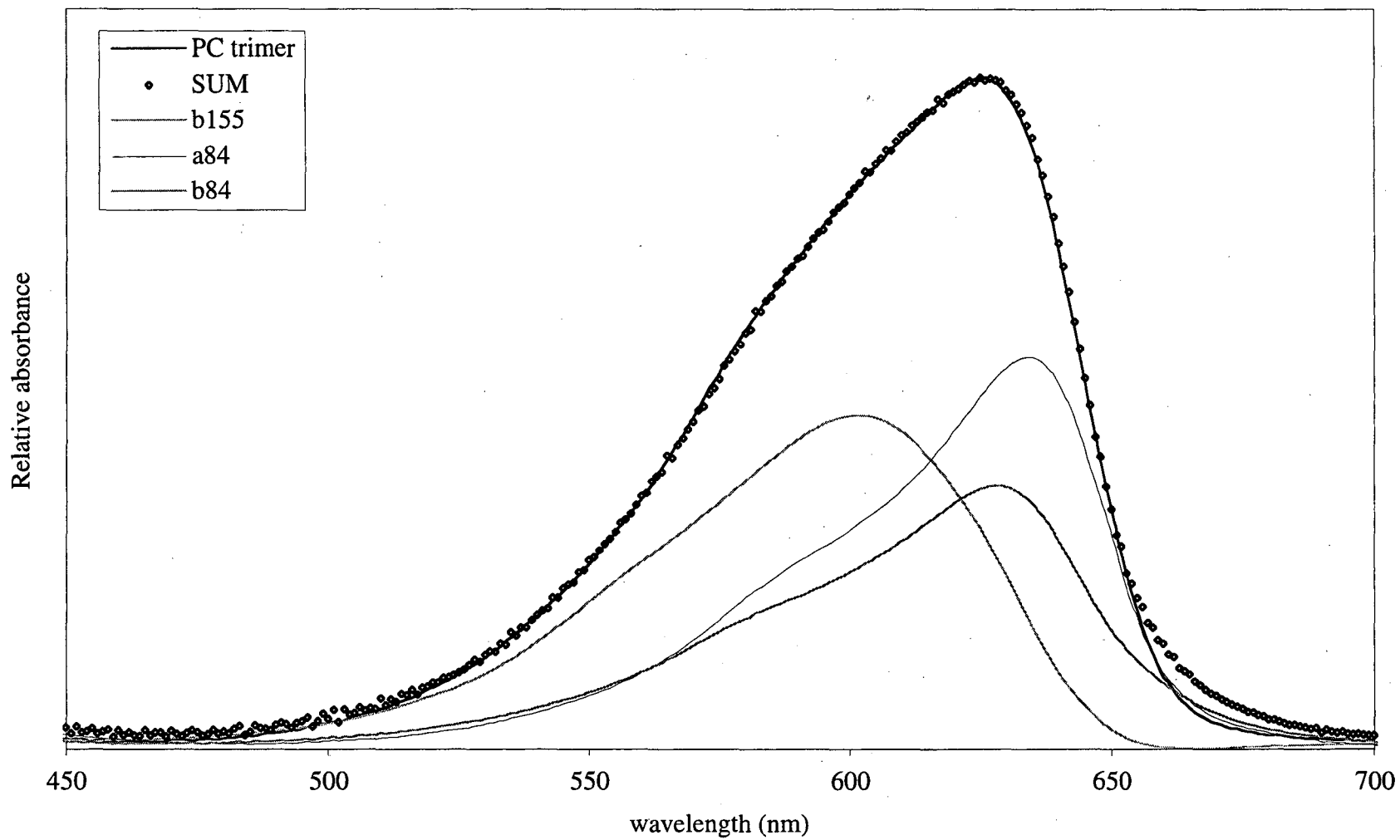
**Figure A3.** Absorbance of PC monomer at 88 K compared with the sum of the  $\alpha$ - and  $\beta$ - subunits at different amplitudes



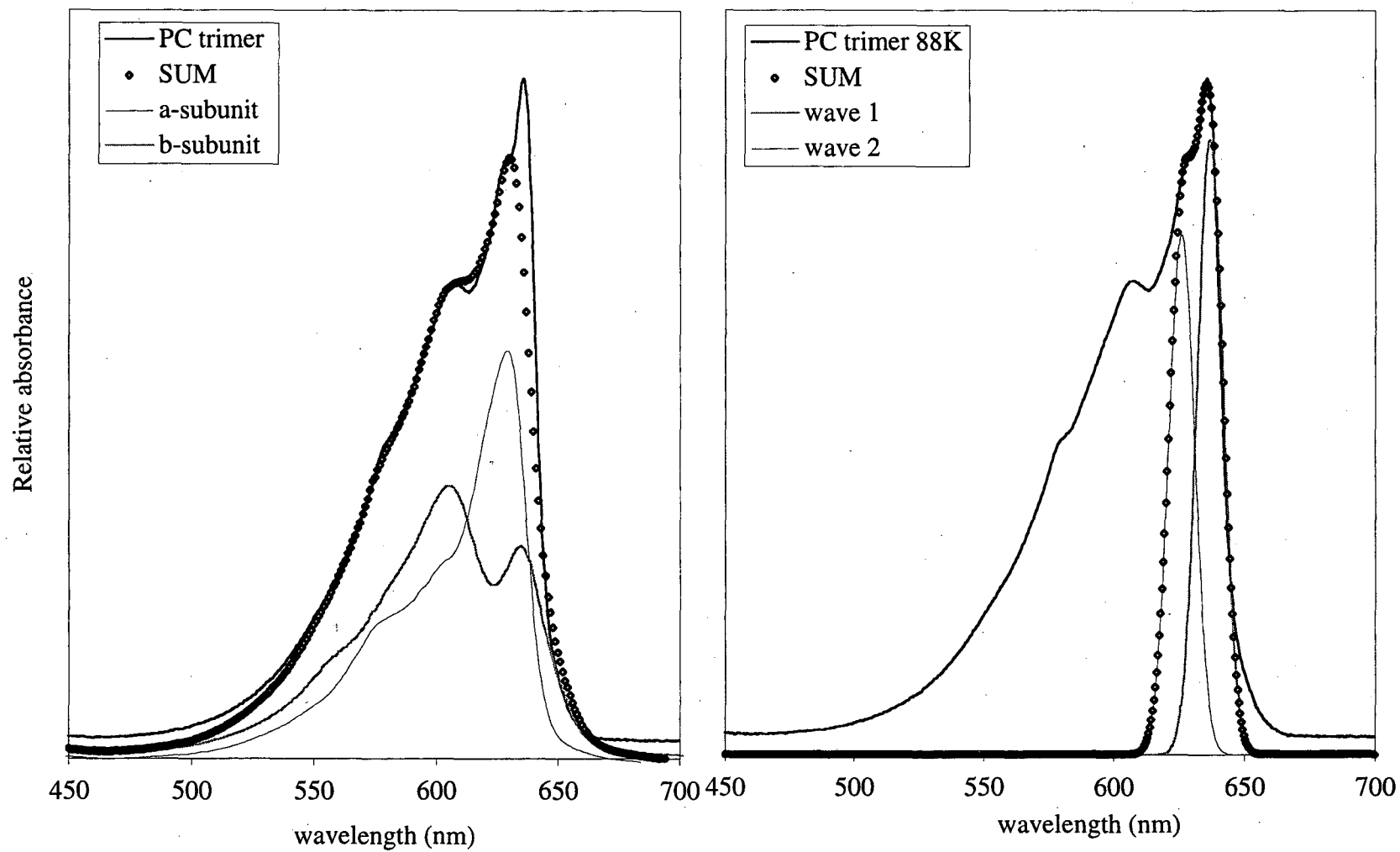
**Figure A4.** Absorbance of  $\beta$ -subunit at 77 K simulated using two " $\alpha$ -subunit" waveforms



**Figure A5.** Absorbance of PC trimer at 298 K compared with the sum of the individual chromophore spectra resolved from the PC monomer



**Figure A6.** Absorbance of **PC trimer** at 88 K simulated using the isolated  $\alpha$ - and  $\beta$ - subunits (77 K) and Gaussian functions



## Chapter 5. Concluding Remarks and Future Directions

This thesis examines the effect of linker association on the spectral properties of PC trimer aggregates with the goal of pinpointing the mechanism of interaction between the linker and the PC chromophores. The modulating effect of linker association upon the spectral properties of the biliproteins has been known since the 1980's, even before the first biliprotein crystal structure<sup>1</sup> appeared in 1987. In 1990, the amino acid sequences of the *Synechococcus* sp. PCC 7002 linker peptides used in this study were resolved<sup>2-4</sup> but their three-dimensional structures remain unknown. This year, we have enjoyed the publication of the first biliprotein/linker crystal structure<sup>5</sup>; APC/L<sub>C</sub><sup>7,8</sup> has provided a guiding light for examination of the amino acid sequences of the linker peptides to find possible interaction sites with the biliprotein chromophores. This study brings together a detailed spectroscopic characterization of each isolated PC/linker complex (Chapter 4) with proteomic analysis of the linker amino acid sequences (Chapter 3) to produce a model for biliprotein/linker interaction.

### Proposed model of biliprotein/linker interaction

Figure 5-1 presents a schematic diagram of the proposed model of biliprotein-linker assembly. The PC trimers assemble face-to-face with each other; therefore, the two exposed surfaces of a hexameric rod disk are symmetrical, and the depth of one disk is ~60Å. The hydrophobic properties of the linker sequences indicate that they would be more stable in an environment that protects them from the solvent, such as the biliprotein central cavity. Studies of rod assembly in *Anabaena variabilis* showed that a significant part of the linker remains outside the biliprotein when associated with only one hexamer,

and this portion is susceptible to digestion by trypsin.<sup>6</sup> The resulting digested product was spectroscopically identical to the untreated complex, but it could no longer connect with other PC hexamer disks. A 32.5 kDa linker was completely protected from proteolysis when assembled with two hexamer stacks but would degrade to 28 kDa if associated with only one hexamer. This suggests that, although only the N-terminal portion of the linker is necessary to affect the biliprotein spectroscopic properties, the rest of the linker is well-protected from the solvent and thus likely inserted in another hexamer. Examination of the amino acid sequence of the *Synechococcus* sp. PCC 7002 L<sub>R</sub><sup>32.3</sup> linker showed that it has two separate regions of shared homology with the other rod linker peptides investigated in this study; its N-terminal region is similar to the N-terminal region of L<sub>RC</sub><sup>28.5</sup> and a smaller area near its C-terminus is homologous with the C-terminal area of L<sub>R</sub><sup>8.9</sup> (Figures 3-1 and 3-2). Because these two conserved regions are not similar to each other, it supports the idea of two distinct biliprotein-binding mechanisms for a single linker: one that produces the observed spectral changes upon linker-biliprotein association and a second that may provide structural support without influencing the chromophore spectra. If the chromophores from only one side of the hexamer disk are “tuned” by linker association towards lower energy, then the energy transfer process is guided in one direction. The linker must then penetrate only as far as the first PC trimer to produce this favorable configuration.

It is unknown whether the linkers interact with each other through the PC hexamer’s central cavity. Such an arrangement would provide optimal structural support to the rod assembly and could thus enhance the transfer of energy. A 28 kDa protein could sufficiently occupy the PC hexamer central cavity space further than the distance



necessary to interact with one PC trimer and attach to another linker through the central cavity. This linker-to-linker binding interaction would produce a strong backbone for the phycobilisome (PBS) rods with minimal additional disturbance of the PC chromophores by the linkers. However, the relatively easy isolation of PC trimer/linker aggregates<sup>7</sup> indicates that, if such a linker-to-linker connection does exist, it is weaker than the linker-to-biliprotein association.

The linker-chromophore interaction can be examined from two perspectives: (1) general effects of linker association that affect the structure of the PC trimer without targeting individual chromophores and (2) specific perturbations of the chromophore conformation or its electrostatic environment by the linker. The C<sub>3</sub>-symmetry of the PC trimer internal cavity is disrupted by the presence of the single linker peptide. From a general perspective, the interaction among the three  $\beta$ 84 chromophores exposed to the cavity would be equally perturbed by placing the linker between them. The  $\beta$ 84 chromophores are located  $\sim 12$  Å from the PC disk surface exposed to solvent so the linker must penetrate this deeply within the central cavity to interfere with the interaction of these central chromophores. The APC/L<sub>C</sub><sup>7,8</sup> crystal structure showed that there is a variability of  $\sim 3$  Å in the insertion depth of this linker within different APC trimer units. L<sub>C</sub><sup>7,8</sup> induced a physical stress on the planar configuration of the APC trimer that served to bring the central  $\beta$ 84 chromophores slightly closer together without disrupting their distance or relative orientation to neighboring  $\alpha$ 84 chromophores. The variability of the linker insertion depth and its consequence on (1) the biliprotein planarity, or (2) disruption of the C<sub>3</sub>-symmetry of the internal cavity, could be a general effect of linker association that uniformly affects the spectral characteristics of the chromophores. The

basic pI of  $L_{RC}^{28.5}$  as well as its relatively high content of hydrophobic residues (Table 3-2) suggest that this linker may penetrate further within the PC internal cavity than  $L_R^{32.3}$ .

It may be that both linkers bind at the same biliprotein site, because each induces a similar decrease in one of the PC trimer CD bands (Figure 4-4), indicating a structural change among coupled chromophores. However, each linker has an opposite effect on the PC trimer excitation anisotropy (Figure 4-5), and the absorbance and fluorescence modulation are characteristically distinct. The spectral properties of the  $PC/L_R^{32.3}$  and  $PC/L_{RC}^{28.5}$  complexes are so different from each other that they must be produced by specific linker-chromophore interactions once the linker is bound to the biliprotein. The  $L_{RC}^{28.5}$  linker dramatically alters the spectral properties of the PC trimer, while  $L_R^{32.3}$  provides only slight changes. The remarkable differences in the absorbance spectra of these two complexes suggest that PC trimer association with  $L_{RC}^{28.5}$  linker not only alters the C3-symmetry of the structure but also seriously perturbs the individual chromophores. Because the 88 K absorbance and fluorescence bands of  $PC/L_{RC}^{28.5}$  are so markedly different from those of the PC trimer (Figures 4-2D and 4-7D), we propose that  $L_{RC}^{28.5}$  is affecting the relation between the  $\beta 84$  and  $\alpha 84$  chromophores to establish one single terminal emitter to the PBS core. A calculated difference spectrum between the PC trimer and  $PC/L_{RC}^{28.5}$  absorbance bands at 88K shows that the linker greatly enhances absorption at wavelengths  $> 636$  nm while suppressing the shorter wavelength absorbing components. In contrast, the spectral modulation of the  $L_R^{32.3}$  linker on the PC trimer is slight. The 88 K  $PC/L_R^{32.3}$  absorbance is very similar to that of the PC

trimer (Figure 4-2D) but the 88 K fluorescence (Figure 4-7D) reveals a shoulder on the long-wavelength side of the band maximum (at ~649 nm) that is absent in the PC trimer. This suggests that, although the  $L_R^{32.3}$  linker does not appear to interrupt established PC trimer interchromophore interactions and overall produces minor changes in the PC trimer spectra, it does specifically influence the properties of the terminal emitter.

The APC/ $L_C^{7.8}$  structure revealed that the  $L_C^{7.8}$  linker binds to only two of the three available APC monomers and that each linker chromophore interaction is different (Figure 3-4). In one APC monomer, the conformation of the  $\beta 84$  chromophore is altered by an aromatic residue extending from an  $\alpha$ -helix segment of the  $L_C^{7.8}$  linker, while in a second APC monomer the partially exposed  $\beta 84$  chromophore binding pocket is covered by an  $L_C^{7.8}$  loop segment that contains a concentration of positively charged residues. The amino acid sequences of both  $L_R^{32.3}$  and  $L_{RC}^{28.5}$  were examined for analogous sites of interaction. Both possess predicted loop segments containing multiple arginine residues near their C-terminal regions that are well conserved across species, as well as especially promising aromatic candidates identified in a predicted  $\alpha$ -helix segment of  $L_{RC}^{28.5}$ . The majority of evolutionarily conserved residues reside in the N-terminal sequences of each of these two linkers, especially the aromatic and charged residues, indicating a more specific function for this segment than for the remainder of the protein. The crystal structures of the APC and PC trimers were also compared and the  $\alpha$ -helix segment that shields the  $\beta 81/\beta 84$  chromophore from the central cavity appears to be well conserved across species and biliproteins, so that it is nearly identical in amino acid composition for APC and PC (Figure 3-8). The conformation of a number amino acid

side chains in this segment extend towards the central cavity; some of these are visibly altered by the  $L_C^{7,8}$  linker and are identified in Figure 3-6.

To summarize, Figure 5-1 shows the proposed model of biliprotein-linker assembly. The linker is expected to be fully inserted in the central cavities of two biliprotein hexamer stacks. Each linker has two functional segments; one that affects the biliprotein spectroscopy and therefore penetrates at least 12 Å (distance from surface to  $\beta 84$ ), and a second that provides structural support to the rods and involves  $\sim 1/4$  of the total peptide sequence (stipulated from the trypsin digestion studies<sup>6</sup>). The linkers may bind to each other via the biliprotein central cavity, but this is unlikely to be a strong interaction because it is relatively easy to isolate PC trimer/linker complexes. The linkers are expected to influence the planarity of the PC trimer structure as well as specifically interact with the PC chromophore conformation or electrostatic environment in a manner analogous to the  $APC/L_C^{7,8}$  structure.

### **Future Directions**

There are many unresolved issues about the linker influence on the chromophore spectra, and a number of future experiments are immediately suggested by the results presented in this study. Now that a biliprotein/linker structure is known, it provides a wealth of potential targets for genetic manipulation. Alteration or deletion of specific residues in the  $L_C^{7,8}$  sequence should help to differentiate the spectroscopic effects of the two types of linker-chromophore interaction found in the  $APC/L_C^{7,8}$  structure (summarized in Table 3-4). The absorbance spectra of PC chromophores can be influenced by both chromophore conformation<sup>8</sup> and electrostatic environment of the

chromophore propionate side chains,<sup>9</sup> and this kind of genetic alteration aids in determining the magnitude of these two effects. Similar manipulations of the *Synechococcus* sp. PCC 7002 linker peptides based on the aromatic and electrostatically charged amino acid residues identified in Chapter 3 will reveal whether these linker peptides are acting analogously to  $L_C^{7.8}$ . Genetic mutations of conserved aromatic amino acids in  $L_{RC}^{28.5}$  should be particularly easy to differentiate spectroscopically if the PC modulation is produced from a specific aromatic interaction between the linker and the  $\beta 84$  chromophore D ring.

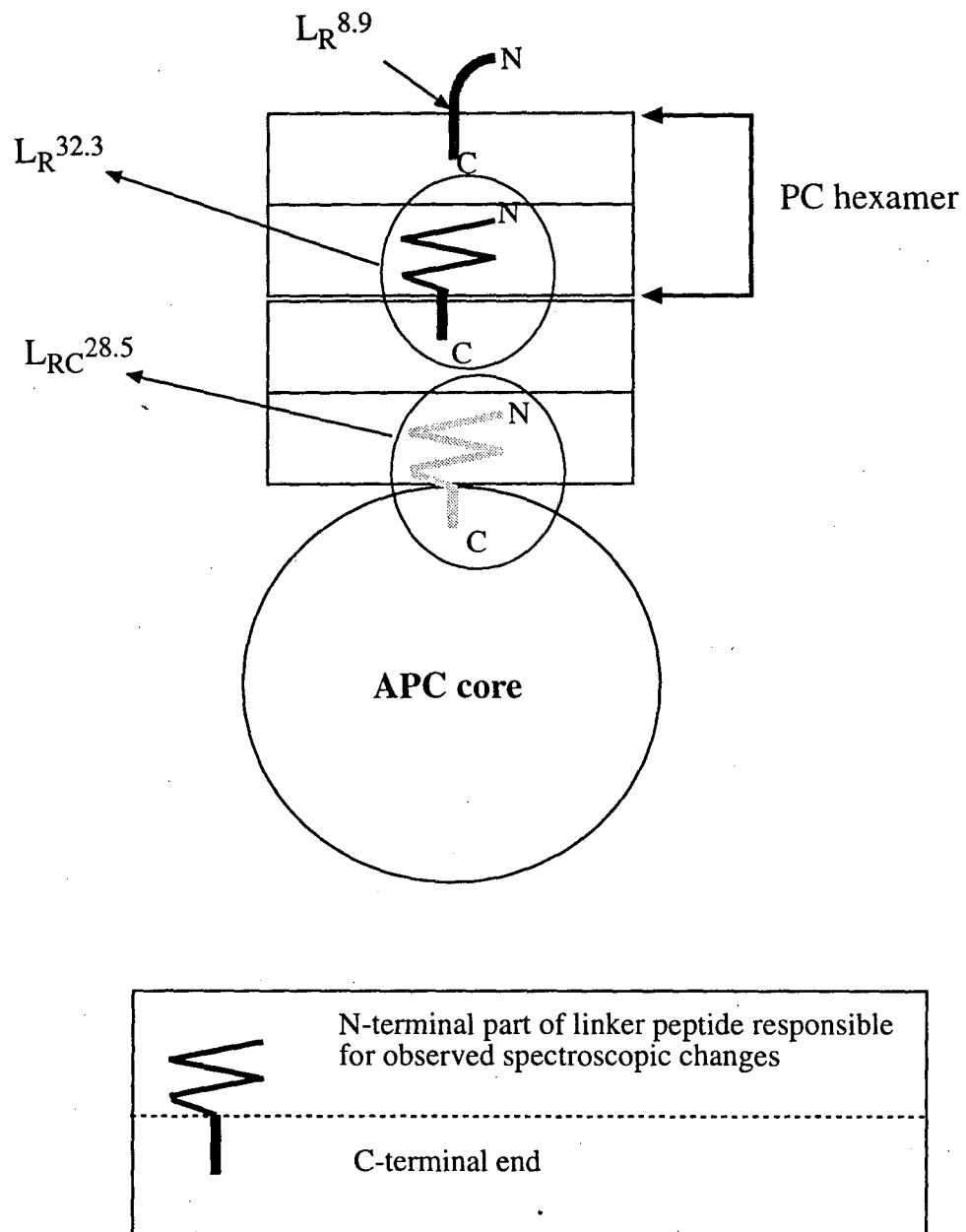
Although the steady-state spectroscopy of these PC/linker genetic mutants would be helpful in identifying the type of linker-chromophore interaction that modulates the PC trimer spectral properties, time-resolved measurements would be more valuable in determining the mechanism of interaction. Time-resolved fluorescence measurements of PC/linker complexes would elucidate the effects of linker association on the  $\alpha 84$ – $\beta 84$  relationship established in the PC trimer structure. If the  $L_{RC}^{28.5}$  linker is selectively disrupting one of the  $\alpha 84$ – $\beta 84$  pairs to guide the energy through one single emitter, then it should change the observed rates of energy transfer in the PC trimer. Debreczeny *et al.*<sup>10</sup> used fluorescence upconversion to measure the rate constants for energy transfer in the *Synechococcus* sp. PCC 7002 PC trimer aggregate; transfer between neighboring  $\alpha 84 \leftrightarrow \beta 84$  chromophores occurred in  $\sim 1$  ps, while energy transfer between the three centrally located  $\beta 84$  chromophores took place in 40 ps. The  $L_C^{7.8}$  linker brings the central  $\beta 84$  chromophores closer together in the APC trimer; according to Förster theory, this shorter distance will accelerate the (40 ps)  $\beta 84 \leftrightarrow \beta 84$  rate of energy transfer. The  $L_C^{7.8}$  linker also specifically affects the conformational geometry of one of the  $\beta 84$

chromophores, and this can also have an effect on the transfer rate constant; if the resulting conformation has a more favorable geometry for energy transfer with another chromophore, then it will result in a faster rate between them. An instrument with femtosecond resolution would be necessary to differentiate between the three different types of  $\alpha 84 \leftrightarrow \beta 84$  interactions observed in APC/ $L_C$ <sup>7,8</sup>. If the linker is inducing a strong excitonic coupling interaction in one of the PC  $\alpha 84 - \beta 84$  pairs, then this should result in a significantly faster observed rate constant than from induced resonance (Förster) energy transfer. In addition, time-resolved fluorescence measurements of the genetic mutants suggested above could identify which type of linker-chromophore interaction (perturbation from aromatic or electrostatically charged residues) has a greater influence on the transfer rate constants.

**References**

1. Schirmer, T., Bode, W., Huber, R. (1987) *J. Mol. Biol.* 196:677-695.
2. De Lorimier R., Bryant D.A., Stevens Jr., S.E. (1990) *Biochim. Biophys. Acta* 1019:29-41.
3. De Lorimier R., Guglielmi G., Bryant D.A., Stevens Jr., S.E. (1990) *Arch. Microbiol.* 153:541-549.
4. Bryant D.A., De Lorimier R., Guglielmi G., Stevens Jr., S.E. (1990) *Arch. Microbiol.* 153:550-560.
5. Reuter, W., Wiegand, G., Huber, R., Than, M.E. (1999) *Proc. Nat. Acad. Sci. USA*, 96:1363-1368.
6. Yu, M.-H., Glazer, A.N. (1982) *J. Biol. Chem.* 257:3429-3433.
7. Yu, M.-H., Glazer, A.N., Williams, R.C. (1981) *J. Biol. Chem.* 256:13130-13136.
8. MacColl, R., Guard-Friar, D. (1987) Phycobiliproteins, CRC Press, Boca Raton, FL.
9. Scharnagl, C., Schneider, S. (1991) *J. Photochem. Photobiol. B: Biol.* 8:129
10. Debreczeny, M.P., Sauer, K.H., Zhou, J., Bryant, D.A. (1995) *J. Phys. Chem.* 99:8420-8431; Debreczeny, M.P. (1994) Ph. D. Thesis, LBL-35672, University of California, Berkeley, CA.

**Figure 5-1.** Proposed model for biliprotein-linker assembly in *Synechococcus* sp. PCC 7002 (vertical cross section of one phycobilisome rod)





ERNEST ORLANDO LAWRENCE BERKELEY NATIONAL LABORATORY  
ONE CYCLOTRON ROAD | BERKELEY, CALIFORNIA 94720