Entwicklung von Hybridmethoden zur Berechnung der Schwingungsspektren von Flavinchromophoren anhand von BLUF Blaulichtrezeptoren

Benjamin Raymond Rieff

München 2011

Entwicklung von Hybridmethoden zur Berechnung der Schwingungsspektren von Flavinchromophoren anhand von BLUF Blaulichtrezeptoren

Benjamin Raymond Rieff

Dissertation
an der Fakultät für Physik
der Ludwig-Maximilians-Universität
München

vorgelegt von
Benjamin Raymond Rieff
aus München

München, 23. August 2011

Erstgutachter: Prof. Dr. Paul Tavan
Zweitgutachter: Prof. Dr. Christian Ochsenfeld
Tag der mündlichen Prüfung: 9. November 2011

für Nini

Zusammenfassung

BLUF (Blue Light sensing Using Flavin) Domänen sind Blaulichtphotorezeptoren, die ein Flavinmolekül als Chromophor in ihrer Bindungstasche enthalten. Durch Photoabsorption am Flavin wird ein Photozyklus angestoßen, der durch eine Konformationsänderung des Proteins den anfänglichen Dunkelzustand in einen Signalzustand überführt. Dabei auftretende Wechselwirkungen zwischen dem Chromophor und dem Apoprotein können vermittels Schwingungsspektroskopie charakterisiert werden. Die resultierenden Schwingungsspektren enthalten strukturelle Informationen in kodierter Form. Computerbasierte Hybridmethoden, die eine dichtefunktionaltheoretische (DFT) Beschreibung eines Chromophors mit einer molekül-mechanischen (MM) Beschreibung seiner Proteinumgebung kombinieren, sollten in der Lage sein, diese Informationen zu entschlüsseln.

Diese Arbeit ist einer genauen Beschreibung der Schwingungsspektren von Flavinen im Allgemeinen und speziell in BLUF Domänen gewidmet. Aus den dazu durchgeführten DFT/MM Simulationsstudien sind drei Publikationen entstanden (1-3), die das Kernstück dieser kumulativen Dissertation darstellen.

Da Schwingungsspektren von Flavinen in der Gasphase experimentell nicht zugänglich sind, befasst sich der erste Artikel mit den Schwingungsspektren von oxidierten Flavinen (Fl) in wässriger Lösung. Es wurde eine DFT/MM Methode entwickelt, die es erlaubt, die Schwingungsfrequenzen im mittleren Infrarotbereich von $1500\,\mathrm{cm}^{-1}$ bis $1800\,\mathrm{cm}^{-1}$ mit einer Genauigkeit von $10\,\mathrm{cm}^{-1}$ zu reproduzieren und vermittels einer entwickelten Routine zur Aufklärung der Normalmodenzusammensetzung detailliert zu analysieren.

In der zweiten Publikation wurde die DFT/MM Methode auf unterschiedliche Redoxzustände von Flavinen in Wasser angewendet. Hier zeigte sich, dass die berechneten Schwingungsspektren für ein anionisches Flavinradikal $\mathrm{Fl}^{\bullet-}$ und ein neutrales Flavinradikal $\mathrm{FlH}^{\bullet}$ die experimentellen Resultate im betrachteten Frequenzbereich mit derselben Genauigkeit wie für ein oxidiertes Flavin vorhersagen können.

Der dritte Artikel beschäftigt sich anhand sechs verschiedener Kristallstrukturen mit den Schwingungsspektren von Flavinen in unterschiedlichen BLUF Domänen. Eine dieser Strukturen erwies sich auf Grund elektrostatischer Wechselwirkungen als dynamisch instabil. Für zwei weitere Strukturen konnte bewiesen werden, dass die ermittelte Kristallstruktur auf Grund der resultierenden DFT/MM Schwingungsspektren nicht mit der solvatisierten Struktur, bei der die experimentellen Schwingungsspektren gemessen wurden, übereinstimmen kann. Nur drei der sechs untersuchten Kristallstrukturen ermöglichten somit eine Vorhersage und Analyse der Schwingungsspektren. Ferner zeigte sich auch eine Kernspinresonanzstruktur in Bezug auf die resultierenden Schwingungsspektren als offensichtlich zu ungenau bestimmt.

Methodisch zeigte sich, dass ein übliches MM Kraftfeld (hier CHARMM22) keine adäquate Berechnung der Schwingungsspektren erlaubt. Erst mit der Einführung ei-

nes polarisierten MM Kraftfeldes für die BLUF Bindungstasche konnten die berechneten DFT/MM Frequenzen die experimentellen Bandenlagen mit großer Genauigkeit vorhersagen.

Verzeichnis der während dieser Arbeit entstandenen Publikationen

(1) B. Rieff, G. Mathias, S. Bauer, and P. Tavan. Density Functional Theory Combined with Molecular Mechanics: The Infrared Spectra of Flavin in Solution. *Photochem. Photobiol.* **87**, 511–523 (2011).

(2) B. Rieff, S. Bauer, G. Mathias, and P. Tavan. IR spectra of flavins in solution: DFT/MM description of redox effects. *J. Phys. Chem. B* **115**, 2117–2123 (2011).

(3) B. Rieff, S. Bauer, G. Mathias, and P. Tavan. DFT/MM description of flavin IR spectra in BLUF domains. submitted to *J. Phys. Chem. B* (2011).

Die angegebenen Arbeiten sind in den Text dieser Dissertation eingearbeitet und dort abgedruckt. Zusätzliche Ergebnisse, die sogenannte „Supporting Information“, die zusammen mit den angegebenen Publikationen veröffentlicht wurden, sind ebenfalls in den Text eingebunden.

Inhaltsverzeichnis

1 Einleitung

Und Gott sprach: Es werde Licht!
Und es ward Licht.
Und Gott sah, dass das Licht gut war.
Da schied Gott das Licht von der Finsternis
und nannte das Licht Tag und die Finsternis Nacht.
Da ward aus Abend und Morgen der erste Tag.

(Genesis, 3-5, Die Bibel)

In den Biowissenschaften stehen Proteine und deren biochemische Reaktionen in wässriger Umgebung schon seit langem im Fokus verschiedenster Untersuchungen. In der Struktur und Funktion von biologischen Zellen zeigt sich die große Bedeutung der Proteine, da sie als elementare Bestandteile in jeder Zelle zu finden sind. Es gibt etwa 10^{10} bis 10^{12} verschiedene natürlich vorkommende Varianten von Proteinen, welche jeweils für eine spezielle biologische Funktion verantwortlich sind [1–4]. So wird die Weitergabe der in der DNS kodierten genetischen Erbinformation über die Steuerung der Transkription und Translation durch Proteine geregelt. Als Bausteine der Haut und der Knochen tragen Proteine zur mechanischen Stabilisierung des Körpers bei. Ferner fungieren sie bei der Immunabwehr als Antikörper oder ermöglichen als Rezeptoren eine Reaktion auf äußere Umweltbedingungen. Dies sind nur einige wenige Beispiele für die Vielfältigkeit von Proteinfunktionen [1–4]. Trotz größter Anstrengungen in der jüngeren Vergangenheit sind jedoch bisher zahlreiche Aspekte der Proteinfunktionen sowohl experimentell als auch theoretisch nicht verstanden.

1.1 Struktur und Dynamik von Proteinen

Auf Grund evolutionärer Prozesse sind Proteine hochoptimierte molekulare Maschinen. Dabei sind sie lediglich lineare Polymere, die in den Zellen aus einer kleinen Zahl chemischer Bausteine, den 20 verschiedenen natürlich vorkommenden Aminosäuren, gebildet werden. [1] Berücksichtigt man die Vielzahl der kombinatorischen Möglichkeiten, mit der sich diese Bausteine dreidimensional zusammensetzen lassen, so wird deutlich, warum die verschiedenen Funktionen der Proteine so mannigfaltig sind.

Um den Aufbau von Proteinen zu verstehen, betrachten wir zunächst deren chemische Bestandteile. Die sogenannten α-Aminosäuren besitzten alle ein zentrales C_α-Atom,

[1] Im Anhang A sind die 20 verschiedenen Aminosäuren sowie die in Biochemie und Biophysik übliche dreibuchstabige und einbuchstabige Kurzbezeichnung angegeben.

an das vier weitere Gruppen kovalent gebunden sind. Neben einem einzelnen Wasserstoffatom (H) sind dies eine Carboxylgruppe (COO^-), eine Aminogruppe (NH_3^+) sowie eine charakteristische Seitengruppe (R), anhand derer sich die einzelnen Aminosäuren unterscheiden lassen. Da die verschiedenen Seitengruppen auch dementsprechend unterschiedliche chemische Eigenschaften aufweisen, können die Aminosäuren beispielsweise als polar, apolar, sauer oder basisch klassifiziert werden [1, 2].

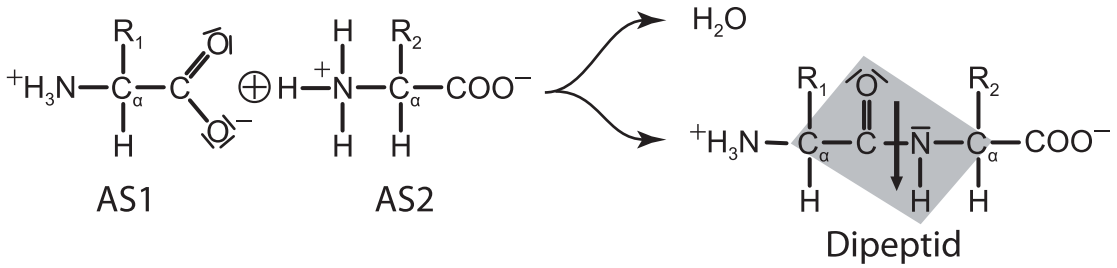


Abbildung 1.1: Synthese von Aminosäuren. Bei der Peptidsynthese werden zwei Aminosäuren (AS1 und AS2) unter Abspaltung eines Wassermoleküls (H_2O) zu einem Dipeptid verbunden. Im Dipeptid bildet sich ein torsionsstabiles Peptidplättchen (grau hinterlegt) aus, welches ein ausgeprägtes statisches Dipolmoment (schwarzer Pfeil) besitzt.

Abbildung 1.1 zeigt schematisch den Vorgang der Peptidsynthese, bei der zwei einzelne Aminosäuren (AS1 und AS2) an ihren beiden Enden (COO^- und NH_3^+) miteinander verbunden werden. Dabei entsteht unter Abspaltung eines Wassermoleküls (H_2O) ein sogenanntes Dipeptid, dessen Teile durch ein torsionsstabiles Peptidplättchen (grau hinterlegt) verbunden sind. Auf Grund der großen Elektronegativität des Stickstoffs (N) und des Sauerstoffs (O) ziehen diese beiden Atome die Elektronen des benachbarten Wasserstoff- bzw. Kohlenstoffatoms an sich heran. Durch diese Ladungsverschiebung entsteht ein ausgeprägtes statisches Dipolmoment (schwarzer Pfeil), welches für das Peptidplättchen charakteristisch ist.

Indem der Vorgang der Peptidsynthese stetig wiederholt wird, werden im Ribosom einer Zelle typischerweise einige 100 bis 1000 Aminosäuren zu einem Protein polymerisiert [1, 2]. Die so erzeugte Aminosäurenkette wird die Primärstruktur des Proteins genannt. Sie ist für jedes Protein eindeutig bestimmt und im genetischen Kode der DNS verschlüsselt. Abhängig von ihrer Primärstruktur bilden Proteine unter natürlichen Umgebungsbedingungen nach der Synthese ihre charakteristische dreidimensionale Struktur aus.

Bei der Bildung der nativen Proteinstruktur entstehen sogenannte Sekundärstrukturmotive, die bei fast jedem Protein zu beobachten sind. In Abbildung 1.2 sind die beiden wichtigsten Sekundärstrukturmotive, die α-Helix und das β-Faltblatt schematisch dargestellt. Wie hier illustriert wird, werden beide Sekundärstrukturmotive durch die elektrostatischen Wechselwirkungen zwischen den Dipolen der einzelnen Peptidplättchen stabilisiert. Bei der spiralförmig gewundenen α-Helix 1.2(a) wird diese Stabilisierung durch parallel ausgerichtete Dipole des i-ten und (i+4)-ten Peptidplättchens erzielt. Das β-Faltblatt 1.2(b) mit seinen parallel nebeneinander liegenden Strängen zeigt bei den Dipolen längs

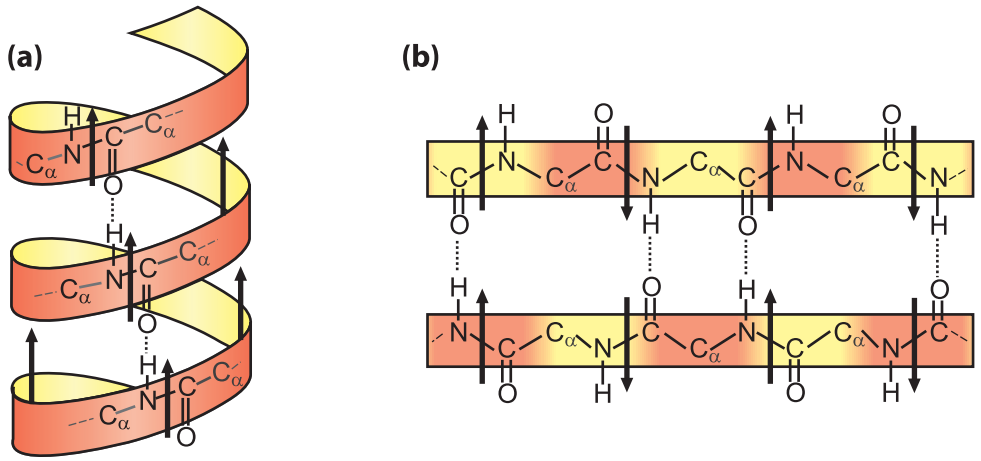


Abbildung 1.2: Sekundärstrukturmotive. Die spiralförmig gewundene α-Helix (a) wird durch die elektrostatischen Wechselwirkungen zwischen den parallel ausgerichteten Dipolen (schwarze Pfeile) benachbarter Peptidplättchen stabilisiert. Bei den parallel nebeneinander liegenden Strängen des β-Faltblatts (b) wird diese Stabilisierung durch kollinear-parallel ausgerichteten Dipole verschiedener Peptidstränge und die axial-antiparallel ausgerichteten Dipole im Strang benachbarter Peptidgruppen erzielt.

der einzelnen Stränge eine antiparallele und zwischen den Strängen die energetisch günstigere parallele Ausrichtung. Sowohl die α-Helix als auch das β-Faltblatt sind relativ rigide Strukturen und tragen somit stark zur definierten Struktur im nativen Zustand bei. Diejenigen Bereiche des Proteins, die eine hohe Flexibilität zeigen und oftmals als offene Schleifen bezeichnet werden, ermöglichen dagegen dynamische Funktionen des Proteins. Der gesamte Prozess bis hin zur Ausbildung der nativen Proteinstruktur wird Proteinfaltung genannt und ist bis heute Gegenstand aktueller Forschung [5–8].

Die sich aus der Proteinfaltung ergebende dreidimensionale Struktur, die sogenannte Tertiärstruktur eines Proteins, und die daraus abgeleitete Beweglichkeit sind nun für die spezifische Funktion der einzelnen Proteine verantwortlich [9–12]. Auch wenn die physiologische Struktur durch die Sequenz der Aminosäurenkette festgelegt ist, so ist man bis heute nicht in der Lage, die Tertiärstruktur oder die damit einhergehende Funktionalität eines gefalteten Proteins aus dieser Aminosäurenkette mit atomarer Auflösung vorherzusagen. Um solche Informationen gewinnen zu können, müssen Proteine in ihrer nativ gefalteten Struktur untersucht werden. Da häufig nur wenige 10 Atome in den eigentlich biologisch aktiven Prozess involviert sind, müssen sowohl experimentelle als auch theoretische Methoden herangezogen werden, um die biologisch aktiven Zentren eines Proteins auf atomarer Ebene zu identifizieren. Bevor jedoch die entsprechenden Methoden zur Untersuchung von Proteinen erläutert werden, wird zunächst ein Blick auf die dynamischen Prozesse der Proteinfaltung geworfen.

Dynamische Prozesse der Proteinfaltung. Um die einzelnen dynamischen Bewegungen und Prozesse, die sich aus dem eben diskutierten Aufbau der Proteine ergeben, zu skizzieren, wird im Folgenden ein Überblick gegeben, der sich angefangen bei den schnellen Bindungsschwingungen bis hin zur Bildung der Sekundärstrukturmotive erstreckt.

	Prozess/Bewegungen	Zeitskala
(i)	Bindungs- und Winkelschwingungen	10 fs ··· 100 fs
(ii)	Bewegungen von Peptid- und Seitengruppen	10 ps ··· 1 ns
(iii)	Kollektive Bewegungen	10 ns ··· 1 ms
(iv)	Bildung von α-Helizes	10 ns ··· 100 ns
(v)	Aufbau von β-Faltblättern	1 μs ··· 10 μs
(vi)	Proteinfaltung m Ganzen	4 μs ··· 1000 s

Tabelle 1.1: Zeitskalen der Proteinfaltung. Die einzelnen Dynamiken umfassen atomare Bewegungen im Femtosekundenbereich bis hin zu Prozessabläufen von einigen Mikrosekunden. Insgesamt erstrecken sich die Vorgänge der Proteinfaltung über einen Zeitbereich von etwa 17 Größenordnungen [13–16].

In Tabelle 1.1 sind die wichtigsten dynamischen Zeitskalen aufgelistet, die folgende in Proteinen stattfindenden Prozesse umfassen:

(i) Kleinskalige Schwingungen und thermische Fluktuationen von einzelnen Atomen oder kleineren Gruppen um ihre gegebenen Gleichgewichtslagen, die durch lokale Bindungen definiert sind. Diese Schwingungen stellen die schnellsten Bewegungen in einem Protein dar und laufen innerhalb einiger Femtosekunden (fs, 10^{-15} s) ab [13].

(ii) Rotationen ganzer Seitengruppen und das Umklappen einzelner Peptidplättchen sind Bewegungen, die auf Zeitskalen von einigen Pikosekunden (ps, 10^{-12} s) bis zu einer Nanosekunde (ns, 10^{-9} s) vonstatten gehen [13].

(iii) Kollektive Bewegungen der Sekundärstrukturmotive, die zum Beispiel mit scharnierartigen Deformationen der Peptidkette verbunden sein können, erstrecken sich von 10 ns bis hin zu Millisekunden (ms, 10^{-3} s) [13, 14].

(iv) Die Ausbildung der helikalen Sekundärstrukturelemente gehen im ns-Bereich vonstatten, wobei schnellere Prozesse auf der ps-Zeitskala diese Abläufe beeinflussen können [15].

(v) Die Formierung eines β-Faltblattes, wobei als Vorstufe zwei Ketten zu einem sogenannten Hairpin gewunden werden, benötigt auf Grund entropischer Barrieren mit einer Zeitkonstante im Bereich einiger Mikrosekunden (μs, 10^{-6} s) relativ lange [15].

(vi) Der gesamte Faltungsvorgang, bei dem ein Protein von seinem anfänglich ungeordneten Zustand zu seiner funktionellen nativen Strukur gefaltet wird, kann sogar einige Minuten beanspruchen. Momentan gelten kleine Proteine, die innerhalb von 4 μs ihre native Form annehmen als die schnellsten, wohingegen zum Beispiel die Faltung großer Eiweißkomplexe deutlich langsamer abläuft [13, 16].

Neben den verschiedenen Zeitskalen unterscheiden sich die dargestellten dynamischen Prozesse auch in Bezug auf die Größenordnung der einhergehenden Geometrieänderungen eines Proteins. So durchläuft ein Protein sehr rasch viele geometrisch ähnliche Strukturen, wenn lediglich die schnellen Freiheitsgrade (i) fluktuieren. Im Gegensatz dazu können die anderen Prozesse zu Änderungen ganzer Strukturbereiche des Poteins führen, welche dann in unterschiedlichen sogenannten Konformationen eines Proteins resultieren. Dabei bezeichnet der Begriff Konformation einen metastabilen Zustand eines Proteins, in welchem es eine gewisse Zeitspanne verweilt und thermisch fluktuiert, bevor es dann wiederum relativ spontan in einen anderen, meist auch geometrisch deutlich unterschiedlichen Zustand übergeht. Eine Konformation entspricht daher einem lokalen Minimum der Freien Energie, das von anderen Konformationen durch Barrieren innerhalb der Freien Energielandschaft abgetrennt ist [17, 18]. Während der Ausübung seiner Funktion wechselt ein Protein oftmals zwischen verschiedenen Konformationen, um die einzelnen funktionalen biochemischen Prozesse zu ermöglichen [1, 13]. Um die strukturelle Beschaffenheit sowie die dynamischen Prozesse eines Proteins zu identifizieren und zu verstehen, bedient man sich mehrerer Untersuchungsmethoden, die im nun folgenden Abschnitt erläutert werden.

1.2 Experimentelle Methoden zur Untersuchung von Biomolekülen

Zu den wichtigsten biophysikalischen Methoden zur Untersuchung von Biomolekülen zählen die Röntgenstrukturanalyse [12, 19], die Kernspinresonanz (NMR, von engl.: nuclear magnetic resonance) [20, 21] und die kurzzeit-spektroskopischen Experimente mit Lasern [15]. Diese Methoden werden eingesetzt, um die Struktur, die Form und die Funktionalität von Peptiden und Proteinen zu bestimmen. Diejenigen experimentellen Untersuchungsmethoden, die für die anfängliche Datenbasis dieser Arbeit von entscheidender Bedeutung sind, sollen nun kurz vorgestellt werden.

1.2.1 Zur Strukturaufklärung

Die Röntgenstrukturanalyse von Kristallstrukturen stellt die verbreitetste Methode zur Analyse von biomolekularen Strukturen dar. Anhand der Beugungsmuster gestreuter Röntgenstrahlung an Einkristallen des zu untersuchenden Biomoleküls wird dabei die Anordnung der Atome dreidimensional ermittelt [19, 22].

Allerdings sind dieser Analysemethode entscheidende Grenzen auferlegt, die nun kurz skizziert werden. Um eine Röntgenstrukturanalyse zu ermöglichen, muss das zu untersu-

chende Biomolekül aus seiner natürlichen Umgebung entnommen werden, damit es kristallisiert und anschließend bei tiefen Temperaturen bestrahlt werden kann [23, 24]. Dabei ist jedoch unklar, ob die Struktur auf Grund der im Kristall veränderten Umgebungsbedingungen unverändert bleibt, das heißt, inwieweit also die Röntgenstruktur dann tatsächlich auch die native Struktur wiedergibt. Ferner besitzt die Röntgenstrukturanalyse durch die Verwendung von Röntgenstrahlen eine endliche Auflösung, sodass nicht alle Atome eines Biomoleküls im Beugungsmuster erkennbar sind. Gerade die Wasserstoffatome, die einen sehr kleinen Wirkungsquerschnitt für Röntgenbeugung aufweisen, können meist nicht detektiert werden. Daher ermöglicht die Röntgenstrukturanalyse keine Aussage über Protonierungszustände einzelner Gruppen und liefert somit beispielsweise auch keine Informationen über Reaktionsmechanismen von Protontransferprozessen, wie etwa im Fall des Bakteriorhodopsins [25–28]. Neben den Wasserstoffatomen ergeben auch flexiblere Bereiche von Biomolekülen, die meist eine funktionale Bedeutung haben [29], oftmals keine verwertbaren Beugungsmuster. Des Weiteren können dynamische Prozesse, wie enzymatisch katalysierte Reaktionen oder Faltungsprozesse von Proteinen, nicht zeitaufgelöst dargestellt werden, da sich entsprechende Röntgenmethoden noch in der Entwicklungsphase befinden [30].

Eine weitere experimentelle Methode zur Strukturauflösung von Biomolekülen ist die NMR Spektroskopie. Bei dieser Methode wird der räumliche Abstand zwischen zwei Atomen anhand der Dipol-Kopplung ihrer jeweiligen Kernspins ermittelt [20, 21, 31]. Im Gegensatz zur Röntgenstrukturanalyse können hier jedoch die Positionen der Wasserstoffatome bestimmt werden. Die vollständige Zuordnung aller paarweise gemessenen Atomabstände wird mit wachsender Größe des zu untersuchenden Moleküls immer schwieriger und reicht ferner nicht aus, um die dreidimensionale Struktur eines Biomoleküls eindeutig zu charakterisieren [32–34]. Daher werden in einem weiteren Strukturverfeinerungsprozess beispielsweise die Topologie von Aminosäuren und die empirisch bekannten Längen und Winkel chemischer Bindungen herangezogen, um eine oder auch mehrere Strukturen zu generieren, die jeweils gleichzeitig die paarweise gemessenen Atomabstände am besten wiedergeben [35, 36].

Strukturen, die für Biomoleküle mit Hilfe der eben geschilderten Methoden ermittelt wurden, sind in der Protein Daten Bank (PDB) [37] hinterlegt und haben somit einen wichtigen Beitrag zum Verständnis der Form und Funktionalität von Biomolekülen geliefert [7, 12]. Da jedoch weder mit der Röntgenstrukturanalyse noch mit der NMR Spektroskopie alleine eine Ableitung der Struktur und Funktionsweise eines Biomoleküls aus den entsprechenden Daten möglich ist, bedarf es weiterer Untersuchungsmethoden, um eine detaillierte Aufklärung zu erreichen.

1.2.2 Spektroskopische Untersuchungen

Eine Vielzahl spektroskopischer Methoden bedient sich einer Kombination aus einem Laserpuls zum Aktivieren einer Reaktion (Pump-Puls) gefolgt von einem zeitlich verzögerten Laserpuls (Proben-Puls), der zur Untersuchung der angeregten Probe verwendet wird [38]. In Abhängigkeit der Wellenlänge des Pump-Pulses können unterschiedliche Anre-

gungen, zum Beispiel Molekülschwingungen im elektronischen Grundzustand oder elektronische Übergänge, erzielt werden. Zur anschließenden Untersuchung der Probe stehen mehrere spektoskopische Techniken zur Verfügung [38].

Verwendet man zum Beispiel Licht des sichtbaren oder des ultravioletten (UV/vis) Bereichs zur Untersuchung der Probe, so können Absorptionsbanden, die von elektronischen Übergängen hervorgerufen werden, analysiert werden [38]. Ein verändertes Absorptionsspektrum, das auf Grund struktureller Konformationsänderungen, der Anregung von Schwingungen oder auf Grund chemischer Reaktionen entsteht, erlaubt es, die verantwortlichen Prozesse mit einer hohen Zeitauflösung zu beschreiben [39–41]. Allerdings ist eine strukturelle Aufklärung mit dieser Methode nicht möglich.

Schwingungsspektroskopische Untersuchungen, bei denen die Absorption oder die Streuung elektromagnetischer Strahlung an Biomolekülen ermittelt werden (z.B. Infra-Rot (IR) oder Resonanz Raman (RR) Spektroskopie), können allerdings einen entscheidenden Beitrag zur Aufklärung sowohl der Struktur als auch dynamischer Prozesse liefern. Hierbei wird die quantenmechanische Natur der Biomoleküle ausgenutzt, auf Grund derer sich ein diskretes Spektrum von Zuständen verschiedener Energien ergibt, die wiederum für jede molekulare Gruppe ganz charakteristische Eigenschaften aufzeigen [42].

Aus einem experimentell gemessenen Schwingungsspektrum können ferner strukturelle Informationen in verschlüsselter Form nicht nur über das Biomolekül selbst, sondern auch über dessen Umgebung gewonnen werden. Das Schwingungsspektrum eines Biomoleküls ist nämlich unter anderem von seiner Polarisation abhängig, die durch die elektrostatischen Wechselwirkungen mit der Umgebung hervorgerufen wird (Solvatochromieeffekte). So zeigen beispielsweise die beiden schon angesprochenen Sekundärstrukturmotive unterschiedliche Charakteristika im IR Spektrum, wodurch entschlüsselt werden kann, ob eine α-Helix oder ein β-Faltblatt (siehe Abb. 1.2) in der Molekülstruktur vorhanden ist [43–46]. Verwendet man darüberhinaus zeitaufgelöste Methoden, so können aus den Spektren auch Informationen über die Dynamik des betrachteten Biomoleküls gewonnen werden.

Das Problem besteht jedoch darin, die dekodierten strukturellen Informationen aus den gemessenen Spektren zu extrahieren. Da ein nichtlineares N-atomiges Biomolekül $3N - 6$ Schwingungsmoden besitzt [47, 48] und die Frequenzen der einzelnen Moden eines großen Moleküls oftmals sowohl sehr nahe nebeneinander liegen als auch breite Banden, wie beispielsweise die sogenannten Amidbanden [43–45], beinhalten, erhält man experimentell ein quasi-kontinuierliches Spektrum.

Die Information über das biologisch aktive Zentrum eines Proteins ist somit meist im Molekülspektrum versteckt, da die mitgemessenen starken Banden der Amidschwingungen, der verschiedenen Seitengruppen und des Lösungsmittels, hier ist vor allem das Wasser zu nennen, die spektral interessanten Bereiche überlagern. Um die Überlagerung zu vermeiden, wird die sogenannte Differenzspektroskopie herangezogen, welche insbesondere in der IR Spektroskopie häufig angewendet wird.

Bei dieser Untersuchungsmethode werden zwei separate Spektren zu unterschiedlichen Zuständen eines Biomoleküls subtrahiert. Durch die Subtraktion werden diejenigen Bereiche, die sich in den verschiedenen Zuständen des Biomoleküls nicht unterscheiden,

aus dem Spektrum eleminiert und im resultierenden Spektrum zeigen sich diejenigen Banden, die sich beim Übergang von dem einen in den anderen Zustand verändert haben. Um unterschiedliche Zustände eines Biomoleküls zu realisieren, werden beispielsweise Isotopenmarkierungen, Mutationen (ein Austauch einer bestimmten Aminosäure durch eine andere) oder chemische Reaktionen verwendet. Je nach Beschaffenheit der Modifikation werden zum Beispiel im Protein eingebettete Kofaktoren [46, 49–53], Teilbereiche eines Biomoleküls [45, 46] oder Reaktanden der vom Biomolekül katalysierten Reaktion [54] untersucht. Die Abhängigkeit der Spektren der betrachteten Bereiche von ihrem chemischen Zustand oder der Struktur ihrer Umgebung wird dann verwendet, um strukturelle Informationen aus den entsprechenden Spektren zu erhalten [46, 52, 55, 56].

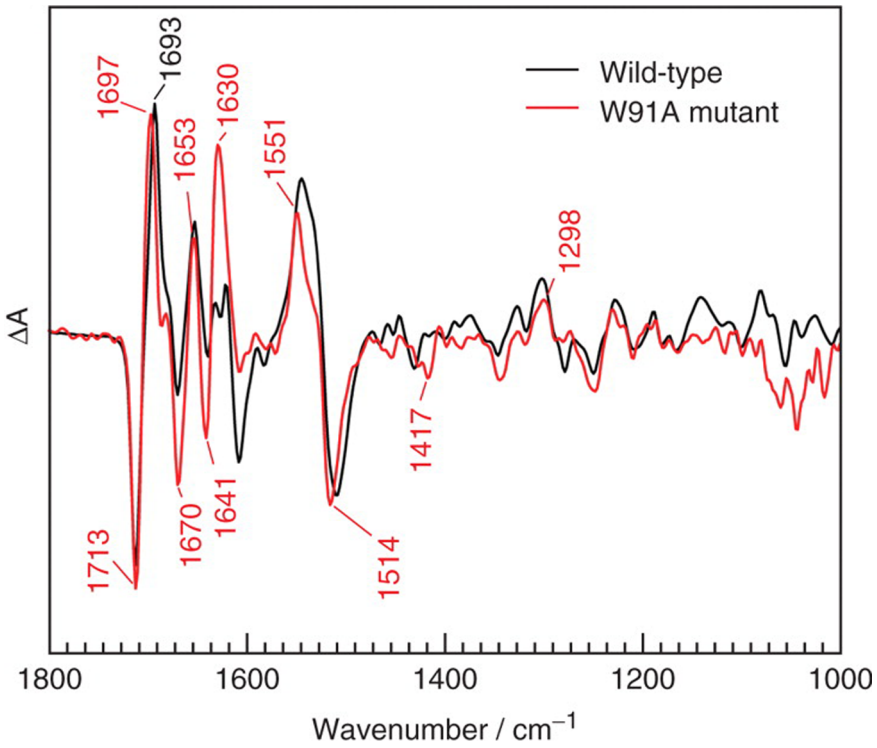


Abbildung 1.3: Differnzspektrum von Slr1694 BLUF. Licht-minus-Dunkel FTIR Differenzspektrum des Slr 1694 BLUF Proteins aus dem Cyanobakterium *Synechocystis* sp. PC6803 (schwarze Linie) sowie eines Slr1694 Mutanten (rote Linie), bei dem das Trp91 durch ein Alanin ersetzt wurde (entnommen aus [57]).

Abbildung 1.3 zeigt ein Beispiel für ein IR Differenzspektrum. Bei diesem Spektrum wurde das ermittelte Fourier-transformierte IR (FTIR) Spektrum vom Dunkelzustand des Slr1694 BLUF Proteins aus dem Cyanobakterium *Synechocystis* sp. PC6803 vom entsprechenden Spektrum des Lichtzustandes abgezogen (schwarze Linie) [57]. Auf Grund unterschiedlicher Schwingungsfrequenzen in den beiden Zuständen können nun Rückschlüsse auf Konformationsänderungen gezogen werden [57]. Weitere Informationen ergeben sich zum Beispiel auch aus dem Vergleich mit einem mutierten Protein, wobei hier die Amino-

säure Trp91 des Slr1694 Proteins durch ein Alanin ersetzt wurde (rote Linie). Im späteren Verlauf dieser Einleitung wird auf dieses Spektrum noch näher eingegangen werden.

Durch den Einsatz von zeitaufgelöster Spektroskopie ist es ferner möglich herauszufinden, wie lange ein beobachteter Prozess benötigt, um vom Eduktzustand in den Produktzustand überzugehen [58]. Des Weiteren können intermediäre Konformationen, das sind Konformationen, die zwischen der anfänglichen und finalen Konformation eingenommen werden, ermittelt werden, sofern sich deren Spektren von denen des Edukts und des Produkts unterscheiden. Allerdings ist dabei eine Aussage über die Struktur der Intermediate allein auf Grund der spektralen Änderungen meist nicht möglich. Hier kommen nun die theoretischen Methoden ins Spiel, die herangezogen werden müssen, um die experimentell gewonnenen Daten zu analysieren und zu verstehen. So können zum Beispiel berechnete Spektren mit experimentell gemessenen Spektren verglichen werden, wodurch man in der Lage ist, gewisse Konformationen eines Biomoleküls bestimmten Spektren zuzuordnen, sofern die experimentellen und berechneten Spektren hinreichend genau bestimmt sind. Daher wird nun ein Blick auf die theoretischen Methoden geworfen, wobei die Dichtefunktionaltheorie (DFT) [59, 60] und die Molekulardynamik (MD) Simulationen [61] im Vordergrund stehen.

1.3 Theoretische Methoden

Da die in dieser Dissertation vorgestellten biophysikalischen Simulationen vornehmlich anwendungsorientiert sind, werden die verwendeten Methoden hier nur kurz skizziert. Für detaillierte Darstellungen der verwendeten Methoden sei daher auf die angeführten Literaturzitate verwiesen [61–67].

Um Proteine näherungsweise zu beschreiben, verwendet man sogenannte molekülmechanische (MM) Kraftfelder $V_K(\mathbf{r}_K)$, die von den Kernkoordinaten $\mathbf{r}_K$ aller Atome eines Biomoleküls und seiner Lösungsmittelumgebung abhängen. Mit Hilfe dieser Kraftfelder ist man in der Lage, experimentelle Röntgen- bzw. NMR-Daten durch Informationen über lokale chemische Strukturen (z.B. Bindungslängen und -winkel) zu erweitern. So können beispielsweise auch diejenigen Atompositionen, die durch eine Röntgenstrukturanalyse nicht eindeutig bestimmt wurden, durch die Verwendung von MM Kraftfeldern im Nachhinein ermittelt werden [68]. Bei der oben schon erwähnten NMR Strukturverfeinerung werden sogenannte „Simulated Annealing“ Verfahren zur Minimierung einer durch Vernachlässigung der Elektrostatik teils vereinfachten und durch Einbeziehung experimenteller Randbedingungen teils erweiterten Potentialfunktion $V_K(\mathbf{r}_K)$ durchgeführt [35, 36]. Die berücksichtigten experimentellen Einschränkungen ermöglichen somit die Einhaltung der NMR generierten, experimentell paarweise gemessenen Atomabstände. Weil dabei die Elektrostatik üblicherweise nicht berücksichtigt wird, können derart berechnete Strukturen größere Fehler aufweisen.

MD Simulationen verwenden dagegen MM Kraftfelder, welche die elektrostatischen Wechselwirkungen näherungsweise berücksichtigen. Durch Integration der resultierenden Newtonschen Bewegungsgleichungen erlauben es MD Simulationen dann auch, ausge-

hend von experimentell ermittelten Strukturen, Informationen über die dynamischen Prozesse von Proteinen auf atomarer Ebene zu gewinnen [69]. Eine erste Anwendung der MD Simulation für Proteine fand im Jahre 1977 am *bovine pancreatic trypsin inhibitor* (BPTI) Protein statt [70]. Nachfolgende Anwendungen haben eine Vielzahl von wertvollen Informationen zur Dynamik und Funktionsweise von Proteinen aufgezeigt. Einige herausragende Beispiele finden sich in den angegebenen Literaturzitaten [8, 34, 71–74].

Selbstverständlich sind auch der MD Simulation Grenzen auferlegt. Bei heutigen MD Simulationsprogrammen [68, 75–77] werden in der Regel Kraftfelder [76–78] verwendet, denen zum Zweck einer möglichst schnellen Berechenbarkeit verschiedenste Näherungsverfahren zu Grunde liegen und die deshalb auch erhebliche Artefakte produzieren können [75, 79, 80]. Insbesondere wird in den üblichen MM Kraftfeldern [76–78] die Polarisierbarkeit der Elektronenhüllen der Atome gänzlich vernachlässigt, sodass deren Reaktion auf äußere elektrische Felder nicht explizit modelliert wird. Weiterhin kann selbst mit modernsten Computern und trotz der eingesetzten Näherungen lediglich eine stark begrenzte Anzahl von Simulationsschritten ausgeführt werden, wenn zum Beispiel ein Protein in realistischer Lösungsmittelumgebung betrachtet wird. Als Folge sind die Zeitskalen von Millisekunden bis Minuten, auf denen die Proteinfaltung vonstatten geht [13–16], nicht zugängig. Es können somit nur schnell ablaufende Prozesse mit MD Simulationen im angemessenen Rahmen untersucht werden. Ein weiterer Nachteil ist, dass beispielsweise chemische Reaktionen oder Schwingungsspektren von Biomolekülen sehr genau berechnet werden müssen, um einen hinreichenden Informationsgewinn zu erhalten. Hierzu reichen jedoch übliche MM Kraftfelder nicht aus. Um die erforderliche Genauigkeit zu erzielen, werden deshalb quantenmechanische Rechnungen eingesetzt, die den Rechenaufwand nochmals drastisch steigern. Mit der DFT sind auf diese Weise Systeme bis zu einer Größe von etwa 100 Atomen handhabbar, was es immerhin erlaubt, die biologisch aktiven Zentren eines Proteins genauer zu untersuchen. Um dabei auch den Einfluss der Umgebung des betrachtenen Bereichs zu beschreiben, wurden sogenannte Hybridmethoden [64, 81] entwickelt, welche die DFT mit einem MM Kraftfeld kombinieren.

Im Rahmen dieser Arbeit wurden sowohl die MD Simulation als auch das eben genannte DFT/MM Verfahren eingesetzt, um eine bestimmte Proteinklasse zu analysieren. Im Abschnitt 1.6 wird auf die dabei auftretenden Fragen und Probleme eingegangen. Zunächst wollen wir uns aber der Klasse der Photorezeptoren zuwenden, zu denen auch die in dieser Arbeit untersuchten Proteine zählen.

1.4 Photorezeptoren

Die Sonne spendet das Licht, welches die Grundlage für das Leben auf der Erde darstellt. Es ist sowohl Energiequelle als auch ökologischer Stimulator, besitzt jedoch auch die Fähigkeit organische Zellen zu beschädigen. Die Aufnahme von Lichtenergie, die Wahrnehmung von Licht sowie die Anpassung an unterschiedliche Lichtverhältnisse sind daher entscheidende Faktoren für Wachstum, zirkadiane Rhythmen und den Sehprozess. In evolutionären Prozessen haben Lebewesen ganz unterschiedliche photosensorische Rezeptoren

hervorgebracht, um auf äußere Lichtverhältnisse zu reagieren. Daher können sie zwischen der Lichtqualität (Wellenlänge), der Lichtquantität (Intensität) und der Einfallsrichtung differenzieren. Dabei sind es photosensorische Proteine, die vermittels unterschiedlicher atomarer Mechanismen die physiologischen Reaktionen auf das Licht steuern [82–84]. Die wohl bekannteste lichtinduzierte Reaktion ist die Photosynthese, bei der Kohlenstoffdioxid in den für atmende Lebewesen notwendigen Sauerstoff umgewandelt wird. Allerdings gibt es eine Vielzahl anderer Reaktionen, deren Wirkungsweise auf eine Entschlüsselung wartet.

Die Aufgabe der einzelnen Photorezeptoren besteht darin, das breite elektromagnetische Spektrum vom ultravioletten bis hin zum infraroten Bereich (150 bis 780 nm Wellenlänge) [85] abzudecken, um dadurch das einfallende Licht wahrnehmen und verarbeiten zu können. Daher unterscheiden sich die einzelnen Rezeptoren nicht nur in ihrem strukturellen Aufbau, sondern bedienen sich auch unterschiedlicher Kofaktoren, welche durch Absorption eines Photons einer ganz bestimmten Wellenlänge angeregt werden. Bis heute sind sechs Photorezeptorfamilien bekannt, welche folgende Mitglieder enthalten: Die Rhodopsine, welche das Retinal als Chromophor verwenden, die Phytochrome, eine Familie der Rotlichtrezeptoren, die alle das Phytochromobilin als Kofaktor besitzen und die Xanthopsine, die sogenannten photoactive yellow proteins, bei denen die Cumarsäure als Farbstoff fungiert [84]. Ferner kennt man drei verschiedene Blaulichtrezeptorklassen, welche alle ein Flavin als Chromophor aufweisen [84]. Da sich diese Arbeit mit einem Blaulichtphotorezeptor befasst, soll zunächst der verwendete Flavinfarbstoff vorgestellt werden, bevor die verschiedenen Blaulichtrezeptorklassen und deren Charakterisierung erläutert werden.

1.4.1 Das Flavinmolekül

Um über die Absorption eines Photons die Wahrnehmung von blauem Licht zu ermöglichen, bedienen sich die Blaulichtphotorezeptoren eines bestimmten Pigments: dem Flavin. Dabei leitet sich der Name vom lateinischen “flavus” (gelb) ab, da eine Lösung bestehend aus oxidierten Flavinen im Blauen absorbiert (Absorptionsmaximum um 450 nm) und dementsprechend eine gelbliche Färbung aufweist [1].

Wie in Abbildung 1.4 dargestellt ist, bildet der sogenannte Isoalloxazinring die Grundstruktur der Flavinmoleküle. In Abhängigkeit der angebundenen Molekülkette ergeben sich daraus die einzelnen Flavinderivate, wobei das Lumiflavin, das Riboflavin (RF), das Flavin-Mononukleotid (FMN) und das Flavin-Adenin-Dinukleotid (FAD) zu den wichtigsten gehören [1].

Im Allgemeinen können Flavine an unterschiedlichen Reaktionen beteiligt sein, da sie in der Lage sind, verschiedene Redox-Zustände einzunehmen und daher mit einer Vielzahl von Edukten reagieren können [1]. Auf Grund ihrer Wasserlöslichkeit werden Flavine zum Beispiel in Bindungstaschen von Proteinen meist nicht-kovalent gebunden. Innerhalb der jeweiligen Bindungstasche erfolgt der Informationsaustausch zwischen dem jeweiligen Flavinderivat und dem bindenden Protein auf atomarer Ebene. Ist ein Flavin als Kofaktor in einer Bindungstasche eines Proteins gebunden, so spricht man üblicher-

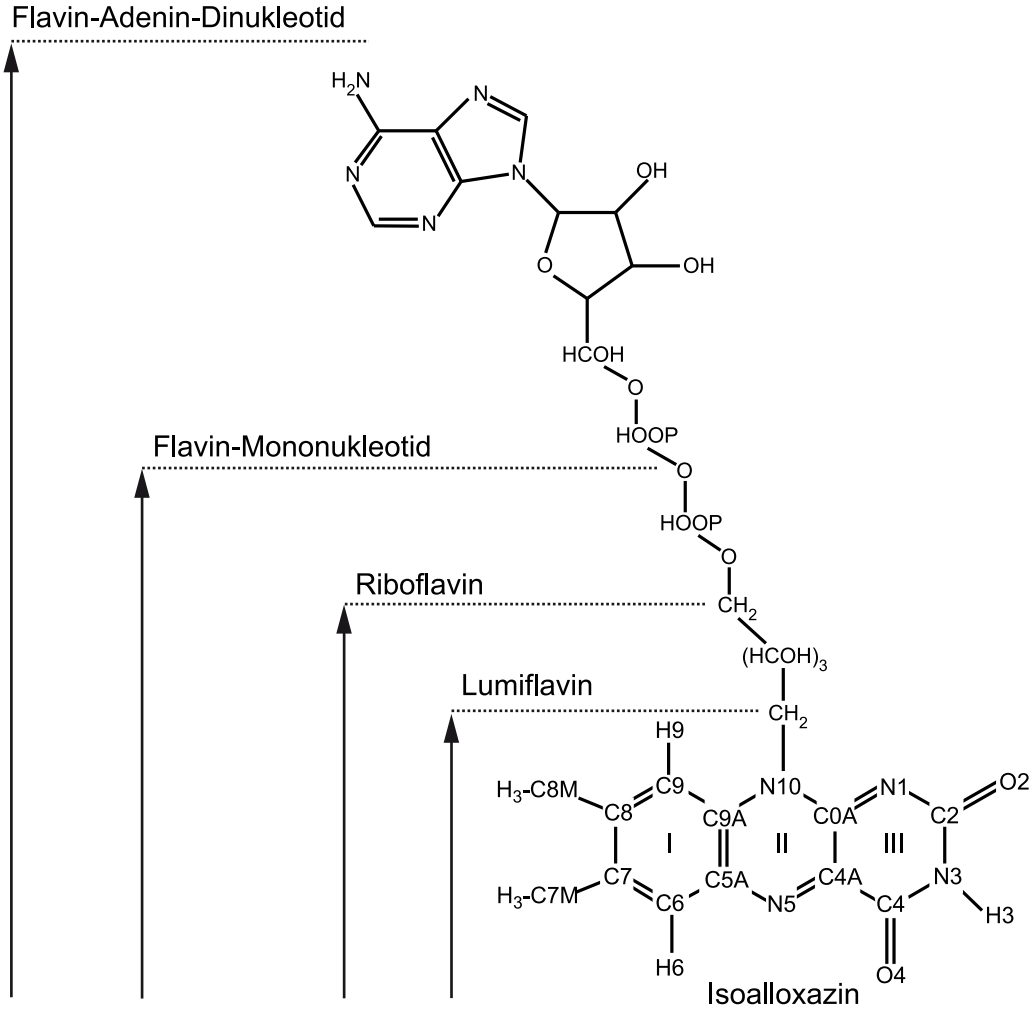


Abbildung 1.4: Die verschiedenen Flavinderivate. Die Grundlage der Flavine bildet ein Ringsystem, das Isoalloxazin. Hier sind die Atombezeichnungen angegeben, die in dieser Arbeit verwendet wurden. Durch die Anbindung unterschiedlicher Molekülketten an das N10 Atom des Isoalloxazins ergeben sich die einzelnen Flavinderivate.

weise von einem Flavoprotein. Da Flavine Elektronentransportketten katalysieren können [1], sind Flavoproteine in biologischen Elektronentransferprozessen, wie zum Beispiel in der mitochondrialen Atmungskette, zu finden. Das Flavin dient dabei als Elektronenüberträger (Elektron-Transfer-Flavoproteine ETF), Elektronendonor (Oxidasen) oder Elektronenakzeptor (Reduktase). Wie jedoch die Wechselwirkungen zwischen dem Flavin und den umgebenden Proteinen genau vonstatten gehen, ist bis heute nicht genau geklärt und ist nach wie vor Gegenstand aktueller Forschung. Da es mehrere Flavin-Protein-Systeme gibt, die alle unterschiedliche Reaktionsmechanismen aufweisen [84], soll nun die Klasse der Blaulichtrezeptoren vorgestellt werden, bevor ein bestimmtes Flavin-Protein-System aus dieser Klasse betrachtet wird.

1.4.2 Blaulichtrezeptoren

Die Klasse der Blaulichtrezeptoren, bei denen ein Flavin (RF, FMN, FAD) meist nichtkovalent als Chromophor gebunden wird, stellt einen großen Anteil der Photorezeptoren dar. Heutzutage werden die Blaulichtphotorezeptoren in drei verschiedene Gruppen unterteilt, welche die Cryptochrome, die sogenannten LOV (*Light-Oxygen Voltage sensing*) und BLUF (*Blue Light sensing Using Flavin*) Domänen umfassen [84].

Die Cryptochrome bilden eine eigene Gruppe der Blaulichtrezeptoren und wurden als erstes in Pflanzen nachgewiesen. Auf Grund ihrer genetischen Homologie wurden sie zu den Photolyasen gezählt [86], welche unter Blaulichtbedingungen die Reparatur von beschädigter DNA katalysieren [87, 88]. Auch wenn beide Proteinsorten an die DNA binden können, so zeigen die Cryptochrome jedoch keinen aktiven Beitrag zur DNA Reparatur. In Kombination mit anderen Proteinen sind Cryptochrome neben der De-Etiolierung, der Pigmentbildung und der Samenkeimung auch an der Steuerung der physiologischen Uhr beteiligt [89, 90]. Da die Cryptochrome auch in Bakterien, Insekten, Vögeln und Säugetieren entdeckt wurden [91], wird vermutet, dass sie als Photopigmente zum Beispiel die zirkadiane Rhythmik von Mäusen aufrecht erhalten [92]. Inwieweit allerdings Cryptochrome hier als Blaulichtrezeptoren fungieren, ist bis dato noch nicht geklärt. So zeigten Untersuchungen an Tauben, dass Cryptochrome, welche in der Retina zu finden sind, anscheinend als magneto-sensorische Rezeptoren das Erdmagnetfeld visualisieren können [93].

Eine weitere Gruppe der Blaulichtphotorezeptoren umfasst die LOV Domänen, von denen das Phototropin die bekannteste ist [94, 95]. In Abhängigkeit der Intensität von blauem Licht induziert das Phototropin in Planzen verschiedene Reaktionen. Zu diesen Reaktionen gehören zum Beispiel Stomata- [96] und Chloroplastenbewegungen [97]. Ferner reguliert das Phototropin den sogenannten Phototropismus [97], welcher allerdings auch durch die eben vorgestellten Cryptochrome beeinflusst wird [98].

Heutzutage ist bekannt, dass während des Photozyklus der LOV Domäne, welcher durch die Absorption eines blauen Photons gestartet wird, einige unterschiedliche molekulare Konformationen durchlaufen werden. Als erstes wird ein Intermediat gebildet, das in spektroskopischen Untersuchungen eine Rotverschiebung zur anfänglichen Struktur aufweist [99, 100] und innerhalb von einigen Millisekunden in ein langlebiges blauverschobenes Intermediat übergeht [99, 100], wobei eine Reduzierung des Flavins stattfindet. Eine Adduktbildung, also die Ausbildung einer kovalenten Bindung, zwischen dem C4A Kohlenstoffatom des Isoalloxazinrings und der Seitengruppe eines konservierten Cysteins der LOV Domäne [99–101] verursacht diese Reduzierung. Das entstandene Addukt ist bei fehlender Lichteinstrahlung vollständig reversibel und führt schließlich in den Dunkelzustand der LOV Domäne zurück [99], womit der Photozyklus abgeschlossen ist. Wie jedoch im einzelnen die Reaktionsmechanismen aussehen, wird noch von mehreren Forschergruppen analysiert.

Die dritte und letzte Gruppe der Blaulichtphotorezeptoren umfasst die BLUF Domänen. Da diese einen zentralen Gegenstand dieser Arbeit darstellen, ist ihnen der folgende Abschnitt gewidmet.

1.5 Die BLUF-Domäne

BLUF Domänen sind das jüngst bekannte Mitglied der Blaulichtphotorezeptorfamilie und wurden in eukariotischen und prokariotischen Mikroorganismen gefunden [84, 102]. Obwohl sie, wie die LOV Domäne, eine redox- und blaulichtregulierte Domäne darstellen, so wurde keine Sequenzhomologie zwischen diesen beiden Domänen festgestellt. Mit Hilfe der Anionen-Austausch-Chromatographie und Dünnschichtchromatographie konnte gezeigt werden, dass in der BLUF Domäne ein oxidiertes Flavinderivat nicht-kovalent eingebettet ist [103].

Zu den bis heute identifizierten BLUF Proteinen gehören das AppA (*Activator of Photopigment and Puc expression A*), das BlrA (*Blue light receptor A*) und das BlrB (*Blue light receptor B*) aus dem Bakterium *Rhodobacter Sphaeroides* [104, 105], das PAC Protein (*photoactivated adenylyl cyclase*) aus dem Einzeller *Euglena gracilis* [106], das Slr1694 (oder auch PixD genannt) und das Tll0078 aus den Cyanobakterien *Synechocystis* sp. PCC6803 und *Thermosynechococcus elongatus* [107, 108]. Dabei haben die einzelnen Proteine unterschiedliche Funktionen. Das Multidomänenprotein AppA, welches das am besten untersuchte BLUF Protein ist, reguliert in Abhängigkeit des zur Verfügung stehenden Sauerstoffs und der vorherrschenden Lichtintensität die Steuerung der Expression des Photosynthese-Apparates [104, 109–112]. Die Funktionalitäten der Proteine BlrA und BlrB sind dagegen bis heute noch nicht entschlüsselt worden [105]. Die Domänen PAC, Slr1694 und Tll0078 ermöglichen photophobische und phototaxische Reaktionen und regulieren so die Bewegung von Zellen in Abhängigkeit des einfallenden Lichts [108, 113].

Um die Funktionsweise der BLUF Proteine verstehen zu können, ist es nötig, deren Aufbau zu kennen. Daher wurden verschiedene Kristall- und NMR-Strukturen von unterschiedlichen BLUF Proteinen ermittelt [105, 114–117]. In den folgenden Abschnitten wird am Beispiel der Proteine AppA und Slr1694 die Struktur der Photorezeptordomäne sowie deren Photozyklus, soweit er bisher entschlüsselt wurde, dargestellt. Die Erkenntnisse bilden die Grundlage für die Modellierung und die Computersimulationen, die im Zuge dieser Arbeit getätigt wurden, um die essentiellen Wechselwirkungen während des Photozyklus der BLUF Proteine weiter zu analysieren und zu verstehen.

1.5.1 Aufbau und Struktur der BLUF Domäne

Der Wildtyp des Multidomänenproteins AppA tritt in nativer Umgebung als Monomer auf [104, 118]. Da das Interesse jedoch auf der Funktion als Photorezptor liegt, wurden reduzierte AppA-Mutanten gezüchtet, die lediglich aus der BLUF Domäne bestehen. Hierzu wurden unter anderen fünf verschiedene Strukturen in der PDB veröffentlicht, die sich auf den ersten Blick nur geringfügig voneinander unterscheiden. Vier stammen aus Röntgenstrukturanalysen [114, 115] (PDB-Kodes: 2IYG, 2IYI und 1YRX) und die fünfte aus einer NMR-Analyse [117] (PDB-Kode: 2BUN). So sind sie, abgesehen vom N- und C-Terminus, nahezu identisch [114, 115, 117] und da sie sich nur in einzelnen Aminosäuren von den BLUF Domänen anderer Proteine unterscheiden, geht man davon aus, dass es sich bei der BLUF Domäne um eine universelle Domäne handelt [105, 116]. Die reduzier-

ten AppA Proteine treten in Lösung allerdings, im Gegensatz zum Wildtyp, nicht mehr als Monomere, sondern als Dimere auf, wie zum Beispiel Untersuchungen durch Gelfiltration gezeigt haben [117, 119].

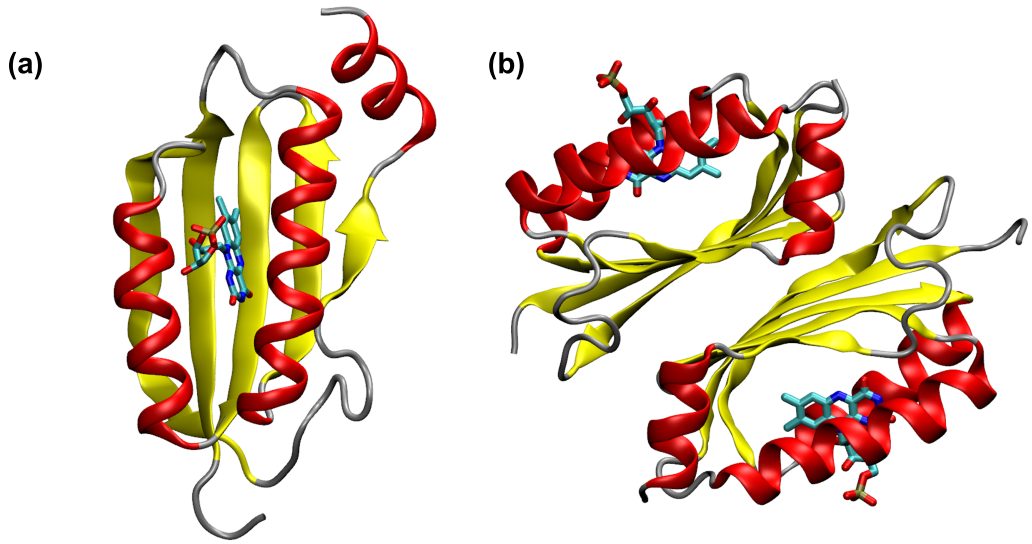


Abbildung 1.5: Struktur der BLUF Domäne. Illustration (a) zeigt den strukturellen Aufbau der AppA BLUF Domäne und (b) die typische Dimerstruktur [115]. In der Proteinbindungstasche zwischen den beiden α-Helizes befindet sich das Flavinderivat FMN.

Abbildung 1.5 zeigt die 3D Kristall-Struktur (2IYG) des reduzierten AppA Monomers (a) und Dimers (b) aus einer Röntgenstrukturanalyse [115]. Der Dimer zeigt eine asymmetrische Konstellation, wobei die Wechselwirkungen zwischen den beiden Monomeren durch hydrophobe Oberflächen der β-Faltblätter vermittelt werden.

Die BLUF Domäne ist ein α-β-Sandwich mit der Topologie $\beta\alpha\beta\beta\alpha\beta\beta$, wobei das fünf-blättrige β-Falltblatt von zwei paralell verlaufenden α-Helizes flankiert wird. Das Flavinderivat, in diesem Fall ein FMN, befindet sich zentral in einer Bindungstasche zwischen den beiden α-Helizes, wobei die unpolare Dimethylphenyl-Seite (Ring I in Abb. 1.4) in einer hydrophoben Proteintasche (bestehend aus den Aminosäuren Ile37, Phe61 und His85) und die polare Pyrimidin-Seite (Ring III in Abb. 1.4) über Wasserstoffbrückenbindungen durch konservierte Aminosäuren (Asn45, Gln63) stabilisiert wird. Die Ribityl-5'-Phosphat Seitenkette des FMN zeigt zur Proteinoberfläche und ist teilweise lösungsmittelexponiert. Die BLUF Domäne kann auch andere Flavinderivate binden, und da der Adeninrest des FAD völlig dem Lösungsmittel exponiert ist, kann er durch eine Röntgenstrukturanalyse nicht identifiziert werden, ist aber beispielsweise in der NMR Struktur detektiert worden [117]. Untersuchungen am AppA Protein haben ergeben, dass die einzelnen Flavinderivate mit unterschiedlichen Wahrscheinlichkeiten gebunden werden. So wird das RF zu 28 %, das FMN zu 35 % und das FAD zu 37 % in der Bindungstasche eingebunden [115].

Die BLUF Domäne wird durch die Absorption eines Blaulicht-Photons am FMN

vom Dunkelzustand in einen langlebigen, transienten Zustand überführt, bei dem sich die Wechselwirkungen zwischen dem Protein und dem Flavin verändern. Dies resultiert in einer Konformationsänderung des Proteins, also in der Bildung eines Signalzustands. Die Aufklärung der Wechselwirkungen zwischen Chromophor und Protein, die vor, während und nach der Photoreaktion vorliegen, ist essentiell, um den Mechanismus der Photoreaktionen auf atomarer Ebene und die damit verbundene molekulare Signalübertragung zu verstehen.

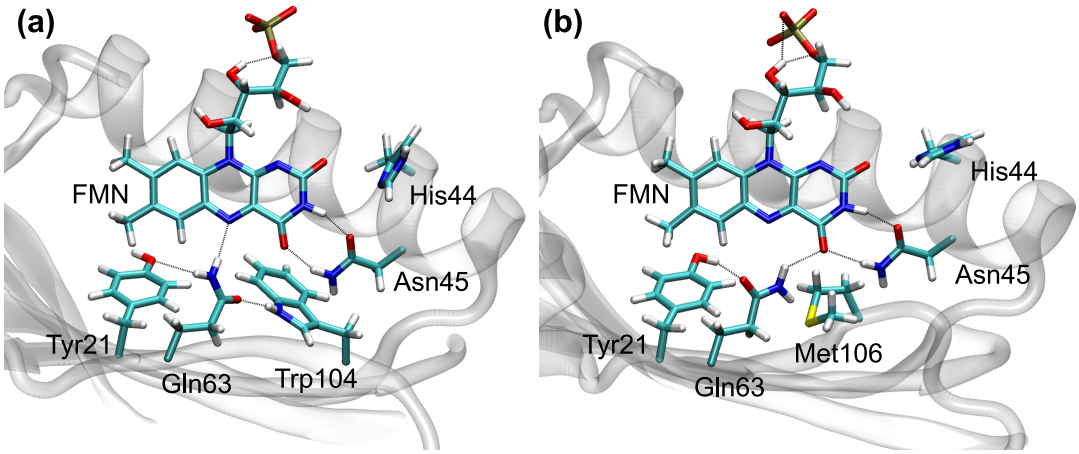


Abbildung 1.6: Konformationen der BLUF Domäne. Im AppA Protein gibt es zwei unterschiedliche Konformationen, die in den Kristallstrukturen gefunden wurden. Dies sind (a) die Trp_{in} [114] und (b) die Met_{in} [115] Konformation. Beide Strukturen wurden mit Hilfe von MD Simulationsprogrammen um die Wasserstoffatome ergänzt. Die gepunkteten Linien markieren Wasserstoffbrückenbindungen, die innerhalb der entsprechenden Konformation bestehen.

Betrachtet man nun die Struktur der BLUF Domäne genauer, so findet man zwei unterschiedliche Konformationen, die sowohl im AppA Protein [114, 115, 117] als auch im Slr1694 [116] vorkommen. Abbildung 1.6 zeigt diese beiden Konformationen für das AppA Protein, welche maßgeblich durch Änderungen im β-Faltblatt 5 charakterisiert sind. Bei der ersten Konformation in Abbildung 1.6(a) befindet sich das Trp104 (Trp94 in Slr1694) aus Faltblatt 5 innerhalb der Bindungstasche und wird deswegen Trp_{in} Konformation bezeichnet. In diesem Fall wird die molekulare Struktur durch Wasserstoffbrückenbindungen zwischen Trp104-N–H und O-Gln63 sowie zwischen Gln63-N–H und O-Tyr21 stabilisiert. Zwischen dem Apoprotein und dem FMN gibt es drei wesentliche Wasserstoffbrückenbindungen, wobei Gln63-N–H an das Atom N5-FMN und Asn45 stark elektrostatisch an den Ring III des Isoalloxazins gebunden ist.

Die zweite Konformation, Abbildung 1.6(b), zeigt, dass sich hier das Met106 (Met96 in Slr1694), welches wie Trp104 zum Faltblatt 5 gehört, in der Nähe des FMN befindet, während das Trp104 an der Proteinoberfläche lokalisiert ist. Diese Konformation wird dementsprechend Met_{in} genannt. Auf Grund des Positionswechsels von Trp_{in} zu Met_{in} sieht die Wasserstoffbrückentopologie etwas anders aus. Wie in der Trp_{in} Konformation existieren zwei starke Wasserstoffbrücken zwischen Asn45 und dem Ring III des Isoal-

loxazins. Die Seitengruppe des Gln63 hingegen ist um 180° rotiert und bildet nun eine Wasserstoffbrücke zum O4-FMN aus. Des Weiteren agiert nun das Gln63 in der Verbindung mit Tyr21 als Wasserstoffbrückenakzeptor.

Es wird vermutet, dass eine der beiden Konformationen den Dunkelzustand der BLUF Domäne beschreibt (z.B. Trp_{in}) und die andere höchstwahrscheinlich den Lichtzustand. Allerdings ist man sich bisher bei der Zuweisung nicht einig. So wurden in Kristall- und NMR-Strukturen von im Dunklen gewachsenen BLUF Proteinen sowohl die Trp_{in} Konformation [114, 116, 117] als auch die Met_{in} Konformation [105, 115] gefunden. Ferner haben Fluoreszenzexperimente am Slr1694 gezeigt, dass das Trp im Lichtzustand Lösungsmittel exponiert zu sein scheint [116], wohingegen solche Experimente am AppA Protein darauf hinweisen, dass sowohl im Dunkel- als auch im Lichtzustand die Trp_{in} Konformation vorliegt [120].

Die Analyse der BLUF Domäne gestaltet sich auf Grund der chemischen, strukturellen und dynamischen Heterogenität des Systems als kompliziert. Experimente zeigen, dass sich die einzelnen Reaktionen des Photozyklus auf verschiedenen Zeitskalen abspielen, die im Pikosekundenbereich anfangen und bis hin zum Minutenbereich reichen. Diese Umstände und eventuelle lichtintensitätsabhängige Seitenreaktionen, die eine Reduktion des Flavins oder eine Auslösung des Kofaktors aus der Bindungstasche verursachen können, gestalten die Analyse des Mechanismus der Photoreaktion zusätzlich schwierig [82]. Trägt man jedoch die bisher gesammelten Daten zusammen, so läßt sich ein möglicher Photozyklus der BLUF Domäne beschreiben.

1.5.2 Der Photozyklus der BLUF-Domäne

Es gibt bisher noch keine einheitliche Theorie darüber, wie der Photozyklus der BLUF Domäne genau vonstatten geht. Der Reaktionszyklus, der durch die Absorption eines Photons am Flavin in den unterschiedlichen BLUF Domänen gestartet wird, wurde dabei nicht nur durch ultraschnelle Spektroskopie im ultravioletten und sichtbaren Bereich (UV/Vis) [40, 41] charakterisiert, sondern auch durch RR und IR Schwingungsspektren [58, 107, 121, 122]. Letztere sind dabei von besonderem Interesse, denn im Gegensatz zu den strukturell unspezifischen UV/Vis Spektren enthalten IR Spektren detaillierte strukturelle Informationen in verschlüsselter Form (siehe Abschnitt 1.2).

Abblidung 1.7 zeigt schematisch den Photozyklus, wie er in dieser Arbeit auf Grund experimenteller Ergebnisse sowohl am AppA [40, 120] als auch am Slr1694 [41, 58] deduziert worden ist. Für die Darstellung des Zyklus ist dabei angenommen, dass es sich bei dem gebundenen Flavinderivat um ein FAD handelt. Unabhängig von der vorliegenden Konformation der BLUF Domäne bildet sich aus dem Dunkelzustand $BLUF_d$ durch Absorption eines blauen Photons ein angeregter Singletzustand FAD*, der, wie am AppA Protein gezeigt wurde [40], in 5 ns in einen Triplettzustand FAD^T übergehen kann und nach weiteren 3 μs wieder direkt in den Dunkelzustand zurückkehrt.

Um den Lichtzustand $BLUF_l$ zu erreichen, zerfällt der angeregte Zustand FAD* multiexponentiell, wobei beim AppA Protein Zeitkonstanten von 90 ps und 570 ps ermittelt wurden [40], bis der Signalzustand erreicht ist. Beim Slr1694 wurde ebenfalls ein multi-

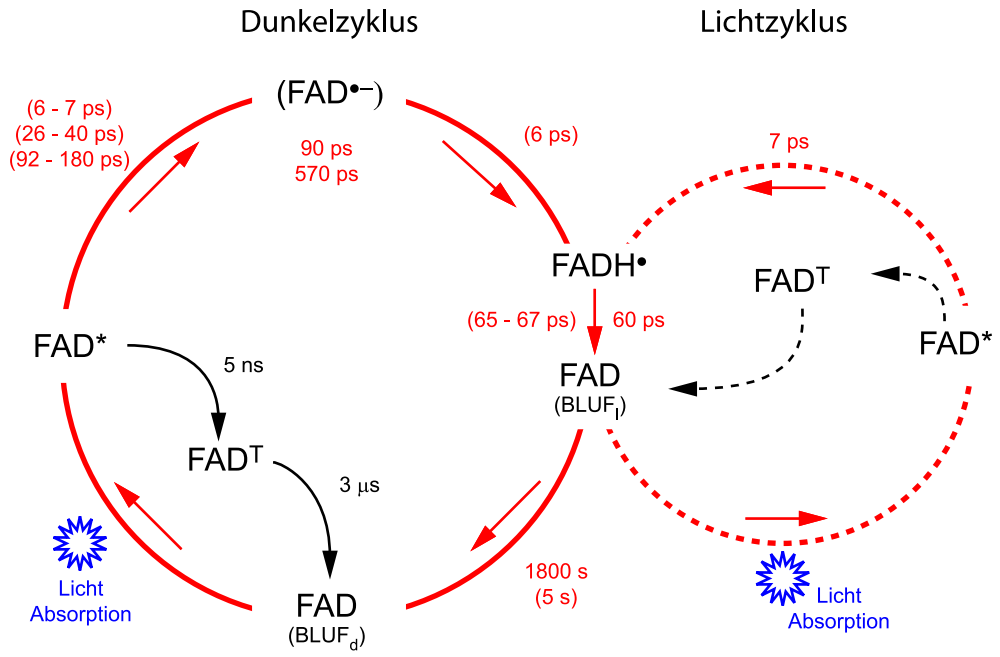


Abbildung 1.7: Photozyklus der BLUF-Domäne. Schematischer Ablauf des Photozyklus der BLUF Domäne, wie er auf Grund verschiedener Experimente am AppA [40, 120] und am Slr1694 [41, 58] Protein hier konstruiert wurde. Die angegebenen Zeitkonstanten stammen von Untersuchungen an AppA Proteinen, in Klammern sind die Zeitkonstanten für das Slr1694 angegeben. Für eine detaillierte Beschreibung siehe Text.

exponentieller Zerfall nachgewiesen, der mit Zeitkonstanten von 6-7 ps, 26-40 ps und 92-180 ps durch einen Elektrontransfer des benachbarten Tyr11 (Tyr21 in AppA) zunächst in ein anionisches Flavinradikal $FAD^{\bullet-}$ überführt wird [41, 58]. Durch einen sich schnell anschließenden Protonentransfer zum $FAD^{\bullet-}$, wobei das Proton ebenfalls vom Tyr11 abgegeben wird, entsteht innerhalb von 6 ps ein neutrales Flavinradikal $FADH^{\bullet}$, welches auch im AppA Photozyklus nachgewiesen wurde, für dessen Lebensdauer jedoch keine Angaben gemacht wurden [123]. Innerhalb von 65-67 ps wandelt sich im Slr1694 dann das neutrale Flavinradikal $FADH^{\bullet}$ durch Deprotonierung und Reoxidation zurück in das oxidierte FAD [41, 58]. Dieser Zustand wird nun mit dem Signalzustand $BLUF_l$ assoziiert [40, 41, 58] und weist eine Lebensdauer von 5 s (Slr1694) beziehungsweise 1800 s (AppA) auf. Im Gegensatz zum Slr1694 wurde das anionische Flavinradikal $FAD^{\bullet-}$ im AppA Photozyklus nicht detektiert, jedoch wird vermutet, dass es auch hier als Intermediat auftritt [120, 123].

Dass das Tyrosin im Photozyklus der BLUF Domäne sowohl Elektron- als auch Protondonor ist, konnte dadurch gezeigt werden, dass dessen Ersetzung durch eine andere Aminosäure zur Folge hat, dass die Bildung des Signalzustandes gehemmt wird. Des Weiteren wird die Annahme des Protonentransfers während der Photoreaktion dadurch unterstützt, dass eine verzögerte Relaxation in den Dunkelzustand nach einem H/D-Austausch und eine beschleunigte Relaxation bei Zugabe von neutralem Imidazol als putativer Pro-

tonendonor erfolgt [118]. Neben dem konservierten Tyr21 (Tyr11 in Slr1694) ist auch das Trp104 maßgeblich an der Photoreaktion beteiligt. Wird das Trp104 durch ein Phenylalanin ausgetauscht, so relaxiert der Signalzustand um etwa ein Faktor 150 schneller in den Dunkelzustand zurück [121, 124].

Letztendlich konnte also gezeigt werden, dass die Bildung des Lichtzustandes $BLUF_l$ nicht auf strukturellen Veränderungen des Flavins selbst basiert, sondern mit einer Neuanordnung der Wasserstoffbrückenbindungen zwischen dem Flavin und den umgebenden Aminosäurenseitenketten als auch durch Konformationsänderungen einhergehen muss. Dies konnte ferner auf der Basis der Bandenzuordnung von Schwingungsmoden des Isoalloxazinrings belegt werden [107, 121, 122].

Der langlebige Signalzustand $BLUF_l$ unterscheidet sich UV/Vis-spektroskopisch durch eine Rotverschiebung des Absorptionsmaximums von etwa 10 nm im Vergleich zum Dunkelzustand $BLUF_d$, weswegen der Lichtzustand oftmals $BLUF_{red}$ genannt wird. Im Lichtzustand $BLUF_l$ liegen nun allerdings veränderte Wechselwirkungsverhältnisse zwischen dem Apoprotein und dem gebundenen Flavin vor, welche für die Signalpropagation zu intramolekularen oder intermolekularen Partnern verantwortlich ist. Die Ähnlichkeiten zu HSCQ (*Heteronuclear Single Quantum Coherence*) Spektren legt dabei nahe, dass sich keine größeren Konformationsänderungen beim lichtinduzierten Zustand ergeben [40, 119, 121]. So stellt sich die Frage, ob das Umschalten zwischen den beiden gefundenen Konformationen (Trp_{in} und Met_{in}) des β-Faltblatts 5 den Übergang vom Dunkelzustand $BLUF_d$ zum Lichtzustand $BLUF_l$ überhaupt beschreiben kann, da es sich hierbei um eine deutliche Konformationsänderung handelt. Ob nun solch eine Konformationsänderung innerhalb der angegebenen Zeit (< 1 ns) vonstatten gehen kann, ist nicht direkt ersichtlich, wenn man die typischen Zeitskalen molekularer Bewegungen in Tabelle 1.1 mit einbezieht.

Auf Grund der langen Lebensdauer des Lichtzustandes $BLUF_l$ stellt sich natürlich auch die Frage der Photoreversibilität. So konnte gezeigt werden, dass die BLUF Domäne nicht photoreversibel ist und auch keinen alternativen Reaktionsweg aufweist, wenn ein weiteres Photon am Flavinderivat absorbiert wird [120]. Wie im Zyklus für den Dunkelzustand kann auch der Lichtzustand $BLUF_l$ gemäß Abbildung 1.7 ein Photon absorbieren. Im Fall des AppA Proteins wird wiederum ein angeregter Singletzustand des Flavins (FAD*) erzeugt, der über einen Triplettzustand FAD^T direkt wieder in den Lichtzustandzustand $BLUF_l$ zurückkehren kann. Für diesen Übergang sind allerdings keine Zeitkonstanten bekannt [120]. Des Weiteren besteht die Möglichkeit, dass der angeregte Zustand FAD* innerhalb von 7 ps in das neutrale Flavinradikal $FADH^\bullet$ und anschließend in 60 ps in den Lichtzustand $BLUF_l$ überführt wird [120]. Ein anionisches Flavinradikal $FAD^{\bullet-}$ wurde dabei nicht nachgewiesen, jedoch vermutet [120]. Ferner wurden auch keine molekularen Veränderungen des Wasserstoffbrückenbindungs-Netzwerkes im Lichtzyklus des AppA Proteins entdeckt, wie sie im Zyklus des Dunkelzustandes vonstatten gehen. Dies ist anscheinend der Grund dafür, dass der erneute Zyklus im Signalzustand auch schneller abläuft als im Dunkelzustand [120].

Die skizzierten Eigenschaften machen die BLUF Domänen zu bisher einzigartigen Photorezeptoren, denn hier liegt sowohl dem Dunkel- als auch dem Lichtzustand der gleiche Chromophor, ein oxidiertes FAD, zu Grunde [84, 120]. Unklar hingegen sind die

molekularen Mechanismen, die verantwortlich sind, den Lichtzustand zu bilden.

1.5.3 Molekulare Mechanismen des Photozyklus

Neben einer Konformationsänderung spielt auch die damit einhergehende Veränderung des Wasserstoffbrückenbindungs-Netzwerkes eine entscheidende Rolle, welche durch den Elektron- und Protontransfer zum und vom Flavin hervorgerufen werden. So ist aus experimentellen Schwingungsspektren, wie zum Beispiel aus dem Differenzspektrum in Abbildung 1.3, ersichtlich, dass sich an der Position C4=O4 des Isoalloxazins eine neue oder auch verstärkte Wasserstoffbrückenbindung ausbilden muss [57, 107, 121, 122]. Im dargestellten FTIR Differenzspektrum [57] des Slr1694 wird nämlich die Bande bei 1713 cm^{-1} des Dunkelzustands beim Übergang zum Lichtzustand zu der Frequenz von 1697 cm^{-1} verschoben, wofür, wie auch in der vorliegenden Arbeit belegt werden wird, eine verstärkte Wasaserstoffbrückenbindung am FMN-O4 verantwortlich sein muss. Die genannte Bande wird der C4=O4 Streckschwingung zugeordnet, und es konnte gezeigt werden, dass auch in weiteren Schwingungsspektren des AppA und des Slr1694 Proteins eine Rotverschiebung von etwa 16 cm^{-1} als Folge des Übergangs in den Signalzustand auftritt [82, 107, 121, 122, 125]. Die Zuordnung der anderen Banden gestaltet sich allerdings als sehr schwierig, da hier die Flavin- und Proteinbanden starke Überlagerungen aufweisen.

Basierend auf Röntgenstrukturen, NMR Daten, UV/Vis Spektren, IR und RR Schwingungsspektren und/oder quantenchemischen Berechnungen wurden verschiedene detaillierte Modelle vorgeschlagen, die auf Grund der Ladungstransfer-Prozesse den Übergang vom Dunkelzustand zum Lichtzustand und die damit einhergehende Reorientierung des Wasserstoffbrückenbindungs-Netzwerkes in BLUF Domänen beschreiben [41, 58, 114, 116, 117, 126–133]. Hierbei stehen diejenigen Aminosäuren, die einer lichtinduzierten Konformationsänderung unterliegen und zu der inneren Bindungstasche gehören (Tyr21, Trp104/Met106) oder direkten Kontakt zum Isoalloxazin besitzen (Gln63) im Fokus (siehe Abb. 1.6). Tyr21 und Gln63 sind dabei wohl die Schlüsselfiguren der ultraschnellen photochemischen Ereignisse, die den Signalzustand erzeugen. Weitere wichtige Aminosäuren sind nur schwer auszumachen, da nicht jede Aminosäure in allen BLUF Domänen konserviert ist.

Die genauen molekularen Mechanismen bei der Bildung des Signalzustandes sind Gegenstand laufender Debatten, was sich daran zeigt, dass derzeit vier unterschiedliche Modelle von verschiedenen Forschergruppen vorgeschlagen wurden. Die vorgeschlagenen Modelle beinhalten einen oder mehrere atomare Mechanismen, die eine Rotation und/oder eine Tautomerisierung der Gln63 Seitenkette, sowie den Flip von Trp_{in} zu Met_{in} oder vice versa umfassen [41, 58, 114, 116, 117, 126–133]. Auch die Umwandlung des Signalzustandes in ein proteineigenes Signal ist derzeit noch unklar.

Insgesamt gilt, dass die experimentellen und theoretischen Resultate es bisher nicht erlauben, eines der vorgeschlagenen Modelle für die molekularen Mechanismen zur Bildung des Signalzustands den anderen Modellen vorzuziehen. Das heißt, im Moment ist es nicht möglich, mit nur einem Modell alle beobachteten Sachverhalte zu verstehen und zu beschreiben. Daher sind weitere Untersuchungen notwendig, um eine exakte Beschrei-

bung der dem BLUF Photozyklus zu Grunde liegenden molekularen Veränderungen zu erreichen. Auf Grund dieser Gegebenheiten wurde diese Arbeit im Rahmen der Forschergruppe 526 iniziiert, um einen weiteren Teil zur Aufklärung beizutragen. Im folgenden Abschnitt werden deshalb die Ziele und die mit der Umsetzung verbundenen Fragen dargestellt.

1.6 Ziele und Überblick

Wie in den vorangegangenen Abschnitten dargestellt wurde, sind nach wie vor einige Fragen in Bezug auf die BLUF Domänen und deren photochemische Prozesse unbeantwortet. Deshalb war es Ziel dieser Arbeit, basierend auf computergestützter Modellbildung und Simulation, weitere Puzzleteile zu liefern, die beitragen können, um das Bild über die BLUF Domänen zu vervollständigen.

Um dieses Ziel mit einer anwendungsorienterten Arbeit zu erreichen, waren allerdings einige vorbereitende Schritte notwendig. Zu Beginn dieser Arbeit war es erforderlich, ein MM Kraftfeld für Flavinmoleküle zu generieren, da in üblichen MM Kraftfeldern wie CHARMM22 [78] für diese Molekülsorte kein Parametersatz zur Verfügung stand. Unter zur Hilfenahme von DFT Rechnungen wurde daher als erstes ein MM Kraftfeld für das Lumiflavin entwickelt, das an die CHARMM22 Topologie angelehnt ist. Um das Lumiflavin zu einem FMN zu erweitern, wurde das entsprechende Kraftfeld mit bereits existierenden Modellen für die angehängte Molekülkette erweitert.

Bevor Flavine und deren Schwingungsspektren in BLUF Domänen untersucht werden konnten, war es nötig zu testen, wie gut der mittlere IR Spektralbereich, und hier insbesonders der Bereich von $1500\,\mathrm{cm}^{-1}$ bis $1750\,\mathrm{cm}^{-1}$, der Schwingungsspektren von Flavinen mit einem DFT/MM Hybridverfahren, im Vergleich zum Experiment, bestimmt werden kann, wenn ein weniger komplexes System betrachtet wird. Üblicherweise kalibriert man solche DFT/MM Berechnungen in der Gasphase und überträgt sie dann erst in die kondensierte Phase. Da aber keine Schwingungsspektren von Flavinen in der Gasphase experimentell bekannt sind, wurde die erforderliche Kalibrierung der Flavinspektren in Wasser realisiert.

Technisch gesehen gibt es mehrere Möglichkeiten, DFT/MM IR Spektren eines Moleküls in kondensierter Phase zu berechnen. Da das Lumiflavin, mit seinen 31 Atomen ein relativ großes Molekül darstellt, war der dynamikbasierte Zugang über die Berechnung der Fouriertransformierten der Zeit-Korrelations-Funktion [134–138] rechentechnisch nicht realisierbar. Aus diesem Grund wurde die Methode der Instantanen Normal-Moden Analyse (INMA) gewählt [52, 134, 139–141], bei der rechenzeit-effiziente MM-MD Simulationen verwendet werden, um den Konfigurationsraum anhand zufällig gewählter Schnappschüsse abzutasten. Die anschließende, DFT/MM basierte Normal-Moden Analyse des Lumiflavins innerhalb der festgehaltenen Lösungsmittelkäfige aller ausgewählten Schnappschüsse liefert dann die Beschreibung eines IR Schwingungsspektrums, das durch die Lagen, die Breiten sowie die Intensitäten der einzelnen Banden gegeben ist.

Die sich daraus ergebende Aufgabe bestand also darin, zu testen, mit welcher Genau-

igkeit die DFT/MM-INMA Methode die experimentellen Schwingungsspektren von Flavinen in wässriger Lösung [142–151] beschreiben kann, wenn man, wie in theoretischer Schwingungsspektroskopie üblich [152], einen Frequenzskalierungsfaktor heranzieht, der hier zu dem verwendeten Wassermodell [153] und dem eingesetzten DFT-Funktional [154–156] gehört. In der ersten Publikation dieses Projektes, die in Abschnitt 2.1 abgedruckt ist, wird gezeigt, dass die DFT/MM-INMA Methode einen präzisen Einblick in die Zusammensetzung der Normal-Moden von oxidierten Flavinen in Wasser liefert, was durch eine Abweichung von nur etwa $10\,\mathrm{cm}^{-1}$ zwischen den experimentellen und berechneten Schwingungsfrequenzen deutlich wird.

In Abschnitt 2.3 ist die zweite Veröffentlichung eingebunden, die sich mit der Fage beschäftigt, ob auch unterschiedliche Redoxzustände der Flavine mit der entwickelten DFT/MM-INMA Methode beschrieben werden können, wenn weder das zu Grunde gelegte MM Wassermodell noch das eingesetzte DFT-Funktional und der ermittelte Frequenzskalierungsfaktor verändert werden. Diese Frage ist deswegen von besonderem Interesse, da während des Photozyklus der BLUF Domäne unterschiedliche Flavinradikale detektiert wurden (siehe Abschnitt 1.5.2) und eine detaillierte Analyse ihrer Schwingungsspektren weitere Informationen liefern können. In dieser Arbeit wird gezeigt, dass auch die DFT/MM Schwingungsspektren von Flavinradikalen in Wasser die experimentellen Beobachtungen [55, 145–149, 157] reproduzieren und erklären können.

Durch die beiden ersten Arbeiten wurden die nötigen Hilfsmittel geschaffen und getestet, um sich schließlich mit den Schwingungsspektren von Flavinen in BLUF Domänen zu befassen. Um eine hochwertige Beschreibung zu ermöglichen, waren allerdings mehrere Fragen zu beantworten. So galt es als erstes herauszufinden, ob die experimentell bestimmten BLUF Strukturen überhaupt mit einer MM-MD Beschreibung verträglich sind. Des Weiteren war unklar, ob die experimentell bestimmten Proteinstrukturen ausreichend genau für die DFT/MM Berechnung der IR Spektren des Flavinchromophors sind. Ferner musste bei Kristallstrukturen geprüft werden, ob sie den solvatisierten Strukturen, bei denen die experimentellen Schwingungsspektren ermittelt wurden, entsprechen.

Abgesehen von der Eignung der verwendeten Proteinstrukturen für DFT/MM Berechnungen der Schwingungsspektren des Chromophors galt es auch methodische Fragen zu beantworten. Da übliche MM Proteinkraftfelder die Polarisierbarkeit der Elektronenhüllen einzelner Atome vernachlässigen, musste überprüft werden, ob die Verwendung eines solchen Kraftfeldes überhaupt eine akkurate Beschreibung der Schwingungsspektren von Flavinen in BLUF Domänen zulässt, oder ob, wie im Fall des Bakteriorhodopsins [139, 158], eine Modellierung der Polarisierbarkeit in das MM Kraftfeld integriert werden muss, um die benötigte Genauigkeit der Spektren zu erreichen. Eine weitere Unklarheit bestand in der Frage, ob der ermittelte Frequenzskalierungsfaktor bei einer Änderung des MM Kraftfeldes, also bei dem Übergang von Wasser zu einer Proteinstruktur, übertragbar ist oder ob er für eine hinreichend genaue Beschreibung der DFT/MM Schwingungsfrequenzen von Flavinen in BLUF Domänen neu angepasst werden muss.

Die dritte und letzte Publikation dieses Projekts, abgedruckt in Kapitel 3, beinhaltet die Ergebnisse zur Untersuchung der dynamischen Stabilität sowie der Schwingungsspektren von Flavinen anhand sieben verschiedener Strukturen von BLUF Domänen, die aus der

PDB entnommen wurden. Sie liefert Antworten auf jene Fragen, mit denen die DFT/MM Beschreibung der Flavinschwingungsspektren in BLUF Domänen konfrontiert war.

Im abschließenden Kapitel 4 werden die Ergebnisse noch einmal kompakt zusammengefasst und ein Ausblick gegeben, wie zukünftige Arbeiten die noch offen gebliebenen Fragen beantworten könnten.

2 DFT/MM Schwingungsspektren von Flavinen in Wasser

Zunächst werden die vorbereitenden Arbeiten dargestellt, die notwendig waren, um die Werkzeuge zu schaffen, die eine hochwertige Beschreibung der Schwingungsspektren von Flavinen in BLUF Domänen ermöglichen. Dazu wurden die Schwingungsspektren von unterschiedlichen Redox-Zuständen des Lumiflavins in Wasser vermittels einer DFT/MM Hybridmethode berechnet. Die erste Arbeit (siehe Abschnitte 2.1 und 2.2) galt dabei der Modellbildung des oxidierten Lumiflavins sowie der Kalibrierung der berechneten IR Spektren in Bezug auf experimentelle Ergebnisse. Im zweiten Schritt (siehe Abschnitte 2.3 und 2.4) wurde das damit entwickelte Verfahren auf Radikalzustände des Lumiflavins übertragen und somit die entsprechenden DFT/MM Schwingungsspektren vorhergesagt und mit experimentellen Resultaten verglichen.

2.1 Oxidiertes Lumiflavin in Wasser

Die nachfolgend abgedruckte Publikation [1]

„Density Functional Theory Combined with Molecular Mechanics:
The Infrared Spectra of Flavin in Solution“
Benjamin Rieff, Gerald Mathias, Sebastian Bauer und Paul Tavan
in *Photochem. Photobiol.* **87**, 511–523 (2011),

die ich zusammen mit Gerald Mathias, Sebastian Bauer und Paul Tavan verfasst habe, befasst sich mit der DFT/MM Berechnung der Schwingungsspektren von oxidiertem Lumiflavin, das in zwei unterschiedliche Wassermodelle eingebettet war. Neben der Ableitung eines MM Kraftfeldes für Flavine aus DFT Rechnungen und der Entwicklung einer automatisierten Methode zur Analyse der Zusammensetzung von Normal-Moden wird in diesem Artikel gezeigt, dass die verwendete DFT/MM-INMA Methode eine hochwertige Beschreibung und Erklärung für die experimentell gemessenen Schwingungsspektren von Flavinen in wässriger Lösung liefert. Diese bezieht sich nicht allein auf die spektrale Bandenlage, sondern erlaubt auch Aussagen zur Breite und Intensität der einzelnen Schwingungsbanden sowie zu den Effekten einzelner Isotopenmarkierungen auf das IR Spektrum.

[1] Mit freundlicher Genehmigung des Wiley Verlags.

Photochemistry and Photobiology, 2011, 87: 511–523

Density Functional Theory Combined with Molecular Mechanics: The Infrared Spectra of Flavin in Solution†

Benjamin Rieff, Gerald Mathias, Sebastian Bauer and Paul Tavan*

Theoretische Biophysik, Lehrstuhl für Biomolekulare Optik, Ludwig-Maximilians Universität, Oettingenstraße 67, München, Germany

Received 15 September 2010, accepted 17 November 2010, DOI: 10.1111/j.1751-1097.2010.00866.x

ABSTRACT

The photophysics and photochemistry of flavin dyes determine the functional dynamics of a series of blue light photoreceptors that include the so-called BLUF (blue light sensors using flavin) domains. To enable molecular dynamics (MD) simulation studies of such signaling processes, we derived molecular mechanics (MM) models of flavin chromophores from density functional theory (DFT). Two 300 K ensembles of lumiflavin (LF) in aqueous solution were generated by extended MM-MD simulations using different MM potentials for the water. In a DFT/MM hybrid setting, in which LF was treated by DFT and the polarizing environment at atomistic resolution by MM, we applied instantaneous normal mode analyses (INMA) to these ensembles. From these data we determined the inhomogeneously broadened solution spectra as mixtures of Gaussian bands using a novel automated procedure for mode classification. Comparisons with vibrational spectra available in the literature on native and isotopically labeled flavins in aqueous solution serve us to determine suitable frequency scaling factors and to analyze the accuracy of our scaled DFT/MM-INMA approach. We show that our approach not only agrees with established computational descriptions but also extends such methods substantially by giving access to inhomogeneous line widths and band shapes.

INTRODUCTION

Flavin dyes play an important role in many enzymatically catalyzed biochemical processes as versatile cofactors (1). They exhibit a series of different redox and protonation states and, therefore, are able to participate in electron and proton transfer processes (2). The basic building block of flavin dyes is isoalloxazine as drawn in Fig. 1. The different kinds of flavins, *e.g.* riboflavin (RF), flavin mononucleotide (FMN) or flavin adenine dinucleotide (FAD), differ by the respective substituents R, which, however, hardly affect the electronic structure of the isoalloxazine ring (1).

The light-induced reactions of many blue light photoreceptors are governed by the photophysics and photochemistry of flavin molecules (3). Prominent examples are the photolyases (4), the so-called light-oxygen-voltage (5), or the "blue light sensoring using FAD" (BLUF) (6) domains. The latter are found in a variety of organisms, where they control cellular processes such as gene expression, nucleotide metabolism and motility.

Electron transfer reactions generating different redox states of flavins in blue light photoreceptors were intensely studied by various techniques and, here, in particular by vibrational spectroscopy (7–12). To identify the modifications of the chromophore properties, which are induced by the respective protein environments, reference spectra were obtained for flavins in different solvents (10–19), because vibrational spectroscopy can monitor the effects of the environment on the intramolecular force field (20). Vibrational bands observed for flavin dyes were experimentally analyzed using isotopically labeled compounds (10,13). The observed bands were assigned to vibrational modes by applying normal mode analyses to isolated flavin molecules (21).

Generally, the identification of flavin bands in the infrared (IR) spectra of flavoproteins is difficult, because apoprotein bands are also found in the corresponding spectral region. Fortunately, however, the stretching vibrations of the two carbonyl groups in the isoalloxazine (see Fig. 1) can be frequently identified in the IR spectra and, therefore, can be used to analyze the chromophore-protein interactions (8,10).

Reliable band assignments require accurate computational methods capable of dealing with flavin molecules embedded in condensed phase environments (22). Such theoretical methods can then help to decode the structural basis of observed spectral band positions. For molecules isolated in the gas phase accurate calculations of vibrational spectra have become readily feasible through the development (23,24) and

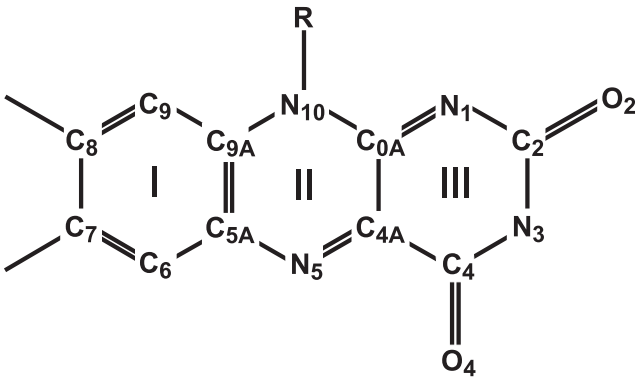


Figure 1. Chemical structure and atom labels of isoalloxazine.

*Corresponding author email: tavan@physik.uni-muenchen.de (Paul Tavan)
†This paper is part of the Symposium-in-Print on "Blue Light Effects."

accessibility (25–27) of density functional theory (DFT) (28–30). In contrast, the computation of vibrational spectra in complex condensed phase environments such as proteins is still neither simple nor standard (22).

For dye molecules dissolved in homogeneous and isotropic solvents a simplified description of the solvent by a dielectric continuum may suffice to model the solvent–solute interactions generating the apparent solvatochromic shifts in the IR spectra of flavin dyes (31). Corresponding continuum models are available in common DFT program packages like Gaussian09 (25) or Turbomole (26) under the names "polarizable continuum model" (32) and "conductor like screening model" (Cosmo) (33), respectively. However, if a dye is embedded in an inhomogeneous and highly structured protein environment, a microscopic and atomistic model of that environment is required for a sufficiently accurate description of the protein–chromophore interactions (22,34).

Corresponding hybrid methods, which combine a quantum mechanical (QM) treatment of a molecule with a simplified molecular mechanics (MM) model of the condensed phase environment, date back to the early work of Warshel and Levitt (35). A further advanced version, which combines the parallelized DFT program package CPMD (27) with the parallelized MM molecular dynamics (MD) program package EGO (36), was suggested by Eichinger *et al.* (37) and has proven its quality for the computation of chromophore IR spectra since then in several applications (see *e.g.* Refs. [34,38,39]).

In such DFT/MM hybrid calculations a given macromolecular simulation system containing a chromoprotein and, possibly, also its surrounding solvent is split into two fragments. One fragment covers the molecule of interest, which is treated quantum mechanically by DFT. The other part of the simulation system is made up of the molecule's condensed phase environment and is usually described by a conventional MM force field (*e.g.* Ref. [49]). A suitable DFT/MM Hamiltonian then accounts for the interactions between these two fragments (37,41,42).

With this contribution we want to initiate a DFT/MM description of the vibrational spectra of flavins embedded in protein environments and, here, particularly in BLUF domains. This description aims at shedding light on the largely unknown structural details of the light-induced reaction cycles in BLUF domains, which lead, on a subnanosecond time scale (9), from the dark-adapted initial states to the light-adapted signaling states.

Various models were suggested for the sequences of structural rearrangements leading from an initial hydrogen bonded network stabilizing the dark-adapted structure of a BLUF domain to a final network defining the light-adapted structure. These models were either based on crystallographic data of putative light- and dark-adapted X-ray structures (43) or on QM/MM model calculations involving putative transition states (44,45). While the quoted hypotheses on the structural basis of the observed dark/light transition are certainly useful for qualitative discussions of various possibilities, there is no clear experimental evidence that would allow preferring one of these suggestions over the others. This situation could change, if one would be able to decode the existing time-resolved spectroscopic data in a reliable fashion.

In fact, the reaction cycles, which follow the flavin's light absorption in several different BLUF domains, have been carefully characterized not only by ultrafast spectroscopy in the visible (46,47) but also by optical-pump/IR-probe spectroscopy (9). The latter data are of particular interest, because IR spectra, in contrast to the structurally quite nonspecific UV/Vis spectra, actually do contain detailed structural information in a coded fashion.

These spectroscopic data suggest that in BLUF domains an electron is transferred on the time scale of a few picoseconds from a neighboring tyrosine to the excited singlet state of the oxidized flavin dye FAD* thus generating a radical pair of two ions. Again within a few picoseconds [*i.e.* 6 ps in Synechocystis Slr1694 BLUF (9)] the anionic $FAD^{\bullet -}$ radical accepts a proton such that the neutral semiquinone radical $FADH^{\bullet}$ is formed. In Slr1694 BLUF the $FADH^{\bullet}$ is subsequently reoxidized and deprotonated within about 60 ps thus restoring the FAD, which, however, is now associated with the signaling state. This state then lives for a few seconds. If one wants to apply DFT/MM methods with the aim to identify the structural rearrangements encoded in the quoted time-resolved IR spectra observed for BLUF domains, one has first to estimate the quality, at which DFT/MM calculations can describe the vibrations of flavins in condensed phase environments. For the thus required calibration of such DFT/MM descriptions it is wise to use a simpler system than a flavin chromophore in a complex protein binding pocket. Flavins in aqueous solution represent such a simplification and have the advantage that many experimental data are available.

Technically, there are several procedures by which one can sample the ensemble IR spectra of a molecule in a condensed phase environment by applying DFT/MM calculations. There are dynamics-based procedures, which sample the atomic fluctuations by extended DFT/MM-MD simulations and compute the IR spectra from Fourier transforms of time-correlation functions (22,48–53). Such techniques give easy access to accurate line widths and shapes, but require a separate simulation for each considered isotope substitution. There are also static procedures, which sample the initial states in Fermi's "golden rule" by selecting snapshots from a computationally inexpensive MM-MD trajectory and apply DFT/MM normal mode analyses to the molecule of interest in its rapidly frozen solvent cages ("instantaneous normal mode analyses," INMA) to obtain the relevant final states (22,34,38,50–52). Due to the neglect of motional narrowing these static techniques overestimate the inhomogeneous line widths, but have the advantage that speedy MM-MD techniques can be applied to the sampling of conformational space. Because flavin chromophores are quite large compounds, the dynamics-based approach is excluded for reasons of lacking computational manageability. Corresponding DFT/MM-MD simulations would have to cover at least 100 ps at a time step of 0.25 fs (53,54), which implies that 400 000 DFT/MM calculations are required to compute an approximately converged IR spectrum of a specific flavin isotopomer in solution. In contrast, the computationally much more efficient INMA technique (22,52) can yield the solution spectra of all isotopomers at about 20% of the computational effort characterized above. Therefore, we had to resort to this method in our attempts to compute the IR spectra of flavins in solution and, subsequently, in BLUF domains.

Thus, it is the specific purpose of this contribution to check for an aqueous solution of lumiflavin (LF), which is the methyl derivative ($R = CH_3$) of the isoalloxazine shown in Fig. 1, at what accuracy DFT/MM calculations can describe the experimental vibrational spectra. Here, two different MM models will be applied to the solvent, which will be simulated at ambient temperature and pressure. Subsequently, we will first give details on the applied methods before presenting and discussing the computational results.

MATERIALS AND METHODS

Lumiflavin was chosen as our flavin model compound, because it contains only $N = 31$ atoms, *i.e.* many atoms less than RF, FMN or FAD and, thus, requires less computation time in DFT/MM calculations. Such calculations were carried out with the hybrid method developed by Eichinger *et al.* (37), which combines the MM-MD program EGO (36) with the plane-wave DFT code CPMD (27).

IR spectra from DFT. For the DFT description of LF we chose the gradient-corrected exchange functional of Becke (55), the correlation functional of Perdew (56) (BP) and the norm-conserving pseudopotentials of Martin and Troullier (MT) (57). The rectangular box containing the grid for the plane-wave expansion of the Kohn-Sham orbitals was centered on the LF in such a way that none of its atoms came closer than 3 Å to one of the faces and that the box volume became minimal. We applied a 70 Ry cutoff to the plane-wave basis set, because test calculations had shown that the vibrational spectrum of the isolated LF most closely resembled the experimental solution spectrum at this cutoff (data not shown). Following earlier work (34,38,39) we denote this DFT approach by "MT/BP." The quality of the MT/BP description obtained with CPMD for the isolated LF molecule was checked by additionally applying the all-electron code Turbomole (26) choosing again the Becke-Perdew functional (55,56) but now combined with the Gaussian triple-zeta basis set including polarization functions called TZVP (58). We call this approach "TZVP/BP."

The minimum energy structure of the isolated LF molecule, the associated Hessian, the normal modes and frequencies were calculated both by MT/BP and TZVP/BP. The dipole gradients required for computing IR intensities were calculated by finite differences displacing the atoms by ± 0.01 Å from their respective equilibrium positions. Normal modes were classified by visual inspection and by calculating potential energy distributions (PED) with the program Gamess (59).

MM force field for LF. Because the CHARMM22 force field (40) does not contain MM models for flavins, we had to derive a new CHARMM22-type force field for this class of dye molecules. For this purpose we adopted the general CHARMM22 force field structure together with many parameters describing analogous chemical motifs. Other parameters were derived from our DFT descriptions of the isolated LF. Here, force constants of atomic bond lengths and angles were calculated from the MT/BP and TZVP/BP Hessians *via* the program Gamess (59). The force constants and the equilibrium values of these internal coordinates, which enter the MM force field of LF as parameters, were chosen by averaging the MT/BP and TZVP/BP results.

Flavins are planar molecules due to the sp^2 hybridization of the carbon and nitrogen atoms making up the isoalloxazine moiety. CHARMM22 parameterizes several planar structures of analogous chemical composition by providing so-called improper dihedral potentials. These parameters and potentials were adopted to ensure isoalloxazine's planarity. Similarly, Lennard-Jones (LJ) parameters were taken from homologous chemical motifs defined in CHARMM22. Also the residue R required to complement the thus established MM model of LF toward FMN was parameterized by homology with molecular structures provided by CHARMM22 (40).

As atomic partial charges we chose electrostatic potential derived (ESP) charges (37,60) calculated from fits to the DFT electrostatic potential in the surface region of LF. Here, the overall neutrality of certain substructures such as methyl groups was imposed. All calculated parameters characterizing the thus obtained MM force field of LF and FMN are shown in Figure S1 and Tables S1–S4 (see Supporting Information) available online.

MM-MD simulation setup. In our DFT/MM calculations on the vibrational spectra of LF in aqueous solution the MM fragment consisted of 1970 rigid water models described either by the "transferable interaction potential using three points (TIP3P)" or by the corresponding four point model (TIP4P) (61). The DFT fragment was confined to LF. To separate the electron density from the partial charges of the MM atoms in the surrounding solvent, each atom in the DFT fragment was surrounded by a LJ sphere (37). For this purpose we adopted, applying homology modeling as described above, for the various atoms of LF certain standard LJ parameters, which are provided by CHARMM22 for similar chemical motifs.

As our simulation system we chose an orthorhombic dodecahedron with periodic boundary conditions. Up to a distance of 18 Å the electrostatic interactions were calculated by means of the fast structure-adapted multipole method (62,63). At larger distances they were covered by the moving boundary reaction field approach introduced in Ref. (36). Here, we employed a dielectric constant of 80 for the distant aqueous continuum. Temperature and pressure were controlled by coupling a Berendsen (64) thermostat (time constant: 500 fs) and barostat (time constant: 5 ps; compressibility: 0.46 GPa^{-1}) to the MM fragment. For the integration of the MD we used the Verlet algorithm (65) with a time step of 1 fs.

In a pure MM-MD setting, *i.e.* by applying the MM force field introduced above to the LF, the LF was first embedded into pre-equilibrated boxes solely containing the respective water models. Here, all water molecules closer than 2 Å from any LF atom were removed. Subsequently the two solvent–solute systems were equilibrated for 1 ns in the *NPT* ensemble at a temperature of 300 K and a pressure of 1 atm. After these equilibrations we selected two sets of 20 snapshots from 200 ps trajectories to compute average ESP atomic partial charges of LF in solution in a DFT/MM setting. With the new solvent adapted partial charges the systems were once again equilibrated for another 200 ps in the *NPT* ensemble. Then the barostats were switched off and the systems were simulated in the *NVT* ensemble for another 500 ps from which $N_S = 50$ snapshots were taken at 10 ps time intervals for the DFT/MM computation of LF's liquid phase IR spectra. Radial distribution functions of the water molecules in the neighborhood of the flavin oxygen atoms O_2 and O_4 were extracted from subsequent 1 ns *NVT* simulations during which the system configuration was sampled every 200 fs.

IR spectra from DFT/MM. The DFT/MM spectra of LF embedded in our TIP3P and TIP4P models of an aqueous solution were calculated by applying the INMA protocols suggested by Schmitz and Tavan (51,52). The resulting spectra and the underlying snapshot ensembles will be labeled by the shortcuts T3 and T4, respectively. As indicated above, these spectra were generated from $N_S = 50$ snapshots chosen from 500 ps MM-MD trajectories at 10 ps intervals. Such intervals are long enough to enable exchanges of the hydrogen bonded water molecules forming the solvent cage (50) and, therefore, guarantee the statistical independence of the snapshots. Keeping the solvent cages fixed the LF molecule's structures of minimum potential energy were calculated by steepest gradient descent. The dipole gradients and Hessians were determined by finite differences like in the MT/BP calculations on the isolated molecule.

Automated mode matching. Due to the large numbers N_S of snapshots and $M = 3N$-6 of LF normal modes, it would be cumbersome to classify all these modes by visual inspection. Such a classification is necessary, because the spectral order and the composition of the normal modes vary from snapshot to snapshot. Therefore, we designed a procedure for an automated mode classification, which we will describe now.

For the ith snapshot ($i = 1, \ldots, N_S$) the energy minimized LF configuration

$$\mathbf{c}_i = (\mathbf{r}_{1,i}, \ldots, \mathbf{r}_{N,i})^T \in \mathbb{R}^{3N} \quad (1)$$

is given by the position vectors $\mathbf{r}_{n,i} \in \mathbb{R}^3$ of the LF atoms ($n = 1,\ldots,N$). From the mass weighted DFT/MM Hessian calculated for snapshot i we obtained the M eigenvectors $\mathbf{q}_{i,\alpha} \in \mathbb{R}^{3N}$, which are associated to the normal modes $\alpha = 1,\ldots, M$ with nonvanishing frequencies $\nu_{i,\alpha} \neq 0$. Because we are mainly interested in the mid-IR spectral region covering the range from 1150 to 1750 cm^{-1}, we restricted our analysis to the $m < M$ normal modes featuring frequencies $\nu_{i,\alpha}$ within that range. Thus, our automated mode classification was restricted to $m = 25$ normal modes α instead considering the complete set ($M = 87$).

If one wants to compare the modes $\mathbf{q}_{j,\beta}$ of another snapshot j with those of the given snapshot i, the configuration $\mathbf{c}_j$ has to be fitted to $\mathbf{c}_i$ by suitably rotating and translating the coordinates $\mathbf{r}_{n,j}$. For the configuration $\mathbf{c}_j$ this transformation may be formally written as

$$\mathbf{c}_j^i = \mathbf{X}_{ij}\,\mathbf{c}_j + \mathbf{t}_{ij}, \tag{2}$$

where $\mathbf{X}_{ij} \in \mathbb{R}^{3N} \times \mathbb{R}^{3N}$ and $\mathbf{t}_{ij} \in \mathbb{R}^{3N}$ rotate and translate, respectively, the coordinates $\mathbf{r}_{n,j}$ of all atoms in an identical fashion. The optimal fit is then given by the requirement that the Euclidean distance

$$d_{ij} = |\mathbf{c}_j^i - \mathbf{c}_i| \tag{3}$$

of the transformed configuration $\mathbf{c}_j^i$ from the reference configuration $\mathbf{c}_i$ becomes minimal. The rotations $\mathbf{X}_{ij}$ may then be used to transform the normalized eigenvectors $\mathbf{q}_{j,\beta}$ into the coordinate system of the best fit through

$$\mathbf{q}_{j,\beta}^i = \mathbf{X}_{ij}\mathbf{q}_{j,\beta}. \tag{4}$$

The transformed eigenvectors $\mathbf{q}_{j,\beta}^i$ can now be compared with the eigenvectors $\mathbf{q}_{i,\alpha}$ associated to the reference configuration $\mathbf{c}_i$ by forming scalar products and computing absolute values, *i.e.* by

$$(V_{ij})_{\alpha,\beta} = |\mathbf{q}_{i,\alpha} \cdot \mathbf{q}_{j,\beta}^i|. \tag{5}$$

If the modes $\mathbf{q}_{i,\alpha}$ and $\mathbf{q}_{j,\beta}^i$ derived for the snapshots i and j are as identical as the frequency orderings, then the $m \times m$ matrix $(V_{ij})_{\alpha,\beta}$ is the unit matrix $\delta_{\alpha,\beta}$. For differing frequency orderings and identical modes it is the matrix $\delta_{\alpha,p(\beta)}$, which derives from $\delta_{\alpha,\beta}$ by permuting the m columns with a permutation p. For similar modes one obtains a matrix with elements $0 \leq (V_{ij})_{\alpha,\beta} \leq 1$, which still resembles one of the permutation matrices just characterized and, in each row α, has a maximum entry $(V_{ij})_{\alpha,\beta(\alpha|i,j)}$ at a location $\beta(\alpha|i,j)$. The pointers $\beta(\alpha|i,j)$ represent a classification of the modes $\mathbf{q}_{j,\beta}^i$, $\beta = 1,\ldots,m$, obtained for the snapshot j in terms of the modes α of snapshot i.

Averaging the maximum entries over all rows yields a score

$$\langle V_{ij}\rangle = \frac{1}{m}\sum_{\alpha}(V_{ij})_{\alpha,\beta(\alpha|i,j)} \tag{6}$$

for the matching of the m modes $\mathbf{q}_{j,\beta}^i$ on the references $\mathbf{q}_{i,\alpha}$. This score can be calculated for all snapshots $j \neq i$ and the global score

$$\langle V_i\rangle = \frac{1}{N_S}\sum_{j\neq i}\langle V_{ij}\rangle, \qquad 0 \leq \langle V_i\rangle \leq 1 \tag{7}$$

then measures, how well the modes $\mathbf{q}_{i,\alpha}$ of snapshot i represent the modes of all other snapshots j. Choosing each of the N_S snapshots i as the reference, evaluating the global scores Eq. (7) and searching for the maximal value $\langle V_{\hat{i}}\rangle$ among the N_S results $\langle V_i\rangle$ yields the snapshot $\hat{i}$, whose modes $\mathbf{q}_{\beta}^{\mathrm{l}} \equiv \mathbf{q}_{\hat{i},\beta}$ feature the best representation of all modes $\mathbf{q}_{j,\beta}^{\hat{i}}$ calculated by DFT/MM for the liquid phase ("l") and matched onto the LF geometry $\mathbf{c}_{\hat{i}}^{\mathrm{l}}$ in snapshot $\hat{i}$. Clearly, we will obtain such a best representative $\hat{i}$ for each liquid model considered.

For each liquid model, the representative modes $\mathbf{q}_{\beta}^{l}$ were finally classified in terms of the DFT reference modes $\mathbf{q}_{\alpha}^{g}$, which had been calculated by MT/BP for the isolated LF and had been assigned to local modes by visualization as well as by calculation of the PED. Here, we identified the position $\beta(\alpha)$ of the maximum entry $(V^{\mathrm{gl}})_{\alpha,\beta(\alpha)}$ in each row α of the overlap matrix

$$(V^{\mathrm{gl}})_{\alpha,\beta} = |\mathbf{q}_{\alpha}^{\mathrm{g}} \cdot \mathbf{q}_{\beta}^{\mathrm{l}}| \tag{8}$$

and took this value $(V^{\mathrm{gl}})_{\alpha,\beta(\alpha)}$ as a measure for the quality, at which the liquid phase modes are classified by our automated procedure. For the evaluation of Eq. (8) the structure $\mathbf{c}_{\hat{i}}^{\mathrm{l}}$ of the liquid phase representative $\hat{i}$ was matched to the structure $\mathbf{c}^{\mathrm{g}}$ from the MT/BP treatment of the isolated LF ("g").

Liquid phase IR spectra. To obtain the liquid phase spectra from the DFT/MM snapshot ensembles we performed mode specific statistics. Based on the classification $\beta(\alpha|\hat{i},j)$ of the modes $\mathbf{q}_{j,\beta(\alpha|\hat{i},j)}$ in terms of the reference modes $\mathbf{q}_{\alpha}^{\mathrm{l}}$ we obtained the frequencies $\nu_{j,\beta(\alpha|\hat{i},j)}$ associated to the reference frequency $\nu_{i,\alpha}$ and calculated for each snapshot ensemble the average frequencies $\langle\nu_\alpha\rangle$ and the corresponding variances $\rho_\alpha^2 = \langle(\nu_\alpha - \langle\nu_\alpha\rangle)^2\rangle$. Furthermore, we evaluated the average mode intensities $\langle I_\alpha\rangle$. Following the arguments and assumptions presented in (22) and (52) and denoting the speed of light by c_0 the liquid phase IR spectrum then is the superposition

$$S(\nu) = \frac{c_0}{6}\sum_{\alpha}\frac{\langle I_\alpha\rangle}{\sqrt{2\pi\sigma_\alpha^2}}\exp\left[\frac{-(\nu - \langle\nu_\alpha\rangle)^2}{2\sigma_\alpha^2}\right] \tag{9}$$

of Gaussian peaks centered at the mode frequencies $\langle\nu_\alpha\rangle$ and inhomogeneously broadened with the scaled standard deviations $\sigma_\alpha = 0.6\rho_\alpha$. The scaled widths σ_α serve to approximately compensate the corresponding overestimate inherent to the INMA approach (52).

RESULTS AND DISCUSSION

As a decisive prerequisite for the intended comparison of our DFT/MM descriptions [cf Eq. (9)] with experimental IR spectra of flavins in solution (10–12,14–19), we have to make sure that we can reliably extract the peak positions $\langle\nu_\alpha\rangle$ of the IR bands from the INMA ensembles, which cover the frequencies $\nu_{i,\beta}$ and normal modes $\mathbf{q}_{i,\beta}$ calculated for $N_S = 50$ snapshots i. Here, the challenge is that, within a snapshot ensemble, the normal mode frequencies and compositions exhibit strong variations, which are caused by differences of the solvent cages.

The automated procedure for normal mode classification introduced above in Materials and Methods represents our way of meeting this challenge and we will first show that it actually enables a high quality mode classification. As our sample data set we chose the INMA results associated with the T3 snapshot ensemble described in Materials and Methods. T3 comprises data on snapshots taken from a 500 ps MM-MD simulation of LF in TIP3P water.

Quality of mode matching

For each snapshot i in T3 we evaluated the score $\langle V_i\rangle$ defined by Eq. (7), which measures the quality, at which the 25 mid-IR normal modes of snapshot i can represent the corresponding normal modes of all other snapshots j. Subsequently, we ordered the snapshots according to decreasing $\langle V_i\rangle$, *i.e.* $[\langle V_i\rangle > \langle V_j\rangle \Rightarrow i < j]$.

Figure 2 shows the resulting scores $\langle V_i\rangle$ for the properly ordered snapshots i in T3 and demonstrates by the small variation of the values $\langle V_i\rangle$ that the modes of all snapshots can represent those of the T3 ensemble at a comparable quality. In

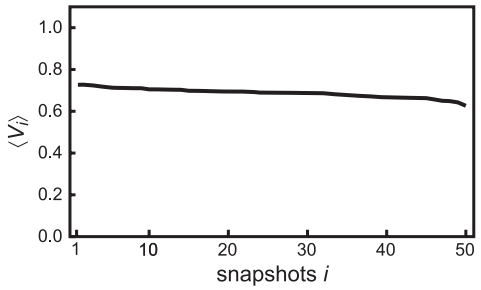


Figure 2. Mode matching scores $\langle V_i\rangle$ as defined by Eq. (7) for the 50 properly ordered snapshots i in T3.

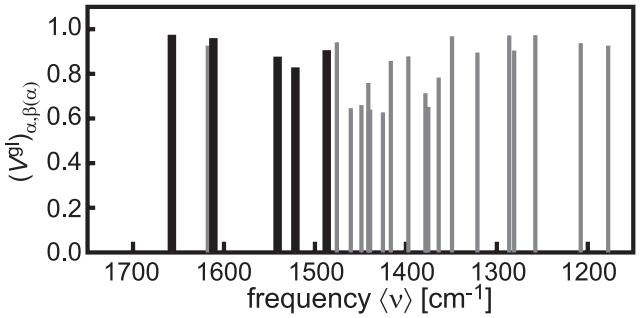


Figure 3. Maximum mode overlaps $(V^{gl})_{\alpha,\beta(\alpha)}$ [cf Eq. (9)] between the normal modes of the isolated lumiflavin and those of the best representative for the snapshot ensemble T3 calculated by DFT and DFT/MM, respectively, and plotted over the liquid phase frequencies $\langle \nu_\alpha \rangle$. Highlighted in black are the mode matchings of the five marker bands used further below in comparisons with experimental data. DFT/MM, density functional theory/molecular mechanics.

fact, the score of the best representative ($i = 1$) is only by 5.7% larger than the average score $\langle\langle V_i \rangle\rangle = 0.6873$.

Next we used the overlap matrix $(V^{gl})_{\alpha,\beta}$ defined by Eq. (8) to compare the normal modes $\mathbf{q}^l_\beta = \mathbf{q}_{l,\beta}$ of the best representative with the reference modes $\mathbf{q}^g_\alpha$ determined from the DFT treatment of the isolated LF. Figure 3 plots the maximum entries $(V^{gl})_{\alpha,\beta(\alpha)}$ found in the $m = 25$ rows α of the matrix over the unscaled average liquid phase frequencies $\langle \nu_\alpha \rangle \in [1150\text{cm}^{-1}, 1750\text{cm}^{-1}]$.

Figure 3 reveals that for most modes α the maximum overlaps $(V^{gl})_{\alpha,\beta(\alpha)}$ are large and close to one. Correspondingly the average value is 0.83. Only in the spectral region between 1350 and 1450 cm^{-1} the overlaps are 14% below the average such that, here, the classification of the liquid phase modes in terms of the DFT results on the isolated molecule is a little less accurate. From the PED classification of the DFT modes one finds (data not shown) that many modes in this region belong to methyl vibrations.

Further below, in comparisons between computed and observed vibrational spectra, we will concentrate on five high frequency marker bands, whose mode matchings are highlighted in Fig. 3 by black bars. The corresponding bands dominate the experimental spectra of oxidized flavins in the mid-IR spectral region and have the additional advantage to be detectable in spectroscopic studies on flavins embedded in protein binding pockets (9,66). For the corresponding modes the overlaps are seen to be quite large indicating that their classification in liquid phase calculations should be straight forward.

The latter statement is corroborated by Fig. 4, which complements the quantitative mode match of Fig. 3, by graphical representations of the atomic displacements characterizing the five marker modes in the isolated LF (column MT/BP) and in the liquid phase representative (column T3), respectively. Line wise comparisons of the mode pictures immediately demonstrate that the mode compositions hardly change upon transfer of LF from the vacuum into the TIP3P liquid. In contrast, the frequencies are quite strongly redshifted by this transfer.

According to Fig. 4 the two highest frequency marker bands belong to stretching vibrations of the two carbonyl groups $C_4=O_4$ and $C_2=O_2$, whereas the three lower frequency

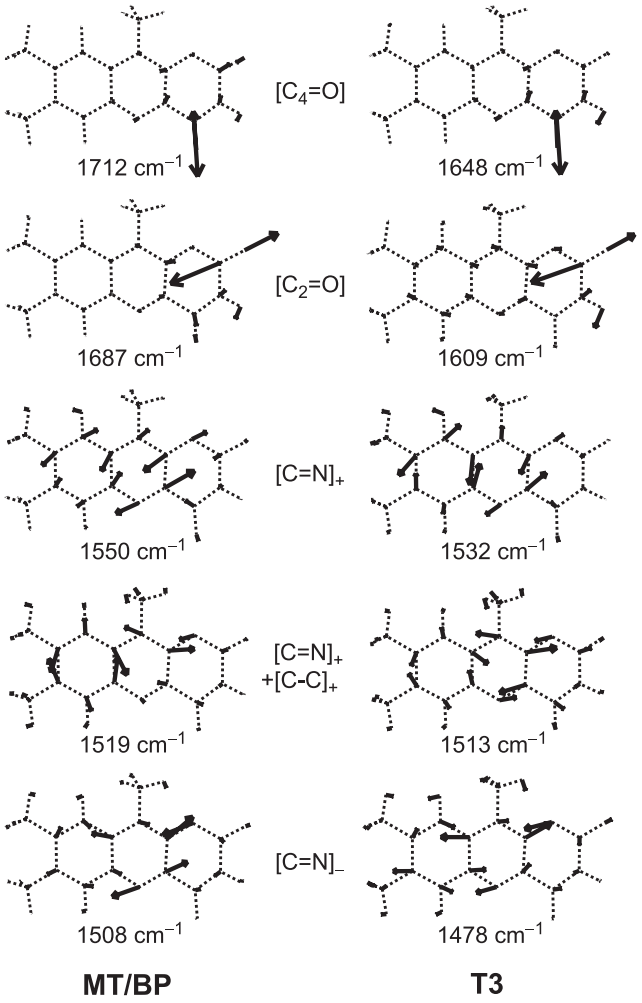


Figure 4. Drawings of the modes computed for the five marker bands. MT/BP: isolated lumiflavin (LF); T3: ensemble representative for LF in TIP3P water. The symbols at the center are the mode labels explained in the text. The frequencies are the DFT results ν_α (MT/BP) and the instantaneous normal mode analyses averages $\langle \nu_\alpha \rangle$ from DFT/MM (T3). DFT/MM, density functional theory/molecular mechanics.

marker bands belong to C=N and C-C stretches localized within the isoalloxazine moiety. In the mode labels listed in the central column of Fig. 4 the subscripts "+" and "−" indicate in-phase and out-of-phase relations, respectively, among the various bond stretches contributing to the normal modes. Note here that Table S5 in Section 3 of the Supporting Information contains a characterization of all those fourteen mid-IR normal modes in terms of PED, for which experimental data and results of a previous normal mode analysis (13,21) are available.

In summary, the above mode matching results have demonstrated that the automated classification of the modes contained in the INMA ensembles is quite reliable. Particularly for the marker bands it allows us to identify the mode compositions and the peak positions $\langle V_\alpha \rangle$ at a high accuracy. Thus it enables a vibrational analysis of the IR spectra observed for fully oxidized flavins in solution (10–12,14–19).

Vibrational analysis of dyes in solution: Role of frequency scaling

For a quantitative comparison of a calculated with an observed vibrational spectrum one generally needs a scaling factor, which serves to minimize the inevitable difference between the two sets of numbers [see *e.g.* Ref. (30)]. Such a scaling factor is method specific. In a DFT treatment of an isolated molecule, for instance, it is valid for a combination of certain functionals with a specific basis set. It is usually determined for a set of small reference molecules with well-known gas phase spectra. Subsequently it enables quite accurate predictions of the vibrational gas phase spectra of other molecules with similar chemical composition and size.

Clearly one needs a suitable factor for frequency scaling also in any attempt to describe the solution spectra of a large dye molecule in a DFT/MM setting. Here, one would preferentially like to use a scaling factor applicable to the prediction of the gas phase spectra by the chosen DFT approach, because then the comparison of DFT/MM spectra for flavins in solution or in protein binding pockets with corresponding observations can enable conclusions on the quality, at which the MM model of the environment can describe the solvatochromic shifts in the vibrational spectra. For flavin dyes, however, vibrational gas phase spectra are (to our best knowledge) unknown and, therefore, no data basis for evaluating a scaling factor for the specific MT/BP approach used by us (see Materials and Methods for details) is available. Because we eventually aim at computing the IR spectra of flavin chromophores in blue light receptor proteins and will need a suitable scaling factor for this purpose, we are now forced to employ the well characterized vibrational spectra of flavins in aqueous solution (10–12,14–18) for the computation of the desired factor.

A scaling factor applicable to a DFT/MM description of solution spectra will not only depend on the applied DFT approach, *e.g.* on the MT/BP method used by us, but also on the specifics of the MM model applied to the solvent environment.

Comparison of unscaled DFT/MM frequencies and IR solution data

To study the issue of proper scaling we decided to use two different and well-known potential functions for the water molecules in the liquid, *i.e.* the Jorgensen's TIP3P and TIP4P models (61). These simple models have dipole moments of 2.35 D and 2.18 D, respectively. At the temperature of 298 K and the density of 0.99 g cm^{-3} a TIP3P liquid has a dielectric constant of about 97 (67,68), which is larger than that of liquid water, which is 78 (69) at these conditions. A TIP4P liquid, in contrast, has a much smaller dielectric constant of about 55 (70–72). On the other hand, the quadrupole moment of the TIP4P model is about 23% larger than that of the TIP3P model and will, therefore, lead to a stronger short range order within the solvation shells of the water molecules in the liquid phase.

To assess the influence of the solvent on the vibrational spectra, we first consider Fig. 5, which compares for the five marker bands characterized in Fig. 4 the unscaled DFT/MM frequencies $\langle \nu \rangle$ of LF in TIP3P (T3) and TIP4P (T4) water, on the one hand, with the vibrational frequencies observed for FMN and FAD in water (H_2O) by resonance Raman (RR) (10,11,14) and IR (16,17) spectroscopy, on the other. Here, the frequencies shown in the column H_2O are averages of the frequencies given in the quoted experimental studies. In addition, Fig. 5 provides IR frequencies for RF in dimethyl sulfoxide (DMSO) (19) and in deuterated acetonitrile (CD_3CN) (73). Note here that the dielectric constant of DMSO is 49 (74) and that of acetonitrile 39 (75). Thus, DMSO is only a little less polar than our TIP4P water model.

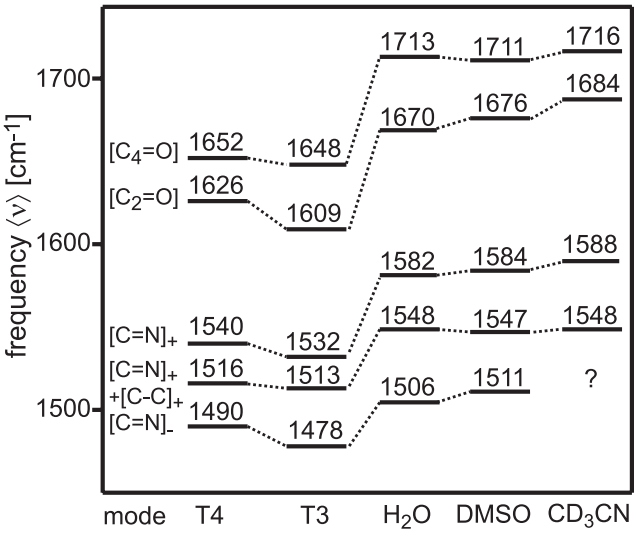


Figure 5. Our unscaled DFT/MM frequencies $\langle \nu \rangle$ calculated for lumiflavin in TIP3P (T3) and TIP4P (T4) water are compared with experimental frequencies observed for fully oxidized FMN/FAD in water (H_2O) (10,11,14,16,17) as well as for riboflavin in dimethyl sulfoxide (DMSO) (19) and acetonitrile (CD_3CN) (73). The experimental H_2O data were calculated by us as averages over the frequencies given in the quoted studies. DFT/MM, density functional theory/molecular mechanics; FMN/FAD, flavin mononucleotide/flavin adenine dinucleotide.

Comparing first in the right part of Fig. 5 the frequencies of the two carbonyl bands measured for the fully oxidized flavins in the protic and highly polar solvent H_2O with the corresponding frequencies observed in the aprotic and increasingly less polar solvents DMSO and CD_3CN, one immediately recognizes blueshifts of the average carbonyl frequencies $\langle \nu_{c=o} \rangle \equiv (\nu_{[c_4=o]} + \nu_{[c_2=o]})/2$ with decreasing polarity of the solvent. In the transition from H_2O to DMSO the shift is 2 cm^{-1} and to CD_3CN it is 8 cm^{-1}.

For the [C_2 = O] band the blueshift, which is associated with the transfer from H_2O to the much less polar acetonitrile, measures 14 cm^{-1} and, thus, is much larger than the 3 cm^{-1} blueshift of the [C_4 = O] band. In the transition from H_2O to DMSO the [C_4 = O] band is even shifted by 2 cm^{-1} to the red. Therefore, the 2 cm^{-1} blueshift of the average carbonyl stretching frequency $\langle \nu_{c=o} \rangle$ is exclusively due to the 6 cm^{-1} blueshift of the [C_2 = O] band. Considering finally the spectral locations of the three bands with strong C = N stretching contributions one recognizes smaller blueshifts upon transition from H_2O to the less polar solvents or even a near invariance ([C = N]$_+$).

Already the first glance at the unscaled DFT/MM frequencies displayed in the left part of Fig. 5 reveals that all computed values are predicted by about 45 cm^{-1} to the red

of the experimental references. Thus, our DFT/MM force field systematically underestimates all stretching force constants indicating that a scaling factor larger than one will be required for quantitative comparisons with experimental frequencies.

Considering next in more detail the influence of the two different solvent models TIP3P and TIP4P, respectively, on the DFT/MM predictions of the spectral band positions one sees that also in our theoretical model the C=O bands are blueshifted upon transfer from a highly polar (T3) to a less polar (T4) solvent. Like in the experimental H_2O to CD_3CN transition also in the transfer of LF from T3 to T4 the predicted blueshift is larger for the [C_2=O] (17 cm^{-1}) than for the [C_4=O] band (4 cm^{-1}). This difference suggests that the $C_2=O_2$ group interacts more strongly with a polar solvent than the $C_4=O_4$ group (we will check this suggestion further below).

The larger influence of the solvent polarity on the [C_2=O] frequency implies that the spectral gap $\Delta\nu_{c=o} \equiv \nu_{[c4=o]} - \nu_{[c2=o]}$ increases with the solvent polarity, *i.e.* in the transition from T4 to T3. Experimentally $\Delta\nu_{c=o}$ increases from 32 cm^{-1} in CD_3CN, to 35 cm^{-1} in DMSO and to 43 cm^{-1} in H_2O, whereas according to our DFT/MM results it increases from 26 cm^{-1} for T4 to 39 cm^{-1} for T3. If one considers the proper description of $\Delta\nu_{c=o}$ as an important goal for a DFT/MM treatment of LF in aqueous solution, then the frequencies displayed in Fig. 5 argue that we should choose the highly polar TIP3P as our solvent model.

For the bands with C=N stretching contributions the DFT/MM model calculations predict blueshifts upon transfer from TIP3P into the less polar model TIP4P, which is in nice agreement with the experimental findings. Here, one should bear in mind, however, that the T3 to T4 transition is one among water models, whose polarities (as measured by the dipole moments) and protic strengths (as measured by the quadrupole moments) differ in opposite ways, whereas the experimental transitions move from a polar and protic solvent composed of small molecules (H_2O) to less polar and aprotic solvents composed of larger molecules.

To finalize the choice of our MM model for water we minimized for the five marker bands the root mean square deviation (RMSD) between the experimental H_2O data and our scaled DFT/MM frequencies calculated for LF within the T3 and T4 settings. We found for T4 that the scaling factor 1.0254 reduces the RMSD to the minimal value 13.1 cm^{-1}, whereas for T3 the minimal RMSD is 12.2 cm^{-1} at the scaling factor 1.0310. Thus, the optimally scaled T3 treatment models the experimental data for all marker bands better than T4, which is why we choose the optimally scaled T3 results from now on in comparisons with vibrational data on flavins in aqueous solution.

Solvation of LF's carbonyl groups

The above suggestion that the $C_2=O_2$ group features stronger interactions with a polar solvent than the $C_4=O_4$ group can be checked by comparing the associated radial distribution functions $g_{OH}(r)$ of the hydrogen atoms H, which are contained in the liquid water near the carbonyl oxygen atoms O. These distribution functions have been extracted from the 1 ns MM-MD simulation of LF in TIP3P water as described in Materials and Methods section.

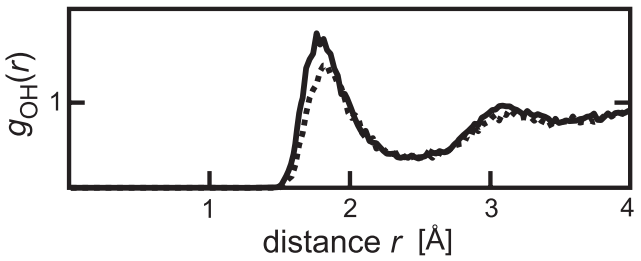


Figure 6. Radial distribution functions $g_{OH}(r)$ of the water hydrogen atoms surrounding the flavin oxygen atoms O_2 (solid line) and O_4 (dashed line) of lumiflavin in TIP3P water.

Figure 6 shows the distribution functions $g_{OH}(r)$ thus obtained for the atoms O_2 (solid line) and O_4 (dashed line) of the LF molecule in TIP3P water. The figure clearly demonstrates that more H atoms are found near the oxygen O_2 than near O_4. Integrating the functions $g_{OH}(r)$ up to a distance of 2.5 Å and, thus, covering the first solvation shell one finds on average 2.6 hydrogen atoms near O_2 but only 2.3 hydrogen atoms in the vicinity of O_4.

The differently strong solvation of the two flavin carbonyl groups is explained by our DFT and DFT/MM results on the isolated and the solvated LF molecule, which comprise ESP atomic charges (37,60). According to these partial charges (see Table S1) the dipole moment $\mathbf{p}_2$ of the $C_2=O_2$ group is by about a factor of two larger than the dipole moment $\mathbf{p}_4$ of the $C_4=O_4$ group. This factor is independent of the environment as both dipole moments are predicted to be 31% larger for the solvated than for the isolated flavin. Correspondingly, the stronger dipole $\mathbf{p}_2$ can bind and order the dipolar water molecules much better than the weaker dipole $\mathbf{p}_4$.

Conversely, the enhanced ordering in the solvation shell of O_2 as compared to that of O_4 will cause a stronger reaction field at O_2 than at O_4. The different strengths of these reaction fields, which are generated by the respective solvation shells, are the reason why the solvation shifts are larger for the [C_2=O] than for the [C_4=O] mode [for a detailed explanation of such an effect see Ref. (54)].

Choice of Lennard-Jones potentials

The variation of the solvent model discussed in connection with Fig. 5 above is by no means the only modeling choice, on which one has to decide in a DFT/MM description of solution spectra. Another subtle issue is the choice of the LJ potentials surrounding LF's carbonyl oxygen atoms, which also may affect the carbonyl solvation. This insight leads to the question, to what extent the calculated solvation patterns and the solution spectra are influenced by this choice.

The connection between LJ parameters and local solvation patterns has been investigated by Freindorf *et al.* (76) in a DFT/MM setting. For a series of small organic molecules described by DFT and embedded in TIP3P water these authors studied the hydrogen bonding between the DFT fragment and the surrounding MM water. From these DFT/MM calculations the authors determined optimized LJ potentials, which keep the atoms of the DFT fragments at proper hydrogen bonding distances and interaction energies with respect to the surrounding MM water molecules. For essentially all atom types occurring in the studied DFT fragments the resulting LJ

parameters were found to differ from the standard values prescribed by the applied AMBER (77) force field.

As we have seen above, LF features two carbonyl groups capable of forming strong hydrogen bonds with the surrounding water. In view of the quoted results we were uncertain, whether the standard LJ parameters suggested by our CHARMM22 force field (40) for the carbonyl oxygen atoms lead to a reasonable description of the hydrogen bonding at these locations and, therefore, to an accurate DFT/MM calculation of the solvatochromic shifts determining the carbonyl frequencies (20,54).

We checked this parameterization issue by calculating additional INMA spectra of LF in TIP3P and TIP4P water using altered LJ parameters for the LF oxygen atoms O_2 and O_4. Here we chose the values, which are suggested in CHARMM22 (40) for oxygens in hydroxyl groups. These atom types feature increased van der Waals radii. As documented by Figures S2 and S3 in Section 2 of the Supporting Information, the increased van der Waals radii of the oxygen atoms O_2 and O_4 actually cause weaker solvations of the two carbonyl groups both in TIP3P and TIP4P water and corresponding blueshifts (10–15 cm^{-1}) of the carbonyl stretching frequencies while leaving the spectral positions of other bands nearly invariant. However, concerning the description of the experimental spectra (10,11,14,16,17) of fully oxidized flavins in water the altered LJ parameter did not seem to render better results. Therefore, we decided to stick to the combination T3 of the TIP3P water model with the original LJ parameters suggested by CHARMM22 for oxygen atoms in C=O groups.

Comparison to a continuum method

The conventional approach to estimate solvation effects in quantum chemical calculations is the use of continuum models such as Cosmo (33). Using this approach, the B3LYP functional (78,79) and the TZVP basis set (58), Klaumünzer *et al.* (31) have recently presented a DFT description of the solvation effects in the IR spectra of RF upon transfer from the gas phase into DMSO. Despite the differences of DFT functionals, basis sets and solvent polarities it is interesting to compare the solvation effects predicted by Klaumünzer *et al.* with a continuum model for the mildly polar DMSO with our explicit and microscopic modeling of a strongly polar aqueous solution.

This comparison is shown in Fig. 7. It is simplified by the availability of experimental spectra on flavins in H_2O and on RF in DMSO (19). To guarantee the fairness of the comparison we have applied analogous scaling procedures to the two computational methods, which implies for the frequencies of the five marker bands calculated by Klaumünzer *et al.* (31) for RF in DMSO that we have minimized the RMSD from the associated IR data (19) and have applied the resulting scaling factor of 0.993 also to their vacuum results.

Concerning the gas phase IR spectra of LF and RF, the scaled frequencies displayed in columns MT/BP and B3LYP of Fig. 7 represent predictions. These two DFT predictions are remarkably similar with a RMSD of only 6.6 cm^{-1} implying that the differences of the residues R attached to the isoalloxazine atom N_{10} and the choices of the DFT descriptions are of minor importance. The qualities, at which

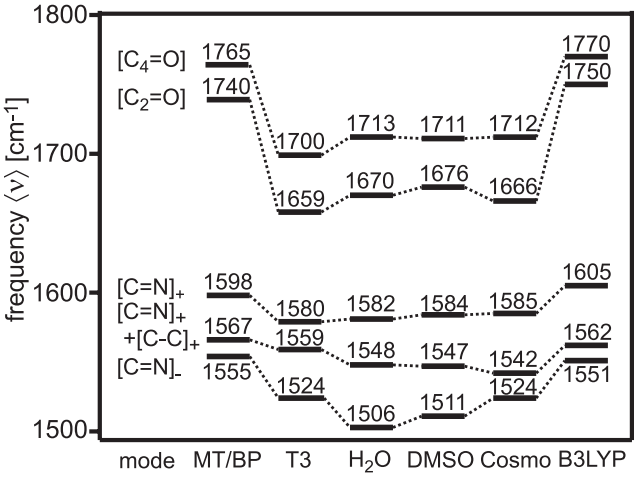


Figure 7. Frequencies of marker bands scaled with a factor of 1.0310, which we calculated for lumiflavin in the vacuum (MT/BP) or embedded in a TIP3P water box (T3), are compared with the experimental data on flavins in water [H_2O, c.f. the caption to Fig. 5]. The marker band frequencies calculated by Klaumünzer *et al.* (31) for isolated riboflavin (RF) (B3LYP) and for RF embedded in a dimethyl sulfoxide (DMSO) continuum with the dielectric constant 46.68 (Cosmo) were scaled with a factor of 0.993 for optimal comparison with infrared data (19) on RF in DMSO.

the two computational approaches describe the respective experimental solution data, can be compared by considering the RMSDs 12.2 cm^{-1} (T3) and 7.7 cm^{-1} (Cosmo) of the calculated from the corresponding experimental frequencies.

Thus, our DFT/MM-INMA approach performs on the spectra of fully oxidized flavins in H_2O slightly worse than B3LYP/Cosmo on the spectra of RF in DMSO. Because both DFT methods describe the spectra of isolated flavins at a similar quality, the suboptimal performance of the DFT/MM approach must be assigned to deficiencies of the applied water model. As we have seen further above, the TIP3P model substantially overestimates the dielectric constant of water [97 *vs* 78]. Therefore, the average carbonyl frequency $\langle \nu_{c=o} \rangle$ predicted by T3 underestimates the experimental H_2O value by 12 cm^{-1}. Correspondingly, the spectral gap between the C=N and the C=O modes is systematically underestimated by 9 cm^{-1}. In contrast, no systematic deviation appears in the Cosmo description of the DMSO data.

On the other hand, if one looks at the spectral gap $\Delta\nu_{c=o}$ between the two carbonyl bands one recognizes that DFT/MM (T3) apparently performs with 41 cm^{-1} better on the experimental data for H_2O (43 cm^{-1}) than B3LYP/Cosmo with 46 cm^{-1} on the IR data for DMSO (35 cm^{-1}). Thus, it may well be that the microscopic modeling of the solvation shells furnished by the DFT/MM approach produces, despite the use of a too polar water model, slightly more realistic solvation shifts at the carbonyl positions than the much simpler continuum method. Nevertheless, one should definitely use a more realistic MM water model in future attempts to describe the IR spectra of molecules embedded in aqueous solutions.

In summary, because the DFT/MM and DFT/Cosmo descriptions of solvation effects are very similar, our DFT/MM approach is certainly at least as much "state of

the art'' for isotropic solutions as well-established continuum approaches. In contrast to the latter, however, it can be applied also to nonisotropic environments such as protein binding pockets. In addition, the DFT/MM approach yields information on the inhomogeneously broadened line widths of the IR bands, which is outside the scope of mean-field type DFT continuum descriptions.

The IR spectrum of LF in water as predicted by DFT/MM

To illustrate the differences between the line spectra typical for DFT treatments and inhomogeneously broadened spectra resulting from DFT/MM-INMA calculations Fig. 8 compares (a) the line spectrum calculated by MT/BP for isolated LF with (b) the prediction Eq. (9) for the IR spectrum of LF in water constructed from the snapshot ensemble T3 and its INMA treatment. Note here, that Section 4 of the Supporting Information provides material [Eq. (S1) and Figure S4] showing that the assumptions underlying Eq. (9) are valid in this case.

Apart from the solvatochromic shifts of the five marker bands, which were already discussed in connection with Fig. 5, the comparison of Fig. 8a,b demonstrates that the C=O stretches exhibit a much stronger inhomogeneous line broadening than the C=N and C-C stretches. Note here that the small sharp peak at 1652 cm^{-1} distorting the broad and intense $[C_2=O]$ band centered at 1659 cm^{-1} belongs to a C-C stretching vibration localized within ring I of isoalloxazine (cf Fig. 1), which generally has a small IR intensity and, therefore, has been omitted in our previous analyses. Furthermore, in the liquid phase the combination mode $[C=N]_+ + [C-C]_+$ (cf Fig. 4) generates the dominant IR peak at 1559 cm^{-1} whereas the higher frequency band at 1580 cm^{-1} (mode $[C=N]_+$), to which the MT/BP description of the isolated LF assigns a similar intensity, becomes relatively small.

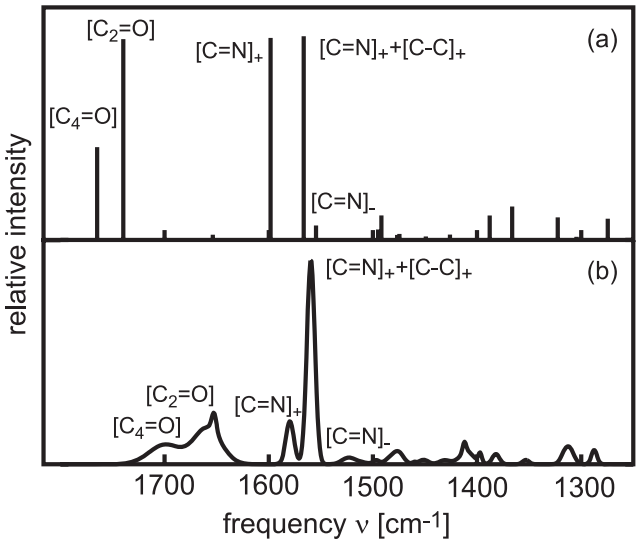


Figure 8. The MT/BP prediction for the line spectrum of isolated lumiflavin (LF) (a) is compared with the DFT/MM description Eq. (9) of the inhomogeneously broadened solution spectrum of LF in water (b). Note that the TIP3P water is transparent and is decoupled from the LF, because TIP3P is a stiff model. DFT/MM, density functional theory/molecular mechanics.

Unfortunately, well-resolved absolute IR spectra of flavins in H_2O or D_2O, to which we could visually compare our INMA spectrum in Fig. 8b, are rare. In Fig. 4 of Ref. (15), however, there is a quite nice IR spectrum of oxidized FAD in D_2O, which exhibits a pattern of marker band intensities that closely resembles our INMA description on LF in aqueous solution. In particular, also here the dominant peak is located at 1546 cm^{-1} and a much smaller peak at 1578 cm^{-1}. Like in Fig. 8b the two C=O bands at 1701 and 1640 cm^{-1} are relatively broad with the lower frequency band carrying a larger intensity and featuring a sharp side peak at 1621 cm^{-1}. The latter becomes much less intense in FMN (15), in which case the similarities between our result and observations should become even more striking. Regrettably, the FMN spectrum was not shown in Ref. (15) but the shown FAD spectrum nevertheless indicates that the intensities and band patterns calculated by us using Eq. (9) seem to be quite realistic.

The visual impressions following from the inspection of Fig. 8 are corroborated by the marker band data in Table 1, which provide information on the band widths through the standard deviations σ and compare the intensities calculated for LF in TIP3P water with those obtained for the isolated chromophore through the quotients $\langle I\rangle/I_g$. Upon transition into the condensed phase the two carbonyl bands are seen to become 2–3 times more strongly broadened than the C=N and C-C bands. Furthermore, they gain a little intensity. This intensity gain of the two carbonyl bands can be nicely explained by the fact that the two carbonyl dipole moments $\mathbf{p}_2$ and $\mathbf{p}_4$ increase by 31% upon transfer of an isolated flavin dye into aqueous solution (cf the discussion of Fig. 6). Furthermore, our above finding that, independently of the environment, $\mathbf{p}_2$ is about two times larger than $\mathbf{p}_4$ explains why already in the isolated compound the intensities I_g of the modes $[C_2=O]$ and $[C_4=O]$ differ by a factor of two (cf Fig. 8a). Finally, the data in Table 1 particularly emphasize that the transition into the liquid implies for the combination mode $[C=N]_+ + [C-C]_+$ an intensity gain by about a factor of two, whereas the mode $[C=N]_+$ suffers an intensity loss of about a factor 1/2. Thus the near equality of intensities predicted for these modes in the isolated compound does not prevail in the aqueous solution.

Quality checks using isotopic shifts

Compared to dynamics-based methods for the computation of solution spectra, the INMA approach has the unique advantage to make the computation of isotope effects on the IR spectra of the solvated compounds relatively straight forward.

Table 1. Instantaneous normal mode analyses results for the five marker bands of lumiflavin in TIP3P water: Average frequencies $\langle\nu\rangle$ scaled by 1.031; Gausssian line widths σ; ratios $\langle I\rangle/I_g$ of the average condensed phase intensities $\langle I\rangle$ to the intensities I_g of the isolated molecule.

Mode	$\langle\nu\rangle$ / cm^{-1}	σ/cm^{-1}	$\langle I\rangle/I_g$
$[C_4=O]$	1700	15.40	1.65
$[C_2=O]$	1659	12.99	1.19
$[C=N]_+$	1580	3.84	0.42
$[C=N]_+ + [C-C]_+$	1559	3.82	1.95
$[C=N]_-$	1524	5.98	1.29

Of course, for such studies our automated procedure for mode classification is a decisive prerequisite, because the data analysis rapidly becomes unmanageable without such a tool. Studies comparing calculated with observed isotopic IR band shifts give deeper insights into the merits of a computational procedure, because they enable the analysis of mode compositions. Fortunately there are many spectral data (10,11,21) on isotopically labeled flavins with which we can compare our descriptions.

To be specific, we did consider LF isotopically labeled with ^{13}C, ^{15}N and ^{2}H nuclei at the positions C_2, $C_{4,0A}$, H_3, $N_{1,3}$ and N_5 within the isoalloxazine moiety, because IR (11,21) and RR (10) spectra are available for these isotopomers of LF and FAD, respectively. When adopting observed band shifts from these studies, we generally took the RR data, because only these measurements refer to a flavin (FAD) in aqueous solution. The IR data, in contrast, were measured for solid state probes of flavins dried on KBr disks. Because the [C$_2$=O] mode is not detected in the RR measurements, we were forced to adopt the solid state IR isotope shifts for this mode. Only for one isotope label, *i.e.* for the deuteration of the N_3H_3 group, there are many IR data on fully oxidized flavins in D_2O (10,12,15,18). To determine an experimental consensus on the spectral shifts induced by deuteration we formed averages of the available shift data. By adding the observed shifts to our consensus H_2O marker band spectrum shown in Fig. 5 we constructed "experimental" marker band spectra of isotopically labeled flavins in aqueous solution.

Figure 9 compares for the five chosen isotopomers the calculated (black) and experimental (gray) frequencies of the marker bands. The numbers at the frequency terms of the isotopomers are the isotope shifts. Note, that the Supporting Information provides in Table S6 a detailed list of the isotope shifts calculated for a series of further flavin bands.

Already the first glance at Fig. 9 demonstrates that the pattern of calculated level shifts (black lines) nicely agrees with the observed pattern. The associated RMSD is 4.5 cm^{-1}. A closer inspection of the details shows certain expected frequency shifts.

The label $^{13}C_2$ mainly shifts the calculated [C$_2$=O] band by 36 cm^{-1} to the red, which closely matches the experimental shift of 32 cm^{-1}. Thus, this band is dominated by the C$_2$=O$_2$ vibration as illustrated by Fig. 4. Similarly, the substitution $^{13}C_{4,0A}$ is predicted to shift the [C$_4$=O] band by 38 cm^{-1} to the red. Here, the experimental value of 51 cm^{-1} is somewhat larger, but theory and experiment agree that this substitution leads to sizable redshifts also of the C-C and C=N stretching bands.

Interesting are the simultaneous redshifts of the carbonyl bands upon deuteration of N_3, because it indicates that the N_3H_3 bending mode couples to the stretches of both neighboring C=O groups in the native compound. Deuteration strongly shifts this bending mode toward lower frequencies and, thus, decouples the bend from the C=O stretches. For the native compound the coupling is clearly revealed by the mode patterns depicted in Fig. 4, according to which the N_3H_3 bend makes a larger contribution to [C$_2$=O] than to [C$_4$=O]. Correspondingly, the [C$_2$=O] band is predicted to shift more strongly (19 cm^{-1}) than the [C$_4$=O] band (10 cm^{-1}). Here the experimental redshifts are 27 cm^{-1} and 13 cm^{-1}, respectively.

On the other hand, the isotopic labels at $N_{1,3}$ and N_5 entail much less dramatic variations of the marker band positions,

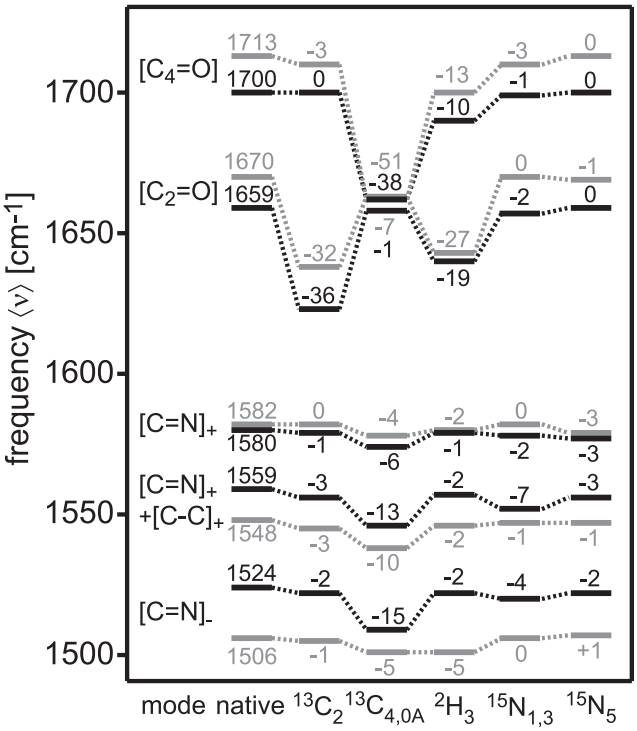


Figure 9. Comparison of scaled DFT/MM (black) and experimental frequencies (gray) for native and isotopically labeled flavins. The experimental reference data were extracted from the literature (10,11,21) as explained in the text. The column labels $^{13}C_2$, $^{13}C_{4,0A}$, $^{2}H_3$, $^{15}N_{1,3}$ and $^{15}N_5$ indicate the isotopic substitutions. The column "native" repeats the scaled DFT/MM-INMA results "T3" from Fig. 7 and the consensus experimental data "H_2O" given in Fig. 5. DFT/MM, density functional theory/molecular mechanics; INMA, instantaneous normal mode analyses.

indicating for the C = N stretches that these modes are strongly delocalized (cf Fig. 4).

SUMMARY AND OUTLOOK

The above discussion of isotopic shifts has demonstrated by the close match of experimental and calculated values, that the DFT/MM description provides a precise insight into the normal mode compositions of oxidized flavins in water. The preceding results have additionally shown that it can, after proper scaling, not only quite accurately predict the frequencies of the IR bands for such flavins, but also their intensities and line-shapes.

With the scaling factor of 1.031 determined in our study, we have established a DFT/MM approach that should be suited to compute the IR spectra of fully oxidized flavins in protein environments like, *e.g.* BLUF domains at a reasonable accuracy. Here, "reasonable accuracy" means that the RMSD between experimental spectra and DFT/MM predictions should not exceed the value of 10 cm^{-1} by a large margin – provided that the molecular model of the solvated protein is close to the solution structure [cf *e.g.* Ref. (38) for a detailed discussion of this issue]. It will be also interesting to see how our scaled DFT/MM-INMA method will perform on the condensed phase IR spectra of flavins in other redox states. If the expected accuracy can be demonstrated to prevail also for such flavins in solution, the application of our scaled

DFT/MM-INMA approach to protein structures and the comparison with time-resolved IR data can provide insights into structural details of the protein–chromophore interactions and their changes during the light-induced functional photo cycles.

Acknowledgements—This work was supported by the Forschergruppe 526 "Sensory Blue Light Receptors" of the Deutsche Forschungsgemeinschaft (DFG/FOR526).

SUPPORTING INFORMATION

Additional Supporting Information may be found in the online version of this article:

On nine pages the Supporting Information provide in four sections a total of four figures (S1–S4), six tables (S1–S6), one equation (S1) and various pieces of text explaining and documenting (1) our CHARMM22 type force field for LF and FMN; (2) the consequences of choosing different LJ potentials for the carbonyl oxygens on the carbonyl solvation and the LF solution spectra; (3) the assignment and isotopic shifts for nine additional vibrational bands of LF in aqueous solution; and (4) the validity of the Gaussian model Eq. (9) for the DFT/MM-IR spectra derived by the INMA procedure. This material can be found as PDF document.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting information supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

REFERENCES

1. Berg, J. M., J. L. Tymoczko and L. Stryer (2003) *Biochemie*, 5th edn. Spektrum Akademischer Verlag GmbH, Heidelberg, Berlin.
2. Mattevi, A. (2006) To be or not to be an oxidase: Challenging the oxygen reactivity of flavoenzymes. *Trends Biochem. Sci.* **31**, 276–283.
3. van der Horst, M. A. and K. J. Hellingwerf (2004) Photoreceptor proteins, star actors of modern times: A review of the functional dynamics in the structure of representative members of six different photoreceptor families. *Acc. Chem. Res.* **37**, 13–20.
4. Sancar, A. (1994) Structure and function of DNA photolyase. *Biochemistry* **33**, 2–9.
5. Christie, J. M., T. E. Swartz, R. A. Bogomolni and W. R. Briggs (2002) Phototropin LOV domains exhibit distinct roles in regulating photoreceptor function. *Plant J.* **32**, 205–500.
6. Gomelsky, M. and G. Klug (2002) BLUF: A novel FAD-binding domain involved in sensory transduction in microorganisms. *Trends Biochem. Sci.* **27**, 497–500.
7. Schelvis, J. P. M., M. Ramsey, O. Sokolova, C. Tavares, C. Cecala, K. Connell, S. Wagner and Y. M. Gindt (2003) Resonance Raman and UV-Vis spectroscopic characterization of FADH* in the complex of photolyase with UV-damaged DNA. *J. Phys. Chem. B* **107**, 12352–12362.
8. Masuda, S., K. Hasegawa and T. Ono (2005) Tryptophan at position 104 is involved in transforming light signal into changes of β-sheet structure for the signaling state in the BLUF domain of AppA. *Plant Cell Physiol.* **46**, 1894–1901.
9. Bonetti, C., T. Mathes, I. H. M. van Stokkum, K. M. Mullen, M.-L. Groot, R. van Grondelle, P. Hegemann and J. T. M. Kennis (2008) Hydrogen bond switching among flavin and amino acid side chains in the BLUF photoreceptor observed by ultrafast infrared spectroscopy. *Biochem. J.* **95**, 4790–4802.
10. Hazekawa, I., Y. Nishina, K. Sato, M. Shichiri, R. Miura and K. Shiga (1997) A Raman study on the C(4)=O stretching mode of flavins in flavoenzymes: Hydrogen bonding at the C(4)=O moiety. *J. Biochem.* **121**, 1147–1154.
11. Nishina, Y., K. Sato, R. Miura, K. Matsui and K. Shiga (1998) Resonance Raman study on reduced flavin in purple intermediate of flavoenzyme: Use of 4-carbonyl-^{18}O-enriched flavin. *J. Biochem.* **124**, 200–208.
12. Kondo, M., J. Neppa, K. L. Ronayne, A. L. Stelling, P. J. Tonge and S. R. Meech (2006) Ultrafast vibrational spectroscopy of the flavin chromophore. *J. Phys. Chem. B* **110**, 20107–20110.
13. Abe, M., Y. Kyogoku, T. Kitagawa, K. Kawano, N. Ohishi, A. Takai-Suzuki and K. Yagi (1986) Infrared spectra and molecular association of lumiflavin and riboflavin derivatives. *Spectrochim. Acta A* **42**, 1059–1068.
14. Copeland, R. and T. Spiro (1986) Ultraviolet resonance Raman spectroscopy of flavin mononucleotide and flavin adenine dinucleotide. *J. Phys. Chem.* **90**, 6654–6657.
15. Birss, V. I., A. S. Hinman, C. E. McGarvey and J. Segal (1994) *In situ* FTIR thin-layer reflectance spectroscopy of flavin adenine dinucleotide at a mercury/gold electrode. *Electrochim. Acta* **39**, 2449–2454.
16. Hellwig, P., D. Scheide, S. Bungert, W. Mäntele and T. Friedrich (2000) FT-IR spectroscopic characterization of NADH:ubiquinone oxidoreductase (Complex I) from *Escherichia coli*: Oxidation of FeS cluster N2 is coupled with the protonation of an aspartate or glutamine side chain. *Biochemistry* **39**, 10884–10891.
17. Wille, G., M. Ritter, R. Friedemann, W. Mäntele and G. Hübner (2003) Redox-triggered FTIR difference spectra of FAD in aqueous solution and bound to flavoproteins. *Biochemistry* **42**, 14814–14821.
18. Nishina, Y., K. Sato, C. Setoyama, H. Tamaoki, R. Miura and K. Shiga (2007) Intramolecular and intermolecular perturbation on electronic state of FAD free in solution and bound to flavoproteins: FTIR spectroscopic study by using C=O stretching vibrations as probes. *J. Biochem.* **142**, 265–272.
19. Wolf, M. M. N., C. Schumann, R. Gross, T. Domratcheva and R. Diller (2008) Ultrafast infrared spectroscopy of riboflavin: Dynamics, electronic structure, and vibrational mode analysis. *J. Phys. Chem. B* **112**, 13424–13432.
20. Siebert, F. and P. Hildebrandt (2007) *Vibrational Spectroscopy in Life Science*. Wiley-VCH, Heidelberg, Berlin.
21. Abe, M. and Y. Kyogoku (1987) Vibrational analysis of flavin derivatives: Normal coordinate treatments of lumiflavin. *Spectrochim. Acta A* **43**, 1027–1037.
22. Schmitz, M. and P. Tavan (2006) On the art of computing the IR spectra of molecules in condensed phase. In *Modern Methods for Theoretical Physical Chemistry of Biopolymers* (Edited by S. Tanaka and J. Lewis) Chapter 8, pp. 157–177. Elsevier, Amsterdam.
23. Hohenberg, P. and W. Kohn (1964) Inhomogeneous electron gas. *Phys. Rev. B* **136**, 864–871.
24. Kohn, W. and L. J. Sham (1965) Self-consistent equations including exchange and correlation effects. *Phys. Rev.* **140**, 1133–1138.
25. Frisch, M. J., G. W. Trucks, H. B. Schlegel, G. E. Scuseria, M. A. Robb, J. R. Cheeseman, G. Scalmani, V. Barone, B. Mennucci, G. A. Petersson, H. Nakatsuji, M. Caricato, X. Li, H. P. Hratchian, A. F. Izmaylov, J. Bloino, G. Zheng, J. L. Sonnenberg, M. Hada, M. Ehara, K. Toyota, R. Fukuda, J. Hasegawa, M. Ishida, T. Nakajima, Y. Honda, O. Kitao, H. Nakai, T. Vreven, J. A. Montgomery Jr, J. E. Peralta, F. Ogliaro, M. Bearpark, J. J. Heyd, E. Brothers, K. N. Kudin, V. N. Staroverov, R. Kobayashi, J. Normand, K. Raghavachari, A. Rendell, J. C. Burant, S. S. Iyengar, J. Tomasi, M. Cossi, N. Rega, J. M. Millam, M. Klene, J. E. Knox, J. B. Cross, V. Bakken, C. Adamo, J. Jaramillo, R. Gomperts, R. E. Stratmann, O. Yazyev, A. J. Austin, R. Cammi, C. Pomelli, J. W. Ochterski, R. L. Martin, K. Morokuma, V. G. Zakrzewski, G. A. Voth, P. Salvador, J. J. Dannenberg, S. Dapprich, A. D. Daniels, O. Farkas, J. B. Foresman, J. V. Ortiz, J. Cioslowski and D. J. Fox (2009) *Gaussian 09, Revision A.0*. Gaussian, Inc., Wallingford, CT.
26. Ahlrichs, R., M. Bär, M. Häser, H. Horn and C. Kölmel (1989) Electronic structure calculations on workstation computers: The program system TURBOMOLE. *Chem. Phys. Lett.* **162**, 165–169.
27. CPMD V3.9, Copyright IBM Corp 1990–2008, Copyright MPI für Festkörperforschung Stuttgart 1997–2001. Available at http://www.cpmd.org.

28. Nonella, M. and P. Tavan (1995) An unscaled quantum mechanical harmonic force field for p-benzoquinone. *Chem. Phys.* **199**, 19–32.
29. Zhou, X., S. J. Mole and R. Liu (1996) Density functional theory of vibrational spectra. 4. Comparison of experimental and calculated frequency of all-trans-1,3,5,7-octatetraene—The end of normal coordinate analysis? *Vib. Spectroscopy* **12**, 73–79.
30. Neugebauer, J. and B. A. Hess (2003) Fundamental vibrational frequencies of small polyatomic molecules from density-functional calculations and vibrational perturbation theory. *J. Chem. Phys.* **118**, 7215–7225.
31. Klaumünzer, B., D. Kröner and P. Saalfrank (2010) (TD-)DFT calculation of vibrational and vibronic spectra of riboflavin in solution. *J. Phys. Chem.* **114**, 10826–10834.
32. Miertus, S., E. Scrocco and J. Tomasi (1981) Electrostatic interaction of a solute with a continuum. A direct utilization of ab initio molecular potentials for the prevision of solvent effects. *Chem. Phys.* **55**, 117–129.
33. Klamt, A. and G. Schürmann (1993) COSMO: A new approach to dielectric screening in solvents with explicit expressions for the screening energy and its gradient. *J. Chem. Soc. Perin Trans.* **2**, 799–805.
34. Nonella, M., G. Mathias, M. Eichinger and P. Tavan (2003) Structures and vibrational frequencies of the quinones in *Rb.* Sphaeroides derived by a combined density functional molecular mechanics approach. *J. Phys. Chem. B* **107**, 316–322.
35. Warshel, A. and M. Levitt (1976) Theoretical studies of enzymic reactions: Dielectric, electrostatic and steric stabilization of the carbonium ion in the reaction of lysozyme. *J. Mol. Biol.* **103**, 227–249.
36. Mathias, G., B. Egwolf, M. Nonella and P. Tavan (2003) A fast multipole method combined with a reaction field for long-range electrostatics in molecular dynamics simulations: The effects of truncation on the properties of water. *J. Chem. Phys.* **118**, 10847–10860.
37. Eichinger, M., P. Tavan, J. Hutter and M. Parrinello (1999) A hybrid method for solutes in complex solvents: Density functional theory combined with empirical force fields. *J. Chem. Phys.* **110**, 10452–10467.
38. Babitzki, G., G. Mathias and P. Tavan (2009) The infra-red spectra of the retinal chromophore in *Bacteriorhodopsin* calculated by a DFT/MM approach. *J. Chem. Phys. B* **113**, 10483–10495.
39. Schropp, B., C. Wichmann and P. Tavan (2010) A spectroscopic polarizable force field for amide groups in polypeptides. *J. Phys. Chem. B* **114**, 6740–6750.
40. MacKerell, A. D., D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin and M. Karplus (1998) All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616.
41. Laio, A., J. VandeVondele and U. Rothlisberger (2002) A Hamiltonian electrostatic coupling scheme for hybrid Car-Parrinello molecular dynamics simulations. *J. Chem. Phys.* **116**, 6941–6947.
42. Laino, T., F. Mohamed, A. Laio and M. Parrinello (2005) An efficient real space multigrid QM/MM electrostatic coupling. *J. Chem. Theory Comput.* **1**, 1176–1184.
43. Jung, A., J. Reinstein, T. Domratcheva, R.-L. Shoeman and I. Schlichting (2006) Crystal structures of the AppA BLUF domain photoreceptor provide insights into blue light-mediated signal transduction. *J. Mol. Biol.* **362**, 717–732.
44. Sadeghian, K., M. Bocola and M. Schütz (2008) A conclusive mechanism of the photoinduced reaction cascade in blue light using flavin photoreceptors. *J. Am. Chem. Soc.* **130**, 12501–12513.
45. Domratcheva, T., B. L. Grigorenko, I. Schlichting and A. V. Nemukhin (2008) Molecular models predict light-induced glutamine tautomerization in BLUF photoreceptors. *Biochem. J.* **94**, 3872–3879.
46. Gauden, M., S. Yeremenko, W. Laan, I. H. M. van Stokkum, J. A. Ihalainen, R. van Grondelle, K. J. Hellingwerf and J. T. M. Kennis (2005) Photocycle of the flavin-binding photoreceptor AppA, a bacterial transcriptional antirepressor of photosynthesis genes. *Biochemistry* **44**, 3653–3662.
47. Gauden, M., I. H. M. van Stokkum, J. M. Key, D. C. Lührs, R. van Grondelle, P. Hegemann and J. T. M. Kennis (2006) Hydrogen-bond switching through a radical pair mechanism in a flavin-binding photoreceptor. *Proc. Natl Acad. Sci. USA* **103**, 10895–10900.
48. Gaigeot, M. P. and M. Sprik (2003) Ab initio molecular dynamics computation of the infrared spectrum of aqueous Uracil. *J. Phys. Chem. B* **107**, 10344–10358.
49. Ramirez, R., T. Lopez-Ciudad, P. Kumar and D. Marx (2004) Quantum corrections to classical time-correlation functions: Hydrogen bonding and anharmonic floppy modes. *J. Chem. Phys.* **121**, 3973–3983.
50. Nonella, M., G. Mathias and P. Tavan (2003) The infrared spectrum of 1,4-benzoquinone in water obtained from a hybrid molecular dynamics simulation. *J. Phys. Chem. A* **107**, 8638–8647.
51. Schmitz, M. and P. Tavan (2004) Vibrational spectra from atomic fluctuations in dynamics simulations. I. Theory, limitations, and a sample application. *J. Chem. Phys.* **121**, 12233–12246.
52. Schmitz, M. and P. Tavan (2004) Vibrational spectra from atomic fluctuations in dynamics simulations. II. Solvent-induced frequency fluctuations at femtosecond time resolution. *J. Chem. Phys.* **121**, 12247–12258.
53. Baer, M., D. Marx and G. Mathias (2010) Theoretical messenger spectroscopy of microsolvated Hydronium and Zundel cations. *Angew. Chem. Int. Ed. Engl.* **49**, 7346–7349.
54. Schultheis, V., R. Reichold, B. Schropp and P. Tavan (2008) A polarizable force field for computing the infrared spectra of the polypeptide backbone. *J. Phys. Chem. B* **112**, 12217–12230.
55. Becke, A. D. (1988) Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **38**, 3098–3100.
56. Perdew, J. P. and W. Yue (1986) Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation. *Phys. Rev. B* **33**, 8800–8802.
57. Troullier, N. and J. L. Martins (1991) Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B* **43**, 1993–2005.
58. Schäfer, A., C. Huber and R. Ahlrichs (1994) Fully optimized contracted Gaussian-basis sets of triple zeta valence quality for atoms LI to KR. *J. Chem. Phys.* **100**, 5829–5835.
59. Schmidt, M. W., K. K. Baldridge, J. A. Boatz, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. J. Su, T. L. Windus, M. Dupuis and J. A. Montgomery (1993) General atomic and molecular electronic-structure system. *J. Comput. Chem.* **14**, 1347–1363.
60. Singh, U. C. and P. A. Kollman (1984) An approach to computing electrostatic charges for molecules. *J. Comput. Chem.* **5**, 129–145.
61. Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey and M. L. Klein (1983) Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935.
62. Niedermeier, C. and P. Tavan (1994) A structure adapted multipole method for electrostatic interactions in protein dynamics. *J. Chem. Phys.* **101**, 734–748.
63. Niedermeier, C. and P. Tavan (1996) Fast version of the structure adapted multipole method—efficient calculation of electrostatic forces in protein dynamics. *Mol. Simul.* **17**, 57–66.
64. Berendsen, H. J. C., J. P. M. Postma, W. F. van Gunsteren, A. DiNola and J. R. Haak (1984) Molecular dynamics with coupling to an external bath. *J. Chem. Phys.* **81**, 3684–3690.
65. Verlet, L. (1967) Computer "experiments" on classical fluids. I. Thermodynamical properties of Lennard-Jones molecules. *Phys. Rev.* **159**, 98–103.
66. Hasegawa, K., S. Masuda and T. Ono (2005) Spectroscopic analysis of the dark relaxation process of a photocycle in a sensor of blue light using FAD (BLUF) protein SLR1694 of the cyanobacterium *Synechocystis* sp. PCC6803. *Plant Cell Physiol.* **46**, 136–146.
67. Höchtl, P., S. Boresch, W. Bitomsky and O. Steinhauser (1998) Rationalization of the dielectric properties of common three-site water models in terms of their force field parameters. *J. Chem. Phys.* **109**, 4927–4937.

68. Richardi, J., P. Fries and C. Millot (2005) Fast hybrid methods for the simulation of dielectric constants of liquids. *J. Mol. Liquids* **117**, 3–16.
69. Guillot, B. (2002) A reappraisal of what we have learnt during three decades of computer simulations on water. *J. Mol. Liq.* **101**, 219–260.
70. Neumann, M. (1986) Dielectric relaxation in water. Computer simulations with the TIP4P potential. *J. Chem. Phys.* **85**, 1567–1579.
71. Alper, H. E. and R. M. Levy (1998) Dielectric properties of water. *J. Chem. Phys.* **91**, 1242–1251.
72. Watanabe, K. and M. L. Klein (1989) Effective pair potentials and the properties of water. *Chem. Phys.* **131**, 157–167.
73. Martin, C. B., M.-L. Tsao, C. M. Hadad and M. S. Platz (2002) The reaction of triplet flavin with indole. A study of the cascade of reactive intermediates using density functional theory and time resolved infrared spectroscopy. *J. Am. Chem. Soc.* **124**, 7226–7234.
74. Martin, D., A. Weise and H.-J. Niclas (1967) Das Lösungsmittel Dimethylsulfoxid. *Angew. Chem.* **79**, 340–357.
75. Pollak, P., G. Romeder, F. Hagedorn and H.-P. Gelbke (2005) *Nitriles*. Wiley-VHC Verlag GmbH & co KGaA, Weilheim. DOI:10.1002/14356007.a17_363.
76. Freindorf, M., Y. Shao, T. R. Furlani and J. Kong (2005) Lennard-Jones parameters for the combined QM/MM method using the B3LYP/6-31 + G*/AMBER potential. *J. Comput. Chem.* **26**, 1270–1278.
77. Pearlman, D., D. Case, J. Caldwell, W. Ross, T. Cheatham III, S. DeBolt, D. Ferguson, G. Seibel and P. Kollman (1995) AMBER, a computer program for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to elucidate the structures and energies of molecules. *Comp. Phys. Comm.* **91**, 1–41.
78. Becke, A. D. (1993) Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **98**, 5648–5652.
79. Lee, C., W. Yang and R. G. Parr (1988) Development of the Colle-Salvetti correlation-energy formula into a functional of the electron density. *Phys. Rev. B* **37**, 785–789.

2.2 Zusatzinformationen zu oxidiertem Lumiflavin in Wasser

Der folgende Abdruck [1]

„Supporting Material to the manuscript:
Density Functional Theory Combined with Molecular Mechanics:
The Infrared Spectra of Flavin in Solution“
Benjamin Rieff, Gerald Mathias, Sebastian Bauer und Paul Tavan

enthält zusätzliche Informationen über die Betrachtungen der DFT/MM generierten IR Schwingungsspektren von oxidiertem Lumiflavin in Wasser, die nicht im Haupttext der ersten Veröffentlichung enthalten sind. Auf neun Seiten, gegliedert in 4 Abschnitte, stellt die Supporting Information insgesamt vier Abbildungen, sechs Tabellen, eine Gleichung und verschiedene Diskussionen bereit, die (i) das CHARMM22 ähnliche MM Kraftfeld für das Lumiflavin und das FMN, (ii) die Konsequenzen von unterschiedlich gewählten Lennard-Jones Parametern der Carbonylsauerstoffe auf die Solvatisierung sowie die IR Schwingungsspektren von Lumiflavin, (iii) die Bandenzuordnung und Isotopenshifts für weitere neun Schwingungsbanden des Lumiflavins und (iv) die Aussagekraft der zu Grunde gelegten Gauss‘schen Modelle für die DFT/MM-INMA Spektren von Lumiflavin dokumentieren und erklären.

[1] Mit freundlicher Genehmigung des Wiley Verlags. Kostenlos einsehbar unter DOI: 10.1111/j.1751-1097.2010.00866.x

Supporting material to the manuscript

Density functional theory combined with molecular mechanics: The infra-red spectra of flavin in solution

Benjamin Rieff, Gerald Mathias, Sebastian Bauer, and Paul Tavan*
Theoretische Biophysik, Lehrstuhl für Biomolekulare Optik, Ludwig-Maximilians-Universität, Oettingenstr. 67, 80538 München, Germany

*corresponding author, email: tavan@physik.uni-muenchen.de,
phone: +49-89-2180-9220, fax: +49-89-2180-9202

1 CHARMM22 type force field for isoalloxazine

The CHARMM22 [1] force field does not cover the isoalloxazine moiety. Because we needed a corresponding parameter set for our MM-MD simulations of LF in aqueous solution, we had to generate such a set. As explained in Section Methods we carried out, for this purpose, DFT calculations on the isolated LF and DFT/MM calculations on LF in TIP3P and TIP4P water to determine many of the required parameters and applied homology modeling for the remaining ones. According to Figure S1 the results are given in the tables S1-S4.

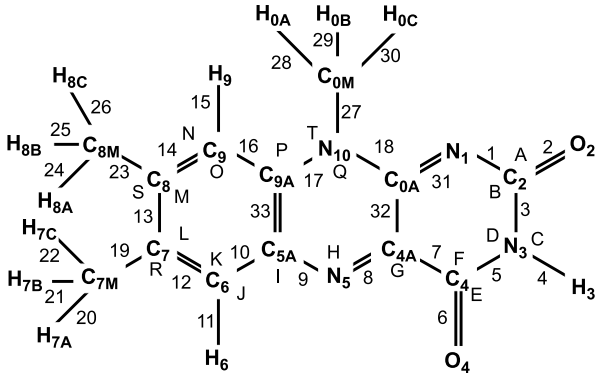


Figure S1: Chemical structure and atom labels of isoalloxazine and a definition of labels for bonds (numbers) and angles (capital letters).

Fig. S1 uses the chemical structure of LF to introduce names for the atoms, labels for the bond lengths (integer numbers) and for the bond angles (capital letters).

atom	type	q_g	q_l	atom	type	q_g	q_l
N_1	NN3A	−0.4291	−0.6632	C_8	CA3	0.1146	0.1272
C_2	CN1A	0.5123	0.7244	C_9	CA4	−0.2497	−0.2561
O_2	ON1	−0.4890	−0.6419	H_9	HP	0.1968	0.2345
N_3	NN2U	−0.2315	−0.2509	C_{9A}	CPTA	−0.1534	−0.1809
H_3	HN2	0.2631	0.3129	N_{10}	NN2G	0.3171	0.3518
C_4	CN1B	0.2558	0.3368	C_{0A}	CN5A	0.0847	0.2462
O_4	ON1	−0.4266	−0.5287	C_{7M}	CT3	−0.2709	−0.3123
C_{4A}	CN5B	0.3475	0.2929	H_{7x}	HAI	0.0903	0.1041
N_5	NN3A	−0.4759	−0.5408	C_{8M}	CT3	−0.3117	−0.3879
C_{5A}	CPTB	0.3114	0.3861	H_{8x}	HAI	0.1039	0.1293
C_6	CA1	−0.2004	−0.2774	C_{0M}	CT3	−0.3078	−0.5520
H_6	HP	0.1339	0.1647	H_{0x}	HAI	0.1026	0.1840
C_7	CA2	0.1184	0.1621				

Table S1: CHARMM22 atom types and partial charges calculated by DFT for isolated LF (q_g) and by DFT/MM for LF in water (q_g) are associated to the atom names defined in Fig. S1.

Table S1 associates certain atom types to the atom names defined in Fig. S1. In CHARMM22

such types serve as selectors for force field parameters. The table extends the set of CHARMM22 standard types CN1, CN5, CPT, and CA by adding capital letters (A,B) or numbers (1,2,3,4) to uniquely define a certain positions within isoalloxazine. The table also contains partial charges applicable to LF in the gas phase (q_{g}) or in water (q_{l}). These charges were calculated by the procedures described in Section Methods. Comparing the gas phase values $|q_{\mathrm{g}}|$ with the liquid phase values $|q_{\mathrm{l}}|$ one recognizes a general increase, i.e. an enhanced polarity of the LF molecule.

CHARMM22 models the bond stretches, the angle deformations, and the stiffnesses of sp^2 hybridized atoms by harmonic potentials specified by force constants and equilibrium values of the associated internal coordinates, which are the bond lengths l_i, the bond angles θ_j, and the so-called improper dihedral angles ϕ_k, respectively.

bond i	k_l[kcal/mol]	l_0[Å]	bond i	k_l[kcal/mol]	l_0[Å]
1	300.2	1.384	15	384.4	1.090
2	749.1	1.227	16	326.7	1.405
3	185.1	1.420	17	300.2	1.388
4	472.2	1.024	18	295.7	1.392
5	310.3	1.386	19	265.9	1.508
6	767.6	1.225	23	265.0	1.507
7	187.6	1.503	27	249.1	1.470
8	463.5	1.307	20-22	359.4	1.100
9	284.9	1.370	24-26	359.4	1.100
10	346.1	1.409	28-30	359.4	1.100
11	379.9	1.092	31	425.9	1.312
12	327.0	1.388	32	295.0	1.461
13	315.6	1.426	33	311.6	1.425
14	354.0	1.396			

Table S2: Force constants k_l and equilibrium values l_0 calculated by DFT for the isoalloxazine bond lengths l_i, whose labels i are defined by the numbers in Fig. S1.

Table S2 lists the parameters calculated by DFT for the harmonic potentials of isoalloxazine's bond lengths l_i. Table S3 displays the corresponding data for the bond angles θ_j.

A dihedral angle ϕ_k is usually characterized by a label k denoting a set of four atoms. For the improper dihedrals ϕ_k used by us to ensure the planarity of isoalloxazine the labels k are listed in Table S4. All associated harmonic potentials have the same force constant k_ϕ = 100 kcal/(mol·rad^2) and equilibrium value $\phi_0 = 180°$.

The description of the isoalloxazine force field is complete, as soon as also the parameters of the Lennard-Jones (LJ) potentials are given. For these parameters we adopted the values specified in the CHARMM22 force field [1] for the basic atom types.

The LF molecule was extended towards FMN by attaching a glycerol chain and phosphate ion to atom N_{10} (cf. Fig. S1). The CHARMM22 force field provides a parametrization for these two chemical motifs, which was adopted.

angle j	k_θ [kcal/mol rad^2]	θ_0 [degrees]	angle j	k_θ [kcal/mol rad^2]	θ_0 [degrees]
A	105.00	123.33	M	77.00	120.14
B	249.75	117.61	N	52.20	118.65
C	70.00	115.51	O	230.00	121.27
D	91.60	127.94	P	52.25	122.71
E	117.00	122.84	Q	422.50	120.99
F	81.00	112.44	R	70.00	120.17
G	33.60	118.60	S	70.50	120.61
H	250.00	118.13	T	81.25	120.16
I	75.00	118.86	methyl at CA2	38.00	110.10
J	53.75	116.62	methyl at CA3	38.00	110.10
K	220.70	122.20	methyl at NN2G	40.50	110.10
L	78.50	118.45	methyl H-C-H	35.50	108.40

Table S3: Force constants k_θ and equilibrium values θ_0 calculated by DFT for the isoalloxazine angles θ_j, whose labels j are given by the capital letters in Fig. S1.

labels k of improper dihedrals ϕ_k				labels k of improper dihedrals ϕ_k			
ON1	NN3A	CN1A	NN2U	CPTB	NN3A	CA1	HP
NN3A	CN1A	NN2U	HN2	NN3A	CPTB	CA1	CA2
CN1A	NN3A	NN2U	CN1B	CA1	CPTB	CA2	CA3
CN1A	NN2U	CN1B	ON1	CA2	CA1	CA3	CA4
NN2U	CN1A	CN1B	CN5B	CA2	CA3	CA4	HP
NN2U	CN1B	CN5B	NN3A	CA3	CA2	CA4	CPTA
CN1B	CN5B	NN3A	CPTB	CA3	CA4	CPTA	NN2G
CN5B	NN3A	CPTB	CA1	CA4	CPTA	NN2G	CN5A

Table S4: The labels k of the improper dihedral angles ϕ_k, which are employed to guarantee the planarity of isoalloxazine, are quadruples of atom types pointing to the four atoms, which are harmonically forced to be coplanar.

2 Choice of Lennard-Jones potentials.

In a microscopic description of a solvent surrounding a molecule one explicitly includes the van der Waals interactions. In addition to the electrostatics, also these interactions affect the solvent's short range order characterized by Fig. 6 for the neighborhood of the flavin carbonyl groups. This insight leads to the question, to what extent the solvation patterns are influenced by the choice of the LJ parameters for the carbonyl oxygens.

Steering the carbonyl solvation. We checked this parameterization issue by calculating additional INMA spectra of LF in TIP3P and TIP4P water using altered LJ parameters for the LF oxygen atoms O_2 and O_4. Here we slightly increased the van der Waals radii from 3.0291 Å to 3.1538 Å and the depths of the LJ potentials from 0.1200 kcal/mol to 0.1521 kcal/mol. In CHARMM22 [1] the former values belong to atom type ON1 and the latter to ON5 (in guanosine, e..g., ON1 models the carbonyl and ON5 the hydroxyl oxygen atoms). Depending on the solvent model, DFT/MM spectra obtained for the modified LJ potentials will be labeled by $T3_{LJ}$ and $T4_{LJ}$, respectively. To assess the steering of the solvation, we have first calculated the radial distribution functions $g_{OH}(r)$ of the TIP3P hydrogen atoms around the modified models

of the LF oxygen atoms O_2 and O_4 from a 1 ns MM-MD simulation.

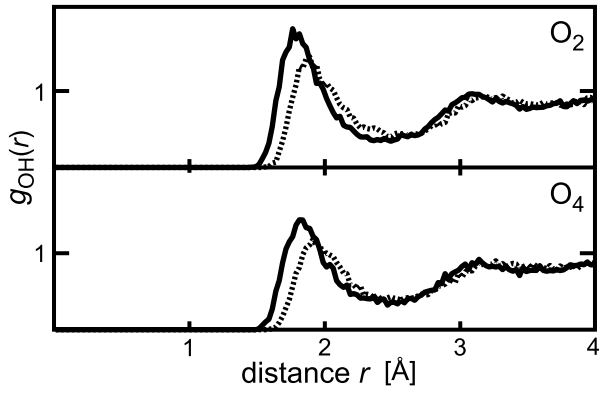


Figure S2: Radial distribution functions $g_{OH}(r)$ of the TIP3P hydrogen atoms surrounding the flavin oxygen atoms O_2 and O_4 using standard (solid lines) and altered (dashed lines) LJ parameters.

Figure S2 compares the distribution functions $g_{OH}(r)$ for the standard (solid lines) and modified (dashed lines) LJ models of the carbonyl oxygens O_2 and O_4. The graphs immediately demonstrate for each of the two carbonyl groups that the larger van der Waals radii of the flavin oxygens assumed in the modified parameterization (dashed lines) shift the distances r_{max} of the first maxima to larger values. Furthermore, the number of hydrogen atoms in the first solvation shell is found to decrease by 5 % at O_2 (from averages 2.56 to 2.43) and by 8 % at O_4 (from averages 2.28 to 2.09). The consequences of the correspondingly weakened solvation of the two LF carbonyl groups on the calculated IR spectra will now be scrutinized. One expects, of course, that this weakened solvation causes a blue shift of the carbonyl bands.

DFT/MM frequencies for modified Lennard-Jones parameters. Figure S3 compares the DFT/MM vibrational frequencies of our five LF marker bands, which were obtained using the standard LJ parameters at the carbonyl oxygens (left part) and two solvent models (T3, T4), with the corresponding frequencies calculated by applying the modified parameters (right part) and the two solvent models ($T3_{LJ}$, $T4_{LJ}$). The central column (H_2O) of Fig. S3 are the experimental consensus data from Fig. 5. For a simplified visual comparison with these experimental data, all DFT/MM frequencies have been scaled by the factor 1.0334. This factor has been chosen as to minimize the RMSD of the four average DFT/MM frequencies $\langle\nu_{C=O}\rangle$ from the corresponding average experimental frequency.

Up to the scaling of the DFT/MM frequencies, the columns T4, T3, and H_2O in Fig. S3 contain the same data as the corresponding columns in Fig. 5. Thus the patterns of band shifts induced by exchanging the mildly (T4) by the strongly (T3) polar solvent model are essentially identical in the two figures. If one now increases the van der Waals radii of the LF carbonyl oxygen atoms and, thus, applies the weaker solvation characterized by Fig. S2 to the two carbonly groups, the DFT/MM calculations predict the expected blue shifts of the cabonyl frequencies for both solvents (cf. columns $T3_{LJ}$ and $T4_{LJ}$ in Fig. S3). In the case of the TIP4P water model the average carbonyl stretching frequency $\langle\nu_{C=O}\rangle$ is blue shifted by 16 cm^{-1} and for TIP3P by

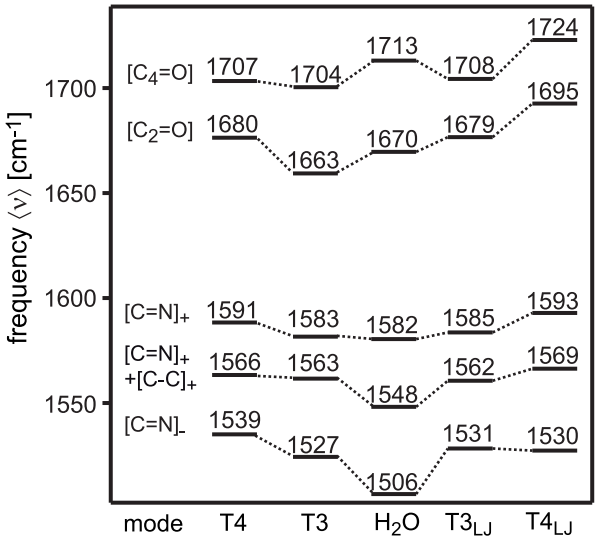


Figure S3: DFT/MM frequencies calculated by INMA for LF in liquid water using different models and scaled with a factor of 1.0334 are compared with the experimental consensus frequencies "H_2O" for fully oxidized flavins in water (see Fig. 5). The scaled DFT/MM results cover the data already shown in Fig. 5 in an unscaled fashion, which were obtained for LF in TIP3P (T3) and TIP4P water (T4) with the standard LJ parameters of CHARMM22 for the carbonyl oxygens. Furthermore, they cover frequencies obtained for the modified LJ parameters and for LF in TIP3P ($T3_{LJ}$) and TIP4P water ($T4_{LJ}$).

10 cm^{-1} by the increased van der Waals radii. In contrast, the average frequencies $\langle\nu_{C=N}\rangle$ of the three C=N marker bands are nearly invariant (± 2 cm^{-1}) upon the change of the LJ potentials at the carbonyl oxygens for both solvents. Thus, a modification of the LJ potentials at the carbonyl oxygens can be used to tune the spectral gap between the C=O and C=N modes. Furthermore, such a modification apparently has only minor effects on the spectral spacings among the C=N modes.

In contrast, the spectral gap $\Delta\nu_{C=O}$ between the two high frequency carbonyl modes can be modified somewhat more strongly by the change of the LJ potentials. In the TIP3P case, for instance, the gap $\Delta\nu_{C=O}$ decreases by 29 % [from 41 cm^{-1} (T3) to 29 cm^{-1} ($T3_{LJ}$)], whereas for TIP4P it increases by 7 % [from 27 cm^{-1} (T4) to 29 cm^{-1} ($T4_{LJ}$)]. As compared to the experimental value of 43 cm^{-1} for the gap, however, solely the original T3 model furnishes a reasonable description (41 cm^{-1}). Therefore, we decided to consider the T3 combination of the TIP3P water model with the standard CHARMM22 parameters for the LJ potentials of the LF carbonyl oxygens as the most promising choice for the DFT/MM description of the IR spectra of flavins in solution or protein environments.

There are, most probably, better choices for the solvent model and for the LJ parameters of flavin dyes in DFT/MM calculations of condensed phase IR spectra. However, a systematic search for an optimal parameterization is excluded by the enormous computational cost posed by any such attempt. Here, we solely wanted to check to what extent two reasonable choices of LJ potentials for oxygen atoms, which are used in the CHARMM22 force field [1] for different chemical motifs, can modify spectral descriptions in a DFT/MM setting and whether, by chance,

we might detect a better description than the CHARMM22 standard, which is T3.

3 Assignment and isotopic shifts of additional flavin bands

Here we present additional data on flavin bands in the spectra region between 1200 cm^{-1} and 1750 cm^{-1} . Table S5 shows IR frequencies for solid state LF dried on KBr discs together with a PED assignment based on an empirical normal mode analysis [2]. These data are compared with our scaled (1.031) DFT/MM-INMA frequencies and PED values, derived from the MT/BP normal mode analysis of the isolated LF by using the program Gamess [3].

No.	exp/cm^{-1}	PED				T3/cm^{-1}	PED			
81	1708	61	$\nu_{C_4=O}$	20	$\nu_{C_2=O}$	1700	79	$\nu_{C_4=O}$	4	$\nu_{C_2=O}$
80	1662	82	$\nu_{C_2=O}$	16	$\nu_{C_4=O}$	1659	77	$\nu_{C_2=O}$	7	δ_{N_3H}
79	1621	31	$\nu_{C_{5A}C_{9A}}$	15	$\nu_{C_{4A}N_5}$	1652	17	$\nu_{C_6C_7}$	17	$\nu_{C_{5A}C_6}$
78	1583	29	$\nu_{N_1C_{0A}}$	19	$\nu_{N_{10}C_{0A}}$	1580	31	$\nu_{C_{4A}N_5}$	18	$\nu_{C_{0A}N_1}$
77	1552	33	$\nu_{N_1C_{0A}}$	18	$\nu_{C_{4A}N_5}$	1559	32	$\nu_{C_7C_8}$	18	$\nu_{C_{0A}N_1}$
76	1513	33	$\nu_{C_{4A}N_5}$	14	$\nu_{C_{9A}N_{10}}$	1524	57	$\nu_{C_{0A}N_1}$	40	$\nu_{C_{4A}N_5}$
68	1461	20	$\nu_{C_7C_8}$	15	ν_{C_8Me}	1451	18	$\nu_{C_8C_9}$	5	$\nu_{C_{5A}C_6}$
64	1425	29	$\nu_{C_7C_8}$	23	$\nu_{C_6C_7}$	1409	35	$\nu_{C_8C_9}$	15	$\nu_{C_7C_8}$
63	1413	71	δ_{N_3H}	8	$\nu_{C_4=O}$	1411	65	δ_{N_3H}	9	$\nu_{C_4=O}$
62	1396	33	$\nu_{C_2N_3}$	33	$\nu_{C_2N_1}$	1382	20	$\nu_{C_7C_8}$	12	$\nu_{C_4C_{4A}}$
61	1346	17	$\nu_{C_{5A}C_6}$	16	$\nu_{N_{10}C_{0A}}$	1353	13	$\nu_{N_{10}C_{0A}}$	10	$\nu_{C_4C_{4A}}$
60	1301	30	$\nu_{N_5C_{5A}}$	14	$\nu_{C_8C_9}$	1312	32	$\nu_{N_{10}C_{0A}}$	7	$\nu_{C_{0A}N_1}$
58	1272	22	$\nu_{C_{4A}C_{0A}}$	17	$\nu_{C_4N_3}$	1288	49	δ_{CH}	11	$\nu_{N_5C_{5A}}$
57	1238	43	$\nu_{C_4N_3}$	23	$\nu_{C_4C_{4A}}$	1236	24	$\nu_{C_4N_3}$	17	$\nu_{C_8C_{Me}}$

Table S5: IR frequencies (exp) observed for solid LF in a KBr disc and assigned by an empirical normal mode analysis to PED's [2] (left) are compared with our scaled DFT/MM-INMA (T3) frequencies of LF and the PED associated to the MT/BP normal modes calculated for the isolated LF. Only the two largest local contributions to a normal mode are given in % .

To illustrate the normal mode compositions we calculated in addition to the PED also DFT/MM-INMA isotopic band shifts for the isotope labels enumerated in the caption to Fig. 9 (omitting, however, the deuterated case and all marker modes). Table S6 compares the isotopic band shifts calculated for LF in TIP3P water with IR data [2, 4] on solid state LF and with RR shifts measured for FAD in water [5].

The mode compositions of the five marker bands α = 81, 80, 78, 77, and 76 are extensively discussed in the paper. At this point we, therefore, solely add a few remarks on the character and compositions of some of the remaining 9 modes additionally covered by the Table S6. According to our DFT/MM results the modes 75-69 and 67-65 are dominated by methyl deformations, which carry mostly negligible intensities. The corresponding bands were experimentally not determined and, therefore, these modes are not listed here.

Inspecting Table S6 clearly demonstrates that the calculated level shifts nicely agree with the observed ones. If we choose as experimental reference, whenever possible, the average shifts resulting from the IR and RR experiments we find that our calculated shifts deviate from the observed shifts by an RMSD of only 2.9 cm^{-1}. This value is well within the limits of the

No.	native	$^{13}C_2$			$^{13}C_{4,0A}$			$^{15}N_{1,3}$			$^{15}N_5$		
	cm^{-1}	T3	IR	RR	T3	IR	RR	T3	IR	RR	T3	IR	RR
79	1652	0		0	−1		−2	0	0	−1	0	0	−1
68	1451	−2	−2	−1	−5	−3	−5	−1	−2	−1	−2	−2	−1
64	1409	−3	0		−3			−6	0		0	0	
63	1411	0	0		−3			−2			−1	0	
62	1382	−5	−13	−1	−8		−2	−4	−12	−2	−1	−1	0
61	1353	−5	−6	−2	−12	−12	−8	−2	−2	0	−2	−2	+1
60	1312	−1	−1		−6			−4	−1		−3	−3	
58	1288	−2	−3		−2	−4		−1	−1		−2	−4	
57	1236	0	−11		−4		−4	−3	−8	−9	−1	−2	+1

Table S6: Comparison of scaled DFT/MM and experimental frequencies representing IR [2, 4] and RR [5] data for native and isotopically labeled flavins. The column labels $^{13}C_2$, $^{13}C_{4,0A}$, $^{15}N_{1,3}$, and $^{15}N_5$ indicate the isotopic substitutions. The column "native" repeats the scaled INMA results of the T3 ensemble. The data for the five marker bands are omitted here, because they are displayed in Fig. 9.

experimental uncertainty, because the RMSD between the isotopic shifts determined by IR and RR, respectively, is 4.5 cm^{-1}.

Also the details of our band assignment agree very well with the observed isotopic shifts. Consider, for instance, mode 79. Applying an empirical normal mode analysis Abe and Kyogoku [2] calculated PEDs (see Table S5), which indicate that this mode is mainly a mixture of the C_{5A}-C_{9A} and C_{4A}=N_5 stretching vibrations. In contrast, our DFT (and DFT/MM) calculations assign only C-C stretches localized in ring I of isoalloxazine but no contribution of the C_{4A}=N_5 stretch to mode 79. Our result is consistent with the spectroscopic data for the isotopic substitution at N_5. In agreement with our DFT/MM result, experimentally this substitution induces shifts of at most 1 cm^{-1}. If mode 79 would contain a substantial contribution of the C_{4A}=N_5 stretch, then one would expect a larger effect of this particular isotope substitution. Thus, for this mode our DFT(/MM) based normal mode analysis provides a more realistic description of the mode compositions than the quoted empirical normal mode analysis.

We leave it to the reader to compare for some of the other modes the details of the observed and calculated isotopic shifts with the predicted mode compositions. Admittedly, we cannot explain some of the discrepancies between the RR and IR data, which are most strikingly visible for mode 62 and may be due to the different experimental conditions with RR referring to aqueous solution and IR to the solid state.

4 Gaussian approximation for IR line shapes obtained by INMA

Eq. (9) assumes Gaussian shapes for the inhomogeneously broadened IR bands of flavin in solution. To check this assumption Fig. S4 compares three different representations of the INMA results on the IR spectrum of LF in TIP3P water.

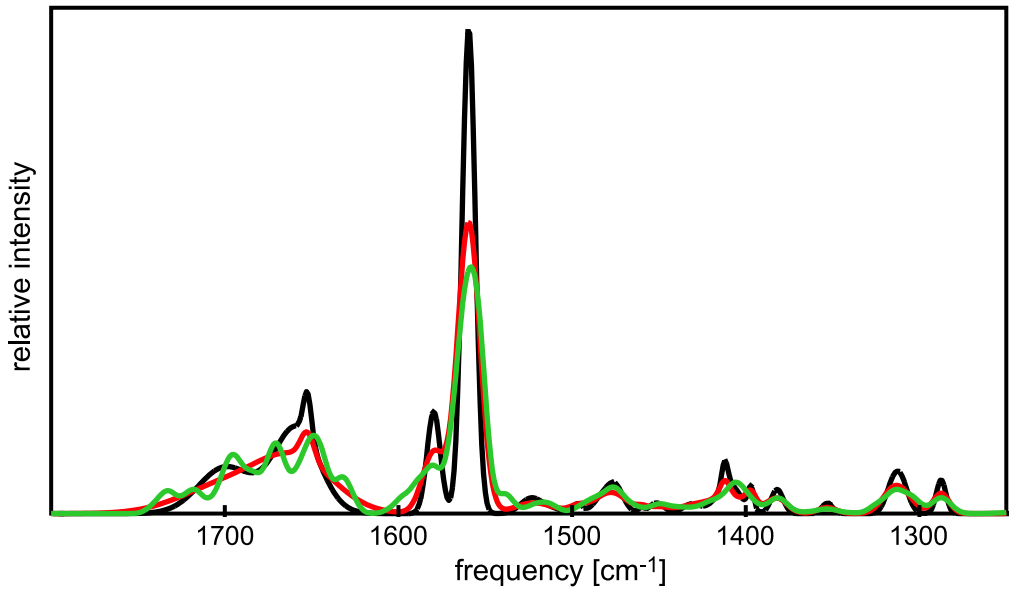


Figure S4: The IR spectrum of oxidized LF generated directly from the INMA snapshot ensemble through Eq. (S1) (green line) is compared with the Gaussian model Eq. (9) (black line) and with a correponding model (red line), which employs instead of the scaled widths $\sigma_\alpha = 0.6\rho_\alpha$ for each mode α the unscaled standard deviations ρ_α of the frequencies $\nu_{\alpha,j}$ within the ensemble of snapshots $j = 1, \ldots, N_S$.

The spectrum drawn in green derives from the expression

$$S(\nu) = \frac{c_0}{6}\frac{1}{N_S}\sum_\alpha \sum_{j=1}^{N_S} \frac{I_{\alpha,j}}{\sqrt{2\pi\kappa^2}} \exp\Big[\frac{-(\nu-\nu_{\alpha,j})^2}{2\kappa^2}\Big], \tag{S1}$$

in which the double sum runs over all modes α and all N_S snapshots j in the ensemble. Thus, the IR spectrum is represented as an ensemble average of normalized Gaussian lines centered at the INMA frequencies $\nu_{\alpha,j}$ and weighted with the INMA intensities $I_{\alpha,j}$. As widths of the Gaussians we chose the small value $\kappa = 4$ cm^{-1}.

The green kernel estimate Eq. (S1) of the IR spectrum is compared with a Gaussian model drawn in red, which derives from Eq. (9), if one replaces the scaled line widths σ_α by the unscaled standard deviations ρ_α of the mode frequencies $\nu_{\alpha,j}$ within the ensemble of snapshots j. Finally, the black spectrum exactly represents Eq. (9), i.e. is constructed using the scaled widths $\sigma_\alpha = 0.6\rho_\alpha$, which serve to approximately correct the neglected effects of motional narrowing.

The red spectrum represents a smoothened version of the green spectrum, which is characterized by sizable fluctuations. Particularly in the high-frequency region of the two carbonyl modes the superposition Eq. S1 is by no means smooth indicating not quite sufficient statistics. On the other hand, in the green and red spectra the line widths are quite similar. Particularly, the two C=O bands ar not easily distinguishable. Upon scaling of the band widths, which tries to correct the systematic overestimate of band widths characteristic for the INMA approach, the two C=O bands become discernible.

References

[1] MacKerell, A. D.; Bashford, D.; Bellott, M.; Dunbrack, R. L.; Evanseck, J. D.; Field, M. J.; Fischer, S.; Gao, J.; Guo, H.; Ha, S.; Joseph-McCarthy, D.; Kuchnir, L.; Kuczera, K.; Lau, F. T. K.; Mattos, C.; Michnick, S.; Ngo, T.; Nguyen, D. T.; Prodhom, B.; Reiher, W. E.; Roux, B.; Schlenkrich, M.; Smith, J. C.; Stote, R.; Straub, J.; Watanabe, M.; Wiorkiewicz-Kuczera, J.; Yin, D.; Karplus, M., All-atom empirical potential for molecular modeling and dynamics studies of proteins, *J. Phys. Chem. B* **1998,** *102,* 3586–3616.

[2] Abe, M.; Kyogoku, Y., Vibrational analysis of flavin derivatives: Normal coordinate treatments of lumiflavin, *Spectrochim. Acta A* **1987,** *43,* 1027-1037.

[3] Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A., General atomic and molecular electronic-structure system, *J. Comput. Chem.* **1993,** *14,* 1347-1363.

[4] Nishina, Y.; Sato, K.; Miura, R.; Matsui, K.; Shiga, K., Resonance Raman study on reduced flavin in purple intermediate of flavoenzyme: Use of 4-carbonyl-^{18}O-enriched flavin, *J. Biochem.* **1998,** *124,* 200-208.

[5] Hazekawa, I.; Nishina, Y.; Sato, K.; Shichiri, M.; Miura, R.; Shiga, K., A Raman study on the C(4)=O stretching mode of flavins in flavoenzymes: Hydrogen bonding at the C(4)=O moiety, *J. Biochem.* **1997,** *121,* 1147-1154.

2.3 Lumiflavinradikale in Wasser

In BLUF Domänen ist ein oxidiertes Flavin eingebettet. Nach der Photonenabsorption wird das gebundene Flavin auf Grund von Elektron- und Protontransfer Prozessen in transiente Radikalzustände überführt und liegt im Lichtzustand wiederum als oxidiertes Molekül vor. Da also Redox-Effekte eine entscheidende Rolle beim Photozyklus der BLUF Domäne spielen, galt es zu ergründen, ob auch Radikalzustände des Lumiflavins mit der in Abschnitt 2.1 eingesetzten Methode adequat beschrieben werden können.

Der nachfolgend abgedruckte Artikel [1]

„IR spectra of flavins in solution: DFT/MM description of redox effects“
Benjamin Rieff, Sebastian Bauer, Gerald Mathias und Paul Tavan
in *J. Phys. Chem. B* **115**, 2117–2123 (2011),

den ich zusammen mit Sebastian Bauer, Gerald Mathias und Paul Tavan veröffentlicht habe, beschäftigt sich deshalb mit der Frage, wie sich unterschiedliche Redoxzustände des Lumiflavins auf das DFT/MM generierte IR Schwingungsspektrum auswirken. Ausgehend von der vorangegangenen Arbeit wurden mit demselben DFT/MM Hybridverfahren, demselben Wassermodell und demselben Frequenzskalierungsfaktor die IR Schwingungsspektren für ein anionisches und ein neutrales Lumiflavinradikal in Wasser berechnet. Auch wenn in diesen Fällen eine deutlich dünnere Basis experimenteller Resultate in der Literatur vorliegt, so kann gezeigt werden, dass die verwendete DFT/MM-INMA Methode auch für diese Radikalzustände eine exzellente Vorhersage der experimentellen Resultate liefert und somit in der Lage ist, die beobachteten Schwingungsspektren zu charakterisieren und zu erklären.

[1] Reproduced with permission. Copyright 2011 American Chemical Society.

THE JOURNAL OF PHYSICAL CHEMISTRY B

ARTICLE

pubs.acs.org/JPCB

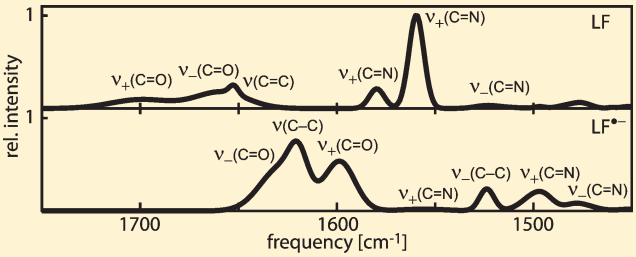
IR Spectra of Flavins in Solution: DFT/MM Description of Redox Effects

Benjamin Rieff, Sebastian Bauer, Gerald Mathias, and Paul Tavan*

Lehrstuhl für Biomolekulare Optik, Ludwig-Maximilians-Universität, Oettingenstr. 67, 80538 München, Germany

Supporting Information

ABSTRACT: The functional reactions in blue light photoreceptors generally involve transiently reduced flavins exhibiting characteristic infrared (IR) spectra. To approach a theoretical understanding, here we apply density functional theory (DFT) to flavin radicals embedded in a molecular mechanics (MM) model of an aqueous solution. Combining a DFT/MM approach with instantaneous normal-mode analyses (INMA), we compute the IR solution spectra of anionic and neutral flavin radicals. For a set of mid-IR marker bands, we identify those changes of spectral locations, intensities, and widths, which are caused by sequentially adding an electron and a proton to the oxidized flavin. Comparisons with experimental IR solution spectra of flavin radicals show the accuracy of our DFT/MM-INMA approach and allow us to assign the observed bands. The room temperature ensembles of solvent cages required for the INMA calculations of the IR spectra are generated in an MM setting from molecular dynamics (MD) simulations. For the solvated flavin radicals, these MD simulations employ MM force fields derived from DFT/MM calculations.



INTRODUCTION

Many enzymatically catalyzed biochemical reactions use flavin dyes as cofactors.[1] These dyes exhibit different redox and protonation states and, therefore, can promote both electron and proton transfer processes.[2] Figure 1 depicts the valence structures of isoalloxazine, which is the structural core motif of flavins, for the fully oxidized form (top) and the neutral semiquinone radical (bottom).[3] The various flavins, e.g., riboflavin (RF), flavin mononucleotide (FMN), or flavin adenine dinucleotide (FAD), are distinguished by the respective substituents R.[1]

The photochemistry of flavin dyes steers the light-induced reactions of blue light photoreceptors[4] such as photolyases,[5] cryptochromes,[6] LOV (light-oxygen-voltage),[7] or BLUF (blue light sensoring using FAD)[8] domains. Electron transfer reactions generating different flavin redox states in blue light photoreceptors were intensely studied by vibrational spectroscopy.[9−14] To identify the spectral signatures of chromophore−protein interactions,[15] reference spectra were measured for flavins in solution[12−14,16−22] including many isotopically labeled compounds. For oxidized flavin, normal-mode analyses were carried out with the aim of assigning observed bands to vibrational modes.[23−25]

Because of overlapping apoprotein bands, the identification of flavin bands in the infrared (IR) spectra of flavoproteins is generally difficult. Frequently, however, the stretching vibrations of the two carbonyl groups in the isoalloxazine moiety (see Figure 1) can be identified. If one can reliably analyze such spectroscopic data by a computational method, then one can decipher the chromophore−protein interactions.[10,12]

To establish such a computational method, we have recently calculated[25] the inhomogeneously broadened IR spectra of oxidized flavins in aqueous solution by a hybrid approach[26]

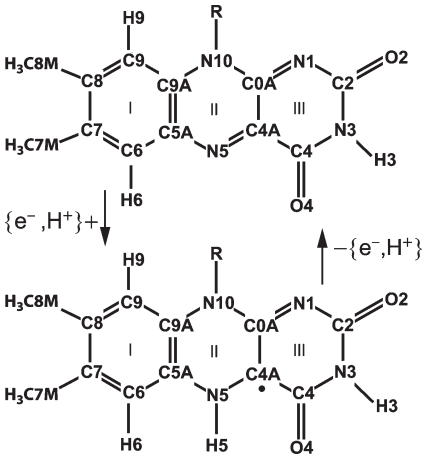


Figure 1. Chemical structures and atom labels of oxidized isoalloxazine (top) and of its neutral semiquinone radical (bottom).

combining a density functional theory (DFT)[27,28] description of the flavin with a molecular mechanics (MM) model of its liquid phase environment. Here, the liquid was tuned to a temperature of 300 K and a pressure of 1 atm by molecular dynamics (MD) simulation. We calculated the IR solution spectra by the protocol[29] for "instantaneous normal mode analysis" (INMA), which meanwhile has been established as a viable tool for the computation of condensed phase IR spectra in several applications.[30−34] The INMA

Received: November 29, 2010
Revised: January 10, 2011
Published: February 10, 2011

protocol was chosen because for large dyes such as flavins it is much more cost-effective than dynamics based methods[31−37] (for explanations, see refs 25, 33, and 34). As a further improvement of the INMA protocol, we additionally suggested and successfully applied an automated method for optimal mode matching,[25] which enables the rapid computation of average mode frequencies and intensities from large structural INMA ensembles generated by MD simulations.

For the frequency scaling factor, which induces for oxidized flavin an optimal match between our DFT/MM results and the spectroscopically observed band positions, we found the value 1.031 and we expressed the expectation that it should remain valid also for other redox states.[25] As is common in theoretical vibrational spectroscopy,[38] scaling factors are method specific. The particular approach used by us combines a DFT description of flavin, which we denote by the shortcut "MT/BP",[30,31,39] with a MM model for the surrounding water, for which we chose the stiff "transferable interaction potential using three points" (TIP3P).[40] The MT/BP method is implemented in the plane wave DFT program CPMD[41] and comprises the gradient-corrected exchange functional of Becke,[42] the correlation functional of Perdew,[43] and the norm-conserving pseudopotentials of Martin and Troullier[44] together with a plane wave cutoff of 70 Ry. Flavin Lennard-Jones parameters, which are additionally required for a complete specification of the DFT/MM interface[26] between the DFT solute and the TIP3P aqueous solvent, were given in our preceding paper[25] together with a complete DFT-derived MM force field for oxidized flavins, which enables pure MM−MD simulations of flavin in condensed phase.

For aqueous solutions of the oxidized flavin and of its isotope derivatives, our properly scaled DFT/MM calculations turned out to explain the shapes and spectral locations of the observed IR bands quite well.[25] Extending our preceding study, it is the purpose of this contribution to check at what quality our scaled DFT/MM approach additionally predicts the IR solution spectra of flavin in other redox states. The DFT/MM description of these IR spectra is of special interest to us, because we eventually want to apply our DFT/MM approach to the analysis of time-resolved IR data[11,45,46] on the light-induced reactions, which in BLUF domains lead from dark-adapted initial states to light-adapted signaling states, and because these reactions involve transiently reduced flavins.[11,46]

Our contribution is organized as follows. First, we will shortly sketch the computational methods subsequently applied to the computation of the IR spectra of reduced flavins in aqueous solution. Like in our preceding study,[25] also here we chose lumiflavin (LF), which is the methyl derivative (R = CH_3) of isoalloxazine, as our model compound, because it is smaller than RF, FMN, or FAD. Thus, we subsequently present the IR spectra calculated by our DFT/MM INMA technique for the anionic radical $LF^{\bullet-}$ and the neutral radical $LFH^{\bullet}$ in TIP3P water at ambient temperature and pressure. These results are compared with our previous DFT/MM results on oxidized flavin[25] and with spectroscopic data.[3,12,13,17,19,20,47]

METHODS

Nearly all aspects of the computational methods applied here to calculate the IR spectra of reduced LFs by DFT/MM and INMA were thoroughly described in our preceding study[25] on oxidized LF solvated in water. Therefore, a condensed sketch of these methods shall suffice except for the differences required by the DFT treatment of the unpaired electron in $LF^{\bullet-}$ and $LFH^{\bullet}$ and associated with the modified MM force fields newly derived here for these two reduced flavins.

The DFT/MM calculations were carried out with the hybrid method developed in ref 26, which provides an interface between the MM-MD program EGO[48] and the plane-wave DFT code CPMD.[41] To account for the unpaired electron in the radicals, we combined the local spin density approximation,[49] as provided by CPMD, with the DFT functionals, pseudopotentials, and plane-wave cutoffs characterizing the MT/BP approach sketched above.

For both LF radicals, a MM force field was derived from MT/BP calculations exactly following the procedures explained in ref 25 for the case of oxidized LF. The resulting MM force fields are compatible with and have been partially adopted from CHARMM22.[50] The parameters calculated for the CHARMM22-type force fields of $LF^{\bullet-}$ and $LFH^{\bullet}$ are listed in Tables S2−S5 of the Supporting Information.

For the generation of the INMA snapshot ensembles of the LF radicals in water, we also adopted the procedures described in the preceding paper.[25] Thus, in a pure MM-MD setting, the LF radicals were solvated in periodic boxes containing 1970 TIP3P water moleceules and were equilibrated in the Gibbs ensemble for 1 ns at the temperature $T = 300$ K and the pressure $p = 1$ atm. Subsequently, for each radical, a set of 20 snapshots was taken at 10 ps time intervals from a 200 ps trajectory, from which solution adapted atomic partial charges were calculated by DFT/MM as average "electrostatic potential" (ESP)[26,51] charges. With the new partial charges, the systems were once again equilibrated for 200 ps in the NpT ensemble. Afterward, the systems were simulated in the NVT ensemble for 500 ps. Every 10 ps, which is long enough to guarantee mutual statistical independence,[31] a total of 50 snapshots were taken representing the INMA ensembles for the DFT/MM computation of the liquid phase IR spectra. Additional 1 ns NVT simulations served to compute radial distribution functions of the water molecules in the neighborhood of the flavin oxygen atoms O2 and O4 (system configurations sampled every 200 fs).

While freezing for all members of the INMA ensembles the solvent cages at the snapshot configurations, the flavin radical structures of minimum potential energy were computed by steepest gradient descent. To evaluate the IR intensities, the dipole gradients were calculated by finite differences displacing each atom symmetrically by ±0.01 Å from its equilibrium position. For each INMA ensemble, the normal modes were classified by the automated mode classification procedure developed in ref 25. The normal modes of the particular snapshot, which was determined to be the best representative for the respective INMA ensemble, were characterized by calculating potential energy distributions (PED) with the help of the program Gamess.[52]

From the DFT/MM snapshot ensembles, we performed mode specific statistics to obtain Gaussian bands whose combination models the liquid phase IR spectra. Applying our automated mode classification, we calculated for both radicals the INMA ensemble average intensities $\langle I_\alpha \rangle$, frequencies $\langle \nu_\alpha \rangle$, and variances $\rho_\alpha{}^2 = \langle (\nu_\alpha - \langle \nu_\alpha \rangle)^2 \rangle$ for all modes α with $\nu_\alpha \in [1250,1750]$ cm^{-1}. Because INMA neglects motional narrowing and, therefore, overestimates the band widths ρ_α,[32] we applied an approximate correction by scaling ρ_α with 0.6. Denoting the speed of light by c_0, the liquid phase IR spectra are then modeled as the superpositions[25,33,34]

$$S(\nu) = \frac{c_0}{6} \sum_\alpha \frac{\langle I_\alpha \rangle}{\sqrt{2\pi\sigma_\alpha{}^2}} \exp\left[\frac{-(\nu - \langle I_\alpha \rangle)^2}{2\sigma_\alpha{}^2}\right] \qquad (1)$$

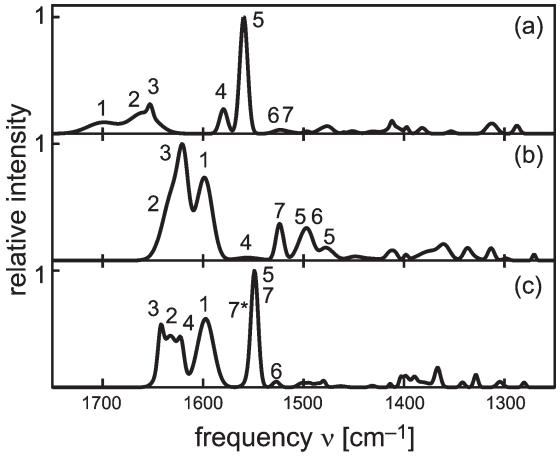


Figure 2. DFT/MM description of the IR spectra of (a) LF,[25] (b) $LF^{\bullet-}$, and (c) $LFH^{\bullet}$ in TIP3P water obtained from the respective INMA ensembles by our automated mode matching[25] and modeled by the Gaussian superposition eq 1. Seven marker bands are labeled by the mode identifiers κ.

of Gaussian peaks centered at the mode frequencies $\langle \nu_\alpha \rangle$ and inhomogeneously broadened with the scaled standard deviations $\sigma_\alpha = 0.6\rho_\alpha$. The validity of the Gaussian superposition model eq 1 is demonstrated in the Supporting Information. Here, Figure S9 compares a suitable representation of the original INMA data with the Gaussian model, and the accompanying text guides through this comparison.

RESULTS AND DISCUSSION

We start the presentation of our DFT/MM results on the IR spectra of the LF radicals by first showing how the sequential addition of an electron and a proton affects the intensities, frequencies, and widths of seven important flavin marker bands $\kappa = 1, \ldots, 7$, which are identfied in and extracted from the respective INMA ensembles by our automated mode matching procedure.[25] For five of the corresponding modes and for the oxidized state, a detailed analysis has been given in ref 25. As mentioned in the Introduction, in this preceding paper, we additionally determined the method specific scaling factor of 1.031. We now applied this factor to all computed INMA frequencies for a hopefully optimal match with experimental data.

IR Spectra of LF Radicals in Water as Predicted by DFT/MM. Figure 2 compares the DFT/MM predictions constructed by eq 1 for the three LF redox states under consideration. Thus, Figure 2 refers (a) to the oxidized LF,[25] (b) to the anionic radical $LF^{\bullet-}$, and (c) to the neutral radical $LFH^{\bullet}$ in TIP3P water (temperature 300 K, pressure 1 atm). The bands of the seven marker modes are labeled in parts a−c by the mode identifiers κ. Each spectrum is normalized to its respective intensity maximum, which has a different value for each of the three LF states. Therefore, the visual comparison of parts a−c of Figure 2 can solely indicate shifts of frequencies and band widths that are caused by sequential electron and proton attachment.

Already the first glance at Figure 2 demonstrates that the IR spectra of LF, $LF^{\bullet-}$, and $LFH^{\bullet}$ are predicted to be drastically different concerning the relative spectral orderings, intensities, and widths of the assigned marker bands. Here, the changes of the relative spectral orderings are visible through the shown permutation of the mode identifiers κ in going from part a to c. Because the understanding of the underlying frequency shifts is greatly simplified by a different representation of the data, which will be given further below, we will first concentrate on the overall differences of spectral shapes.

Table 1. Shapes of Marker Bands in LF Radicals[a]

	$LF^{\bullet-}$		$LFH^{\bullet}$	
κ[b]	$\langle I \rangle / \langle I_{LF} \rangle$[c]	σ/σ_{LF}[d]	$\langle I \rangle / \langle I_{LF} \rangle$[c]	σ/σ_{LF}[d]
1	1.59	0.50	1.59	0.52
2	1.25	0.81	0.82	0.65
3	4.83	2.05	1.67	1.08
4	0.14	3.02	0.43	0.75
5	0.06	1.71	0.43	0.91
6	2.20	1.09	0.40	0.55
7	6.50	0.55		

[a] Band shapes are measured by shapes in oxidized LF. [b] Mode identifier obtained by automated matching.[25] [c] INMA ensemble average intensity $\langle I \rangle$. [d] Scaled standard deviation σ of INMA frequencies.

While the INMA model for the IR spectrum of LF (see Figure 2a) is dominated by the sharp and intense band 5 at 1560 cm^{-1} and additionally features two broad high-frequency bands at 1700 and 1659 cm^{-1}, respectively, the IR spectrum predicted for the anionic radical $LF^{\bullet-}$ in part b lacks any high-frequency bands. Instead, it is characterized by an intense and very broad double peak near 1615 cm^{-1} and three lower intensity peaks in the spectral region around 1500 cm^{-1}. Similarly, also spectrum c calculated for $LFH^{\bullet}$ lacks high-frequency bands and shows a somewhat narrower and, therefore, partially structured double peak near 1620 cm^{-1}. However, this spectrum is dominated by a sharp and intense peak at 1549 cm^{-1}, which is labeled by the mode identifiers 5, 7, and 7* and closely resembles the dominant peak 5 of the LF spectrum. Here, the new mode identifier 7* points to the in-plane vibration of the hydrogen atom H5 (cf. Figure 1) additionally present in $LFH^{\bullet}$.

Table 1 lists the changes of the average intensities $\langle I \rangle$ and line widths σ induced by the attachment of an electron and a proton by relating values calculated for the LF radicals to corresponding values in the oxidized LF.[25] According to our DFT/MM results, the electron transfer, which generates the anionic radical $LF^{\bullet-}$, increases the intensities of most bands. This effect is particularly pronounced for the bands belonging to modes 3 and 7. A notable exception, however, is the prominent LF band 5, whose intensity strongly decreases upon electron attachment. Note that bands 1, 2, and 7 become sharper and the other marker bands broader in this transition. Adding a proton yields the neutral radical $LFH^{\bullet}$. Compared to LF, our DFT/MM treatment predominantly predicts for $LFH^{\bullet}$ decreased marker band intensities (the exceptions are bands 1 and 3) and smaller band widths (an exception is band 3). As a result, the IR spectrum of $LFH^{\bullet}$ should be less intense and much more strongly structured than that of LF, whereas the partially sharpened and partially broadened spectrum of $LF^{\bullet-}$ should be more intense, if one disregards the sharp and dominant peak 5 in the IR spectrum of LF.

The compositions of those normal modes, which give rise to spectral features, are scrutinized in the Supporting Information. Here, Figure S8 graphically displays for LF, $LF^{\bullet-}$, and $LFH^{\bullet}$ and for the seven marker modes κ the atomic displacements calculated by DFT/MM for the three snapshots, which our automated mode classification[25] determined to be the best representatives of the respective INMA ensembles. The graphs in Figure S8 clearly reveal both the partial changes and partial invariances of the

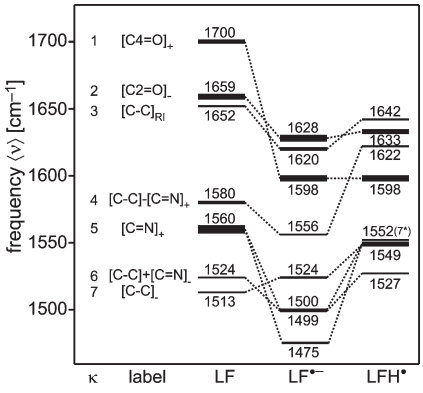


Figure 3. Average INMA frequencies $\langle\nu\rangle$ calculated by DFT/MM for the seven marker bands κ of LF, $LF^{\bullet-}$, and $LFH^{\bullet}$ in TIP3P water, and scaled by 1.031. The mode intensities $\langle I\rangle$ are visualized by the thicknesses of the bars marking the frequencies. Here, thicker bars code larger intensities. The assignment of the normal modes resulting from our automated classification[25] is given by the mode labels [X−Y], which point to contributing stretches of bonds X−Y (see the text), and by the dotted lines connecting the frequency levels.

normal mode compositions that are caused by sequentially adding an electron and a proton to the oxidized LF. The Supporting Information additionally contains three tables (Tables S6−S8), which list the PEDs calculated for 13 high-frequency normal modes including, of course, the seven marker modes. These data enable us to characterize the various modes by their most important local contributions, which, in the case of the marker modes, are stretches [X—Y] of certain bonds connecting atoms X and Y. In the following analysis of frequency shifts, we will use corresponding mode labels.

For a better insight into the spectral band shifts, Figure 3 displays the average and scaled INMA frequencies $\langle\nu\rangle$ for the oxidized LF, the anionic radical $LF^{\bullet-}$, and the neutral radical $LFH^{\bullet}$ in TIP3P water. Furthermore, the average intensities $\langle I\rangle$ are indicated by the thicknesses of the bars. The subscripts "+" and "−" at the mode labels [X—Y] denote in-phase and out-of-phase relations, respectively, among the stretches of bonds X—Y contributing to the normal modes (cf. Figure S8 in the Supporting Information). The subscript "RI" points to ring I of isoalloxazine (cf. Figure 1).

Figure 3 demonstrates that nearly all LF bands get more or less strongly red-shifted upon electron attachment with the sole exception of the out-of-phase C−C single bond stretching mode 7, which experiences a 11 cm^{-1} blue-shift. Adding a proton and thereby localizing the added electron then shifts nearly all vibrational frequencies to the blue (except for mode 1, whose frequency remains invariant). Like the red-shifts of the LF → $LF^{\bullet-}$ transition, also the $LF^{\bullet-}$ → $LFH^{\bullet}$ blue-shifts are of varying size, which we now consider in some more detail.

In the oxidized LF, the highest frequency band 1 at 1700 cm^{-1} belongs to the $[C4{=}O]_+$ stretching vibration. This band is strongly red-shifted by as much as 102 cm^{-1} in both radical states. Concurrently, mode 1 gains a considerable admixture of the C2=O2 stretch, which is larger for $LFH^{\bullet}$ than for $LF^{\bullet-}$ (cf. the Supporting Information). According to our DFT/MM treatment the $[C2{=}O]_-$ band, which in LF is calculated at 1659 cm^{-1}, is red-shifted by 31 cm^{-1} in $LF^{\bullet-}$ and by 26 cm^{-1} in $LFH^{\bullet}$, respectively. As compared to LF, mode 2 features enhanced contributions of the C4=O4 stretching and N3—H bending vibrations in both radical states (cf. the Supporting Information). Here, it now represents the carbonyl stretch of highest frequency.

The dominant band in the IR spectrum of LF belongs to mode 5 and is calculated at 1560 cm^{-1}. Mode 5 mainly consists of the in-phase combination $[C{=}N]_+$ of the C=N double bonds in isoalloxazine (cf. Figure 1). According to Figure 3, in $LF^{\bullet-}$, band 5 is split into two differently red-shifted and much weaker bands at 1500 and 1475 cm^{-1}, respectively, which contain additional contributions of C−C stretches and methyl deformations (see the Supporting Information). Protonation shifts this band again sizably to the blue to 1549 cm^{-1} and increases its intensity, such that its spectral position and intensity in $LFH^{\bullet}$ resemble the corresponding LF values (up to a 11 cm^{-1} red-shift and a somewhat smaller intensity). In $LFH^{\bullet}$, the underlying mode exhibits large admixtures of mode 7 and is spectrally close to a mode (7*), which is a combination of mode 7 with the in-plane bending vibration of the hydrogen atom added at N5. Note that mode 4 shows even a larger blue-shift than mode 5 in the $LF^{\bullet-}$ → $LFH^{\bullet}$ transition. As a result, the corresponding band appears at 1622 cm^{-1}, i.e., way above the lowest frequency carbonyl vibration 1 at 1598 cm^{-1}.

Most remarkable among the many spectral differences between LF and the two radicals, which we have discussed above, are perhaps the differently large red-shifts of the two carbonyl bands and the resulting reversal of their spectral ordering. To provide a physical interpretation of these findings, we will now discuss the changes of carbonyl solvation induced by electron and proton attachment.

Solvation of the Carbonyl Groups for Different LF Redox States. The solvation of the two LF carbonyl groups can be described by the radial distribution functions $g_{OH}(r)$, which provide statistics on the distances r between the carbonyl oxygens O and the water hydrogens H occurring in the surrounding liquid. Such distribution functions were extracted from 1 ns MM-MD simulations of different LF redox states embedded in TIP3P water (see the Methods section).

Figure 4 shows the distribution functions $g_{OH}(r)$ thus obtained for the atoms O2 (a) and O4 (b) of LF (solid lines), $LF^{\bullet-}$ (dashed lines), and $LFH^{\bullet}$ (dotted lines) in TIP3P water at 300 K and ambient pressure. A close inspection of the figure indicates that for all three redox states more H atoms are found near O2 than near O4. Integrating the functions $g_{OH}(r)$ up to a distance of 2.5 Å and, thus, covering the first solvation shells, one finds for LF on average 2.6 hydrogens near O2 and 2.3 hydrogens near O4. Thus, in the oxidized state, O2 is more strongly hydrogen bonded than O4. In the LF → $LF^{\bullet-}$ transition, the number of hydrogens in the first solvation shell increases by 15% at O2 and by 19% at O4. In the LF → $LFH^{\bullet}$ transition, the increases are slightly smaller, measuring 14% at O2 and 16% at O4.

The differently strong solvation of the two carbonyl groups can be explained by our DFT/MM results, which also comprise ESP atomic charges.[26,51] According to these partial charges, which are listed for the two radicals in Table S2 of the Supporting Information and are given for LF in ref 25, the dipole moments of the two C=O bonds differ for all LF redox states. For instance, in LF, the dipole moment p_2 of the C2=O2 bond is 4.0 D and, therefore, much larger than the dipole moment p_4 = 2.6 D of the C4=O4 bond, which nicely explains why the C2=O2 group is more strongly solvated than the C4=O4 group. The larger solvation of O2 and the correspondingly stronger reaction field generated by the water and polarizing the C2=O2 bond[53] are the reasons why the solvatochromic shift experienced by the LF band

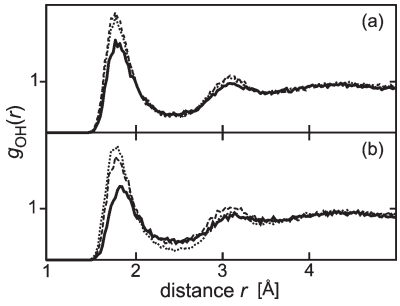


Figure 4. Radial distribution functions $g_{OH}(r)$ of the water hydrogen atoms surrounding the flavin oxygen atoms O2 (a) and O4 (b) of LF (solid lines), $LF^{\bullet-}$ (dashed lines), and $LFH^{\bullet}$ (dotted lines) in TIP3P water.

2 upon transfer from the gas phase into water is much larger than the shift of the LF band 1 (see ref 25). One may therefore assume that the different increases of solvation noted above for the two radicals find a correspondence in different increases of the dipole moments p_2 and p_4.

In the LF → $LF^{\bullet-}$ transition, the dipole moment p_2 increases by 11%, whereas p_4 grows by as much as 37%, explaining the larger increase of solvation at O4 than at O2 (see above). For the LF → $LFH^{\bullet}$ transition, the corresponding numbers are 9% for p_2 and 38% for p_4. Therefore, in the transition to the radicals, the reaction fields polarizing the C═O bonds become much more strengthened at O4 than at O2, explaining why the $[C4{=}O]_+$ band experiences larger red-shifts than the $[C2{=}O]_-$ band in the transition to the radical states (cf. Figure 3). Note here that p_2 is slightly smaller for $LFH^{\bullet}$ than for $LF^{\bullet-}$. Correspondingly, the $[C2{=}O]_-$ band becomes slightly blue-shifted upon protonation of $LF^{\bullet-}$.

Comparison to Experimental Vibrational Spectra. As we have seen above, many spectral shifts associated with the transition from LF to its singly reduced relatives $LF^{\bullet-}$ and $LFH^{\bullet}$ can be consistently explained in terms of the altered charge distributions within the chromophores and the correspondingly modified interactions with the highly polar aqueous environments. While such physical discussions may be intellectually appealing, the question remains whether the underlying descriptions are actually true.

For an answer to this question, a comparison with experimental evidence is necessary. Here, one has to resort to a few FTIR difference spectra between oxidized FMN and FAD and their anionic radicals in aqueous solution.[19,20] Note that one cannot easily extract intensity estimates from such difference spectra. More extended data on the vibrational spectra of flavin radicals, mainly on the ionic states $F^{\bullet-}$, are provided by resonance Raman (RR) spectroscopy.[12,13,17,47] Only one author[47] presents RR data also for a neutral flavin radical $FH^{\bullet}$ in H_2O, namely, for that of RF. For this neutral radical, some of the IR intensity information, which is missing in the RR data, can be extracted from time-resolved IR difference spectroscopy of $RFH^{\bullet}$ in acetonitrile.[3]

Despite the quite sparse data basis, we have tried to extract from the quoted experimental spectra the frequencies and intensities of the marker bands in the radicals as good as possible. The results of this analysis are the "experimental" term schemes shown in Figures 5 and 6 and labeled by the capital letters "E". Correspondingly, the DFT/MM results are marked by the letters "T".

To provide a first estimate of the quality at which our computational results match available experimental data, Figure 5 compares the scaled average INMA frequencies $\langle\nu\rangle$ of LF[25] and

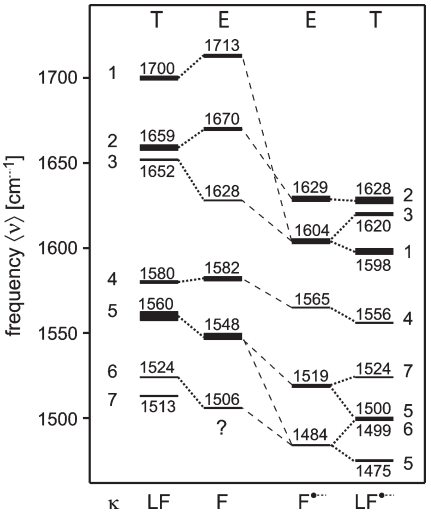


Figure 5. Average INMA frequencies $\langle\nu\rangle$ calculated by DFT/MM (T) for LF[25] and $LF^{\bullet-}$ and scaled by 1.031 are compared to experimental frequencies (E) of F and $F^{\bullet-}$ extracted by us from vibrational spectra.[12,13,17,19,20,47] The dashed lines connecting the experimental frequencies of F and $F^{\bullet-}$ indicate mode assignments, which we obtained from applications of our automated mode matching procedure to the INMA ensembles of LF and $LF^{\bullet-}$. The dotted lines mark our assignment of observed bands to calculated modes.

$LF^{\bullet-}$ in TIP3P water with the vibrational frequencies observed for the flavins (F) FMN and FAD, respectively, and for the associated anionic radicals ($F^{\bullet-}$) in aqueous solution.

The IR spectra[19,20] of the anionic radical $F^{\bullet-}$ show a double peak consisting of two intense bands at 1629 and 1604 cm^{-1}, respectively. Inspection of Figure 2 demonstrates that also our DFT/MM spectrum features a double peak with maxima at 1625 and 1598 cm^{-1}. According to the term scheme in Figure 5, these two peaks belong to the $[C2{=}O]_-$ and $[C4{=}O]_+$ stretching modes calculated for $LF^{\bullet-}$ at 1628 cm^{-1} (mode 2) and 1598 cm^{-1} (mode 1), respectively.[54] As indicated by the dashed lines in Figure 5, we assign the $[C{-}C]_{RI}$ stretching vibration within ring I (mode 3) also to the second peak, because the corresponding frequency is overestimated by our calculations for the oxidized flavin and because we assume that a corresponding frequency overestimate is likely to occur also in the anionic radical. As a result, the somewhat weaker band calculated by us at 1620 cm^{-1} should be buried under the intense peak observed at 1604 cm^{-1} in the FTIR difference spectra.[19,20]

The assignment of mode 4 calculated at 1556 cm^{-1} for $LF^{\bullet-}$ to the experimental band at 1565 cm^{-1} is obvious considering the assignment of this mode in LF and the pattern of calculated and observed frequency shifts resulting from electron attachment. Mode 5, in contrast, is less easily assigned. Our calculations predict that the prominent LF band 5 splits into two weaker bands upon electron transfer. Therefore, the associated IR band of F at 1548 cm^{-1} could also undergo such a splitting into the $F^{\bullet-}$ bands at 1519 and 1484 cm^{-1}.[19,20] In the IR spectrum, the weak bands 6 and 7 are likely to be buried under the more intense bands resulting from the splitting.

In summary, for the anionic radical in water, the root-mean-square deviation (rmsd) between DFT/MM and experimental band positions is 10.6 cm^{-1}. This value is even a little smaller than the optimal rmsd of 12.2 cm^{-1}, which was achieved[25] for five marker modes of oxidized flavin in water with a scaling factor

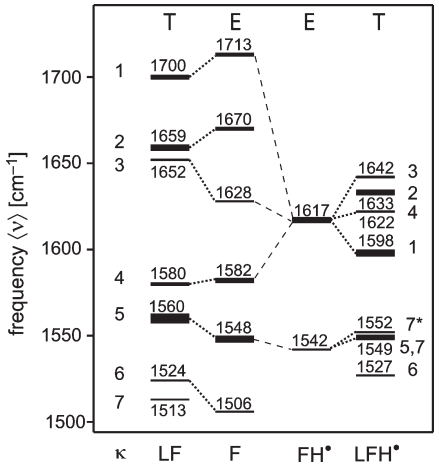


Figure 6. Average INMA frequencies $\langle\nu\rangle$ calculated by DFT/MM (T) for LF[25] and LFH$^{\bullet}$ and scaled by 1.031 are compared to the experimental frequencies (E) of F and FH$^{\bullet}$. See the caption to Figure 5 for further explanations.

of 1.031. This result nicely confirms our expectation that the carefully chosen scaling factor[25] should be specific solely with respect to the applied computational method but should otherwise be applicable to DFT/MM descriptions of all kinds of flavins in the condensed phase. Clearly, one favorable example could be considered as rather weak evidence. Therefore, we additionally consider the neutral flavin radical in water, although here the experimental data basis is much smaller.

Figure 6 compares the scaled average INMA frequencies $\langle\nu\rangle$ calculated by DFT/MM (T) for LFH$^{\bullet}$ in TIP3P water with vibrational frequencies (E) observed for the neutral RF radical FH$^{\bullet}$ in aqueous solution by RR[47] spectroscopy. The latter RR spectrum exhibits two large and broad bands at 1617 and 1542 cm^{-1}. Interestingly, the FTIR spectrum of FH$^{\bullet}$ in the much less polar and protic solvent acetonitrile features an intense band at 1660 cm^{-1}, a nearby weaker band at 1612 cm^{-1}, and two equally weaker bands at 1532 and 1508 cm^{-1}, respectively. Because the intense 1660 cm^{-1} band, which likely belongs to C═O vibrations, should experience a sizable red-shift upon transfer into the strongly polar and protic solvent water, the RR band at 1617 cm^{-1} must also be assigned to such modes. Similarly, the two acetonitrile bands above 1500 cm^{-1} could contribute to the broad and unstructured RR band observed in H_2O at 1542 cm^{-1}. If one looks at Figure 2c and imagines the application of a strongly smoothening low-pass filter, one ends up with two broad bands, one peaking at 1618 cm^{-1} and the other at 1540 cm^{-1}, which is in nearly perfect agreement with the RR spectrum.[47] The following points may explain the difference between the structured and rather sharp IR bands in Figure 2c and the broad and unstructured bands observed in the corresponding RR spectrum: (i) the band widths in the IR spectra calculated by the INMA procedure and the kernel estimate eq 1 may be underestimated because of insufficient statistics and of too strongly scaled frequency standard deviations ρ_α; (ii) intensities and band shapes in IR and RR spectra will generally differ. Thus, although the information on the experimental spectra of FH$^{\bullet}$ in H_2O is quite sparse, also in this case our DFT/MM approach enables not only the assignment of observed bands to normal modes but also a nearly quantitative prediction of peak positions.

Summary and Outlook. In a previous contribution,[25] we established a scaled DFT/MM-INMA approach that quite accurately *models* the IR spectra of oxidized flavins in aqueous solution and additionally assigns normal mode compositions to observed bands. By applying this computational method to the two flavin radicals F$^{\bullet-}$ and FH$^{\bullet}$, respectively, we have now clearly demonstrated with the results presented and discussed above that the method can equally well *predict* the IR solution spectra of flavins in other redox states. Here, we were able to predict a set of complex spectral changes that are caused by electron and proton attachment and cover the spectral locations, intensities, widths, and mode compositions of the IR bands. Because our calculated frequencies and band shapes turned out to excellently match the available experimental data, we were able to assign the observed bands to normal modes and to explain their spectral locations for the given aqueous solvent.

The apparent success of our DFT/MM-INMA method in predicting the solution IR spectra of reduced flavins and in modeling these spectra in the oxidized case now nourishes the hope that this method should be capable of predicting the IR spectra of flavins in protein environments at a comparable accuracy. For this purpose, it will be necessary to convert crystallographic data on the respective flavoprotein into a hopefully complete structural MM model, which in particular should contain all embedded water molecules and the correct protonation states of the titratable groups. For a correspondingly realistic MM model of the protein, our DFT/MM-INMA method should then be able to predict, e.g., the positions of the flavin vibrational bands with a rmsd of at most 15 cm^{-1} between computational and spectroscopic results.

In a next step, we therefore intend to apply our DFT/MM-INMA procedure to BLUF domains, for which several X-ray structures are available,[55−57] because comparisons of calculated IR spectra with time-resolved IR data[11,46] can help to decipher structural details of the chromophore−protein interactions and their changes caused by the functional dynamics of the light induced photocycles.[11,45,46] Furthermore, such comparisons may even help to decide which of the existing structural models for these functional dynamics are correct.[11,45,46,58,59]

■ ASSOCIATED CONTENT

Ⓢ Supporting Information. The supporting material contains figures, tables, equation, and various pieces of text explaining and documenting (i) our CHARMM22-type force field for LF and FMN radicals, (ii) the results of the automated mode matching for the radical states, (iii) the PED assignment for the seven discussed and six additional vibrational bands of LF radicals in aqueous solution, and (iv) the validity of the Gaussian model eq 1 for the DFT/MM-IR spectra derived by the INMA procedure. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: tavan@physik.uni-muenchen.de. Phone: +49 (0)89 2180 9220. Fax: +49 (0)89 2180 9202.

■ ACKNOWLEDGMENT

This work was supported by the Forschergruppe 526 "Sensory Blue Light Receptors" (DFG/FOR526) and by the

Sonderforschungsbereich 749 "Dynamics and Intermediates of Molecular Transformations" (DFG/SFB749-C4) of the Deutsche Forschungsgemeinschaft.

REFERENCES

(1) Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemie*, 5th ed.; Spektrum Akademischer Verlag GmbH: Heidelberg, Berlin, Germany, 2003.
(2) Mattevi, A. *Trends Biochem. Sci.* **2006**, *31*, 276–283.
(3) Martin, C. B.; Tsao, M.-L.; Hadad, C. M.; Platz, M. S. *J. Am. Chem. Soc.* **2002**, *124*, 7226–7234.
(4) van der Horst, M. A.; Hellingwerf, K. J. *Acc. Chem. Res.* **2004**, *37*, 13–20.
(5) Sancar, A. *Biochemistry* **1994**, *33*, 2–9.
(6) Ahmad, M.; Cashmore, A. R. *Nature* **1993**, *366*, 162–166.
(7) Christie, J. M.; Swartz, T. E.; Bogomolni, R. A.; Briggs, W. R. *Plant J.* **2002**, *32*, 205–500.
(8) Gomelsky, M.; Klug, G. *Trends Biochem. Sci.* **2002**, *27*, 497–500.
(9) Schelvis, J. P. M.; Ramsey, M.; Sokolova, O.; Tavares, C.; Cecala, C.; Connell, K.; Wagner, S.; Gindt, Y. M. *J. Phys. Chem. B* **2003**, *107*, 12352–12362.
(10) Masuda, S.; Hasegawa, K.; Ono, T. *Plant Cell Physiol.* **2005**, *46*, 1894–1901.
(11) Bonetti, C.; Mathes, T.; van Stokkum, I. H. M.; Mullen, K. M.; Groot, M.-L.; van Grondelle, R.; Hegemann, P.; Kennis, J. T. M. *Biochem. J.* **2008**, *95*, 4790–4802.
(12) Hazekawa, I.; Nishina, Y.; Sato, K.; Shichiri, M.; Miura, R.; Shiga, K. *J. Biochem.* **1997**, *121*, 1147–1154.
(13) Nishina, Y.; Sato, K.; Miura, R.; Matsui, K.; Shiga, K. *J. Biochem.* **1998**, *124*, 200–208.
(14) Kondo, M.; Neppa, J.; Ronayne, K. L.; Stelling, A. L.; Tonge, P. J.; Meech, S. R. *J. Phys. Chem. B* **2006**, *110*, 20107–20110.
(15) Siebert, F.; Hildebrandt, P. *Vibrational Spectroscopy in Life Science*; Wiley-VCH: Berlin, 2007.
(16) Abe, M.; Kyogoku, Y.; Kitagawa, T.; Kawano, K.; Ohishi, N.; Takai-Suzuki, A.; Yagi, K. *Spectrochim. Acta, Part A* **1986**, *42*, 1059–1068.
(17) Copeland, R.; Spiro, T. *J. Phys. Chem.* **1986**, *90*, 6654–6657.
(18) Birss, V. I.; Hinman, A. S.; McGarvey, C. E.; Segal, J. *Electrochim. Acta* **1994**, *39*, 2449–2454.
(19) Hellwig, P.; Scheide, D.; Bungert, S.; Mäntele, W.; Friedrich, T. *Biochemistry* **2000**, *39*, 10884–10891.
(20) Wille, G.; Ritter, M.; Friedemann, R.; Mäntele, W.; Hübner, G. *Biochemistry* **2003**, *42*, 14814–14821.
(21) Nishina, Y.; Sato, K.; Setoyama, C.; Tamaoki, H.; Miura, R.; Shiga, K. *J. Biochem.* **2007**, *142*, 265–272.
(22) Wolf, M. M. N.; Schumann, C.; Gross, R.; Domratcheva, T.; Diller, R. *J. Phys. Chem. B* **2008**, *112*, 13424–13432.
(23) Abe, M.; Kyogoku, Y. *Spectrochim. Acta, Part A* **1987**, *43*, 1027–1037.
(24) Klaumünzer, B.; Kröner, D.; Saalfrank, P. *J. Phys. Chem. B* **2010**, *114*, 10826–10834.
(25) Rieff, B.; Mathias, G.; Bauer, S.; Tavan, P. *Photochem. Photobiol.*; published online 17 Dec 2010, DOI: 10.1111/j.1751-1097.2010.00866.x.
(26) Eichinger, M.; Tavan, P.; Hutter, J.; Parrinello, M. *J. Chem. Phys.* **1999**, *110*, 10452–10467.
(27) Hohenberg, P.; Kohn, W. *Phys. Rev. B* **1964**, *136*, 864–871.
(28) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, 1133–1138.
(29) Nonella, M.; Mathias, G.; Eichinger, M.; Tavan, P. *J. Phys. Chem. B* **2003**, *107*, 316–322.
(30) Babitzki, G.; Mathias, G.; Tavan, P. *J. Chem. Phys. B* **2009**, *113*, 10483–10495.
(31) Nonella, M.; Mathias, G.; Tavan, P. *J. Phys. Chem. A* **2003**, *107*, 8638–8647.
(32) Schmitz, M.; Tavan, P. *J. Chem. Phys.* **2004**, *121*, 12233–12246.
(33) Schmitz, M.; Tavan, P. *J. Chem. Phys.* **2004**, *121*, 12247–12258.
(34) Schmitz, M.; Tavan, P. In *Modern methods for theoretical physical chemistry of biopolymers*; Tanaka, S., Lewis, J., Eds.; Elsevier: Amsterdam, The Netherlands, 2006; Chapter 8, pp 157−177.
(35) Geigeot, M. P.; Sprik, M. *J. Phys. Chem. B* **2003**, *107*, 10344–10358.
(36) Ramirez, R.; Lopez-Ciudad, T.; Kumar, P.; Marx, D. *J. Chem. Phys.* **2004**, *121*, 3973–3983.
(37) Baer, M.; Marx, D.; Mathias, G. *Angew. Chem., Int. Ed.* **2010**, *49*, 7346–7349.
(38) Neugebauer, J.; Hess, B. A. *J. Chem. Phys.* **2003**, *118*, 7215–7225.
(39) Schropp, B.; Wichmann, C.; Tavan, P. *J. Phys. Chem. B* **2010**, *114*, 6740–6750.
(40) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
(41) CPMD V3.9, Copyright IBM Corp 1990−2008, Copyright MPI für Festkörperforschung Stuttgart 1997−2001, see also www.cpmd.org.
(42) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
(43) Perdew, J. P.; Yue, W. *Phys. Rev. B* **1986**, *33*, 8800–8802.
(44) Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993–2005.
(45) Gauden, M.; Yeremenko, S.; Laan, W.; van Stokkum, I. H. M.; Ihalainen, J. A.; van Grondelle, R.; Hellingwerf, K. J.; Kennis, J. T. M. *Biochemistry* **2005**, *44*, 3653–3662.
(46) Gauden, M.; van Stokkum, I. H. M.; Key, J. M.; Lührs, D. C.; van Grondelle, R.; Hegemann, P.; Kennis, J. T. M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 10895–10900.
(47) Su, Y.; Tripathi, G. N. R. *J. Am. Chem. Soc.* **2010**, *116*, 4405–4407.
(48) Mathias, G.; Egwolf, B.; Nonella, M.; Tavan, P. *J. Chem. Phys.* **2003**, *118*, 10847–10860.
(49) Parr, R. G.; Yang, W. *Density functional theory of atoms and molecules*; Oxford Sience Publications: New York, 1989.
(50) MacKerell, A. D.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
(51) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129–145.
(52) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347–1363.
(53) Schultheis, V.; Reichold, R.; Schropp, B.; Tavan, P. *J. Phys. Chem. B* **2008**, *112*, 12217–12230.
(54) The intense IR band at 1629 cm^{-1} is absent in the RR spectra.
(55) Jung, A.; Domratcheva, T.; Tarutina, M.; Wu, Q.; Ko, W.; Shoeman, R.; Gomelsky, M.; Gardner, K.; Schlichting, I. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 12350–12355.
(56) Jung, A.; Reinstein, J.; Domratcheva, T.; Shoeman, R.-L.; Schlichting, I. *J. Mol. Biol.* **2006**, *362*, 717–732.
(57) Yuan, H.; Anderson, S.; Masuda, S.; Dragnea, V.; Moffat, K.; Bauer, C. *Biochemistry* **2006**, *45*, 12687–12694.
(58) Stelling, A.; Ronayne, K.; Nappa, J.; Tonge, P.; Meech, S. *J. Am. Chem. Soc.* **2007**, *129*, 15556–15564.
(59) Toh, K. C.; van Stokkum, I. H. M.; Hendriks, J.; Alexandre, M. T. A.; Arents, J. C.; Perez, M. A.; van Grondelle, R.; Hellingwerf, K. J.; Kennis, J. T. M. *Biochem. J.* **2008**, *95*, 312–321.

2.4 Zusatzinformationen zu Lumiflavinradikalen in Wasser

Der folgende Abdruck [1]

„Supporting Information to the manuscript:
IR spectra of flavins in solution: DFT/MM description of redox effects“
Benjamin Rieff, Sebastian Bauer, Gerald Mathias und Paul Tavan

enthält die zusätzlichen Informationen über die Betrachtungen der DFT/MM generierten IR Schwingungsspektren von Lumiflavinradikalen in Wasser, die nicht im Haupttext der zweiten Veröffentlichung enthalten sind. Auf elf Seiten, gegliedert in 4 Abschnitte, stellt die Supporting Information insgesamt drei Abbildungen, sieben Tabellen, eine Gleichung und verschiedene Diskussionen bereit, die (i) das CHARMM22 ähnliche MM Kraftfeld für Lumiflavinradikale, (ii) die Resultate der automatisierten Normal-Moden Analyse der Radikalzustände, (iii) die Bandenzuordnung der sieben diskutierten sowie sechs weiterer Schwingungsbanden der Lumiflavinradikale in Wasser und (iv) die Aussagekraft der zu Grunde gelegten Gauss‘schen Modelle für die DFT/MM-INMA Spektren von Lumiflavinradikalen dokumentieren und erklären.

[1] Reproduced with permission. Copyright 2011 American Chemical Society. Kostenlos einsehbar unter http://pubs.acs.org.

Supporting information to the manuscript

IR spectra of flavins in solution:

DFT/MM description of redox effects

Benjamin Rieff, Sebastian Bauer, Gerald Mathias, and Paul Tavan*

Lehrstuhl für Biomolekulare Optik, Ludwig-Maximilians-Universität,

Oettingenstr. 67, 80538 München, Germany

E-mail: tavan@physik.uni-muenchen.de

Phone: +49 (0)89 2180 9220. Fax: +49 (0)89 2180 9202

*To whom correspondence should be addressed

CHARMM22 type force field

The CHARMM22[1] force field does not cover flavins. Because we needed corresponding parameter sets for our MM-MD simulations of the LF radicals in aqueous solution, we had to generate such sets. As explained for oxidized LF in our previous publication[2] we carried out, for the purpose of parameterization, DFT calculations on the isolated LF radicals and DFT/MM calculations on the LF radicals in TIP3P water to determine many of the required parameters and applied homology modeling to the remaining ones.

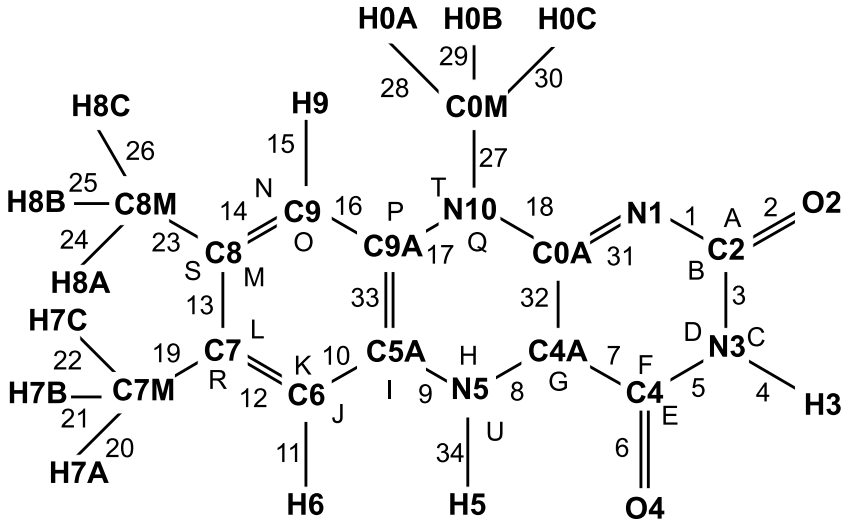


Figure S7: Chemical structure and atom labels of the neutral radical LFH$^{\bullet}$ and a definition of labels for bonds (numbers) and angles (capital letters).

The nomenclature used for the labeling of the atoms, bond lengths, and bond angles in LF is displayed in Figure S7, where its is attached to the structure of the neutral radical. Employing this nomenclature, the CHARMM22-type parameter sets obtained for the two radicals from our DFT and DFT/MM calculations are given in the Tables S2-S5.

Table S2: Atom types[a] and partial charges[b] of $LF^{\bullet -}$ and $LFH^{\bullet}$

atom	type	$q_{LF^{\bullet -}}$	$q_{LFH^{\bullet}}$	atom	type	$q_{LF^{\bullet -}}$	$q_{LFH^{\bullet}}$
N1	NN3A	−0.6997	−0.6519	C8	CA3	0.0358	0.0700
C2	CN1A	0.7173	0.7481	C9	CA4	−0.2900	−0.0292
O2	ON1	−0.7719	−0.7371	H9	HP	0.1894	0.1949
N3	NN2U	−0.3601	−0.4064	C9A	CPTA	−0.1429	−0.1901
H3	HN2	0.3288	0.3580	N10	NN2G	0.2134	0.2730
C4	CN1B	0.4671	0.4925	C0A	CN5A	0.2524	0.2618
O4	ON1	−0.6991	−0.6824	C7M	CT3	−0.2335	−0.3701
C4A	CN5B	0.1330	0.0224	H7X	HAI	0.0778	0.1234
N5	NN3A	−0.6746	−0.2742	C8M	CT3	−0.2299	−0.3645
C5A	CPTB	0.3605	0.1797	H8X	HAI	0.0767	0.1215
C6	CA1	−0.3186	−0.2904	C0M	CT3	−0.4024	−0.3952
H6	HP	0.1363	0.3012	H0X	HAI	0.1342	0.1318
C7	CA2	0.1228	0.1958	H5	HP	—	0.1644

[a]CHARMM22 atom types are associated to the atom names defined in Figure S7
[b]ESP partial charges q from DFT/MM calculations in TIP3P water averaged over 20 snapshots and given in units of e.

Table S2 associates certain atom types to the atom names defined in Figure S7. In CHARMM22 such types serve as selectors for force field parameters. The table extends the set of CHARMM22 standard types CN1, CN5, CPT, and CA by adding capital letters (A,B) or numbers (1,2,3,4) to uniquely define certain positions within isoalloxazine. The table also contains partial charges applicable to the anionic radical $LF^{\bullet -}$ ($q_{LF\bullet -}$) and to the neutral radical $LFH^{\bullet}$ ($q_{LFH\bullet}$) in TIP3P water. These charges were calculated by the procedures described in section Methods.

CHARMM22 models the bond stretches, the angle deformations, and the stiffnesses of sp^2 hybridized atoms by harmonic potentials specified by force constants and equilibrium values of the associated internal coordinates, which are the bond lengths l_i, the bond angles θ_j, and the so-called improper dihedral angles ϕ_k, respectively.

Table S3: Force constants[a] and equilibrium values[b] of radical bond lengths l_i

	$LF^{\bullet-}$		$LFH^{\bullet}$	
bond i [c]	k_l [a]	l_0 [b]	k_l [a]	l_0 [b]
1	325.9	1.371	287.0	1.382
2	666.7	1.245	691.5	1.229
3	207.0	1.410	176.0	1.429
4	478.1	1.021	459.8	1.022
5	259.5	1.404	281.2	1.384
6	691.3	1.242	667.7	1.244
7	248.5	1.466	262.6	1.450
8	336.2	1.358	355.8	1.365
9	318.6	1.356	297.3	1.379
10	331.1	1.421	362.7	1.402
11	376.1	1.094	375.5	1.094
12	332.1	1.395	341.7	1.395
13	315.4	1.415	312.9	1.417
14	336.6	1.410	353.5	1.402
15	377.9	1.092	382.1	1.090
16	338.8	1.400	334.8	1.402
17	289.1	1.400	285.9	1.402
18	268.7	1.407	270.1	1.404
19	259.6	1.512	265.1	1.507
23	263.3	1.508	266.9	1.507
27	272.6	1.455	260.2	1.463
20-22	349.5	1.100	356.5	1.100
24-26	349.5	1.100	356.5	1.100
28-30	349.5	1.100	356.5	1.100
31	369.8	1.334	381.2	1.320
32	331.6	1.429	362.5	1.416
33	306.0	1.436	322.8	1.424
34	—	—	459.3	1.027

[a]Force constants k_l in kcal/mol calculated by DFT;
[b]Equilibrium values l_0 in Å calculated by DFT;
[c]Bond lengths l_i, whose labels i are defined by the numbers in Figure S7.

Table S3 lists the parameters calculated for the anionic radical $LF^{\bullet-}$ and the neutral radical $LFH^{\bullet}$ by DFT for the harmonic potentials of isoalloxazine's bond lengths l_i. Table S4 displays the corresponding data for the bond angles θ_j of both radicals.

Table S4: Force constants[a] and equilibrium values[b] of radical bond angles θ_j

	$LF^{\bullet-}$		$LFH^{\bullet}$	
angle j [c]	k_θ [a]	θ_0 [b]	k_θ [a]	θ_0 [b]
A	220.00	124.12	250.00	123.75
B	249.75	116.85	230.00	117.47
C	90.00	116.16	105.00	115.95
D	91.60	127.96	90.00	126.44
E	400.00	120.21	225.00	124.66
F	81.00	112.97	100.00	112.38
G	145.00	118.85	110.00	118.06
H	250.00	117.40	250.00	122.10
I	85.00	119.73	80.00	121.96
J	53.75	116.20	53.75	118.46
K	220.70	123.10	250.00	121.72
L	78.50	119.19	75.00	118.68
M	77.00	118.93	77.00	119.56
N	52.20	118.26	75.00	118.45
O	230.00	121.96	230.00	122.24
P	55.00	122.37	52.25	122.22
Q	422.50	120.62	350.00	121.14
R	15.00	130.07	70.00	120.23
S	70.50	121.30	70.50	120.72
T	81.25	119.81	81.25	120.07
U	—	—	53.75	115.50
methyl at CA2	38.00	110.10	38.00	110.10
methyl at CA3	38.00	110.10	38.00	110.10
methyl at NN2G	40.50	110.10	40.50	110.10
methyl H-C-H	35.50	108.40	35.50	108.40

[a]Force constants k_θ in kcal/(mol·rad^2) calculated by DFT;
[b]Equilibrium values θ_0 in degrees calculated by DFT;
[c]Angles θ_j, whose labels j are given by the capital letters in Figure S7.

Table S5: Labels[a] of the radical improper dihedral angles ϕ_k

labels k^a				labels k^a			
ON1	NN3A	CN1A	NN2U	CPTB	NN3A	CA1	HP
NN3A	CN1A	NN2U	HN2	NN3A	CPTB	CA1	CA2
CN1A	NN3A	NN2U	CN1B	CA1	CPTB	CA2	CA3
CN1A	NN2U	CN1B	ON1	CA2	CA1	CA3	CA4
NN2U	CN1A	CN1B	CN5B	CA2	CA3	CA4	HP
NN2U	CN1B	CN5B	NN3A	CA3	CA2	CA4	CPTA
CN1B	CN5B	NN3A	CPTB	CA3	CA4	CPTA	NN2G
CN5B	NN3A	CPTB	CA1	CA4	CPTA	NN2G	CN5A
only for neutral radical LFH$^\bullet$				HN2	NN3A	CN5B	CN1B

[a]Labels k which are employed to guarantee in both radical states the planarity of isoalloxazine, are quadruples of atom types pointing to the four atoms, which are harmonically forced to be coplanar.

A dihedral angle ϕ_k is usually characterized by a label k denoting a set of four atoms. The labels k of the improper dihedrals ϕ_k, which are used by us to ensure the planarity of the two radicals, are listed in Table S5. All associated harmonic potentials have the same force constant $k_\phi = 100$ kcal/(mol·rad^2) and equilibrium value $\phi_0 = 180^\circ$.

The description of the LF radical force fields is complete, as soon as also the parameters of the Lennard-Jones potentials are given. For these parameters we adopted the values specified in the CHARMM22 force field[1] for the basic atom types.

The LF radicals were extended towards FMN radicals by attaching a glycerol chain and phosphate ion to atom N10 (cf. Figure S7). Because the CHARMM22 force field provides a parametrization for these two chemical motifs, we adopted these parametrizations to finish the modelling of FMN radicals.

Results of the automated mode classification

In our previous contribution[2] we have demonstrated that our automated classification of the modes contained in a DFT/MM-INMA snapshot ensemble is reliable. It allowed us to identify mode compositions and peak positions $\langle \nu \rangle$ at a high accuracy. Therefore it enabled a vibrational analysis of the solution IR spectra of oxidized LF in the spectral range from 1200 cm^{-1} to 1750 cm^{-1}.

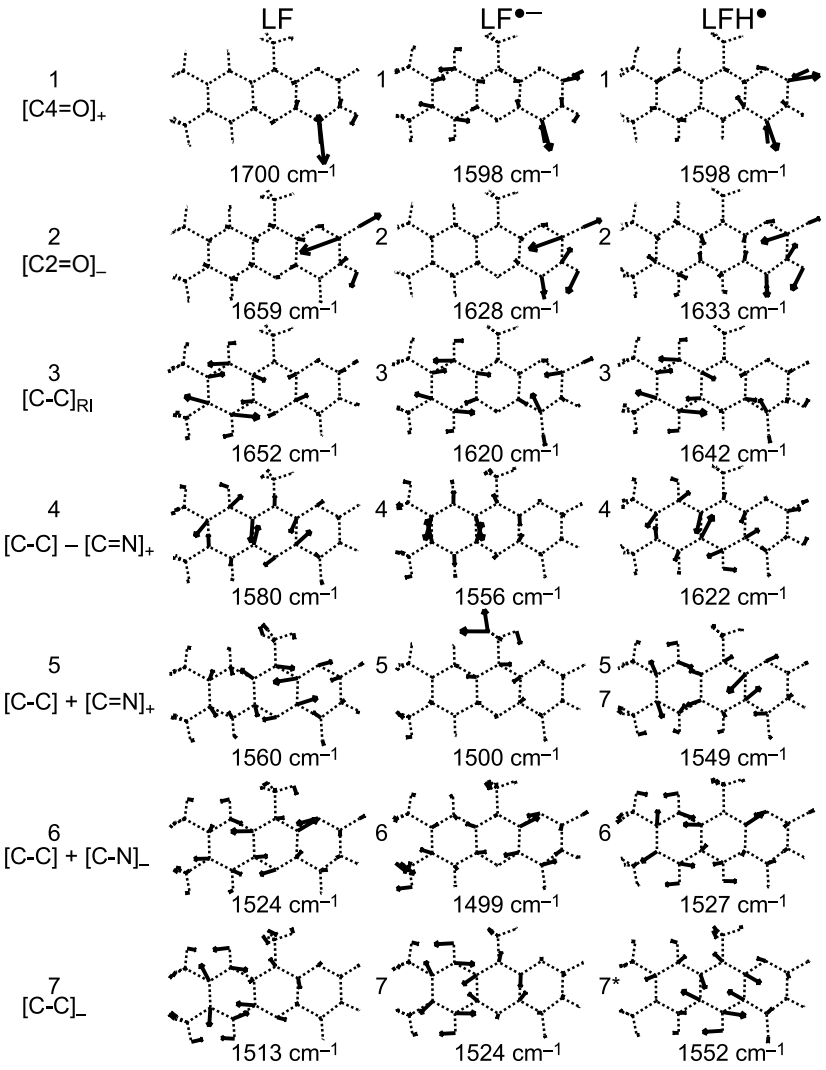


Figure S8: Drawings of the modes given by the best representative of the TIP3P snapshot ensemble of the oxidized LF, the anionic radical $LF^{\bullet-}$, and the neutral radical $LFH^{\bullet}$ computed for the seven observed bands. The symbols in the first column are the mode labels explained in the text. The frequencies are the scaled INMA averages $\langle \nu \rangle$ from our DFT/MM calculations.

Figure S8 displays graphical representations of the atomic displacements characterizing the seven observed modes given by the best representative of the TIP3P snapshot ensemble for the oxidized LF, the anionic radical $LF^{\bullet-}$, and the neutral radical $LFH^{\bullet}$. In the mode labels listed in the first column of Figure S8 the subscripts "+" and "−" indicate in-phase and out-of-phase relations, respectively, among the various bond stretches contributing to the normal modes. The subscript "RI" points to ring I of isoalloxazine.

Line wise comparisons of the mode pictures immediately demonstrate that, indeed, the mode compositions change upon the electron and proton transfer to the oxidized LF. However, despite these changes the main contributions to the modes by the respective bond stretches are retained. Note here that, in the neutral radical LFH$^{\bullet}$, the modes 5 and 7 of the oxidized LF are represented only by a single mode (line 5). However, due to the proton transfer to the N5 position a new band in the spectrum of the neutral radical semiquinone appears, which is denoted by 7* and is displayed in line 7 of the respective column.

PEDs for oxidized LF and its radicals

Here we present additional data on the vibrational bands of the LF radicals in the spectral region between 1200 cm^{-1} and 1750 cm^{-1}. Table S6 shows our scaled (1.031) DFT/MM-INMA frequencies for fully oxidized LF in TIP3P water together with a PED assignment derived from DFT/MM normal mode analysis of the snapshot, which is determined to be the best representative of the TIP3P snapshot ensemble by using the program Gamess[3] (see Ref 2).

Table S6: Average frequencies $\langle\nu\rangle$ and PEDs for LF

κ^a	ν [b]	PED for LF [c]							
1	1700	75	$\nu_{C4=O}$	3	ν_{C4-C4A}	5	δ_{N3-H}		
2	1659	59	$\nu_{C2=O}$	7	ν_{N1-C2}	15	δ_{N3-H}		
3	1652	27	ν_{C6-C7}	11	$\nu_{C8=C9}$	10	ν_{C9-C9A}	10	ν_{C5A-C6}
4	1580	16	$\nu_{C5A=C9A}$	13	$\nu_{C4A=N5}$	12	$\nu_{C8=C9}$	7	$\nu_{C4A-C0A}$
5	1560	17	$\nu_{C0A=N1}$	16	$\nu_{C4A=N5}$	13	$\nu_{C0A-N10}$	6	ν_{C9-C9A}
6	1524	27	$\nu_{C0A=N1}$	11	ν_{C9-C9A}	10	$\nu_{C9A-N10}$	8	$\nu_{C6=C7}$
7	1513	13	ν_{C7-C8}	11	ν_{C5A-C6}	7	ν_{N5-C5A}	14	δ_{C9-H}
8	1451	10	$\nu_{C9A-N10}$	7	$\nu_{C8=C9}$	6	$\nu_{C4A-C0A}$	25	δ_{C0M-Me}
9	1411	6	$\nu_{C4=O}$	6	ν_{N1-C2}	4	$\nu_{C2=O}$	31	δ_{N3-H}
10	1353	16	$\nu_{C0A-N10}$	14	ν_{C9-C9A}	7	ν_{N1-C2}	5	ν_{C2-N3}
11	1312	19	$\nu_{C9A-N10}$	8	ν_{C5A-C6}	7	$\nu_{C0A-N10}$	6	$\nu_{C4A-C0A}$
12	1288	5	ν_{N5-C5A}	5	$\nu_{C8=C9}$	36	δ_{C9-H}	18	δ_{C6-H}
13	1236	19	ν_{C8-C8M}	9	ν_{C7-C7M}	6	ν_{C7-C8}	22	δ_{C6-H}

[a]Mode identifier κ; [b]Frequencies $\langle\nu\rangle$ in cm^{-1};
[c]PED for the best representative of the snapshot ensemble for LF in TIP3P water in %.

Table S7: Average frequencies $\langle\nu\rangle$ and PEDs for $LF^{\bullet-}$

κ^a	ν [b]	PED for $LF^{\bullet-}$ [c]							
1	1598	32	$\nu_{C4=O}$	11	$\nu_{C2=O}$	8	$\nu_{C6=C7}$		
2	1628	38	$\nu_{C2=O}$	8	$\nu_{C4=O}$	5	ν_{N1-C2}	31	δ_{N3-H}
3	1620	12	$\nu_{C4=O}$	12	$\nu_{C6=C7}$	11	ν_{C9-C9A}		
4	1556	21	ν_{C7-C8}	17	$\nu_{C5A=C9A}$	6	$\nu_{C8=C9}$		
5	1500	3	ν_{C9-N10}	2	ν_{C9-C9A}	2	$\nu_{C0A-N10}$	18	δ_{C0M-Me}
6	1499	14	$\nu_{C0A=N1}$	4	ν_{C4-C4A}	4	ν_{N3-C4}	9	δ_{C7M-Me}
7	1524	8	ν_{C7-C8}	7	ν_{C9-C9A}	19	δ_{C9-H}	13	δ_{C6-H}
8	1450	16	$\nu_{C4A-C0A}$	8	$\nu_{C8=C9}$	7	$\nu_{C9A-N10}$	4	ν_{N1-C2}
9	1373	20	$\nu_{C2=O}$	19	$\nu_{C4=O}$	49	δ_{N3-H}		
10	1337	31	$\nu_{C0A-N10}$	13	$\nu_{C0A=N1}$	8	$\nu_{C4A-C0A}$		
11	1299	12	$\nu_{C9A-N10}$	10	ν_{C5A-C6}	6	$\nu_{N10-C0M}$	14	δ_{C9-H}
12	1271	12	ν_{N5-C5A}	7	$\nu_{C4A=N5}$	7	$\nu_{C8=C9}$	20	δ_{C9-H}
13	1229	14	ν_{C8-C8M}	9	$\nu_{C0A=N1}$	8	ν_{N3-C4}	17	δ_{C6-H}

[a]Mode identifier κ; [b]Frequencies $\langle\nu\rangle$ in cm^{-1};
[c]PED for the best representative of the snapshot ensemble for $LF^{\bullet-}$ in TIP3P water in %.

In the Tables S7 and S8 the corresponding data of the best representatives are listed for $LF^{\bullet-}$ and $LFH^{\bullet}$, respectively.

Table S8: Average frequencies $\langle\nu\rangle$ and PEDs for $LFH^{\bullet}$

mode	ν/cm^{-1}	PED for $LFH^{\bullet}$ / %							
1	1598	29	$\nu_{C4=O}$	28	$\nu_{C2=O}$	11	ν_{C4-C4A}	9	ν_{N1-C2}
2	1633	17	$\nu_{C2=O}$	13	$\nu_{C4=O}$	4	ν_{N1-C2}	29	δ_{N3-H}
3	1642	20	$\nu_{C6=C7}$	13	ν_{C5A-C6}	11	ν_{C9-C9A}	8	$\nu_{C8=C9}$
4	1622	16	$\nu_{C5A=C9A}$	11	ν_{C5A-C6}	10	$\nu_{C4A=N5}$	6	ν_{C7-C8}
5,7	1549	16	$\nu_{C0A=N1}$	9	$\nu_{C4A-C0A}$	9	ν_{C7-C8}	9	ν_{C9-C9A}
6	1527	12	$\nu_{C0A=N1}$	11	ν_{C9-C9A}	7	ν_{C7-C8}	11	δ_{C9-H}
7*	1552	11	ν_{N5-C5A}	9	$\nu_{C4A=N5}$	8	$\nu_{C8=C9}$	20	δ_{N5-H}
8	1463	9	$\nu_{C4A-C0A}$	5	$\nu_{C8=C9}$	4	ν_{N5-C5A}	5	δ_{C9-H}
9	1383	15	$\nu_{C2=O}$	14	$\nu_{C4=O}$	6	$\nu_{C8=C9}$	38	δ_{N3-H}
10	1366	12	$\nu_{C8=C9}$	11	$\nu_{C0A=N1}$	7	$\nu_{C2=O}$	12	δ_{N3-H}
11	1342	17	$\nu_{C0A-N10}$	16	ν_{N5-C5A}	10	$\nu_{C9A-N10}$	5	$\nu_{C4A-C0A}$
12	1280	15	ν_{C8-C8M}	8	ν_{C7-C7M}	7	ν_{N1-C2}	11	δ_{N5-H}
13	1237	30	ν_{N3-C4}	14	ν_{N1-C2}	17	δ_{C6-H}	14	δ_{C9-H}

[a]Mode identifier κ; [b]Frequencies $\langle\nu\rangle$ in cm^{-1};
[c]PED for the best representative of the snapshot ensemble for $LFH^{\bullet}$ in TIP3P water in %.

Gaussian approximation for IR line shapes obtained by INMA

Eq. (1) assumes Gaussian shapes for the inhomogeneously broadened IR bands of flavin radicals in solution. To check this assumption Figure S9 compares three different representations of the INMA results on the IR spectrum of (a) the anionic radical $LF^{\bullet -}$ and (b) the neutral radical $LFH^{\bullet}$ in TIP3P water.

The spectra drawn in green derive from the expression

$$S(\nu) = \frac{c_0}{6}\frac{1}{N_S}\sum_{\alpha}\sum_{j=1}^{N_S}\frac{I_{\alpha,j}}{\sqrt{2\pi\delta^2}}\exp\Big[\frac{-\big(\nu-\nu_{\alpha,j}\big)^2}{2\delta^2}\Big], \tag{S2}$$

in which the double sum runs over all modes α and all N_S snapshots j in the respective ensemble. Thus, the IR spectra are represented as an ensemble average of normalized Gaussian lines centered at the INMA frequencies $\nu_{\alpha,j}$ and weighted with the INMA intensities $I_{\alpha,j}$. As widths of the Gaussians we chose the small value $\delta = 4$ cm^{-1}.

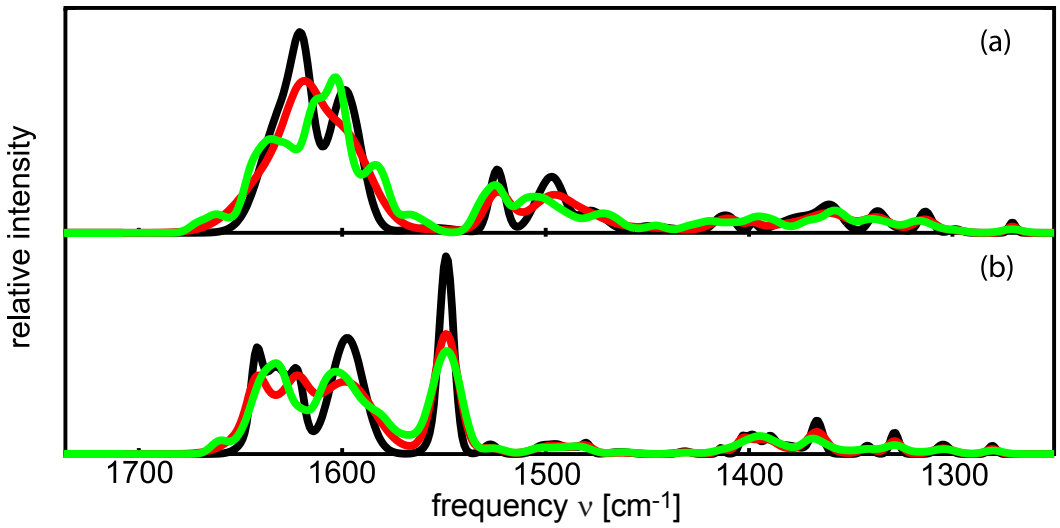


Figure S9: The IR spectra of (a) $LF^{\bullet -}$ and (b) $LFH^{\bullet}$ generated directly from the INMA snapshot ensemble through Eq. (S2) (green lines) are compared with the Gaussian model Eq. (1) (black lines) and with correponding models (red lines), which employ instead of the scaled widths $\sigma_\alpha = 0.6\rho_\alpha$ for each mode α the unscaled standard deviations ρ_α of the frequencies $\nu_{\alpha,j}$ within the respective ensemble of the snapshots $j = 1,\ldots,N_S$.

Figure S9 compares the green kernel estimates Eq. (S2) of the IR spectra with red Gaussian models, which derive from Eq. (1), if one replaces the scaled line widths σ_α by the unscaled standard deviations ρ_α of the mode frequencies $\nu_{\alpha,j}$ within the ensemble of snapshots j. Finally, the black spectra exactly represent Eq. (1), i.e. are constructed using the scaled widths $\sigma_\alpha = 0.6\rho_\alpha$, which serve to approximately correct the neglected effects of motional narrowing.

The red spectra represent a smoothened version of the green spectra, which are characterized by sizable fluctuations. Particularly in the high-frequency region of the two carbonyl modes the superpositions Eq. (S2) are by no means smooth indicating not quite sufficient statistics. On the other hand, in the green and red spectra the line widths are quite similar. Particularly, the two C=O bands are not easily distinguishable. Upon scaling of the band widths, which tries to correct the systematic overestimate of band widths characteristic for the INMA approach, the two C=O bands become discernible.

References

(1) MacKerell, A. D. et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(2) Rieff, B.; Mathias, G.; Bauer, S.; Tavan, P. *Photochem. Photobiol.* **2010**, *(in press)*.

(3) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S. J.; Windus, T. L.; Dupuis, M.; Montgomery, J. A. *J. Comput. Chem.* **1993**, *14*, 1347–1363.

3 DFT/MM Schwingungsspektren von Flavinen in BLUF Domänen

Wie die vorangegangenen Arbeiten gezeigt haben, lassen sich die Schwingungsspektren von Flavinen in wässriger Lösung mit der angewendeten DFT/MM-INMA Methode sehr gut beschreiben und erklären. Somit stehen die erforderlichen Mittel zu Verfügung, um die Schwingungsspektren nun auch in Proteinumgebungen zu berechnen.

Der nachfolgend abgedruckte Artikel [1]

„DFT/MM description of flavin IR spectra in BLUF domains“
Benjamin Rieff, Sebastian Bauer, Gerald Mathias und Paul Tavan
in *J. Phys. Chem. B* **115**, 11239–11253 (2011)

und die dazugehörige Supporting Information

„Supporting Information to the manuscript:
DFT/MM description of flavin IR spectra in BLUF domains“
Benjamin Rieff, Sebastian Bauer, Gerald Mathias und Paul Tavan,

die ich zusammen mit Sebastian Bauer, Gerald Mathias und Paul Tavan veröffentlicht habe, diskutieren neben den DFT/MM generierten Flavin Schwingungspektren in BLUF Domänen auch die methodische Vorgehensweise, mit der im Allgemeinen Schwingungsspektren von Chromophoren in Proteinumgebungen adequat beschrieben werden können. So zeigt sich, dass von den sieben verwendeten BLUF Strukturen, die in der PDB veröffentlicht und in dieser Arbeit untersucht wurden, lediglich drei eine Erklärung der experimentell gemessenen Schwingungsspektren erlauben. Wie ferner gezeigt wird, ist diese Erklärung jedoch nur möglich, wenn zum einen ein modifiziertes MM Proteinkraftfeld, bei dem die elektronische Polarisierbarkeit der Elektronenhüllen geeignet einbezogen wurde, verwendet wird und zum anderen der Einfluss von einzelnen Wassermolekülen, die in Folge der verwendeten MD Simulationen eine Wasserstoffbrückenbindung zum Flavin ausbilden, berücksichtigt wird.

[1] Reproduced with permission, submitted for publication. Unpublished work copyright 2011 American Chemical Society.

THE JOURNAL OF PHYSICAL CHEMISTRY B

ARTICLE

pubs.acs.org/JPCB

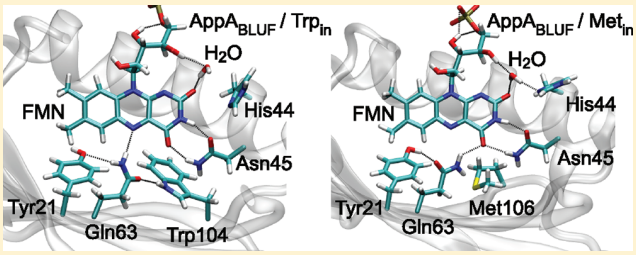
DFT/MM Description of Flavin IR Spectra in BLUF Domains

Benjamin Rieff, Sebastian Bauer, Gerald Mathias, and Paul Tavan*

Lehrstuhl für Biomolekulare Optik, Ludwig-Maximilians-Universität, Oettingenstr. 67, 80538 München, Germany

Supporting Information

ABSTRACT: A class of photoreceptors occurring in various organisms consists of domains that are blue light sensing using flavin (BLUF). The vibrational spectra of the flavin chromophore are spectroscopically well characterized for the dark-adapted resting states and for the light-adapted signaling states of BLUF domains in solution. Here we present a theoretical analysis of such spectra by applying density functional theory (DFT) to the flavin embedded in molecular mechanics (MM) models of its protein and solvent environment. By DFT/MM we calculate flavin spectra for seven different X-ray and NMR structures of BLUF domains occurring in the transcriptional antirepressor AppA and in the blue light receptor B (BlrB) of the purple bacterium *Rb. Sphaeroides* as well as in the phototaxis photoreceptor Slr1694 of the cyanobacterium *Synechocystis*. By considering the dynamical stabilities of associated all-atom simulation models and by comparing calculated with observed vibrational spectra, we show that two of the considered structures (both AppA) are obviously erroneous and that specific features of two further crystal structures (BlrB and Slr1694) cannot represent the states of the respective BLUF domains in solution. Thereby, the conformational transitions elicited by solvation are identified. In this context we demonstrate how hydrogen bonds of varying strengths can tune in BLUF domains the C═O stretching frequencies of the flavin chromophore. Furthermore we show that the DFT/MM spectra of the flavin calculated for two different AppA BLUF conformations, which are called Trp_{in} and Met_{in}, fit very well to the spectroscopic data observed for the dark and light states, respectively, if (i) polarized MM force fields, which are calculated by an iterative DFT/MM procedure, are employed for the flavin binding pockets and (ii) the calculated frequencies are properly scaled. Although the associated analysis indicates that the Trp_{in} conformation belongs to the dark state, no clear light vs dark distinction emerges for the Met_{in} conformation. In this connection, a number of methodological issues relevant for such complex computations are thoroughly discussed showing, in particular, why our current descriptions could not decide the light vs dark question for Met_{in}.



INTRODUCTION

Biological photoreceptors mediate the responses of living organisms to environmental light conditions. Whereas photoreceptors such as the rhodopsins, phytochromes, and xanthopsins employ cofactors featuring light-induced isomerizations,[1] a class of blue-light receptors, which covers the photolyases,[2] the cryptochromes,[3] the so-called LOV[4] (light-oxygen-voltage sensing), and BLUF[5] (blue light sensing using flavin) domains, utilizes the stiff flavin dyes. These flavin chromophores, whose chemical core motif[6] is depicted in Figure 1c, generate the respective biological function through light-induced redox reactions.[1] The reactions can be monitored in a temporally resolved fashion by vibrational spectroscopy.[7] The corresponding spectra can provide information on the flavin redox states[8] and on the conformational changes in the flavin binding pockets, if they can be decoded by a sufficiently accurate computational method. Such a method must be capable of relating the chromoprotein structures, which have to be known at atomic resolution from X-ray crystallography or multidimensional nuclear magnetic resonance (NMR) measurements, to the chromophore spectra determined by infrared (IR) or resonance Raman (RR) spectroscopy.

Computational hybrid methods,[9,10] which combine a density functional theory (DFT)[11,12] description of the respective chromophore with a molecular mechanics (MM) model of its protein environment, are accurate enough to master the task specified above,[13,14] if the applied MM force field is of sufficient quality. Unfortunately, however, the latter condition is not necessarily guaranteed, as long as standard nonpolarizable MM force fields such as CHARMM22,[15] AMBER,[16] or GROMOS[17] are used to describe the condensed phase environment of the chromophore.

In the case of bacteriorhodopsin (BR), for instance, the quite rigid hydrogen bonded network within the binding pocket of the retinal chromophore turned out to be unstable in molecular dynamics (MD) simulations with the quoted standard force fields (see the discussion on pp 10484−6 in ref 18). Correspondingly, DFT/MM calculations of the chromophore IR spectra yielded quite inaccurate descriptions with CHARMM22.[19] However, BR became structurally stable in MD simulations, which employed a new BR specific polarized force field that had been

Received: May 10, 2011
Revised: July 15, 2011
Published: September 02, 2011

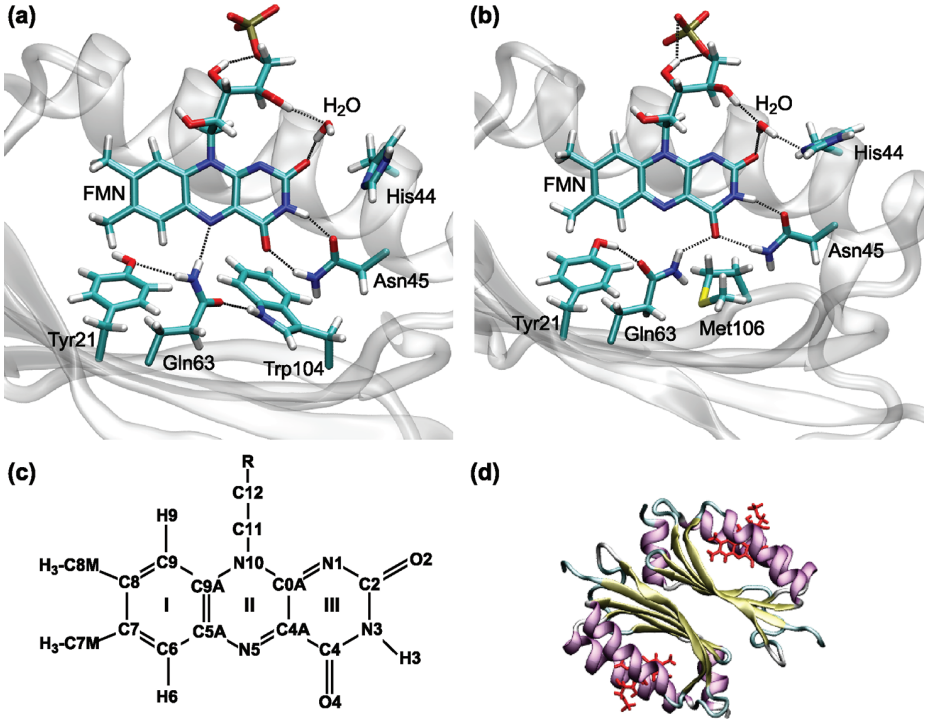


Figure 1. All-atom structures of the active sites in AppA BLUF domains derived by strongly restrained MD simulations (see the Methods section) from (a) the so-called Trp_{in}[21] and (b) the Met_{in}[39] X-ray structural models. The black dotted lines indicate hydrogen bonds between the flavin chromophore and residues of the apoprotein. Panel c depicts the chemical structure and the atom labels of oxidized isoalloxazine, which is the core motif of all flavins,[6] and panel d illustrates the dimeric structure of the AppA BLUF domain.[21,39] The illustrations in a, b, and d were generated with VMD.[40]

generated by iteratively computing[20] new partial charges for the binding pocket with a DFT/MM technique.[18] Furthermore, also the IR spectrum calculated by DFT/MM for the chromophore closely reproduced the observations as soon as the polarized force field was used to model the charge distribution in the chromophore binding pocket.[19]

Taking the BLUF domains as a further example, it is the purpose of this contribution to check whether the quoted failure of standard force fields concerning the DFT/MM description of a chromophore vibrational spectrum was just bad luck (and, thus, restricted to BR) or whether similar failures should be generally expected. Concurrently, we want to pave the way toward high-quality descriptions of the IR spectra of BLUF chromophores. In combination with observed IR spectra such descriptions should be capable to provide (i) answers to the question to what extent given crystal structures represent the solution structures of the IR measurements and (ii) possibly even insights into the details of the conformational changes leading from the initial dark-adapted states to the final light-adapted signaling states of BLUF domains.

Partially supported by X-ray structural models, NMR data, vibrational or UV/vis spectra, and/or quantumchemical calculations various detailed models have been suggested for the dark-light transition in BLUF domains.[21−33] As a common feature these models describe reorganizations of the hydrogen bonded networks, which are supposedly caused by light-induced redox reactions within the flavin binding pockets. The reactions involve the electronically excited flavin as the primary electron acceptor.[22,30] A subsequent proton transfer results in the transient formation of a neutral flavin radical. These charge transfer processes cause a rearrangement of the initial hydrogen bonded network. In particular, the final signaling state features a new or strengthened hydrogen bond to the flavin O4 atom[34−36] as is witnessed by a 17 cm^{-1} redshift of the flavin C4═O4 stretching band.[34−38]

For the AppA (activation of photopigment and Puc expression A) protein from the purple bacterium *Rb. Sphaeroides*, which contains the most intensely studied BLUF domain, two significantly different molecular models were determined by X-ray crystallography and are deposited in the Protein Data Bank (PDB, http://www.rcsb.org). One pertains to the wild-type[21] (PDB-code: 1YRX) and is shown in Figure 1a[40] and the other to the C20S mutant[39] [PDB-code: 2IYG, Figure 1b[40]]. In the wild-type AppA BLUF model,[21] the conserved residue Trp104 is buried in the flavin binding pocket. Therefore, the corresponding BLUF conformation is called "Trp_{in}" [Figure 1a]. In contrast, in the C20S mutant, Trp104 is found at the protein surface,[39] and Met106 takes its position close to the flavin ["Met_{in}", Figure 1b]. Additionally, the orientation and hydrogen bonding pattern of Gln63 is different in the two conformations. Assuming that one of the conformations represents the dark state (e.g., Trp_{in}) and, possibly, the other the light state various intermediate states and mechanistic transformation models have been postulated for the dark-light adaptation process.[21−33] Up to now there is no experimental evidence that clearly favors one of these suggestions.

On the other hand, for various BLUF domains the vibrational spectra of the dark and light states have been characterized in several steady state studies.[34−38] Furthermore, also the kinetics of the dark-light conversion has been carefully investigated by time-resolved UV/vis-pump/IR-probe spectroscopy.[22,30] If one should be able to decode these spectra in terms of structure by reliable DFT/MM calculations one might even be able to assign

larger or smaller degrees of plausibility to the various experimental BLUF structures and theoretical mechanistic suggestions and this is, as mentioned further above, the BLUF specific aim of our endeavor.

In two preceding DFT/MM studies on the flavin IR spectra,[8,41] we have recently presented some of the basic preparatory work, which is required to reach this aim. These studies addressed the IR spectra of oxidized[41] and reduced[8] flavins in aqueous solution, for which a substantial experimental data basis is available in the literature.[42−51] Here, we have predicted the solvatochromic shifts, which should be found in the flavin vibrational spectra upon transfer from the gas phase into a highly polar and isotropic condensed phase environment. For oxidized flavin in water, we have derived a frequency scaling factor applicable to the specific DFT/MM approach used by us which brings the calculated IR frequencies close to the experimental ones.[41]

As expected, this scaling factor turned out to be transferable from the oxidized to the reduced flavins in water, because neither the DFT approach nor the MM force field were changed in this transition. When extending the DFT/MM description to the IR spectra of flavins embedded in protein binding pockets, however, the transferability of the frequency scaling becomes uncertain, because the MM force field for a protein is not necessarily compatible with a given MM model for water. Thus, it remains to be seen to what extent the frequency scaling factor derived from flavins in MM water actually can be extended to flavins in MM models of proteins. Further below we will initially assume transferability, which will be critically addressed later.

The vibrational spectra of flavins solvated in water[45−49] are strikingly similar to those of flavins in dark-adapted BLUF domains.[34−38] For our computational DFT/MM study this finding implies that an oxidized flavin in water, whose DFT/MM spectra and charge distribution were previously calculated by us,[41] should represent a reasonable model for an oxidized flavin mononucleotide (FMN) in dark-adapted BLUF domains.

In the following DFT/MM study on the IR spectra of oxidized flavin in BLUF domains, we will not only consider the two X-ray structures of AppA BLUF quoted above but also two further structural models for this protein[25,39] (PDB-codes: 2IYI and 2BUN) as well as X-ray structures of the BLUF proteins BlrB[52] (Blue light receptor B, PDB-code: 2BYC) from *Rb. Sphaeroides* and Slr1694[24] (PDB-code: 2HFO) from the cyanobacterium *Synechocystis* sp. PCC6803. Before presenting and discussing our computational results, we will first sketch the applied methods.

METHODS

Like in our preparatory work,[8,41] we used for the MM-MD simulations the program package EGO,[53] which treats the electrostatics by the fast structure-adapted multipole method[54,55] combined with a reaction field correction[53] thus implementing toroidal boundary conditions.[56] Furthermore, EGO can speed up the integration by a multiple time step algorithm[57,58]and, here, we applied a basic time step of 1 fs. Berendsen thermo- and barostats[59] were employed to control the temperature T and pressure p, respectively. For MD simulations in the *NVT* ensemble we chose T = 300 K with a coupling time of 0.5 ps, in the *NpT* ensemble we additionally steered p to the target value of 1 bar with a coupling time of 5 ps and an isothermal compressibility of 0.46 GPa^{-1}.

As previously,[8,41] the DFT/MM calculations were carried out with the hybrid method originally suggested in ref 9, which provides an interface between the MM-MD program EGO[53] and the plane-wave DFT code CPMD.[60] Following earlier work[8,13,19,41] we denote the DFT/MM approach applied to FMN in BLUF domains by "MT/BP", because it combines the norm-conserving pseudopotentials of Martin and Troullier[61] with the gradient-corrected exchange functional of Becke[62] and the correlation functional of Perdew.[63] For the MT/BP approach a plane-wave cutoff of 70 Ry has been established as a reasonable choice for computations of flavin vibrational spectra.[41]

Whereas our preparatory DFT/MM calculations on flavins in water were restricted to lumiflavin[8,41] (LF), the BLUF domains contain the somewhat larger FMN molecules as chromophores. To keep the size of the DFT fragment in the DFT/MM treatments of BLUF domains at the reduced size of the LF molecule, we have cut the covalent bond between the atoms C11 and C12 within FMN's ribityl-5′-phosphate side chain [see Figure 1c] using the so-called "scaled position link atom method"[9] (SPLAM). Thus, the DFT fragment was effectively a LF molecule connected through SPLAM to the remainder of the ribityl-5′-phosphate side chain. Note here that a MM force field for the FMN chromophore is given in ref 41. The rectangular box containing the grid for the plane-wave expansion of the Kohn−Sham orbitals was placed around the DFT fragment in such a way that no DFT atom came closer than 3 Å to one of the faces and that the box volume became minimal.

The IR spectra of FMN in BLUF domains were calculated using the protocol[13] for "instantaneous normal mode analysis" (INMA), which has been established as a viable tool for the DFT/MM computation of condensed phase IR spectra in several applications.[8,14,19,41,64−66] Here, a normal-mode analysis is applied to the DFT fragment (LF) embedded in a MM environment (solvated BLUF protein), which is kept fixed at a given structural snapshot. Thus, the FMN can move during the DFT/MM energy minimization in the rigid binding pocket of a BLUF protein and the DFT/MM normal-mode analysis is executed at the optimized geometry. Because a minimum of the potential energy surface is reached in the configurational subspace spanned by the coordinates of the atoms in the DFT fragment, all frequencies are positive like in the usual normal-mode analysis of isolated molecules. For each IR spectrum the normal modes were identified by an automated mode classification procedure.[41]

As is common in theoretical vibrational spectroscopy,[67−69] the calculated DFT/MM frequencies were scaled. Initially we used the scaling factor 1.031. This factor had been derived[41] by comparing vibrational frequencies measured for flavins in aqueous solution with the MT/BP description of LF in a MM water environment, which had been modeled by Jorgensen's "transferable interaction potential using three points" (TIP3P).[70]

Setup of Simulation Models for BLUF Domains. To generate MD simulation models of BLUF domains solvated in TIP3P water, we extracted from the PDB the experimental BLUF structures listed in Table 1. The table characterizes the structures by their PDB codes and by the names of the various proteins enumerated at the end of the Introduction. It distinguishes Trp_{in} from Met_{in} conformations (conf.). Figure 1, panels a and b, provides examples for these conformations. Furthermore, the table gives the assignment of the respective authors, whether their structure is supposed to represent the dark-adapted resting state or the light-adapted signaling state. Finally, the table introduces new names for the various structures, which encode the associated protein name (A/B/S for AppA/BlrB/Slr1694), the conformation (T/M for Trp_{in}/Met_{in}), and the state

Table 1. Experimental BLUF Structures.[a]

PDB code	protein	conf.	state	name
1YRX[21]	AppA	Trp_{in}	dark	ATD
2BUN[25]	AppA	Trp_{in}	dark	ATD*
2IYG[39]	AppA	Met_{in}	dark	AMD
2IYI[39]	AppA	Met_{in}	light	AML
2BYC[52]	BlrB	Met_{in}	dark	BMD
2HFO[24]	Slr1694	Met_{in}	light	SML

[a] 2BUN is from NMR, the others from X-ray; all crystal structures are chain A of the respective sets.

(D/L for dark/light). Note that the new name chosen for the single NMR structure 2BUN[25] is highlighted by an attached star, i.e., this structure is denoted as ATD*.

All PDB entries contain multiple copies of BLUF structures, because the proteins crystallized as multimers. We generally chose chain A. Specifically in the case of entry 2IYI this choice was motivated by a discussion in the original publication of Jung et al.,[39] which claimed that in this case chain A, which structurally differs from chain B, is a more likely model for the light-adapted state. In the case of the NMR data, we have arbitrarily chosen the first structure from the set of 20 geometries deposited under the name 2BUN in the PDB. We consider this structure as a typical representative for the NMR results.

The atom positions for all hydrogens and for other atoms, which are missing in the chosen PDB data sets, were added with the help of the MD program X-PLOR.[71] Here the acidic (Asp, Glu) and basic (Arg, Lys) residues in the various BLUF structures were chosen to be ionized, because they are located close to the water exposed protein surfaces.

In the AppA structures solely the protonation state of His44, which is close to the C2═O2 group of FMN, seemed to be questionable. While the NMR solution structure 2BUN contains a protonated His44, the authors of this structure claim in the associated description[25] that "no positively charged side chains are near C2═O2 in the solution structure of AppA" (p 192), which is an obvious contradiction. Furthermore, in the BlrB structure[52] His44 is replaced by Arg32, which should be positively charged, whereas the Slr1694 structure[24] solely features the polar Asn34 at this position. Due to this uncertainty we generated for AppA structures two alternative all-atom simulation models, one with a positively charged and one with a neutral His44. The models with a protonated His44 are distinguished by appending a "P" to the respective name (cf. Table 1 for the names, to which a "P" is appended).

Next, we embedded the all-atom MD simulation models of the BLUF proteins into periodic water boxes pre-equilibrated in the *NpT* ensemble (T = 300 K, p = 1 bar, see further above) by removing overlapping water molecules. The boxes were shaped as orthorhombic dodecahedra with inner radii of about 40 Å and contained 12 428 rigid TIP3P water molecules before the embedding. Note here that (for both His44 protonation states) the AppA structures ATD, AMD, and AML were modeled as dimers to preserve the hydrophobic interfaces between the β-sheets of the monomeric BLUF units [see Figure 1d]. Only the NMR structure ATD* was modeled as a monomer, because a dimeric structure is not available in this case.

To check whether the choice of monomeric and dimeric simulation models affects the predictions of the FMN vibrational spectra we additionally embedded an AMD monomer into the TIP3P box. Also BMD and SML were constructed as monomers, because here the hydrophobic β-sheet surfaces of the BLUF domains are separated by two additional α-helices from the solvent.[24,52]

If not explicitly stated otherwise we applied the CHARMM22 force field[15] to the BLUF MD simulation systems. In the MM fragments the bond lengths of all hydrogen atoms were kept fixed by the M-SHAKE[72] algorithm.

Restrained Equilibrations of the MD Simulation Models. To preserve the experimental BLUF structures as much as possible during the generation of equilibrated MD simulation models we adopted the following procedure. First the hydrogen and the other modeled protein atoms were adjusted to the experimental structures by short energy minimizations during which the positions of the crystallographic atoms were kept fixed. Coupling the thermostat *exclusively* to the solvent,[73] like in all subsequent MD simulations, and keeping the BLUF domains rigid the solvent was allowed to adjust to the protein volume during 100 ps MD simulations in the *NVT* ensemble. The pressure of the solvent surrounding the rigid BLUF models was subsequently relaxed toward its 1 bar target value in 300 ps *NpT* simulations.

Then all heavy BLUF atoms were softly restrained to the experimental coordinates $\mathbf{r}_c$ by harmonic potentials $k_c(\mathbf{r} - \mathbf{r}_c)^2$. Depending on the location of the respective heavy atom within the protein structure we used two different harmonic force constants. For atoms in the backbone the force constant k_c was set to 6 kcal/(mol Å^2) representing a quite stiff restraint. A value of 2 kcal/(mol Å^2) was chosen for the atoms in the side chains and in FMN, which is a much softer restraint. At the target temperature T = 300 K these potentials restrict the atomic thermal fluctuations to root-mean-square-deviations (RMSDs) of about 0.32 and 0.55 Å, respectively. The softly restrained and solvated protein structures were then simulated for 600 ps in the *NpT* ensemble. Here, measuring the uncontrolled temperature of the BLUF domains served as a check, whether the heat bath (represented by the solvent) properly thermalized the solute protein.[73] Finally the barostats were switched off and further 200 ps were simulated in the *NVT* ensemble for data acquisition (storage every 10 ps). From the resulting 20 snapshots we selected for each BLUF model the snapshot featuring the best match with the respective experimental BLUF structure.

Figure 1, panels a and b, shows for ATD and AMD, respectively, the best matching structures resulting from these restrained MD (rMD) equilibrations. Note that the hydrogen bonds indicated in this figure were defined by the following conditions: the donor−acceptor (D-A) distance is shorter than 3.3 Å and the D-H-A angle is 130° or larger.

Polarized Force Fields for the BLUF Models. As explained in the Introduction the X-ray structure of BR was recently shown to be unstable in extended MD simulations, which applied standard (nonpolarizable) MM force fields.[18] Interestingly such a structural decay was observed once again by us in extended and unrestrained CHARMM22-MD simulations of an AMD dimer in TIP3P water. To speed up the sampling, we had applied for 10 ns a replica-exchange solute tempering approach,[74,75] which showed a large scale motion of Tyr21 within the binding pocket away from its original position [cf. Figure 1b] toward a seemingly much more stable alternative position (data not shown). Note that a similar decay of the hydrogen bonded network stabilizing the flavin binding pocket of BLUF domains was found in

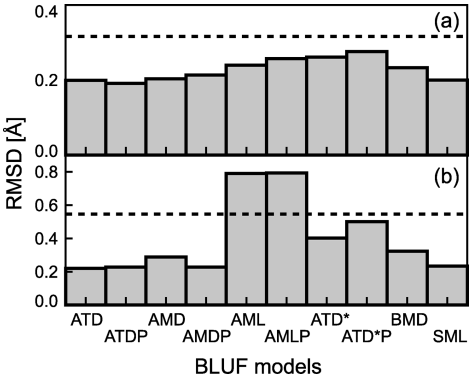


Figure 2. RMSDs between experimental and best matching rMD equilibrated structures for (a) the backbone and (b) side chains in the protein binding pocket for the BLUF models studied here (nomenclature see Table 1). The black dotted lines show the deviations expected at T = 300 K for thermal motion within the soft potentials restraining the positions of the respective heavy atoms.

previous 500 ps CHARMM27-MD simulations of a BlrB model.[32] Here the authors tried to fix the problem by simply adding a water molecule.

Assuming instead that the lacking stability of a hydrogen bonded network in BLUF domains has the same cause as in BR,[18] where it was identfied as the lacking electronic polarizability of standard protein force fields, we decided to copy the same healing strategy that was proven to be successful for BR.[18] Correspondingly we generated polarized force fields (PFFs) for the considered BLUF models by iteratively recalculating "electrostatic potential" (ESP)[76] partial charges of the amino acid side chains surrounding the isoalloxazine in the best matching rMD equilibrated model structures through the DFT/MM approach characterized further above. When the residues Tyr21, Ser41, His44, Asn45, Phe61, Gln63, Asp82, Arg84, His85, and Trp104/Met106 as defined by the AppA numbering for polarization were chosen, the respective DFT fragments consisted of the side chains whose covalent linkages to the backbone were cut at the $C_\alpha - C_\beta$ bonds using SPLAM. We additionally included the special water molecule, which in the AppA BLUF models is hydrogen bonded to FMN-O2 [cf. panels a and b in Figure 1], into the set of polarized molecular groups. The iterative DFT/MM calculation of the ESP partial charges was finished after an approximate self-consistency, as specified by the convergence criterion of ESP partial charge changes smaller than 0.01 e, had been reached for the chosen set. In this paper, the PFFs are mainly utilized in our DFT/MM calculations on the vibrational spectra of BLUF chromophores, where the respective polarized BLUF models are labeled by the subscript "p", i.e., as ATD_p, AMD_p, and so on. We note, however, that unrestrained PFF-MD simulations of the AMD_p dimer showed an enhanced structural stability of the FMN binding pocket (data not shown).

RESULTS AND DISCUSSION

Before calculating the IR spectra of FMN in the various BLUF models by DFT/MM, we checked to what extent the best matching simulation models, which were generated by the careful rMD equilibrations described above, retained the geometries of the experimental parent structures. As measures we calculated RMSDs for the experimentally defined positions of protein atoms. Whenever the electrostatic and van der Waals interactions added by the CHARMM22 force field are incompatible with the experimentally suggested BLUF geometry, the deviations of the rMD equilibrated model from its parent structure will be larger than the expectation values associated with the softly restraining potential at the given temperature T = 300 K (values given above) .

Stability Analysis of the rMD Equilibrated BLUF Structures. We measured the compatibility of the experimental BLUF structures with the all-atom simulation models as described by the CHARMM22 force field using two different RMSDs, one for the heavy atoms in the backbone and the other for the heavy atoms in the side chains of the important residues Tyr21, His44, Asn45, Gln63, and Trp104/Met106 (AppA numbering) determining the structure of the FMN binding pocket [see Figure 1, panels a and b].

Figure 2a shows the RMSDs for the backbones and Figure 2b for the side chains enumerated above. The black dotted lines are the deviations expected for free thermal fluctuations within the respective restraining potentials. According to Figure 2a the backbone geometries of all experimental BLUF models comply with CHARMM22 as is witnessed by the fact that all backbone RMSDs are smaller than the thermally expected deviation. For the side chains Figure 2b demonstrates that most of the best matching rMD equilibrated models remain quite close to their experimental parents. Notable exceptions are the AML and AMLP models,[39] which exhibit RMSDs larger than the thermal limit indicating that they are destabilized by strong forces. Judging from these RMSDs one might additionally argue that the NMR structures[25] ATD*(P) are only marginally stable.

The physical nature of the forces destabilizing the rMD equilibrated models AML and AMLP is analyzed in the Supporting Information (SI). Here we provide in Figure S11 for ATD(P), AMD(P), and AML(P) visual comparisons between the experimental and the rMD equilibrated structures. An accompanying text discusses the structural changes, which are visible in these figures and are caused by the rMD equilibrations. The discussion concludes that the X-ray model AML,[39] as opposed to the other two X-ray structures ATD and AMD, is incompatible with the electrostatics. Therefore we excluded the implausible X-ray structure AML from further consideration.

At this point one may ask whether the structurally different chain B of the PDB entry 2IYI[39] would have been a better choice than the dynamically instable chain A ("AML"). Denoting the all-atom simulation model of chain B as AML_B we found that this model is dynamically as stable as AMD (data not shown). On the other hand, it is also structurally extremely similar to AMD, particularly with respect to the hydrogen bonded network stabilizing the FMN in the AppA-BLUF binding pocket [cf. Figure 1b]. Therefore one is led to assume that IR spectra calculated by the DFT/MM technique for AMD and AML_B will be extremely similar. We checked this issue and found that the C═O stretching frequencies of FMN in AMD and AML_B differed by at most 1 cm^{-1}(data not shown). Therefore, we took AMD as the representative for the type of hydrogen bonded network found at the FMN in the AMD and AML_B models of AppA-BLUF.

Note here that panels a and b of Figure 1 depict the best matching rMD equilibrated models ATD and AMD, respectively. In both cases the rMD representatives are nearly identical to the respective parent X-ray structures (see the discussion of Figure S11 in the SI for further details).

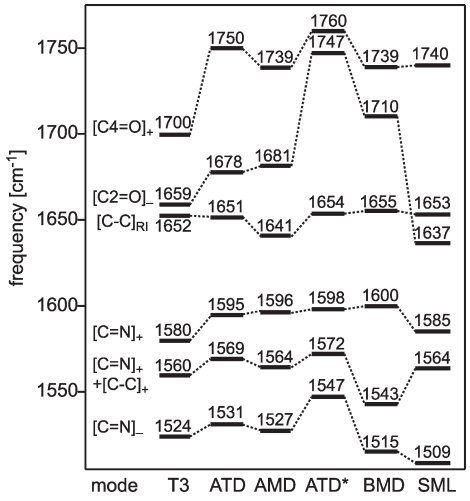


Figure 3. DFT/MM vibrational frequencies of FMN in the rMD equilibrated BLUF proteins ATD, AMD, ATD*, BMD, and SML as described by the CHARMM22 force field.[15] In the AppA models His44 is chosen as unprotonated. As a reference the DFT/MM frequencies of LF in TIP3P water[41] are shown in column T3. All frequencies are scaled by 1.031.[41]

Vibrational Frequencies of FMN in rMD Equilibrated BLUF Models. As mentioned in the Introduction, the spectral positions of the flavin vibrational bands hardly change upon transfer from aqueous solution[45−49] into dark-adapted BLUF domains.[34−38] Computing from the quoted publications average frequencies for all those vibrational flavin bands, which are observed in the 1500−1750 cm^{-1} spectral range, we determined that the two sets of experimental average frequencies (water vs dark-adapted BLUF) deviate by a rmsd of only 1.4 cm^{-1}. Therefore one is lead to expect that such a similarity should show up also in the IR spectra, which are computed by DFT/MM for flavins in TIP3P water and in CHARMM22 models of BLUF domains, respectively. Based on this expectation we initially took our previous DFT/MM results on the IR spectra of fully oxidized LF[41] in TIP3P water as a reference for our new computational results on the IR spectra of FMN in the rMD equilibrated BLUF models.

Figure 3 displays DFT/MM derived and scaled normal-mode frequencies of FMN in the rMD equilibrated BLUF models of AppA (ATD,[21] AMD,[39] ATD*,[25] and unprotonated His44), BlrB (BMD[52]), and Slr1694 (SML[24]). Column T3 depicts the scaled reference spectrum of LF in TIP3P water.[41] The subscripts "+" and "−" at the mode labels [X−Y] denote in-phase and out-of-phase relations, respectively, among the stretches of bonds X−Y contributing to the normal modes. The subscript "RI" points to ring I of isoalloxazine [cf. Figure 1c].

As a most conspicuous feature, all BLUF models show a flavin [C4═O]$_+$ frequency, which is strongly blueshifted with respect to the T3 reference. Somewhat smaller blueshifts are predicted for all [C═N]$_+$ frequencies. In contrast the shifts of the [C2═O]$_-$ frequency exhibit strong differences among the various models. For ATD and AMD the DFT/MM calculations predict blueshifts of only about 20 cm^{-1}, for ATD* and BMD blueshifts of more than 50 cm^{-1}, and for SML even a 22 cm^{-1} redshift. With a few exceptions the remaining stretching modes show minor and varying frequency shifts upon transfer of an oxidized flavin from TIP3P water into the rMD equilibrated BLUF models.

For the depicted frequencies the rmsd from the reference spectrum T3 is smallest at SML (19.8 cm^{-1}), slightly larger at AMD (20.0 cm^{-1}), at ATD (23.2 cm^{-1}), and at BMD (28.6 cm^{-1}). However, at ATD* the rmsd is exceptionally large as it measures as much as 45.4 cm^{-1}. All these RMSDs (see Table S2 in the SI) exceed the expected deviations of at most 10 cm^{-1} by at least a factor of 2, where we have based this accuracy expectation on previous experiences with DFT/MM descriptions of condensed phase IR spectra.[8,14] Despite our care in transforming the experimental structures into all-atom simulation models, certain aspects of these models must be grossly erroneous.

This conclusion is valid even if one doubts whether the quite large[19,64] DFT/MM frequency scaling factor of 1.031 is actually transferable from TIP3P water to CHARMM22 protein models. A CHARMM22 adapted and most likely smaller scaling factor will solely introduce a homogeneous compression of the frequency spectra and, therefore, cannot heal the strikingly different computational results shown in Figure 3 for models of several BLUF domains (AppA, BlrB, and Slr1694), whose chromophores have very similar vibrational spectra.[34−38] Thus disregarding the open question on the transferability of the DFT/MM frequency scaling factor, there remain several possible sources for the apparent misdescriptions. These error sources may vary among the various existing structural models and, therefore, will be discussed by sequentially considering the various structures.

In this context we will present several computer experiments, in which the electrostatics of BLUF residues featuring hydrogen bonds (H-bonds) with the C═O groups of the FMN chromophore are specifically modified. These computer experiments serve to reveal the mechanisms, by which a BLUF protein can steer the frequencies of FMN's important C═O stretching modes.

Why BLUF Structures Do Not Fit to Vibrational Spectra. *NMR Model ATD*.* We start the discussion with the NMR structure ATD*,[25] because this model shows the exceptionally large rmsd of 45.4 cm^{-1} from the T3 reference. According to Figure 3 this rmsd is mainly due to enormous blueshifts of the two C═O stretching frequencies. Figure 3 furthermore demonstrates that the spectrum calculated for ATD* shows substantial deviations (rmsd: 29.3 cm^{-1}) also from the spectrum calculated for the X-ray structure ATD[21] of the same BLUF domain in the same conformation. The large RMSDs, by which the spectrum calculated for the solution NMR structure ATD* deviates from the T3 reference and from the spectrum obtained for the crystal X-ray structure ATD, require an explanation.

An obvious difference between ATD* and ATD is that ATD* is provided in the PDB as a monomer (2BUN) instead of a dimer. Like ATD (1YRX) also the X-ray structure AMD[39] (2IYG) of the same BLUF domain in the other (Met$_{in}$) conformation is a dimer. To check whether the monomeric modeling causes the much larger rmsd of ATD*, we additionally calculated the IR frequencies for FMN in a solvated AMD monomer. However, the spectra of the AMD monomer and dimer turned out to differ by a rmsd of only 7.2 cm^{-1} (data not shown). Furthermore the spectra of the AMD and ATD dimers differ only by an rmsd of 6.7 cm^{-1}. Therefore, the monomeric character of ATD* cannot be the cause for its much too large RMSDs of 29.3 cm^{-1} from the ATD and of 45.4 cm^{-1} from the T3 references.

Another reason for the huge 45.4 cm^{-1} rmsd observed for ATD* could be the assumption that His44 is deprotonated (cf. our remark in Methods on an apparent contradiction in ref

25 concerning this issue). However, the protonation of His44 does not reduce the rmsd of ATD* very much. It is still 40.6 cm^{-1} (cf. Table S2 in the SI) and, thus, by far too large. In the case of the other AppA BLUF models ATD and AMD the protonation even increases the associated RMSDs a little (Table S2), which is the reason why we omitted the DFT/MM spectra calculated for ATDP and AMDP in Figure 3. However, for the sake of completeness the shifts introduced by the protonation of His44 in the DFT/MM vibrational spectra of the various AppA BLUF models are illustrated by Figure S12 in the SI.

The facts that for ATD* the rmsd from the T3 reference is about two times larger than for ATD and that, consequently, ATD* deviates by a large rmsd from ATD, must be caused by substantial structural differences between the underlying experimental models, although they should map one and the same protein conformation. Ideally, ATD* should represent the structure of AppA BLUF in solution and, hence, should belong to conditions silimar to those used in vibrational spectroscopy, whereas ATD might be modified by crystal artifacts.

Figure S13 in the SI demonstrates that ATD and ATD* actually assign drastically different structures to the FMN binding pocket of AppA BLUF (Trp_{in} conformation). The associated quite detailed discussion explains the differences between the spectra calculated for ATD and ATD* as consequences of different protein-chromophore interactions, which are associated with different local structures. In particular, the discussion argues that the missing H-bond FMN−O4···H−N−Asn45 in the NMR structure ATD* must lead to strongly blueshifted C═O stretching frequencies as seen in Figure 3. FMN−O4 is strongly H-bonded in water[41] and in other AppA BLUF structures (cf. Figure 1, panels a and b, and Figure S11 in the SI), which therefore feature red-shifted C═O stretches. The discussion in the SI concludes that the NMR structure ATD* with its hardly H-bonded FMN-O4 is incompatible with the observed vibrational spectra.

The much smaller rmsd calculated by us for ATD as compared to ATD* now argues that ATD as depicted in Figure 1a is closer to the AppA BLUF structure present and observed in vibrational spectroscopy than ATD*. Because we want to analyze these spectra, we have excluded the obviously error-prone NMR structure ATD* from further considerations.

X-ray Model SML. According to Figure 3 the spectrum calculated for the X-ray model SML[24] exhibits an unusually red-shifted $[C2{=}O]_-$ stretching frequency. Not even in TIP3P water, where FMN-O2 is quite strongly H-bonded to the surrounding solvent,[41] the H-bonding interactions manage to shift the $[C2{=}O]_-$ stretch below the $[C{-}C]_{RI}$ stretching vibration of FMN's ring I [cf. Figure 1(c)]. In SML there must be a very special structural reason for the unusually low $[C2{=}O]_-$ stretching frequency.

If one now looks at the structure of SML depicted in Figure 4 and compares this structure with those of the other BLUF domains in Figures 1a,b and 5, one actually detects a unique feature. Only in SML a charged residue (Arg68) forms an additional H-bond with the atom FMN-O2. In the other BLUF domains the charged Arg68 of Slr1694 is replaced by the polar His78 (AppA) or by the unpolar Leu66 (BlrB), which are located at the protein surfaces. Looking at the crystal structure of SML in a representation highlighting the protein surface (picture not shown) one recognizes that the side chain of Arg68 is buried with its charged tip in the protein interior where it forms the H-bond with FMN-O2 shown in Figure 4, top. This special structure

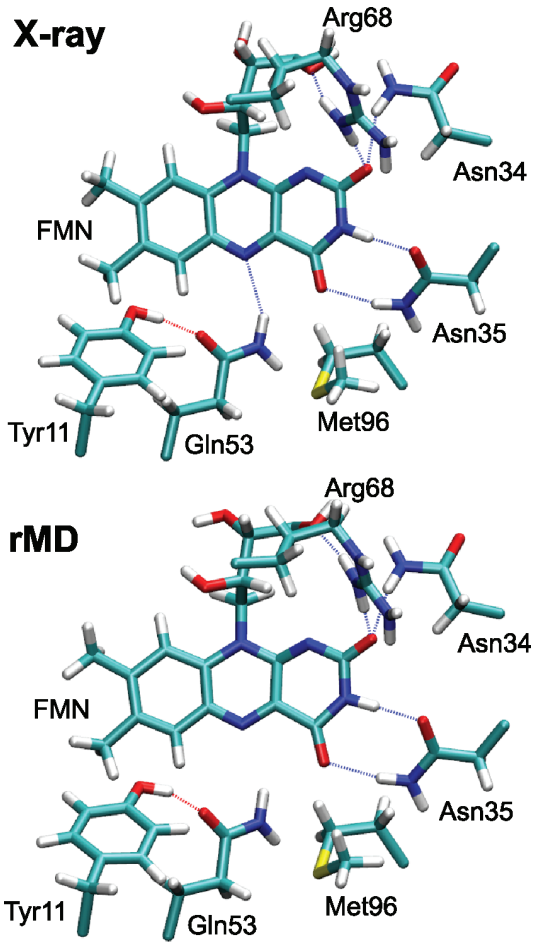


Figure 4. Comparison of the Slr1694[24] X-ray structure (top) with the best matching representative SML (bottom) in the rMD equilibrated structural ensemble. Quite obviously the X-ray structure is closely preserved by the rMD representative apart from the formation of somewhat tighter H-bonds. Furthermore the H-bonding pattern connecting FMN with Asn35, Gln53, and Tyr11 closely resembles the corresponding pattern in AMD [cf. Figure 1b].

immediately explains the 20 cm^{-1} redshift of the $[C2{=}O]_-$ stretching frequency calculated for SML as compared to T3.

To check this simple explanation we carried out a computer experiment. Here, we repeated the DFT/MM calculation of FMN's vibrational spectrum for a model SML* differing from SML solely in the partial charges of the Arg68 side chain. All of these partial charges were reduced by a factor 1/4 thus transforming Arg68 into an essentially unpolar residue. With the unpolar "Arg68" the $[C2{=}O]_-$ stretching frequency experienced a huge 79 cm^{-1} blueshift toward 1716 cm^{-1}, whereas all other FMN stretching frequencies showed much smaller shifts of about 10 cm^{-1}. As a result, the upshifted $[C2{=}O]_-$ frequency was close to the value of 1710 cm^{-1} calculated for BMD, in which FMN-O2 is lacking a clear H-bond (see below), and far above ($\sim$36 cm^{-1}) the frequencies found for ATD and AMD, where FMN-O2 forms a H-bond with a water molecule that has entered the protein during the rMD equilibrations [cf. Figure 1, panels a and b].

Spectroscopically, however, the $[C2{=}O]_-$ band of flavin is found at the same frequency of 1670 cm^{-1} both for solubilized

dark-adapted Slr1694[34,37] and for aqueous solution[45−49] demonstrating that the H-bond between FMN-O2 and Arg68 must be absent in the solubilized protein Slr1694 studied by vibrational spectroscopy. Instead one must expect that the ionic side chain of Arg68 becomes solvent exposed in the solvated protein. Here, FMN-O2 may either form a slightly stronger H-bond with Asn34 (cf. Figure 4) or with water molecules entering the protein and replacing Arg68, which both could explain the observed spectral position of the [C2═O]_ band.

Our calculations definitely exclude, however, the extremely strong H-bond of FMN-O2 with Arg68 for solubilized Slr1694, because that would argue for a [C2═O]_ frequency near 1630 cm^{-1}. For this estimate we took the 40 cm^{-1} shift, by which the extremely strong H-bond of FMN-O2 with Arg68 in SML shifts the [C2═O]_ frequency to the red of the 1680 cm^{-1} calculated for the moderatly strong H-bonds of FMN-O2 in ATD and AMD (cf. Figure 3). This estimate is valid because the structures of SML and AMD are very similar in many respects with the key exception of the H-bonding at FMN-O2.

Thus, we conclude that the H-bond between FMN-O2 and the charged tip of Arg68 is an artifact of the SML crystal, which is caused by a reduced polarity of the Arg68 environment within the crystal lattice and will disappear upon solvation of the protein in aqueous solution. Therefore, the X-ray structure of SML is inappropriate for comparisons of calculations with observed solution spectra, which is the reason why we have excluded it from further considerations.

We would like to note that the soft restraints on the heavy atoms applied in our rMD simulations were strong enough to preserve the electrostatically unfavorable buried position of Arg68 present in the low-dielectric crystalline environment. We expect, however, that extended and unrestrained MD simulations with a sufficiently accurate MM force field will produce solution structures of SML featuring a solvent exposed Arg68, which then could qualify for DFT/MM calculations of the vibrational spectra.

X-ray Model BMD. Similar to the case of SML just discussed also the spectrum calculated for the X-ray model BMD[52] shows an unexpected feature. But instead of an unusual redshift of the [C2═O]_ stretch frequency one here notes an unusual blueshift of that frequency. Figure 3 shows that this frequency is calculated 51 cm^{-1} above the T3 reference and about 30 cm^{-1} above the [C2═O]_ frequency obtained for the AppA BLUF models ATD and AMD. Following the arguments given above concerning the influence of H-bonds at FMN-O2 on the [C2═O]_ frequency, we are led to conclude that sufficiently strong H-bonds resembling those found in liquid water must be absent in the X-ray structure of BlrB.[52]

In fact, if one looks at the X-ray structure and at the structure of its best matching rMD representative BMD depicted in Figure 5, one sees that FMN-O2 has no H-bonding partner. In BMD the His44 side chain in AppA BLUF is replaced by the longer side chain of Arg32, which in the crystal forms H-bonds with the phosphate group of FMN's ribityl-5′-phosphate side chain. Our rMD equilibration preserves this structural arrangement, which prevents the approach of water molecules from the surrounding solvent toward FMN-O2 and the formation of H-bonds like in the rMD models of ATD and AMD [cf. Figure 1, panels a and b]. In an extended unrestrained simulation with a sufficiently accurate MM force field, however, as well as in a solution structure one should expect that the zwitterionic structure, which is present in the crystal and consists of the phosphate and Arg32,

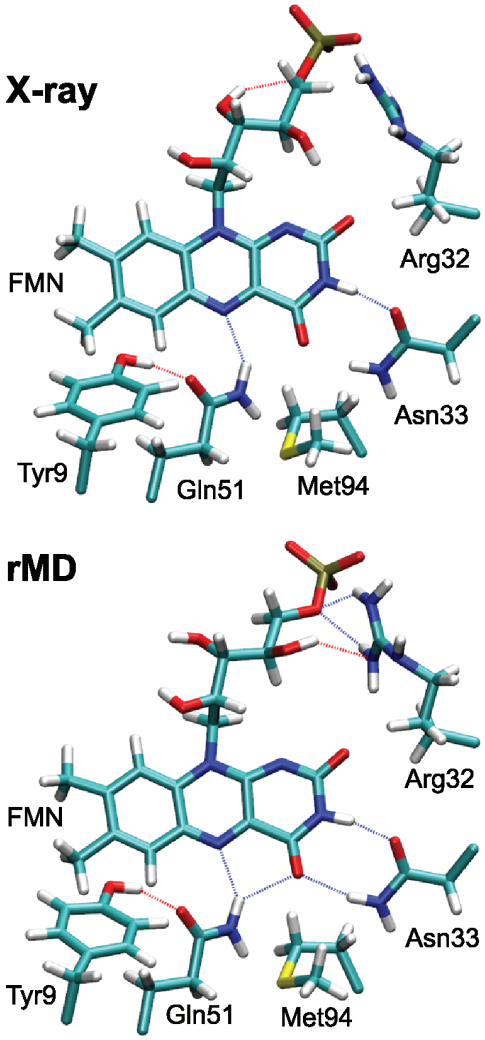


Figure 5. Comparison of the BlrB[52] X-ray structure (top) with the best matching representative BMD (bottom) in the rMD equilibrated structural ensemble. The two structures are highly similar and resemble those of AMD in Figure 1b and of SML in Figure 4, if one disregards the obvious replacements of Arg32 by His44 (AMD) and Asn34/Arg68 (SML) near FMN-O2 as well as the additional water molecule generated by the rMD equilibration in AMD.

is unstable, because both ions should prefer to become solvated in the adjacent aqueous phase. Such a solvated structure will then show H-bonds between water molecules of the solvent and FMN-O2, which can easily reduce the [C2═O]_ frequency by about 30 cm^{-1}.

Because of the lacking H-bonds at FMN-O2 and of the associated artificial blueshift of the [C2═O]_ frequency also the X-ray structure of BlrB[52] and its rMD model BMD cannot serve as computational tools to describe the vibrational spectra[34−38] of FMN embedded in solvated BLUF domains. The spectral positions of the C═O stretching bands in these spectra are close to those in water and, therefore, clearly signify strong H-bonds at FMN-O2 and FMN-O4.

Although the X-ray structure of BlrB may possibly reflect the reality of the crystal, it nevertheless may not represent the solution structure of this BLUF domain, because a lacking H-bond at FMN-O2 should leave clear footprints also in the

optical spectra of the chromophore, which have not been observed, however.[52] Whether a mobile water molecule, which could be present in the BlrB crystal, has been overlooked during refinement could be decided, if in addition to X-ray data also chromophore vibrational spectra would have been recorded for the crystals. Similarly, whether FMN-O2 is H-bonded, if BlrB is solvated in water, could be decided by corresponding IR spectra. Because of the lacking data on the chromophore vibrational spectra of BlrB both in the crystal and in solution we have to exclude the model BMD from further analysis.

Intermediate Summary and Discussion. The above analysis employed rMD simulations of various structural models for BLUF domains and first DFT/MM computations of the FMN vibrational spectra with the aim to check, whether these experimentally derived models qualify for explanations of FMN's vibrational spectra that were recorded for solubilized BLUF domains.

Luckily we found the two structural models ATD and AMD, which passed our initial test. We note in passing that these structures would not have passed this test either, if the rMD equilibration had not placed a H-bonded water molecule close to the atom FMN-O2 [cf. Figure 1, panels a and b], which is missing in the crystal structures. The reason for the expected failure is that a H-bonding strength similar to that in aqueous solution is necessary to move the $[C2{=}O]_-$ frequency from non-H-bonded values[41] above 1740 cm^{-1} toward a value (1670 cm^{-1}) similar to the one found in water.

Other X-ray models either lacked such a H-bond at FMN-O2 (BMD), because the X-ray structure does not contain a water molecule at the required location and because the rMD simulation did not manage to heal this likely error, or showed a much too strong H-bond at FMN-O2 (SML), because in the crystal an arginine side chain preferred to become buried in the protein interior instead of becoming solvated in the surrounding solvent. For these crystals a knowledge of FMN's vibrational spectra could lead to immediate checks of the correctness of the X-ray structures. In both cases the solution structures will most likely differ from the respective crystal structures, because charged residues at the protein surfaces will become solvated. Then, both solution structures will closely resemble the structure AMD shown in Figure 1b.

Despite the thus established structural similarity of the solubilized BLUF proteins AMD, BMD, and SML, two of these Met_{in} structures (AMD and BMD) were interestingly associated with the dark-adapted resting state,[39,52] whereas one (SML) was assigned to the light-adapted signaling state.[24] The latter assignment was based on the quenching of the Trp fluorescence after light excitation of Slr1694 BLUF and the empirical rule that polar solvents tend to quench this fluorescence.

For the NMR structure ATD* a spectral comparison with the X-ray structure ATD clearly revealed gross inaccuracies of atomic coordinates, which excluded this model from further analysis. Finally, the X-ray structure AML (chain A in the PDB entry 2IYI[39]) did not even pass the test of dynamical stability. Upon the rMD equilibrations it decayed toward a structure resembling that of AMD. On the other hand, the model AML_B derived from chain B of 2IYI was dynamically stable but featured a H-bonded network so similar to that of AMD that the C=O stretching frequencies calculated for AMD and AML_B deviated by at most 1 cm^{-1}, which is why the choice of structure AMD seemed to be sufficient for the further characterization of the Met_{in} conformation of AppA BLUF.

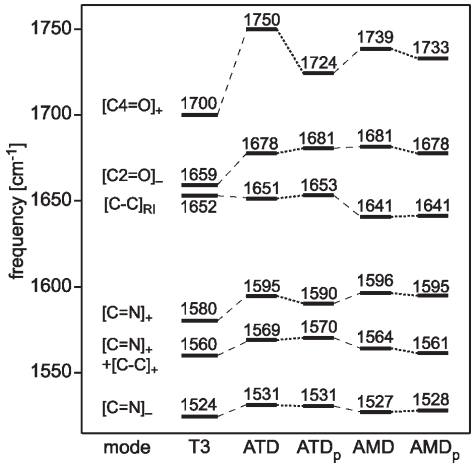


Figure 6. Polarization effects on the scaled DFT/MM vibrational frequencies calculated for FMN in the rMD equilibrated BLUF models ATD and AMD. The effects are due to replacing in the FMN binding pockets the standard partial charges of CHARMM22[15] with the self-consistent ESP charges of the respective structure adapted PFFs (subscript "p") derived from iterative DFT/MM calculations (see the text and the SI for details). Also given is the reference spectrum T3.[41]

X-ray Models ATD and AMD. As a result, we are left with the models ATD and AMD drawn in Figure 1, panels a and b, and with their descendants ATDP and AMDP distinguished by a protonated His44. As we pointed out above in our discussion of Figure 3 and Table S2 in the SI, also for these models the DFT/MM spectra of the FMN chromophore calculated at the respective best matching rMD equilibrated structures show RMSDs from the T3 reference spectrum exceeding the expected value[14] of 10 cm^{-1} by a factor of 2.

Following the arguments of Babitzki et al.,[19] a possible reason for the unexpectedly large RMSDs obtained for these AppA BLUF models could be the use of the standard MM force field CHARMM22 for the description of the BLUF proteins surrounding the DFT fragment, i.e., the flavin chromophore. The neglect of the electronic polarizability is inherent to such a standard MM force field and could lead to an ill-structured electrostatic potential within the FMN binding pocket, which could be witnessed, e.g., by an insufficiently stable H-bonded network. To check this issue, we calculated for the binding pockets of the four rMD equilibrated AppA BLUF models new ESP[76] partial charges by the iterative DFT/MM procedure outlined in Methods. The resulting structure adapted partial charges are listed in Tables S3−S13 of the SI, and the corresponding PFFs were used to recalculate the IR spectra of FMN in the AppA BLUF models ATD and AMD.

Polarization Effects on the IR Spectra of FMN in AppA BLUF. Figure 6 illustrates the spectral changes which are induced, if one replaces the CHARMM22 force field by the respective structure adapted PFFs in the rMD equilibrated simulation models ATD and AMD. For both models the $[C4{=}O]_+$ frequency is seen to be red-shifted upon exchanging in the rMD equilibrated models the standard CHARMM22 partial charges by those of the structure adapted PFFs. The frequencies of the other FMN stretching modes show only minor changes.

As a result the rmsd from the T3 reference spectrum decreases by 23% for ATD_p and by 19% for AMD_p. Thus, accounting for

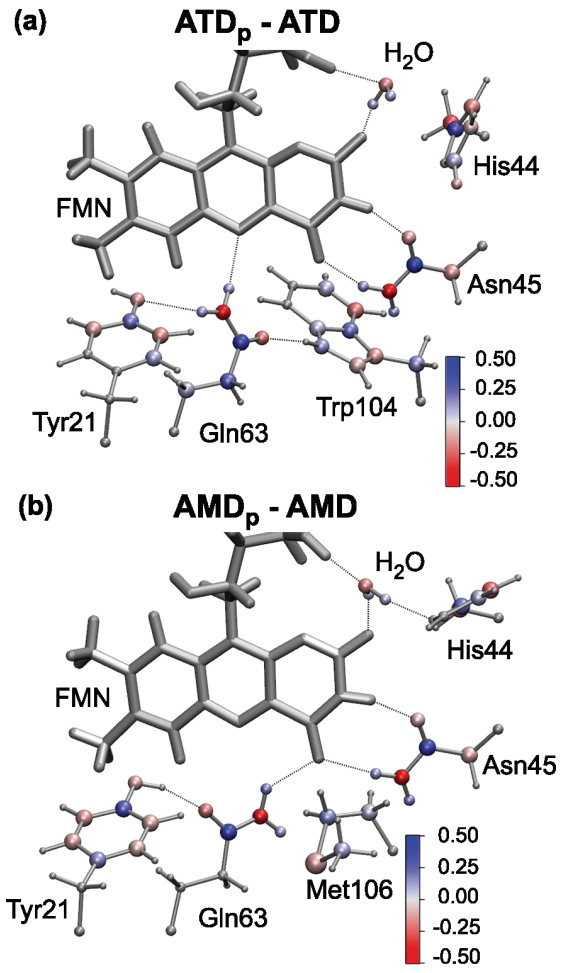


Figure 7. DFT/MM derived differences $\Delta q_i = q_i(p) - q_i(c)$ between the atomic partial charges q_i characterizing the PFF models $p \in \{ATD_p, AMD_p\}$ and the CHARMM22 models $c \in \{ATD, AMD\}$. Negative Δq_i are colored in red and positive Δq_i in blue.

the electronic polarization somewhat reduces the unexpectedly large RMSDs calculated with CHARMM22 (cf. Tables S2 and S14 in the SI). To identify the reasons, why the DFT/MM spectra of the models ATD_p and AMD_p are closer to the T3 reference, we consider the changes of the charge distributions induced by the polarization of the FMN binding pocket in AppA BLUF.

Polarization of the FMN Binding Pocket in AppA BLUF. Figure 7 partially illustrates the polarization effects Δq_i on the partial charges q_i in the FMN binding pockets of (a) ATD and (b) AMD (effects omitted in the drawings are documented by Tables S3−S13 in the SI). For both models one immediately recognizes that the electronic polarization makes the C═O and NH_2 groups of Asn45 and Gln63 much more polar than assumed by CHARMM22. In particular, the H-bonds FMN−O4···H−N−Asn45 and FMN−N3−H···O−Asn45 become stronger by 44% and 18%, respectively, upon inclusion of the electronic polarizability. Interestingly also the strongly bound water molecule, which forms a H-bond with FMN-O2, becomes more polar. Its dipole moment is in both PFFs by about 20% larger than assumed by TIP3P (cf. ref 18).

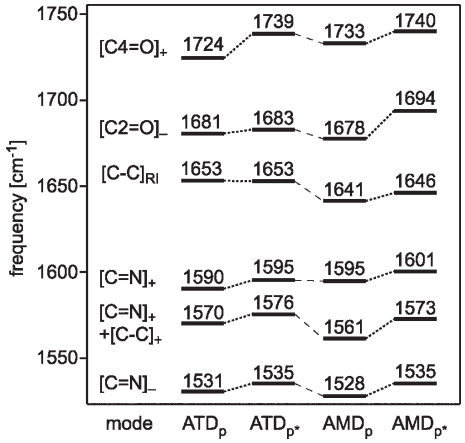


Figure 8. Vibrational frequencies calculated by DFT/MM for FMN embedded in the BLUF models ATD_p and AMD_p are compared with those of ATD_{p^*} and AMD_{p^*}, respectively, whose PFF were modified by adopting for Asn45 the less polar CHARMM22 model instead of the highly polar Asn45 model calculated by DFT/MM.

Spectral Role of the Asn45 Polarity in AppA BLUF. The strongly increased polarity of Asn45 in the two polarized AppA BLUF models should be the cause for the polarization-induced redshifts of the $[C4{=}O]_+$ frequency shown in Figure 6. To check this issue we carried out another computer experiment. Here, we repeated the calculation of the vibrational spectra for modified PFF models ATD_{p^*} and AMD_{p^*} of AppA BLUF, in which solely the highly polar residue Asn45 resulting from the DFT/MM calculations was replaced by its much less polar CHARMM22 model (cf. Table S6 in the SI).

Inspection of Figure 8 immediately shows that the decreased polarity of Asn45 entails a blueshift of essentially all FMN stretching frequencies. In particular, this polarity reduction shifts the average C═O stretching frequency, which we define as $\bar{\nu}_{C=O} \equiv (\nu_{C4=O} + \nu_{C2=O})/2$, by 9 and 12 cm^{-1} to the blue as compared to ATD_p and AMD_p, respectively. Because the polarities of many residues in the FMN binding pocket were changed during the iterative DFT/MM calculations of the PFFs (cf. Figure 7 and Tables S3−S13 in the SI), the partial restoration of the CHARMM22 force field at Asn45 does not lead back to the CHARMM22/DFT spectra ATD and AMD of Figure 6. In the case of the $[C4{=}O]_+$ frequency of ATD, for instance, the PFF polarity of Asn45 explains only 58% of the 26 cm^{-1} redshift calculated for the transition toward the PFF model ATD_p. As a rule of thumb we may nevertheless state that the strength of the H-bond FMN−O4···H−N−Asn45 as given by the polarity of Asn45 in our DFT/MM setting steers the spectral position of the average C═O stretching frequency $\bar{\nu}_{C=O}$ in general and of the $[C4{=}O]_+$ frequency in particular. Thus, the common explanation that the 17 cm^{-1} redshift of the $[C4{=}O]_+$ frequency observed in the dark-light transition of BLUF domains[34−38] signifies an increased H-bonding strength at FMN-O4, is clearly supported by our second computer experiment.

We would like to stress at this point that the three structures ATD, ATD_p, and ATD_{p^*} obtained by DFT/MM minimization from the associated rMD equilibrated model are nearly identical as we have checked by graphical inspection. A likewise striking structural similarity implying movements of FMN atoms of the order of only 0.05 Å holds for the DFT/MM optimized structures AMD, AMD_p, and AMD_{p^*}. Therefore, the spectral

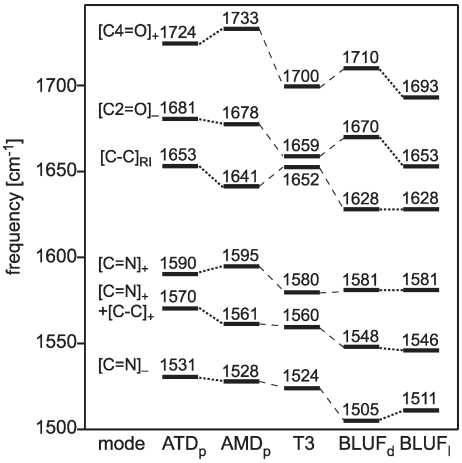


Figure 9. Frequencies calculated by DFT/MM for oxidized flavin in the ATD_p and AMD_p models of AppA BLUF and in the TIP3P model of liquid water (T3), which were scaled by 1.031, are compared with experimental consensus frequencies[34−38] for flavin in the dark ($BLUF_d$) and light ($BLUF_l$) states of BLUF domains.

shifts seen in Figures 6 and 8 among the various spectra for ATD and AMD, respectively, are almost exclusively caused by the differences among the partial charges used to model the electrostatics of the protein environments.

Together with the results of our calculations on SML and BMD, which showed that the [C2═O]_ frequency can be selectively tuned by H-bonding interactions at FMN-O2, the picture emerges that the frequencies of FMN's important C═O stretches are mainly steered by the strengths of local H-bonds. These simple rules agree, of course, with the results of our previous calculations on LF in water[41] and with expectations common[34−36] among spectroscopists.

Transferability of Frequency Scaling. As we have seen, the use of PFFs for the AppA BLUF models ATD and AMD has shifted the C═O stretching frequencies toward the desired values. Unfortunately, the H-bonding interaction of FMN with the highly polar Asn45 provided by the PFFs did not suffice to bring the average C═O frequency $\bar{\nu}_{C=O}$ of 1698 (ATD_p) and of 1706 cm^{-1} (AMD_p) down to the value of 1680 cm^{-1} predicted by the reference calculation T3 or to the somewhat larger value of 1690 cm^{-1} observed[34−38] for dark-adapted BLUF domains.

To illustrate this issue we now compare the scaled (1.031) DFT/MM frequencies with spectroscopic data on BLUF domains. We extracted these data as averages from available IR and RR spectra on flavins in the dark- and light-adapted states of BLUF domains in Slr1694[34,37] and AppA.[35,36,38] Here, we cannot exclusively rely on spectra for AppA, because the [C2═O]_ frequency has not been determined for this protein. On the other hand, the spectra for AppA and Slr1694 are very similar for the remaining bands.

Figure 9 demonstrates that the vibrational frequencies calculated for FMN in the rMD equilibrated AppA BLUF models ATD_p and AMD_p are strongly and about homogeneously blueshifted compared to the experimental data represented in the columns $BLUF_d$ and $BLUF_l$. In contrast, the LF spectrum T3, which is calculated with the TIP3P model of an aqueous solution, covers about the same spectral range as the experimental $BLUF_d$ reference spectrum, which closely resembles the vibrational spectrum observed for oxidized LF in aqueous solution (data not shown).[41]

Note, however, that the two C═O frequencies in column T3 of Figure 9 are red-shifted compared to the experimental $BLUF_d$ reference, whereas many of the other frequencies show sizable blueshifts. This indicates that the TIP3P scaling factor of 1.031 does not quite succeed to shift the C═O frequencies, which are extremely sensitive to the strengths of local H-bonds,[41] to sufficiently large values, while it shifts most other frequencies, which are less sensitive to H-bonding, too far to the blue. The stated imbalance of scaling-induced frequency shifts shows that a TIP3P environment considerably overestimates the H-bonding strength of real water. With a smaller scaling factor many T3 frequencies would match the observations much better, but the H-bonding sensitive C═O frequencies would then represent gross underestimates. We note that the stated overstimate of the H-bonding strength by the TIP3P water model leads to a dielectric constant,[77−79] which is by about 25% larger than that of real water.

The nearly homogeneous overestimates of the experimental data in columns $BLUF_d$ and $BLUF_l$ of Figure 9 by the computational results in columns ATD_p and AMD_p demonstrate that the application of the large TIP3P scaling factor is inappropriate and that a smaller factor must be chosen. Thus, our initial assumption that the TIP3P adjusted scaling factor of 1.031 should be transferable to CHARMM22 or PFF models for protein environments was incorrect. This conclusion is supported by earlier DFT/MM descriptions of chromophore vibrational spectra,[19,64] which also combined the MT/BP approach with CHARMM22 or PFF models of the respective protein environments but used much smaller scaling factors of 1.0122 and 1.005, respectively. We note that the assumption of a protonated His44 in AppA BLUF does not change this conclusion. On the contrary, as documented in the SI by Figures S14 and S15 and as explained in the accompanying text, the protonated models describe the experimental spectra generally even worse than the unprotonated models (cf. Table S14).

DFT/MM Results for FMN in ATD and AMD. To compute a scaling factor, which is more appropriate for the DFT/MM description of FMN in PFF models of BLUF domains, we have matched the spectra ATD_p and AMD_p to the spectra $BLUF_d$ and $BLUF_l$ by minimizing the total rmsd for all pairwise comparisons. This procedure was chosen, because we did not want introduce a prejudice for the assignment of ATD_p and AMD_p to the dark or light states of BLUF domains. We found the strongly reduced scaling factor of 1.0187. It entails RMSDs of 7.2 and 6.6 cm^{-1} matching ATD_p and AMD_p, respectively, to $BLUF_d$. For the light state spectrum $BLUF_l$ the corresponding deviations are 7.6 (ATD_p) and 8.7 cm^{-1} (AMD_p) (cf. Table S15 in the SI). Because of the slightly better matches with the dark state spectrum $BLUF_d$ it may seem that both AppA conformations, i.e., Trp_{in} (ATD_p) and Met_{in} (AMD_p), represent the dark state.

Figure 10 illustrates the quality, at which the properly rescaled vibrational frequencies, which were calculated by DFT/MM for FMN in the rMD equilibrated AppA BLUF models ATD_p and AMD_p, explain the spectroscopic data[34−38] on the dark- and light-adapted states of BLUF domains, respectively. As already indicated by the RMSDs, the frequencies calculated for ATD_p and AMD_p are seen to fit better to the experimental dark state frequencies $BLUF_d$. Most of the dark state frequencies are close or nearly identical to those measured for the light state ($BLUF_l$) with the important exceptions of the two C═O frequencies, which are both red-shifted by 17 cm^{-1}. According to the

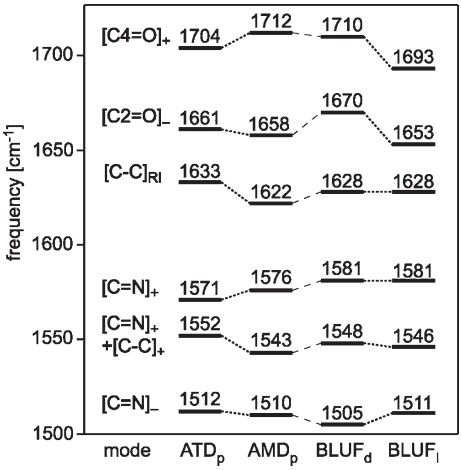


Figure 10. Vibrational frequencies, which were calculated by DFT/MM for FMN in the BLUF models ATD_p and AMD_p, are compared after application of the new scaling factor 1.0187 with spectroscopic data[34−38] on the dark ($BLUF_d$) and light ($BLUF_l$) states of BLUF domains.

computational results presented further above, these redshifts indicate that the H-bonding interactions at FMN-O4 and FMN-O2 must be stronger in the light than in the dark state. In contrast, the calculated spectra depicted in Figure 10 for ATD_p and AMD_p do not reveal clearly stronger H-bonds at FMN's carbonyl groups in one of these structures suggesting again that both models belong to the dark state.

Can We Assign Conformations to Functional States? At this point readers interested in a clear-cut assignment of the experimentally known BLUF structures to one of the two functional states of BLUF domains, i.e., the dark-adapted resting and the light-adapted signaling state, respectively, may ask to what extent the above conclusion that both conformations Trp_{in} (ATD_p) and Met_{in} (AMD_p or AML_{Bp}) belong to the dark state is actually solid. Here, we have to admit that we are quite uncertain concerning a certain part of this question, i.e., the one concerning the Met_{in} conformation, and we will now give arguments why this uncertainty cannot be avoided because of insufficient statistics.

First, our above analysis has been based solely on single though hopefully representative structures. However, our previous study on LF in TIP3P water has shown that thermal fluctuations of the H-bonds cause large fluctuations (∼15 cm^{-1}) of the C═O frequencies.[41] Correspondingly, other structural snapshots from the rMD equilibrated ensemble could have given somewhat different spectra fitting to different assignments.

If one assumes, for instance, that the $[C4{=}O]_+$ frequency of AMD_p would have been calculated at 1702 cm^{-1} instead of 1712 cm^{-1}, the rmsd from the experimental dark state spectrum $BLUF_d$ would increase from 6.6 to 7.3 cm^{-1} and the rmsd from $BLUF_l$ decrease from 8.1 to 5.4 cm^{-1}. These RMSDs would then indicate that the Met_{in} conformation AMD belongs to the light state and not to the dark state.

Furthermore, one can easily imagine conditions, which could bring about such a 10 cm^{-1} redshift of the $[C4{=}O]_+$ frequency. According to the computer experiment documented by Figure 8 it can be caused by H-bonding interactions at FMN-O4, which are somewhat stronger than those present in the DFT/MM optimized structure AMD_p. In fact, the DFT/MM optimized structure of $AMDP_p$, in which His44 is protonated, shows stronger H-bonds at FMN-O4. Thus, the $[C4{=}O]_+$ frequency is by 8 cm^{-1} smaller than in AMD_p (cf. Figures S16 and S17, Table S16, and the associated text in the SI). It may well be that similarly strong H-bonds can be found for many other snapshots of AMD's rMD ensemble. Then this Met_{in} ensemble would be classified as "light state".

Finally, it could even be that an additional water molecule, which has not been identified in the X-ray structures of the Met_{in} conformers allegedly representing the light state (e.g., SML or AML_B) is H-bonded to FMN-O4 in that state. If this should be the case, the 17 cm^{-1} red shift observed in the dark-light transition might be explained. But then none of the experimental structures studied by us would comply with the light state vibrational spectrum and all of them would have to be classified as belonging to the dark state.

As a check we carried out a computer experiment, in which we added to the model AML_{Bp} a TIP3P water molecule H-bonded to FMN-O4, because AML_B as opposed to AMD features a cavity near the FMN due to a movement of Met106.[39] This water molecule stayed attached to FMN-O4 during 150 ps of rMD simulation. We found by DFT/MM for an arbitrary snapshot that the C═O frequencies experienced an average red shift of 8 cm^{-1}. As a result the computed spectrum resembled more closely the experimental $BLUF_l$ data than the $BLUF_d$ data. However, like the spectra of unperturbed AMD_p, AML_{Bp}, and ATD_p also this modeling effect needs an extended statistics for a precise DFT/MM characterization within the INMA approach.

As a result we see that the missing large scale statistics over the rMD snapshot ensembles is the reason, why we cannot surely assign the Met_{in} conformation to the light or to the dark state of AppA BLUF. In contrast, the Trp_{in} structure does not provide any chance for stronger H-bonds at FMN-O4 and, therefore, belongs quite surely to the "dark state". Here, the snapshot chosen for the computation of the PFF even happened to render a somewhat too strongly polar residue Asn45. In fact, we have checked that 85% of the snapshots in the rMD equilibrated ATD ensemble yield less polar H_2N-Asn45 groups. Referring again to the computer experiment illustrated by Figure 8, a reduced polarity would entail for ATD a further blueshift of the $[C4{=}O]_+$ frequency well fitting to $BLUF_d$.

Note that the still possible assignment of the Met_{in} conformation to the light state and the quite certain association of the Trp_{in} conformation to the dark state of BLUF domains explained above agree with interpretations of the altered Trp fluorescence quenching upon light-adaptation[24] and with other experimentally derived arguments.[21,22,25,26,80,81]

Sufficient Statistics? The above DFT/MM analysis of the BLUF chromophore vibrational spectra was carried out in a classical quantum chemistry style by considering single albeit carefully chosen structures instead of thermalized structural ensembles, which are much more appropriate, if one wants to compute room temperature properties of soft-condensed matter. An ensemble approach has been used by us in our preparatory DFT/MM work on the IR spectra of flavins in solution.[8,41] Correspondingly, we obtained estimates on the widths and shapes of the various flavin IR bands, which are inaccessible in the single snapshot scenario employed in the work presented above.

In the given case the desirable ensemble approach posed an effort, which was unmanageable with the means presently available to us. Within such an approach one should compute

ensemble average PFFs by DFT/MM in a self-consistent manner for all considered BLUF models instead of computing just single PFFs for selected snapshots. Here, self-consistency requires to recalculate the rMD equilibrated ensembles iteratively with ensemble average PFFs until they become stationary. Given completely automated procedures for all required computational steps and large computational resources such a task will become manageable. To us it was precluded, because only parts of the procedures could be automated in the present project and because our computational resources were too small.

Correspondingly, the second step, which would have been the DFT/MM computation of vibrational spectra for thermal snapshot ensembles by the INMA approach, remains to be tackled. Note that the INMA approach could be chosen in our preceding studies on the vibrational spectra of LF in solution,[8,41] because here only a few models for the environment had to be considered, whereas now we had to deal with seven BLUF models, with the unknown protonation state of His44 in AppA BLUF, and with the necessity to compute PFFs in addition to the standard CHARMM22 force field.

Because we considered only single structures, we could not surely decide whether the Met_{in} conformation as represented by the model AMD (or AML_B, which was not discussed in detail here) belongs to the light or to the dark state of AppA BLUF. For this reason we also cannot provide estimates on the statistical errors associated with our computational results. We have solely checked for selected amino acids like Asn45, how strongly its PFF varies within the given rMD equilibrated conformational ensemble, which then strengthened our assignment of ATD to the dark state. As a check for the validity of our speculative discussion on the structure of the dark and light states of BLUF domains we would have to know the ensemble average frequencies of the flavin C═O stretching vibrations instead of snapshot values. Unfortunately the lacking statistics on the PFFs and on the vibrational spectra has precluded a more reliable conclusions concerning the assignment of AMD (or possibly AML_B) to the dark or the light state of AppA BLUF.

Summary and Outlook. Despite the statistical uncertainties addressed above, certain results definitely prevail, because they reflect large, simple, and plausible electrostatic effects. Examples are the dynamical instability of the X-ray model AML[39] (cf. Figure 2 and Figure S11 in the SI), the apparent inaccuracy of the NMR structure ATD*,[25] or the steering of the $[C2{=}O]_-$ frequency by H-bonds at FMN-O2. This steering became apparent in the computations of vibrational spectra for the X-ray structures SML[24] of Slr1694 from the cyanobacterium *Synechocystis* and BMD[52] of BlrB from *Rb. Sphaeroides*. In these cases H-bonds, which are present (SML) or absent (BMD) in the crystals and are most likely absent (SML) or present (BMD) in the solution structures of these proteins, lead to the firm conclusion that the vibrational spectra calculated for the crystal structures are definitely incompatible with observed solution spectra. Also the steering of the average C═O frequency and, in particular, of the $[C4{=}O]_+$ frequency by H-bonding interactions at FMN-O4 is one of the unambigouos results, which supports well-known empirical rules.[34−36] Finally, the IR spectra calculated for the Trp_{in} conformation of AppA BLUF (ATD) fitted quite clearly to the spectroscopic data on the dark state, which agrees with previous assignments.[21,22,80]

Unfortunately, our lacking computational resources and the still incomplete automation of PFFs by iterative DFT/MM calculations and rMD equilibrations lead to statistical uncertainties, which precluded us to assign the Met_{in} conformation as represented by AMD or AML_B[39] definitely to the dark or light state of AppA BLUF. Based on our insights gained from the many calculations of FMN spectra for BLUF domains we nevertheless continue to speculate that the light state might feature a Met_{in} conformation.[26,81]It would be nice, if we had the means to substantiate our speculation better than we could do here.

Concerning the DFT/MM methodology our current study has underlined the necessity of a careful and cautious modeling in the conversion of X-ray structures into all-atom simulation systems. Using rMD equilibrations, which enable careful checks to what extent given experimental structures are compatible with force field descriptions, one can find out, whether these structures are trustworthy starting points for further computations. Based upon such a modeling, comparisons of vibrational spectra calculated for a crystal structure with spectra observed for solubilized proteins can provide clear evidence on conformational changes, which must occur upon transfer of the protein from the crystalline lattice into solution (see the cases of the BLUF domains Slr1694 and BlrB).

Furthermore, our investigation has emphasized the caveats resulting from an earlier study on the retinal chomophore of BR,[19] which state that the quality of standard nonpolarizable force fields is too poor for sufficiently accurate computations of protein chromophore vibrational spectra. By using PFFs calculated with DFT/MM techniques for BLUF domains the spectra computed for the FMN chromophore could be brought much closer to the observations.

The investigation has additionally shown that the electrostatics coded into the simple TIP3P model[70] for water is incompatible with the protein force field CHARMM22 or DFT/MM derived PFFs. Therefore, the frequency scaling factor (1.031) obtained from computing flavin vibrational spectra in TIP3P water[8,41] had to be reduced to 1.0187 to bring stretching frequencies calculated for FMN in BLUF domains described by PFFs close to observed values.

Finally our study rendered suggestions how one should include the effects of electronic polarization into the MM fragments of DFT/MM simulation systems. Here one may either choose a mean field approch by computing for a given protein conformation an ensemble average PFF or one may explicitly apply a polarizable MM force field. To make the mean field approach accessible, the lacking automated procedures should be generated and a faster DFT/MM approach should be used. The DFT program CPMD[60] used by us is definitely suboptimal for this purpose. More cost-effective calculations could become available through the program Quickstep,[82] for instance, once a suitable interface to a MM code like EGO[53] has been established. However, for a most cost-effective inclusion of the electronic polarization a new hybrid method combining the DFT description of a chromophore with a polarizable MM force field modeling its environment should be constructed and applied.

■ ASSOCIATED CONTENT

Supporting Information. Additional figures (S11−S17), tables (S2−S16), and various pieces of text explaining and documenting structural properties of the rMD equilibrated BLUF models, specifically for AppA BLUF domains the effects of His44 protonation on the chromophore vibrational spectra, the PFFs iteratively calculated by DFT/MM for the FMN

binding pockets of the rMD equilibrated AppA BLUF models, etc. This material is available free of charge via the Internet at http://pubs.acs.org.

AUTHOR INFORMATION

Corresponding Author

*E-mail: tavan@physik.uni-muenchen.de. Phone: +49 (0)89 2180 9220. Fax: +49 (0)89 2180 9202.

ACKNOWLEDGMENT

This work was supported by the Forschergruppe 526 "Sensory Blue Light Receptors" (DFG/FOR526) and by the Sonderforschungsbereich 749 "Dynamics and Intermediates of Molecular Transformations" (DFG/SFB749-C4) of the Deutsche Forschungsgemeinschaft.

REFERENCES

(1) van der Horst, M. A.; Hellingwerf, K. J. *Acc. Chem. Res.* **2004**, *37*, 13–20.
(2) Sancar, A. *Biochemistry* **1994**, *33*, 2–9.
(3) Ahmad, M.; Cashmore, A. *Nature* **1993**, *366*, 162–166.
(4) Christie, J. M.; Swartz, T. E.; Bogomolni, R. A.; Briggs, W. R. *Plant J.* **2002**, *32*, 205–500.
(5) Gomelsky, M.; Klug, G. *Trends Biochem. Sci.* **2002**, *27*, 497–500.
(6) Berg, J. M.; Tymoczko, J. L.; Stryer, L. *Biochemie*, 5th ed.; Spektrum Akademischer Verlag GmbH: Heidelberg, Germany, 2003.
(7) Siebert, F.; Hildebrandt, P. *Vibrational Spectroscopy in Life Science*; Wiley-VCH: Berlin, 2007.
(8) Rieff, B.; Bauer, S.; Mathias, G.; Tavan, P. *J. Phys. Chem. B* **2011**, *115*, 2117–2123.
(9) Eichinger, M.; Tavan, P.; Hutter, J.; Parrinello, M. *J. Chem. Phys.* **1999**, *110*, 10452–10467.
(10) Laio, A.; VandeVondele, J.; Rothlisberger, U. *J. Chem. Phys.* **2002**, *116*, 6941–6947.
(11) Hohenberg, P.; Kohn, W. *Phys. Rev. B* **1964**, *136*, 864–871.
(12) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, 1133–1138.
(13) Nonella, M.; Mathias, G.; Eichinger, M.; Tavan, P. *J. Phys. Chem. B* **2003**, *107*, 316–322.
(14) Schmitz, M.; Tavan, P. In *Modern methods for theoretical physical chemistry of biopolymers*; Tanaka, S., Lewis, J., Eds.; Elsevier: Amsterdam, 2006; Chapter 8, pp 157−177.
(15) MacKerell, A. D.; et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.
(16) Pearlman, D.; Case, D.; Caldwell, J.; Ross, W.; Cheatham, T., III; DeBolt, S.; Ferguson, D.; Seibel, G.; Kollman, P. *Comput. Phys. Commun.* **1995**, *91*, 1–41.
(17) van Gunsteren, W.; Daura, X.; Mark, A. *Encycl. Comput. Chem.* **1998**, *2*, 1211–1216.
(18) Babitzki, G.; Denschlag, R.; Tavan, P. *J. Chem. Phys. B* **2009**, *113*, 10496–10508.
(19) Babitzki, G.; Mathias, G.; Tavan, P. *J. Chem. Phys. B* **2009**, *113*, 10483–10495.
(20) Gascon, J. A.; Leung, S. S. F.; Batista, E. R.; Batista, V. S. *J. Chem. Theory Comput.* **2006**, *2*, 175–186.
(21) Anderson, S.; Dragnea, V.; Masuda, S.; Ybe, J.; Moffat, K.; Bauer, C. *Biochemistry* **2005**, *44*, 7998–8005.
(22) Gauden, M.; van Stokkum, I. H. M.; Key, J. M.; Lührs, D. C.; van Grondelle, R.; Hegemann, P.; Kennis, J. T. M. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 10895–10900.
(23) Unno, M.; Masuda, S.; Ono, T. A.; Yamauchi, S. *J. Am. Chem. Soc.* **2006**, *128*, 5638–5639.
(24) Yuan, H.; Anderson, S.; Masuda, S.; Dragnea, V.; Moffat, K.; Bauer, C. *Biochemistry* **2006**, *45*, 12687–12694.
(25) Grinstead, J.; Hsu, S.-T.; Laan, W.; Bonvin, A.; Hellingwerf, K.; Boelens, R.; Kaptein, R. *ChemBioChem* **2006**, *7*, 187–193.
(26) Masuda, S.; Tomida, Y.; Ohta, H.; Takamiya, K. *J. Mol. Biol.* **2007**, *368*, 1223–1230.
(27) Stelling, A.; Ronayne, K.; Nappa, J.; Tonge, P.; Meech, S. *J. Am. Chem. Soc.* **2007**, *129*, 15556–15564.
(28) Takahashi, R.; Okajima, K.; Suzuki, H.; Nakamura, H.; Ikeuchi, M.; Noguchi, T. *Biochemistry* **2007**, *46*, 6459–6467.
(29) Obanayama, K.; Kobayashi, H.; Fukushima, K.; Sakurai, M. *Photochem. Photobiol.* **2008**, *84*, 1003–1010.
(30) Bonetti, C.; Mathes, T.; van Stokkum, I. H. M.; Mullen, K. M.; Groot, M.-L.; van Grondelle, R.; Hegemann, P.; Kennis, J. T. M. *Biochem. J.* **2008**, *95*, 4790–4802.
(31) Domratcheva, T.; Grigorenko, B. L.; Schlichting, I.; Nemukhin, A. V. *Biophys. J.* **2008**, *94*, 3872–3879.
(32) Sadeghian, K.; Bocola, M.; Schütz, M. *J. Am. Chem. Soc.* **2008**, *130*, 12501–12513.
(33) Sadeghian, K.; Bocola, M.; Schütz, M. *Phys. Chem. Chem. Phys.* **2010**, *12*, 8840–8846.
(34) Masuda, S.; Hasegawa, K.; Ishii, A.; Ono, T. A. *Biochemistry* **2004**, *43*, 5304–5313.
(35) Masuda, S.; Hasegawa, K.; Ono, T. A. *Plant Cell Physiol.* **2005**, *46*, 1894–1901.
(36) Unno, M.; Sano, R.; Masuda, S.; Ono, T. A. *J. Phys. Chem. B* **2005**, *109*, 12620–12626.
(37) Hasegawa, K.; Masuda, S.; Ono, T. *Plant Cell Physiol.* **2005**, *46*, 136–146.
(38) Person, B. Resonanz-Raman-Spektroskopie zur Untersuchung der Photoreaktionen biologischer Blaulichtrezeptoren. Ph.D. thesis; Universität Bielefeld, Germany, 2007.
(39) Jung, A.; Reinstein, J.; Domratcheva, T.; Shoeman, R.-L.; Schlichting, I. *J. Mol. Biol.* **2006**, *362*, 717–732.
(40) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics* **1996**, *14*, 33–38.
(41) Rieff, B.; Mathias, G.; Bauer, S.; Tavan, P. *Photochem. Photobiol.* **2010**, *87*, 511–523.
(42) Abe, M.; Kyogoku, Y.; Kitagawa, T.; Kawano, K.; Ohishi, N.; Takai-Suzuki, A.; Yagi, K. *Spectrochim. Acta A* **1986**, *42*, 1059–1068.
(43) Birss, V. I.; Hinman, A. S.; McGarvey, C. E.; Segal, J. *Electrochim. Acta* **1994**, *39*, 2449–2454.
(44) Kondo, M.; Neppa, J.; Ronayne, K. L.; Stelling, A. L.; Tonge, P. J.; Meech, S. R. *J. Phys. Chem. B* **2006**, *110*, 20107–20110.
(45) Hazekawa, I.; Nishina, Y.; Sato, K.; Shichiri, M.; Miura, R.; Shiga, K. *J. Biochem.* **1997**, *121*, 1147–1154.
(46) Nishina, Y.; Sato, K.; Miura, R.; Matsui, K.; Shiga, K. *J. Biochem.* **1998**, *124*, 200–208.
(47) Copeland, R.; Spiro, T. *J. Phys. Chem.* **1986**, *90*, 6654–6657.
(48) Hellwig, P.; Scheide, D.; Bungert, S.; Mäntele, W.; Friedrich, T. *Biochemistry* **2000**, *39*, 10884–10891.
(49) Wille, G.; Ritter, M.; Friedemann, R.; Mäntele, W.; Hübner, G. *Biochemistry* **2003**, *42*, 14814–14821.
(50) Nishina, Y.; Sato, K.; Setoyama, C.; Tamaoki, H.; Miura, R.; Shiga, K. *J. Biochem.* **2007**, *142*, 265–272.
(51) Wolf, M. M. N.; Schumann, C.; Gross, R.; Domratcheva, T.; Diller, R. *J. Phys. Chem. B* **2008**, *112*, 13424–13432.
(52) Jung, A.; Domratcheva, T.; Tarutina, M.; Wu, Q.; Ko, W.; Shoeman, R.; Gomelsky, M.; Gardner, K.; Schlichting, I. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 12350–12355.
(53) Mathias, G.; Egwolf, B.; Nonella, M.; Tavan, P. *J. Chem. Phys.* **2003**, *118*, 10847–10860.
(54) Niedermeier, C.; Tavan, P. *J. Chem. Phys.* **1994**, *101*, 734–748.
(55) Niedermeier, C.; Tavan, P. *Mol. Simul.* **1996**, *17*, 57–66.
(56) Allen, M. P.; Tildesley, D. *Computer Simulations of Liquids*; Clarendon: Oxford, 1987.
(57) Eichinger, M.; Grubmüller, H.; Heller, H.; Tavan, P. *J. Comput. Chem.* **1997**, *18*, 1729–1749.
(58) Grubmüller, H.; Tavan, P. *J. Comput. Chem.* **1998**, *19*, 1534–1552.
(59) Berendsen, H. J. C.; Postma, J. P. M.; van Gunsteren, W. F.; DiNola, A.; Haak, J. R. *J. Chem. Phys.* **1984**, *81*, 3684–3690.

(60) CPMD V3.9, Copyright IBM Corp 1990−2008, Copyright MPI für Festkörperforschung Stuttgart 1997−2001, see also www.cpmd.org.
(61) Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993–2005.
(62) Becke, A. D. *Phys. Rev. A* **1988**, *38*, 3098–3100.
(63) Perdew, J. P.; Yue, W. *Phys. Rev. B* **1986**, *33*, 8800–8802.
(64) Nonella, M.; Mathias, G.; Tavan, P. *J. Phys. Chem. A* **2003**, *107*, 8638–8647.
(65) Schmitz, M.; Tavan, P. *J. Chem. Phys.* **2004**, *121*, 12233–12246.
(66) Schmitz, M.; Tavan, P. *J. Chem. Phys.* **2004**, *121*, 12247–12258.
(67) Rauhut, G.; Pulay, P. *J. Phys. Chem.* **1995**, *99*, 3093–3100.
(68) Nonella, M.; Tavan, P. *Chem. Phys.* **1995**, *199*, 19–32.
(69) Neugebauer, J.; Hess, B. A. *J. Chem. Phys.* **2003**, *118*, 7215–7225.
(70) Jorgensen, W. L.; Chandrasekhar, J.; Madura, J. D.; Impey, R. W.; Klein, M. L. *J. Chem. Phys.* **1983**, *79*, 926–935.
(71) Brünger, A. X-PLOR Manual. The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University: New Haven, 1992.
(72) Kräutler, V.; van Gunsteren, W. F.; Hünenberger, P. *J. Comput. Chem.* **2001**, *22*, 501–508.
(73) Lingenheil, M.; Denschlag, R.; Reichold, R.; Tavan, P. *J. Chem. Theory Comput.* **2008**, *4*, 1293–1306.
(74) Liu, P.; Kim, B.; Friesner, R. A.; Berne, B. J. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 13749.
(75) Denschlag, R.; Lingenheil, M.; Tavan, P.; Mathias, G. *J. Chem. Theory Comput.* **2009**, *5*, 2847–2857.
(76) Singh, U. C.; Kollman, P. A. *J. Comput. Chem.* **1984**, *5*, 129–145.
(77) Höchtl, P.; Boresch, S.; Bitomsky, W.; Steinhauser, O. *J. Chem. Phys.* **1998**, *109*, 4927–4937.
(78) Richardi, J.; Fries, P.; Millot, C. *J. Mol. Liq.* **2005**, *117*, 3–16.
(79) Guillot, B. *J. Mol. Liq.* **2002**, *101*, 219–260.
(80) Grinstead, J.; Avila-Perez, M; Hellingwerf, K.; Boelens, R.; Kaptein, R. *J. Am. Chem. Soc.* **2006**, *128*, 15066–15067.
(81) Unno, M.; Kikuchi, S.; Masuda, S. *Biophys. J.* **2010**, *98*, 1949–1956.
(82) VandeVondele, J.; Krack, M.; Mohamed, F.; Parrinello, M.; Chassaing, T.; Hutter, J. *Comput. Phys. Commun.* **2005**, *167*, 103–128.

Supporting information for

DFT/MM description of flavin IR spectra in BLUF domains

Benjamin Rieff, Sebastian Bauer, Gerald Mathias, and Paul Tavan*

Lehrstuhl für Biomolekulare Optik, Ludwig-Maximilians-Universität,

Oettingenstr. 67, 80538 München, Germany

E-mail: tavan@physik.uni-muenchen.de

Phone: +49 (0)89 2180 9220. Fax: +49 (0)89 2180 9202

*To whom correspondence should be addressed

Instability of AML in rMD equilibrations

Figure 2(b) in the main text of the paper demonstrates for key side chains in the FMN binding pocket that the best matching rMD equilibrated models AML and AMLP deviate much more strongly from their parent AppA BLUF X-ray structure[1] than is to be expected from Boltzmann statistics for the restraining potentials. This figure furthermore shows that for AML and AMLP the deviations are much larger than for all other best matching rMD equlibrated AppA BLUF simulation models. The unexpectedly large size of these deviations indicates that the AML and AMLP simulation models are destabilized by strong forces.

To enable insights into the physical nature of these forces Figure S11 visually compares for the FMN binding pockets of ATD,[2] AMD,[1] and AML[1] the strucures of the experimental parents (left column), which are extended by the missing hydrogen atoms, with those of the best matching rMD equilibrated models containing an unprotonated (central column) and a protonated His44 (right column), respectively.

According to the structure drawn at the bottom (left) of Figure S11, the X-ray model AML features a repulsive "H-bond" between the isoalloxazine N3–H group and one of the H–N groups of Asn45. Another repulsion of this type involves the Tyr21-O–H group and the H_2N-Gln63 group. While the former repulsion is an inevitable consequence of the X-ray structure, the latter is the result of the modeling required to convert a X-ray structure into an all-atom simulation model. Here, the automatic procedure for adding hydrogen atoms[3] happened to generate an energetically unfavorable orientation of the Tyr21-O–H group. In such a case the rMD equilibration should heal the modeling artifact by either enforcing a reorientation of the Tyr21-O–H group away from the H_2N-Gln63 group or, alternatively, by reorienting the H_2N-Gln63 group away from the Tyr21-O–H group.

In fact, the rMD equilibrations plausibly identify the two side group arrangements discussed above as highly unstable. Independently of the His44 protonation state, near Asn45 they transform the initial AML structure toward the H-bonding arrangement found both in ATD(P) and AMD(P). This arrangement is characterized by two H-bonds between the FMN and Asn45 (N3–H$\cdots$O=C, C4=O$\cdots$H–N), which fix the relative orientation of these two molecular groups. Near Tyr21, instead of reorienting the Tyr21-O–H group, the rMD equlibrations lead to nearly identical reorientations of the Gln63 side chain, upon which the H_2N group moves toward the FMN. Whereas in AML the O=C-Gln63 group newly becomes an H-bond acceptor (Tyr21-O–H$\cdots$O=C-Gln63) such that the structure is similar to that of AMD, the criteria for identifying this H-bond are missed in the best matching rMD equilibrated representative AMLP. During the rMD equilibrations the soft restraining potentials, which harmonically bind the terminal O and N atoms of Gln63 to the crystallographic AML positions, apparently prevent full 180° rotations around the C_γ–C_δ bond, which would transform the AML(P) conformations of Gln63 completely to the AMD conformation.

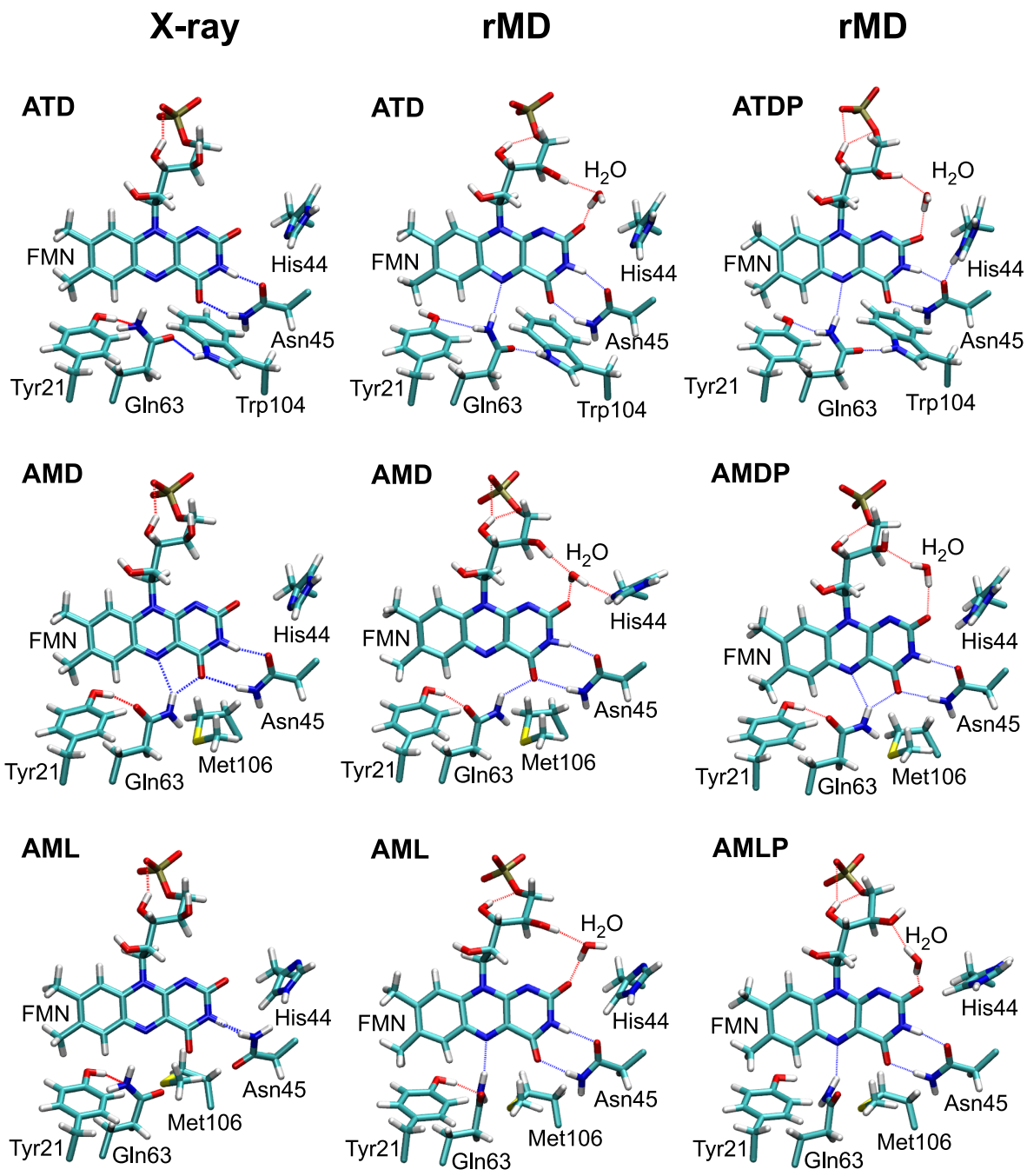


Figure S11: Comparison of AppA BLUF X-ray structures[1,2] (left column) with the best matching representatives ATD, AMD, and AML (central column, His44 unprotonated) and ATDP, AMDP, and AMLP (right column, His44 protonated) in the rMD equilibrated structural ensembles.

In contrast to AML(P), the structures of ATD(P) and AMD(P) remain essentially stable upon rMD equilibrations. According to Figure S11 (central row) the rMD equilibrated models AMD and AMDP are strikingly similar to their parent X-ray structure, whereas the two rMD equilibrated

models ATD and ATDP (top row of Figure S11) slightly differ from their X-ray parent near Tyr21 and Gln63. This local difference is most likely the result of the same modeling artifact identified further above for the X-ray structure AML. In both cases the automated addition of the hydrogen atom to the Tyr21-O–H group has lead to an energetically unfavorable orientation of this group (Figure S11, column X-ray). Independent of the His44 protonation, the rMD equilibration has reoriented the Tyr21-O–H group of ATD and ATDP in such a way that the oxygen became the acceptor of a H-bond, for which the H_2N-Gln63 group is the donor (top row in Figure S11). However, apart of this local difference the best matching rMD structures ATD and ATDP are quite similar to each other and to their parent all-atom X-ray model. Independent of the His44 protonation the two rMD equilibrated ATD structures exhibit the H-bond Gln63-C=O$\cdots$H–N-Trp104, which stabilizes the binding pocket, in general, and the conformation of the Gln63 side chain, in particular. A corresponding stabilization is absent in the Met_{in} structures AMD and AML. Furthermore, both rMD equilibrated ATD structures exhibit the two H-bonds between the FMN and Asn45, which are also found in the AMD structures and in the rMD equilibrated AML models.

A slight, but as we will see quite important, difference between the rMD equilibrated structures ATD and ATDP can be detected in the relative arrangement of FMN, His44, and Asn45. Before protonation the distance $d(\text{Asn45-O}, \text{His44-N})$ is 3.48 Å, whereas after protonation it is only 2.82 Å signifying the formation of a new H-bond between the two residues. Due to this new H-bond the two residues slightly move toward each other. As a result, the angle between the atoms FMN-C4, FMN-O4, and Asn45-N reduces from 130° to 116°, which points toward a somewhat altered H-bonding interaction between the FMN and Asn45.

Note that a water molecule appears in all rMD equilibrated AppA BLUF models as a H-bonding donor at the FMN-C2=O group. In AppA BLUF this water molecule always enters a small vacancy near the protein surface during the rMD equilibration, whereas none of the X-ray models features a crystallographic water molecule near this position.

Summarizing one may state that the AML structure decays in the rMD equilibration to structures resembling that of AMD for the obvious reasons of awkward initial H-bonding interactions, i.e. this structural model is incompatible with the electrostatics. Therefore, we consider the X-ray structure AML as highly implausible whereas the other X-ray structures ATD and AMD of the AppA BLUF protein do not show any signs of incompatibility with a force field description. Furthermore we have seen that the results of the rMD equilibrations are largely independent of the His44 protonation state. Finally we would like to remark that the best matching representatives chosen for the rMD ensembles are actually typical for the respective ensembles (data not shown).

Deviations of calculated FMN vibrational frequencies from reference data — CHARMM22/DFT

To globally measure the quality, at which the DFT/MM results on the IR spectra of FMN in BLUF models (described by the CHARMM22 force field) match the TIP3P reference spectra[4] or the experimental dark state BLUF data,[5–9] we calculated RMSDs of the FMN stretching frequencies shown in Figure 3. Table S2 lists these data for all rMD equilibrated BLUF models. The RMSDs, which all exceed the expected value[10] of about $10\,\mathrm{cm}^{-1}$, are used and discussed in the main text of the paper.

Table S2: RMSDs[a] between vibrational frequencies calculated by DFT/MM for flavin in TIP3P water (T3) and in BLUF proteins modeled by CHARMM22. Also given are RMSDs from experimental[c] dark state BLUF frequencies (exp). DFT/MM frequencies were scaled by 1.031.

model	T3[b]	exp[c]
ATD	23.2	24.2
ATDP	30.0	26.7
AMD	20.0	18.7
AMDP	23.3	25.2
ATD*	45.4	44.1
ATD*P	40.6	37.8
BMD	28.6	27.7
SML	19.8	22.0

[a]RMSDs given in cm^{-1} [b]Ref. 4 [c]Refs. 5–9

Effect of His44 protonation on the IR spectra of FMN in AppA BLUF proteins

Figure S12 compares FMN vibrational frequencies calculated by DFT/MM for the rMD equilibrated AppA BLUF models ATD, AMD, and ATD* with those of their counterparts ATDP, AMDP, and ATD*P, respectively, containing a protonated His44.

The spectral effects of the His44 protonation are different for the three rMD equilibrated AppA BLUF models considered here. Since the residue His44 is positioned directly next to the C2=O group of FMN, one would naively expect for this stretching frequency a redshift upon introducing a nearby positive charge. However, as is demonstrated by Figure S12, the protonation of His44 induces blueshifts of the $[\mathrm{C2{=}O}]_-$ band in ATD and AMD thereby moving its frequency away from the value calculated for TIP3P water. Solely in the case of ATD* the $[\mathrm{C2{=}O}]_-$ frequency shows the expected redshift ($\approx 5\,\mathrm{cm}^{-1}$). In ATD and AMD the $[\mathrm{C4{=}O}]_+$ band experiences blueshifts upon His44 protontation moving it away from the reference frequency calculated with TIP3P water, whereas in ATD* it is redshifted. There is no simple electrostatics argument, which could

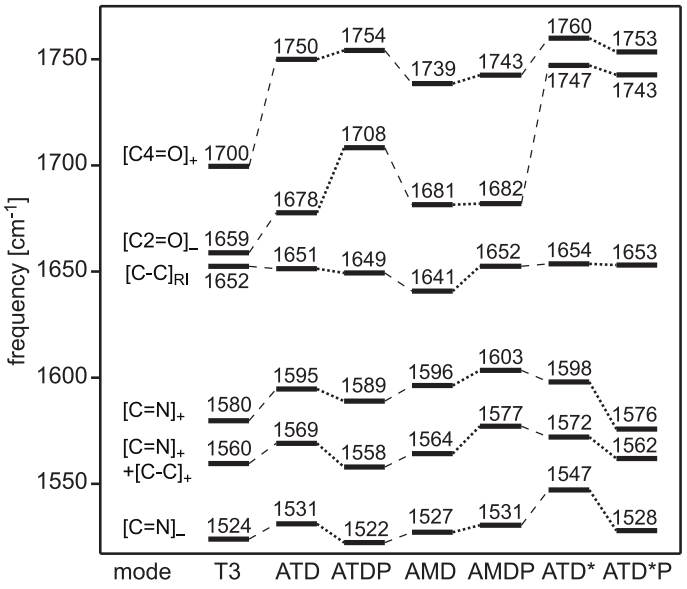


Figure S12: Protonation (P) of His44 in the rMD equlibrated AppA BLUF structures ATD, AMD, and ATD* shifts DFT/MM frequencies (scaled by 1.031) of FMN stretching vibrations.

explain these shifts. Clearly the positive charge can modify the orientation and position of the water molecule, which is H-bonded to FMN-O2 in most rMD equilibrated AppA BLUF models (cf. Figures S11 and S13) and is solely missing in ATD*P. Thereby the added charge can directly or indirectly change the coupled C=O stretching frequencies. Also for remaining protonation induced frequency shifts documented in the figure there is no obvious explanation.

Differences between the AppA BLUF models ATD and ATD*

As pointed out in the main text and documented by Table S2, the DFT/MM spectra calculated for the NMR structure ATD*[11] deviate much more strongly from the reference spectra than those calculated for the X-ray structure ATD.[2] Therefore it was suggested that the two experimental structures should significantly differ.

Figure S13 visually compares the experimental AppA BLUF models ATD[2] (X-ray) and ATD*[11] (NMR) with each other (left column). Furthermore the figure provides insights into the rMD equilibrated structures, which show the best match with their respective experimental parent structures, assuming His44 to be unprotonated (central column) or protonated (right column). One immediately recognizes substantial differences between the two experimental structures.

Whereas ATD apparently exhibits a strong H-bond between FMN-O4 and H_2N-Asn45 as is signified by the distance $d_{O4\cdots N} = 2.9$ Å, this H-bond is broken in ATD* ($d_{O4\cdots N} = 3.5$ Å). Although, the rMD equilibration and the DFT/MM minimization partially restores this H-bond in

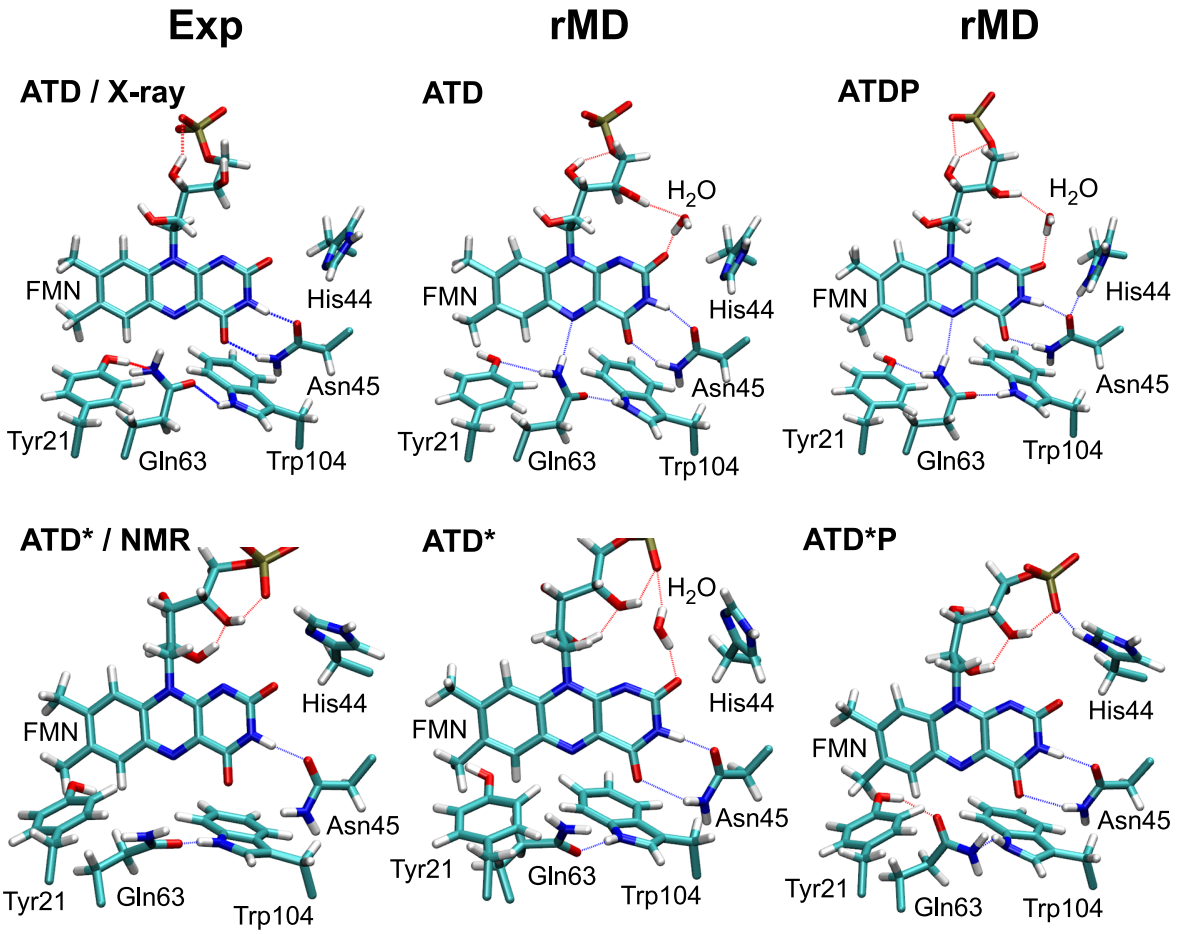


Figure S13: Comparison of the X-ray AppA BLUF structure ATD[2] with its NMR relative ATD*[11] (left column). Also shown are the respective best matching rMD equilibrated representatives containing an unprotonated His44 (central column) or a protonated His44 (right column).

ATD* ($d_{O4\cdots N} = 3.3$ Å), these simulation steps do not suffice for a complete restoration toward the distance $d_{O4\cdots N} = 2.8$ Å obtained for ATD after DFT/MM minimization of the best matching rMD equilibrated structure.

As is shown in the paragraph entitled "The spectral role of the Asn45 polarity in AppA BLUF" of the main text, the H-bond FMN-O4$\cdots$H–N-Asn45 can steer the C=O stretching frequencies shifting them to the red with increasing strength. Thus, the absence of this H-bond in ATD* and its presence in ATD partially explain why the average C=O stretching frequency is blueshifted by 35 cm^{-1} for the DFT/MM model of ATD* as compared to that of ATD (cf. Figure 3). In view of the fact that the other AppA BLUF models ATD(P) and AMD(P) already overestimate the average experimental C=O stretching frequency[5–9] by more than 20 cm^{-1}, although they all feature an FMN-O4$\cdots$H–N-Asn45 H-bond in the DFT/MM optimized structures ($d_{O4\cdots N} < 3.0$ Å), a model like ATD* becomes highly unlikely, because it exhibits an additional 35 cm^{-1} blueshift of the

average C=O stretching frequency.

Conversely, the value of 1690 cm^{-1} spectroscopically found for the average C=O stretching frequency in dark-adapted BLUF domains[5–9] and its similarity to the value of 1691 cm^{-1} measured for water[12–16] argue that FMN-O4 must be as strongly H-bonded in BLUF domains as it is in water. A structure, in which no H-bond at all can be detected at this position, cannot be compatible with vibrational spectroscopy. Thus the NMR structure ATD* is definitively inappropriate for computations of vibrational spectra and is incompatible with the data from vibrational spectroscopy.

Also other details of ATD* significantly differ from ATD. In ATD the group H_2N-Gln63, for instance, features an H-bond with Tyr21, whereas in ATD* the different orientation of Gln63 suggests that this H-bond is absent. The overall impression of ATD as compared to ATD* is that it is much better optimized with respect to electrostatics and H-bonding. This may be a result of the limited accuracy of the NMR structure analysis and of the usual neglect of electrostatics during the MD-based refinement, which can lead to atomic coordinates far off from those at which the spectra were recorded.

We would like to remark that the errors apparent in ATD* were partially healed during the rMD equilibrations despite the soft restraints binding the heavy atoms to the experimental coordinates. In the ensembles calculated with protonated and unprotonated His44, for instance, the H-bond FMN-O4$\cdots$H–N-Asn45 is usually present as is documented by the ensemble average distances $d_{O4\cdots N} \approx 2.90 \pm 0.12$ Å. However, the requirement to select from the rMD ensembles those structures, which best match the experimental parent ATD*, led to the structures shown in the bottom row of Figure S13 (center and right), at which the DFT/MM calculations of the spectra were initialized. Both structures feature distances $d_{O4\cdots N} \approx 3.1$ Å indicating a weakened or broken H-bond.

Polarized force fields for the rMD BLUF models

As described in section Methods, we generated PFFs for several side chains contained in the binding pockets of the rMD equilibrated BLUF models. Here we present the DFT/MM derived ESP partial charges for the side chains Tyr21, Ser41, His44, Asn45, Phe61, Gln63, Asp82, Arg84, His85, and Trp104/Met106 as defined by the AppA numbering. We additionally included the special water molecule, which in the AppA BLUF models is hydrogen bonded to FMN-O2 [cf. Figures 1(a) and 1(b)], into the set of polarized molecular groups.

The following Tables S3–S12 list for the various rMD equilibrated BLUF models the calculated ESP partial charges. As a reference also the CHARMM22[17] charges are given and denoted by C22.

Table S3: DFT/MM derived ESP[a] partial charges for Tyr21.

atom[b]	C22	ATD	ATDP	AMD	AMDP
CB	−0.1800	−0.1218	−0.2592	−0.1332	−0.1181
HB1	0.0900	0.0803	0.0986	0.0461	0.0574
HB2	0.0900	0.0803	0.0986	0.0461	0.0574
CG	0.0000	−0.0107	0.3150	0.1470	0.0687
CD1	−0.1150	−0.0079	−0.3532	−0.2247	−0.0878
HD1	0.1150	0.1006	0.1979	0.0966	0.0731
CD2	−0.1150	−0.0920	−0.2853	−0.1920	−0.1250
HD2	0.1150	0.1286	0.1657	0.1467	0.1536
CE1	−0.1150	−0.2651	−0.0458	−0.2418	−0.3232
HE1	0.1150	0.1333	0.1131	0.1600	0.1195
CE2	−0.1150	−0.2393	−0.2225	−0.2118	−0.2235
HE2	0.1150	0.1213	0.1484	0.1563	0.1585
CZ	0.1100	0.3055	0.2622	0.3656	0.3644
OH	−0.5400	−0.6688	−0.6615	−0.6473	−0.6924
HH	0.4300	0.4556	0.4280	0.4858	0.5171

[a]in units of e; [b]C22 atom names.[17]

Table S4: DFT/MM derived ESP[a] partial charges for Ser41.

atom[b]	C22	ATD	ATDP	AMD	AMDP
CB	0.0500	0.0503	0.1929	0.1623	0.1458
HB1	0.0900	0.0397	0.0030	0.0060	0.0165
HB2	0.0900	0.0397	0.0030	0.0060	0.0165
OG	−0.6600	−0.6084	−0.6942	−0.6809	−0.7182
HG	0.4300	0.4785	0.4952	0.5064	0.5392

[a]in units of e; [b]C22 atom names.[17]

Table S5: DFT/MM derived ESP[a] partial charges for His44.

atom[b,c]	C22	ATD	AMD	atom[b,d]	C22	ATDP	AMDP
NE2	−0.3600	−0.4413	−0.2730	ND1	−0.5100	−0.3343	−0.3285
HE2	0.3200	0.3625	0.3090	HD1	0.4400	0.4181	0.4209
CD2	−0.0500	−0.2221	−0.2781	NE2	−0.5100	−0.1911	−0.1018
HD2	0.0900	0.1145	0.1577	HE2	0.4400	0.4429	0.3709
ND1	−0.7000	−0.6576	−0.6693	CE1	0.3200	0.0623	−0.0415
CG	0.2200	0.5310	0.4267	HE1	0.1800	0.1782	0.2511
CE1	0.2500	0.3468	0.2095	CD2	0.1900	−0.2560	−0.2497
HE1	0.1300	0.0180	0.1191	HD2	0.1300	0.2267	0.2315
CB	−0.0800	−0.3608	−0.2574	CG	0.1900	0.4332	0.3756
HB1	0.0900	0.1544	0.1279	CB	−0.0500	−0.3398	−0.2733
HB2	0.0900	0.1544	0.1279	HB1	0.0900	0.1798	0.1724
				HB2	0.0900	0.1798	0.1724

[a]in units of e; [b]C22 atom names;[17] [c]of the unprotonated His44; [d]of the protonated His44.

Table S6: DFT/MM derived ESP[a] partial charges for Asn45.

atom[b]	C22	ATD	ATDP	AMD	AMDP
CB	−0.1800	−0.3125	−0.3324	−0.2736	−0.2930
HB1	0.0900	0.0886	0.1257	0.0975	0.1102
HB2	0.0900	0.0886	0.1257	0.0975	0.1102
CG	0.5500	0.9798	0.9483	0.8901	0.9308
OD1	−0.5500	−0.6703	−0.8307	−0.6224	−0.7548
ND2	−0.6200	−1.0574	−0.9228	−1.1008	−0.9644
HD21	0.3200	0.4504	0.4500	0.4466	0.4201
HD22	0.3000	0.4329	0.4360	0.4650	0.4407

[a]in units of e; [b]C22 atom names.[17]

Table S7: DFT/MM derived ESP[a] partial charges for Phe61.

atom[b]	C22	ATD	ATDP	AMD	AMDP
CB	−0.1800	−0.2312	−0.2285	−0.2607	−0.2935
HB1	0.0900	0.0684	0.0856	0.1074	0.0798
HB2	0.0900	0.0684	0.0856	0.1074	0.0798
CG	0.0000	0.2848	0.2105	0.3353	0.3640
CD1	−0.1150	−0.1714	−0.1035	−0.2317	−0.2559
HD1	0.1150	0.1225	0.0833	0.1377	0.1408
CD2	−0.1150	−0.1609	−0.0829	−0.2706	−0.2938
HD2	0.1150	0.1249	0.1246	0.1487	0.1666
CE1	−0.1150	−0.2135	−0.2237	−0.0647	−0.0500
HE1	0.1150	0.1092	0.0964	0.0863	0.0546
CE2	−0.1150	−0.1739	−0.2079	−0.0372	−0.0023
HE2	0.1150	0.1200	0.1164	0.0635	0.0756
CZ	−0.1150	−0.0362	−0.0294	−0.2573	−0.2017
HZ	0.1150	0.0888	0.0733	0.1358	0.1358

[a]in units of e; [b]C22 atom names.[17]

Table S8: DFT/MM derived ESP[a] partial charges for Gln63.

atom[b]	C22	ATD	ATDP	AMD	AMDP
CB	−0.1800	−0.0788	−0.1646	−0.1667	−0.1688
HB1	0.0900	0.0334	0.0520	0.0737	0.0686
HB2	0.0900	0.0334	0.0520	0.0737	0.0686
CG	−0.1800	0.0672	−0.1952	−0.2056	−0.2759
HG1	0.0900	0.0209	0.0728	0.0905	0.1191
HG2	0.0900	0.0209	0.0728	0.0905	0.1191
CD	0.5500	0.9058	0.9782	0.9074	0.9443
OE1	−0.5500	−0.7486	−0.7194	−0.7355	−0.7653
NE2	−0.6200	−1.0598	−1.0737	−1.0234	−0.9952
HE21	0.3200	0.4274	0.4546	0.4884	0.4887
HE22	0.3000	0.4199	0.4252	0.4070	0.3965

[a]in units of e; [b]C22 atom names.[17]

Table S9: DFT/MM derived ESP[a] partial charges for Asp82.

atom[b]	C22	ATD	ATDP	AMD	AMDP
CB	−0.2800	−0.2897	−0.4805	−0.2860	−0.3859
HB1	0.0900	0.0766	0.1628	0.0987	0.1056
HB2	0.0900	0.0766	0.1628	0.0987	0.1056
CG	0.6200	0.9535	1.0315	0.8907	1.0441
OD1	−0.7600	−0.8613	−0.9165	−0.9829	−1.0222
OD2	−0.7600	−0.9558	−0.9600	−0.8192	−0.8471

[a]in units of e; [b]C22 atom names.[17]

Table S10: DFT/MM derived ESP[a] partial charges for Arg84.

atom[b]	C22	ATD	ATDP	AMD	AMDP
CB	−0.1800	−0.1330	−0.1284	−0.2412	−0.0579
HB1	0.0900	0.0166	0.0284	0.0635	0.0055
HB2	0.0900	0.0166	0.0284	0.0635	0.0055
CG	−0.1800	0.0180	0.0718	0.2018	0.0268
HG1	0.0900	0.0337	0.0183	0.0000	0.0360
HG2	0.0900	0.0337	0.0183	0.0000	0.0360
CD	0.2000	0.3952	0.3348	0.2591	0.1822
HD1	0.0900	0.0017	0.0000	0.0000	0.0324
HD2	0.0900	0.0017	0.0000	0.0000	0.0324
NE	−0.7000	−0.8996	−0.8091	−0.7881	−0.7579
HE	0.4400	0.5773	0.5307	0.5296	0.5173
CZ	0.6400	0.9704	0.9397	0.9648	1.0218
NH1	−0.8000	−0.9881	−0.9314	−1.0007	−0.9677
HH11	0.4600	0.4953	0.4240	0.4948	0.4784
HH12	0.4600	0.4917	0.4760	0.4952	0.4815
NH2	−0.8000	−0.9607	−1.0256	−1.1042	−1.1412
HH21	0.4600	0.4853	0.5318	0.5193	0.5427
HH22	0.4600	0.4437	0.4920	0.5423	0.5258

[a]in units of e; [b]C22 atom names.[17]

Table S11: DFT/MM derived ESP[a] partial charges for His85.

atom[b]	C22	ATD	ATDP	AMD	AMDP
NE2	−0.3600	−0.3010	−0.2871	−0.2830	−0.2685
HE2	0.3200	0.3651	0.3597	0.3360	0.3498
CD2	−0.0500	−0.2887	−0.3229	−0.4181	−0.1931
HD2	0.0900	0.1757	0.2074	0.2021	0.1355
ND1	−0.7000	−0.7325	−0.6923	−0.7174	−0.2824
CG	0.2200	0.5666	0.5502	0.6443	0.1471
CE1	0.2500	0.2589	0.2401	0.3195	0.0571
HE1	0.1300	0.0701	0.0532	0.0634	0.1473
CB	−0.0800	−0.4049	−0.3806	−0.4440	−0.2932
HB1	0.0900	0.1452	0.1361	0.1485	0.1001
HB2	0.0900	0.1452	0.1361	0.1485	0.1001

[a]in units of e; [b]C22 atom names.[17]

Table S12: DFT/MM derived ESP[a] partial charges for Trp104/Met106.

atom[b,c]	C22	ATD	ATDP	atom[b,d]	C22	AMD	AMDP
CB	−0.1800	−0.0294	−0.0452	CB	−0.1800	−0.1081	−0.1941
HB1	0.0900	0.0266	0.0374	HB1	0.0900	0.0354	0.0677
HB2	0.0900	0.0266	0.0374	HB2	0.0900	0.0354	0.0677
CG	−0.0300	−0.1603	−0.1812	CG	−0.1400	0.0159	0.2137
CD2	−0.0200	0.1860	0.1466	HG1	0.0900	0.0518	0.0099
CD1	0.0350	−0.0614	0.0415	HG2	0.0900	0.0518	0.0099
HD1	0.1150	0.1261	0.1159	SD	−0.0900	−0.2043	−0.2608
NE1	−0.6100	−0.4732	−0.6050	CE	−0.2200	−0.1240	−0.1635
HE1	0.3800	0.4415	0.4973	HE1	0.0900	0.0820	0.0831
CE2	0.1300	0.2028	0.3105	HE2	0.0900	0.0820	0.0831
CE3	−0.1150	−0.3169	−0.2690	HE3	0.0900	0.0820	0.0831
HE3	0.1150	0.1342	0.1227				
CZ2	−0.1150	−0.1611	−0.2696				
HZ2	0.1150	0.1059	0.1521				
CZ3	−0.1150	−0.0299	−0.1603				
HZ3	0.1150	0.0592	0.0763				
CH2	−0.1150	−0.1618	−0.0667				
HH2	0.1150	0.0849	0.0590				

[a]in units of e; [b]C22 atom names;[17] [c]of Trp104; [d]of Met106.

Table S13: DFT/MM derived ESP[a] partial charges for the H-bonded H_2O.

atom[b]	C22	ATD	ATDP	AMD	AMDP
OH2	−0.8340	−1.0188	−0.9278	−0.9964	−0.8980
H1	0.4170	0.5094	0.4639	0.4982	0.4490
H2	0.4170	0.5094	0.4639	0.4982	0.4490

[a]in units of e; [b]C22 atom names.[17]

RMSDs of calculated FMN vibrational frequencies from references – PFF

To measure the quality, at which the DFT/MM results on the IR spectra of FMN in AppA BLUF models (described by the PFFs) match the T3 reference[4] or the experimental dark state BLUF data,[5–9] we calculated RMSDs of the FMN stretching frequencies shown in Figure 3. Table S14 lists these data for all rMD equilibrated BLUF models. The RMSDs are discussed in the main text of the paper.

Table S14: RMSDs[a] between vibrational frequencies calculated by DFT/MM for flavin in AppA BLUF proteins modeled by PFFs and in TIP3P water (T3) or experimentally[c] measured for the BLUF dark state (exp). DFT/MM frequencies were scaled by 1.031.

model	T3[b]	exp[c]
ATD_p	14.8	19.7
$ATDP_p$	31.6	31.0
AMD_p	17.4	16.6
$AMDP_p$	17.8	22.1

[a]RMSDs given in cm^{-1} [b]Ref. 4 [c]Refs. 5–9

Polarization effects on vibrational spectra of FMN in AppA-His44-P

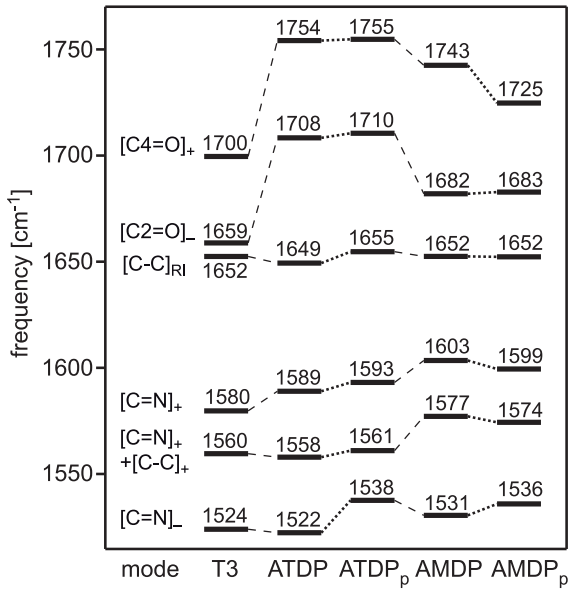


Figure S14: Polarization effects on the scaled (1.031) DFT/MM vibrational frequencies calculated for FMN in the rMD equilibrated BLUF models ATDP and AMDP featuring a protonated His44. Here, CHARMM22 standard partial charges[17] were replaced by those of the respective structure adapted PFFs (subscript "p"). Also given is the reference T3.

Figure S14 complements the data shown in Figure 6 by the polarization effects predicted by DFT/MM for the vibrational frequencies of FMN embedded in AppA BLUF proteins ATDP and AMDP featuring a protonated His44. The inclusion of the electronic polarization is seen to cause only slight frequency shifts for most modes. Exceptions are the $[C{=}N]_-$ mode, which is blueshifted in both models, and the $[C4{=}O]_+$ stretch of AMDP, which experiences a $18\,cm^{-1}$ redshift. These frequency shifts lead for $AMDP_p$ to a somewhat smaller deviation of the vibrational frequencies from the T3 reference with the RMSD decreasing to $17.8\,cm^{-1}$, whereas for ATDP the RMSD even increases to $31.6\,cm^{-1}$ upon inclusion of the electronic polarization into the MM fragment of the simulation system (cf. Tables S2 and S14).

RMSDs of rescaled FMN frequencies for polarized AppA BLUF

Table S15: RMSDs[a] of rescaled (1.0187) vibrational frequencies calculated by DFT/MM for FMN in PFF models of AppA BLUF proteins from experimental[b] data for $BLUF_d$ and $BLUF_l$.

model	$BLUF_d$	$BLUF_l$
ATD_p	7.2	7.6
$ATDP_p$	14.8	20.8
AMD_p	6.6	8.7
$AMDP_p$	7.6	8.1

[a]RMSDs given in cm^{-1}; [b]Refs. 5–9

Table S15 lists for all polarized AppA BLUF models considered by us the RMSDs of FMN's properly scaled DFT/MM frequencies from the vibrational consensus frequencies determined by IR and RR spectroscopy[5–9] for $BLUF_d$ and $BLUF_l$, respectively. These data are used and discussed in the main text of the paper.

Visualization of the PFFs for the FMN binding pockets in AppA BLUF

Figure S15 illustrates how the iterative DFT/MM procedure of computing structure adapted PFFs changed the partial charges q_i at the residues Tyr21, His44, Asn45, Gln63 and Trp104/Met106. Here the changes Δq_i are the differences $q_i(p) - q_i(c)$ of the atomic partial charges used in the various MM models p and c (see Tables S3–S13). Negative Δq_i are colored in red and positive Δq_i in blue.

The left column of Figure S15 is identical to Figure 7 in the main text, where it has been discussed. Thus, only the right column is of importance here, which demonstrates the changes introduced in the force fields ATD_p and AMD_p by protonation of His44. Apart of a different polarity of Tyr21 in $ATDP_p$ all polarity changes are quite small and probably less important for

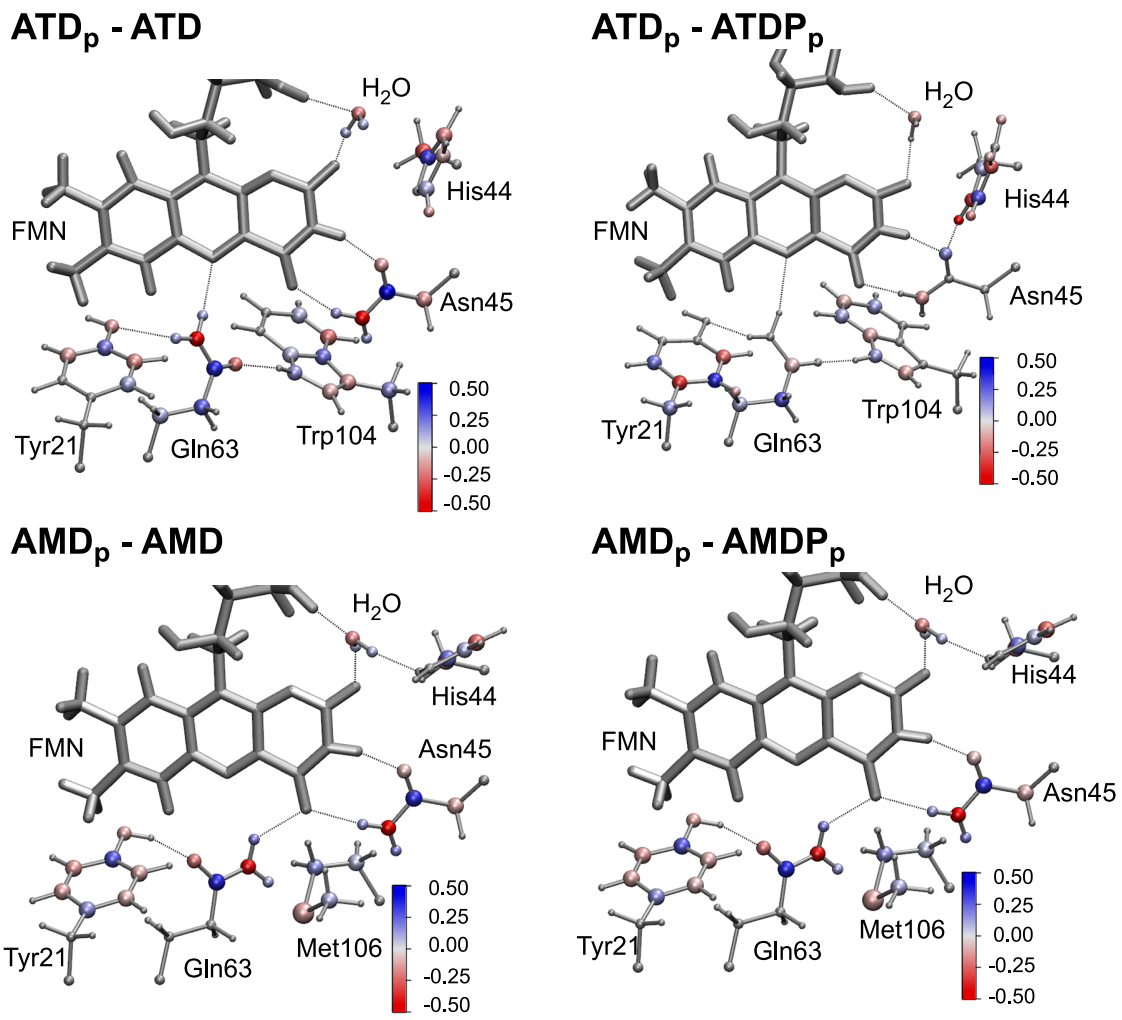


Figure S15: DFT/MM derived changes $\Delta q_i = q_i(p) - q_i(c)$ of the atomic partial charges q_i used in the various MM models $p \in \{\mathrm{ATD_p}, \mathrm{AMD_p}\}$ and $c \in \{\mathrm{ATD}, \mathrm{AMD}, \mathrm{ATDP_p}, \mathrm{AMDP_p}\}$. Negative Δq_i are colored in red and positive Δq_i in blue.

the computed spectra than the slight changes of geometry induced by adding a proton to His44. Exceptions are the H-bonded water molecules, whose polarities decrease by about 12 % upon protonation of His44. As a result, the bound water molecules are about as polar as MT/BP water molecules in a TIP3P bulk MM water environment.[18]

Movements of FMN during DFT/MM minimizations

Following the established INMA procedures[4] the DFT/MM computation of a vibrational spectrum is preceded by an energy minimization during which the FMN chromophore can move within the rigid protein binding pocket. Accordingly the various H-bonding distances will change.

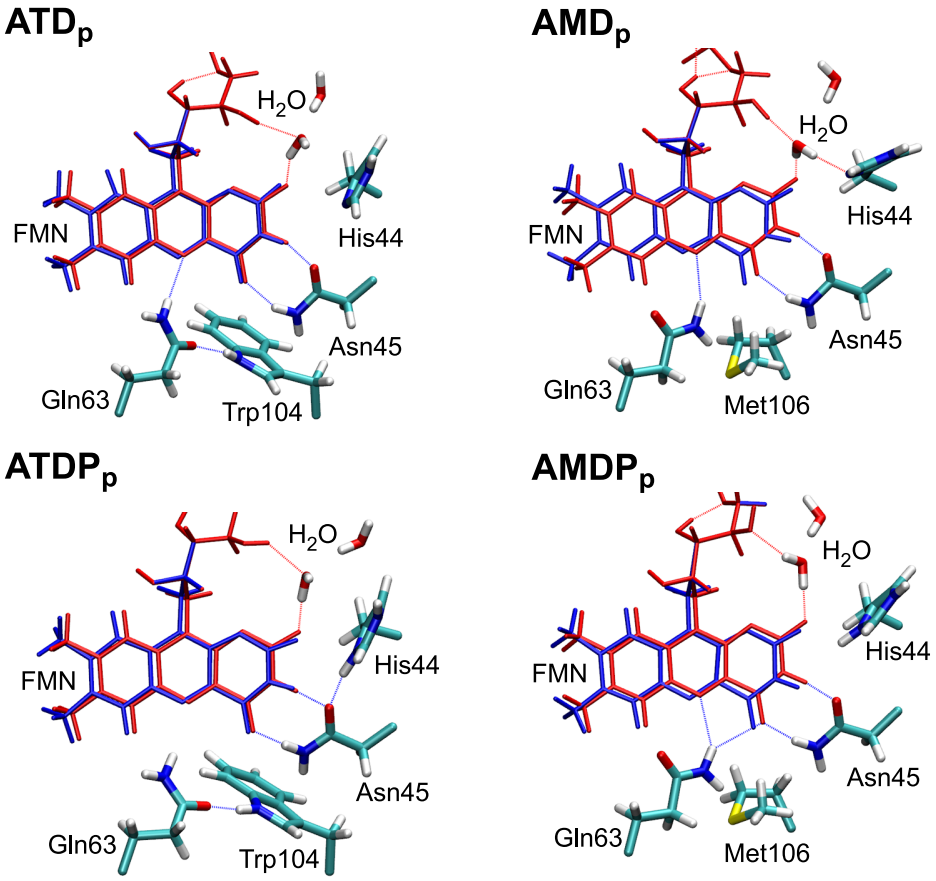


Figure S16: Visualization of the FMN movements in the polarized and rigid binding pockets of AppA BLUF models during the DFT/MM minimizations starting at the best matching representatives ATD, ATDP, AMD, and AMDP in the rMD equilibrated structural ensembles. The initial position of the FMN is drawn in blue and the final position in red.

Figure S16 illustrates these movements for the polarized models ATD$_p$, ATDP$_p$, AMD$_p$, and AMDP$_p$ of the AppA BLUF domain. According to this figure and Table S16, in nearly all models the FMN is seen to slightly translate toward the Asn45 thereby decreasing the associated H-bonding distances. In the case of AMD$_p$ this translation apparently breaks the H-bond FMN-O4$\cdots$H–N-Gln63, which is originally present in the best matching rMD equlibrated representative [cf. Figure 1(b)]. In the case of AMDP$_p$ however, the DFT/MM minimization preserves this H-bond.

Table S16: H-bonding distances[a] in polarized AppA BLUF models.

H-bond			ATD			AMD		
			X-ray[b]	rMD	DFT/PMM	X-ray[b]	rMD	DFT/PMM
FMN-O2	···	H_2O	—	2.68	2.50	—	2.62	2.41
FMN-O4	···	Asn45-N	2.89	2.90	2.74	3.21	2.94	2.68
FMN-O4	···	Gln63-N	3.98	3.97	4.15	2.78	2.90	3.59
FMN-N3	···	Asn45-O	2.74	2.81	2.67	2.82	2.84	2.65
FMN-N5	···	Gln63-N	3.05	3.07	3.20	3.22	3.33	3.26
Trp104-N	···	Gln63-O	3.29	2.71	2.71	—	—	—
H-bond			ATDP			AMDP		
			X-ray[b]	rMD	DFT/PMM	X-ray[b]	rMD	DFT/PMM
FMN-O2	···	H_2O	—	2.70	2.40	—	2.88	2.49
FMN-O4	···	Asn45-N	2.89	2.72	2.55	3.21	2.92	2.61
FMN-O4	···	Gln63-N	3.98	3.99	4.19	2.78	2.82	3.27
FMN-N3	···	Asn45-O	2.74	2.74	2.59	2.82	2.75	2.59
FMN-N5	···	Gln63-N	3.05	3.29	3.37	3.22	3.19	3.24
Trp104-N	···	Gln63-O	3.29	2.73	2.73	—	—	—

[a]given in Å; [b]values of the various X-ray structures;[1,2]

Protonation effects in rescaled FMN spectra for polarized AppA BLUF

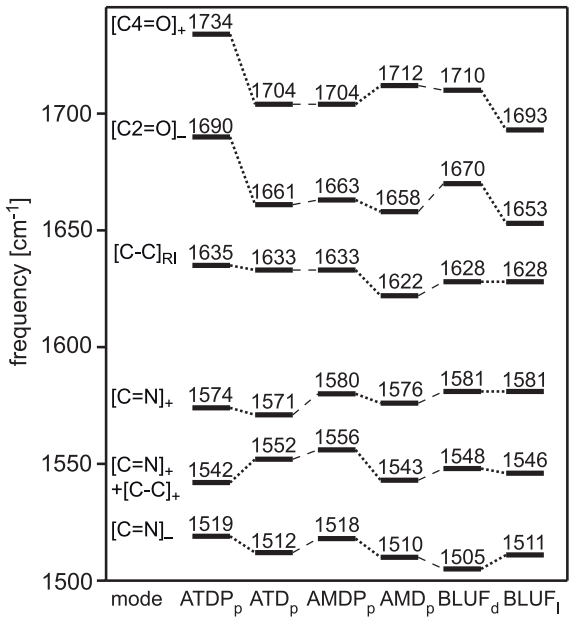


Figure S17: DFT/MM frequencies of FMN in the BLUF models $ATDP_p$, ATD_p, $AMDP_p$, and AMD_p, and scaled with 1.0187. As a reference we also give experimental consensus data[5–9] on the dark ($BLUF_d$) and light ($BLUF_l$) states of BLUF domains.

The visualization of the rescaled vibrational frequencies in Figure 10 is complemented by Figure S17, which shows the properly rescaled (1.0187) frequencies calculated by DFT/MM for FMN embedded in the polarized BLUF models $ATDP_p$, and $AMDP_p$. For comparison the Figure repeats the data on ATD_p, AMD_p, $BLUF_d$, and $BLUF_l$.

The properly scaled DFT/MM frequencies displayed in Figure S17 for the model $ATDP_p$ containing a protonated His44 exhibts much larger deviations from the experimental data $BLUF_d$ (RMSD 14.8 cm^{-1}) and $BLUF_l$ (RMSD 20.8 cm^{-1}) than the computational results for $AMDP_p$ with RMSDs of 7.6 cm^{-1} ($BLUF_d$) and 8.1 cm^{-1} ($BLUF_l$) and for the models with the unprotonated His44 (cf. Table S15). The large deviations calculated for $ATDP_p$ exclude in our view the Trp_{in} conformation of AppA BLUF with a protonated His44 as a model for the dark state and, even more so, for the light state of AppA BLUF. In section Methods of the main text we have shortly discussed apparent contradictions concerning the protonation state of His44 in ATD (or rather in the NMR solution structure ATD*) in Ref. 11.

Figure S17 reveals for the Met_{in} conformation of AppA BLUF an interesting spectral effect of the His44 protonation. In $AMDP_p$ the $[C4{=}O]_+$ frequency is by 8 cm^{-1} redshifted as compared to AMD_p. A possible reason for this redshift may be the fact that FMN-O4 features only one H-bond in the DFT/MM optimized structure AMD_p but two H-bonds in $AMDP_p$. This difference has been mentioned above in connection with Figure S16.

References

(1) Jung, A.; Reinstein, J.; Domratcheva, T.; Shoeman, R.-L.; Schlichting, I. *J. Mol. Biol.* **2006**, *362*, 717–732.

(2) Anderson, S.; Dragnea, V.; Masuda, S.; Ybe, J.; Moffat, K.; Bauer, C. *Biochemistry* **2005**, *44*, 7998–8005.

(3) Brünger, A. X-PLOR Manual. The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University: New Haven, 1992.

(4) Rieff, B.; Mathias, G.; Bauer, S.; Tavan, P. *Photochem. Photobiol.* **2010**, *87*, 511–523.

(5) Masuda, S.; Hasegawa, K.; Ishii, A.; Ono, T. A. *Biochemistry* **2004**, *43*, 5304–5313.

(6) Hasegawa, K.; Masuda, S.; Ono, T. *Plant Cell Physiol.* **2005**, *46*, 136–146.

(7) Unno, M.; Sano, R.; Masuda, S.; Ono, T. A. *J. Phys. Chem. B* **2005**, *109*, 12620–12626.

(8) Masuda, S.; Hasegawa, K.; Ono, T. A. *Plant Cell Physiol.* **2005**, *46*, 1894–1901.

(9) Person, B. Resonanz-Raman-Spektroskopie zur Untersuchung der Photoreaktionen biologischer Blaulichtrezeptoren. Ph.D. thesis, Universität Bielefeld, Germany, 2007.

(10) Schmitz, M.; Tavan, P. In *Modern methods for theoretical physical chemistry of biopolymers*; Tanaka, S., Lewis, J., Eds.; Elsevier: Amsterdam, 2006; Chapter 8, pp 157–177.

(11) Grinstead, J.; Hsu, S.-T.; Laan, W.; Bonvin, A.; Hellingwerf, K.; Boelens, R.; Kaptein, R. *ChemBioChem* **2006**, *7*, 187–193.

(12) Hazekawa, I.; Nishina, Y.; Sato, K.; Shichiri, M.; Miura, R.; Shiga, K. *J. Biochem.* **1997**, *121*, 1147–1154.

(13) Nishina, Y.; Sato, K.; Miura, R.; Matsui, K.; Shiga, K. *J. Biochem.* **1998**, *124*, 200–208.

(14) Copeland, R.; Spiro, T. *J. Phys. Chem.* **1986**, *90*, 6654–6657.

(15) Hellwig, P.; Scheide, D.; Bungert, S.; Mäntele, W.; Friedrich, T. *Biochemistry* **2000**, *39*, 10884–10891.

(16) Wille, G.; Ritter, M.; Friedemann, R.; Mäntele, W.; Hübner, G. *Biochemistry* **2003**, *42*, 14814–14821.

(17) MacKerell, A. D. et al. *J. Phys. Chem. B* **1998**, *102*, 3586–3616.

(18) Babitzki, G.; Denschlag, R.; Tavan, P. *J. Chem. Phys. B* **2009**, *113*, 10496–10508.

4 Schlussfolgerung und Ausblick

Vielleicht ist es so,
dass es am Ende niemals Antworten gibt,
sondern immer nur neue Fragen.

(M. Geyer, Der Spiegel 11/2010)

In der ersten Publikation dieser Arbeit (siehe Abschnitte 2.1 und 2.2) konnte eine DFT/MM-INMA Methode entwickelt werden, die in akkurater Weise die IR Schwingungsspektren von oxidiertem Lumiflavin in Wasser beschreibt und darüberhinaus, basierend auf der in dieser Arbeit neu entwickelten automatisierten Modenklassifizierung, eine Entschlüsselung der Normal-Moden-Zusammensetzung der experimentell gemessenen Banden [142–151] liefert. Dabei konnte nicht nur das beobachtete IR Spektrum im Bereich von $1000\,\mathrm{cm}^{-1}$ bis $1750\,\mathrm{cm}^{-1}$ reproduziert werden, sondern auch sehr gute Vorhersagen der Isotopeneffekte erreicht werden, welche die Qualität der Beschreibung nochmals unterstreichen.

Eine Anwendung der nun entwickelten DFT/MM-INMA Methode auf ein anionisches und ein neutrales Lumiflavinradikal (siehe Abschnitte 2.3 und 2.4) zeigte, dass diese Methode die IR Schwingungsspektren von solvatisierten Flavinen in anderen Redox-Zuständen [55, 145–149, 157] mit derselben Genauigkeit wie für ein oxidiertes Lumiflavin vorhersagen kann. So war es auch in diesen Fällen möglich, die Schwingungsspektren, welche auf Grund der Flavin Reduzierung komplexen Veränderungen unterworfen sind, detailliert zu erklären und eine Zuordnung der beobachteten Banden zu entsprechenden Normal-Moden vorzunehmen.

Im Gegensatz zu den für Flavine in Wasser berechneten Schwingungsspektren, die durch Mittelung einer Vielzahl einzeln betrachteter Schnappschüsse bestimmt wurden, konnte aus Rechenzeitgründen für die Schwingungsspektren des Chromophors von BLUF Proteinen nur ein einziger Schnappschuss für jede der sechs betrachteten Proteinstrukturen verwendet werden.

Um die experimentell vorgegebene Struktur der einzelnen BLUF Proteine möglichst genau zu erhalten, wurden MD Simulationen eingesetzt, die alle schweren Atome vermittels harmonischer Potentiale an die experimentell bestimmten Atomorte fixierten. Anschließend wurde für jedes BLUF Protein derjenige Schnappschuss ausgewählt, der die experimentelle Struktur am besten wiedergab. Daher können für die Resultate aus der dritten Publikation (siehe Kapitel 3) selbstverständlich keine statistischen Fehlerabschätzungen angegeben werden.

Trotz dieser statistischen Unsicherheit haben jedoch die meisten Resultate definitiv Bestand, da sie einfache und plausible Elektrostatikeffekte widerspiegeln. Hierzu gehört die teilweise dynamische Instabilität der AppA Lichtzustandsstruktur [115] aus einer Röntt-

genstrukturanalyse. Bei dieser Struktur, die ein AppA Dimer mit zwei strukturell unterschiedlichen Monomeren enthält, ist das eine Monomer (PDB-Code: 2IYI, chain A), welches von den Autoren als die plausiblere Struktur betrachtet wird [115], dynamisch instabil, weil es nicht mit den elektrostatischen Wechselwirkungen in der Flavin-Bindungstasche kompatibel ist. Das zweite Monomer (PDB-Code: 2IYI, chain B), welches dynamisch stabil ist, ist allerdings mit der Struktur des Dunkelzustandes, die in derselben Publikation veröffentlicht wurde (PDB-Code: 2IYG, chain A/B), im Bereich der Bindungstasche nahezu identisch (siehe Abb. 1.6). Lediglich die Lage des Met106 ist dort um etwa 1 Å verschoben.

Ein weiteres sicheres Resultat ist die offensichtliche Ungenauigkeit der AppA Struktur aus einer NMR Analyse [117]. Ferner wurde die empirisch bekannte Regel [107, 121, 122] bestätigt, dass Wasserstoffbrückenbindungen die spektralen Lagen der C=O Streckschwingungen deutlich beeinflussen. Darauf aufbauend konnte gezeigt werden, dass in den Kristallstrukturen der Proteine BlrB [105] und Slr1694 [116] Wasserstoffbrückenbindungen entweder fehlen oder zusätzlich existieren, welche die Bandenlage der C2=O Streckschwingung gegenüber dem solvatisierten Zustand stark verschieben. So wurden in beiden Fällen C2=O Schwingungsfrequenzen berechnet, die sich um mehr als $30\,\mathrm{cm}^{-1}$ von den für die solvatisierten BLUF Domänen experimentell bestimmten Frequenzen unterscheiden. Diese Resultate und die eingehende Untersuchung der zu Grunde liegenden Proteinstrukturen zeigt, dass die Kristallstrukturen von BlrB [105] und Slr1694 [116] nicht den solvatisierten Proteinstrukturen entsprechen können. Die Bestimmung von experimentellen Schwingungsspektren dieser Kristallstrukturen könnten diese Hypothese bestätigen, da sie in Bezug auf die C2=O Schwingungsfrequenzen deutlich von den Schwingungsfrequenzen solvatisierter Proteine abweichen sollten. Leider ist es jedoch nicht üblich, dass bei der Bestimmung einer Kristallstruktur auch ein Schwingungsspektrum dieser Struktur ermittelt wird, sodass der Vergleich mit einer entsprechenden experimentellen Schwingungsanalyse am Proteinkristall nicht möglich ist.

Methodisch konnte mit dieser Arbeit gezeigt werden, dass, wie im Fall des BR, auch im Fall der BLUF Domänen übliche nicht-polarisierbare MM Proteinkraftfelder zu ungenau sind, um eine hinreichend genaue Beschreibung von Schwingungsspektren eines Chromophors in einer Proteinumgebung zu erreichen. Erst mit der DFT/MM basierten Berechnung von BLUF-spezifischen polarisierten MM (PMM) Kraftfeldern konnten Schwingungsspektren von Flavinen sowohl in der $\mathrm{Trp}_{\mathrm{in}}$ Konformation (PDB-Code: 1YRX) [114] als auch in der $\mathrm{Met}_{\mathrm{in}}$ Konformation (PDB-Codes: 2IYG und 2IYI) [115] des AppA Proteins berechnet werden, die näher an die experimentellen Spektren herankamen. Ferner haben die Untersuchungen an den BLUF Domänen auch deutlich gemacht, dass die Elektrostatik des TIP3P Wassermodells [153], das in den vorangegangenen Arbeiten verwendet wurde, nicht zu den elektrostatischen Gegebenheiten in Proteinstrukturen passt. Deshalb musste der anfänglich für TIP3P generierte Frequenzskalierungsfaktor (1.031) für eine mit CHARMM22/PMM modellierte BLUF Proteinumgebung deutlich reduziert werden (1.0187), um eine sehr gute Übereinstimmung zwischen den berechneten DFT/PMM und den beobachteten Flavin Schwingungsspektren zu erreichen.

Die fehlende Rechenkapazität und die unvollständige Automatisierung zur Erzeugung

von ensemblegemittelten PMM Proteinkraftfeldern vermittels iterativer DFT/MM Rechnungen führten zu der schon angesprochenen statistischen Unsicherheit. Deshalb war es nicht möglich, anhand des Vergleichs mit beobachteten Schwingungsspektren eine konkrete strukturelle Zuweisung des Lichtzustandes der BLUF Domänen zu erreichen. Auf Grund der Ergebnisse aus den DFT/PMM Schwingungsspektren von Flavinen in BLUF Domänen konnte jedoch gezeigt werden, dass die Trp_{in} Konformation [siehe Abb. 1.6(a)] mit hoher Wahrscheinlichkeit den Dunkelzustand repräsentiert. Im Gegensatz dazu ist die Zuweisung der Met_{in} Konformation [siehe Abb. 1.6(b)] unklar. Gemäß den Ergebnissen könnte sie sowohl den Dunkel- als auch den Lichtzustand darstellen. Wenn die Met_{in} Konformation den Lichtzustand zugewiesen werden könnte, so würde sich die Meinung anderer [41, 114, 127, 159, 160] bestätigen, dass die Trp_{in} Konformation den Dunkelzustand und die Met_{in} Konformation den Signalzustand repräsentiert. Unklar bliebe jedoch, ob das Umschalten zwischen diesen beiden Konformationen innerhalb der angegebenen Zeit (< 1 ns) [40, 41, 58, 120] tatsächlich vonstatten gehen kann (siehe Abschnitt 1.5.2).

Wenn allerdings die Met_{in} Konformation ebenfalls den Dunkelzustand darstellt, würde dies bedeuten, dass keine der bislang untersuchten BLUF Strukturen für den Signalzustand in Frage kommt. Ferner wäre das Umschalten zwischen der Trp_{in} und Met_{in} Konformation nicht der Mechanismus, welcher der Funktion der BLUF Domäne zu Grunde liegt. Vielmehr müsste ein anderer Mechanismus für die Bildung des Signalzustandes verantwortlich sein. Wie in der dritten Publikation dieser Arbeit beispielhaft aufgezeigt wurde, könnte das Eindringen eines Wassermoleküls in die Bindungstasche an der Bildung des Signalzustandes beteiligt sein. Die Bewegung von Met106 um etwa 1 Å, die innerhalb der AppA Strukturen [115] gefunden wurde, könnte dafür den erforderlichen Platz schaffen. Wie in einer 150 ps Equilibrierung gezeigt werden konnte, verbleibt ein an diese Stelle modeliertes Wassermolekül in der Nähe des FMN-O4 Atoms und bildet eine Wasserstoffbrücke aus, welche dann die Frequenz der C4=O Streckschwingung ins Rote verschiebt. Eine solche Rotverschiebung ist bekanntlich die entscheidende spektrale Änderung beim lichtinduzierten Übergang von BLUF Domänen in den Signalzustand. Allerdings ist auch diese Möglichkeit mit statistischen Ungenauigkeiten verbunden, da einerseits nur eine kurze Equilibrierung und andererseits nur ein einzelner Schnappschuss für das DFT/PMM Schwingungsspektrum herangezogen wurden. Natürlich wäre es schön gewesen, wenn die Mittel zur Verfügung gestanden hätten, um die vorliegenden Ergbnisse und Hypothesen auch statistisch zu untermauern.

Um die erforderliche Statistik zu erreichen, müssten im nächsten Schritt, wie im Fall der Flavin Schwingungsspektren in Wasser, die Flavin Schwingungsspektren in BLUF Domänen nicht von einzelnen Schnappschüssen, sondern von einem ganzen Schnappschuss Ensemble berechnet werden. Auf Grund der hier vorgeleisteten Arbeit könnte man die Anzahl der betrachteten BLUF Proteinstrukturen auf die drei Strukturen reduzieren, die sich als geeignet erwiesen haben. Diese Maßnahme würde den erforderlichen Rechenaufwand schon deutlich reduzieren. Des Weiteren müssten ensemblegemittelte PMM Kraftfelder in selbst-konsistenter Art und Weise bestimmt werden. Im Gegensatz zu den hier verwendeten PMM Kraftfeldern, die nur auf einem Schnappschuss pro Protein basieren, bedeutet dies eine Vielzahl weiterer DFT/MM Iterationen. Selbstkonsistenz erfordert

dabei eine iterative Neubestimmung der MD generierten Ensembles mit ensemblegemittelten PMM Kraftfeldern, bis auch hier ein gewisses Maß an Stationarität erreicht ist. Dies macht deutlich, dass die Bewältigung des erforderlichen Rechenaufwandes im Rahmen dieser Arbeit nicht zu realisieren war.

Abschließend kann man festhalten, dass die elektronische Polarisierbarkeit in MM Kraftfeldern berücksichtigt werden muss. Um dies zu realisieren, bieten sich zwei Varianten an. Entweder man wählt einen „Mean-Field“ Ansatz, bei dem für eine gegebene Proteinstruktur ein ensemblegemitteltes PMM Kraftfeld ermittelt wird, oder man verwendet *polarisierbare* MM (PMM-II) Kraftfelder [161, 162]. Um die Mean-Field Methode zugänglich zu machen, sollten die fehlenden automatisierten Routinen entwickelt werden und zusätzlich schnellere DFT/MM Methoden verwendet werden. Das in dieser Arbeit verwendete DFT Programm CPMD [66] ist dafür zu langsam. Eine kosteneffizientere Berechnung würde durch modernere Programme, wie zum Beispiel `cp2k`, welches den QUICKSTEP Algorithmus [163] verwendet, ermöglicht, sobald ein passendes Interface zu MM-MD Programmen, wie dem hier verwendeten EGO [75], etabliert wurde. Für eine noch recheneffizientere Einbindung der elektronischen Polarisierbarkeit sollte allerdings eine neue Hybridmethode, die eine DFT Beschreibung eines Chromophors mit einer PMM-II Beschreibung seiner Umgebung verknüpft, entwickelt und schließlich angewendet werden.

A Die 20 natürlichen Aminosäuren

Die folgende Tabelle A.1 gibt einen Überblick über die 20 im genetischen Kode verschlüsselten Standardaminosäuren. Für alle Aminosäuren ist die in der Biochemie und Biophysik übliche dreibuchstabige sowie die einbuchstabige Kurzbezeichnung angegeben.

Aminosäure	3-Buchstabenkode	1-Buchstabenkode
Alanin	Ala	A
Arginin	Arg	R
Asparagin	Asn	N
Asparaginsäure	Asp	D
Cystein	Cys	C
Glutamin	Gln	Q
Glutaminsäure	Glu	E
Glycin	Gly	G
Histidin	His	H
Isoleucin	Ile	I
Leucin	Leu	L
Lysin	Lys	K
Methionin	Met	M
Phenylalanin	Phe	F
Prolin	Pro	P
Serin	Ser	S
Threonin	Thr	T
Tryptophan	Trp	W
Tyrosin	Tyr	Y
Valin	Val	V

Tabelle A.1: Die 20 natürlichen Aminosäuren und ihre Abkürzungen im Drei- und Ein-Buchstabenkode.

Abbildungsverzeichnis

Tabellenverzeichnis

Literaturverzeichnis

[1] Berg, J. M., J. L. Tymoczko und L. Stryer. *Biochemie*. Spektrum Akademischer Verlag GmbH, Heidelberg - Berlin, 5 Auflage (2003).

[2] Creighton, T. E. *Proteins*. W. H. Freeman and Company, San Francisco (1984).

[3] Schulz, G. E. und R. H. Schirmer. *Principles of Protein Structure*. Springer-Verlag, New York (1979).

[4] Branden, C. und J. Tooze. *Introduction to Protein Structure*. Garland Publishing, New York (1991).

[5] Baldwin, R. L. The Nature of Protein-Folding Pathways – the Classical Versus the New View. *J. Biomol. NMR* **5**, 103–109 (1995).

[6] Dill, K. und H. Chan. From Levinthal to pathways to funnels. *Nat. Struct. Biol.* **4**, 10–19 (1997).

[7] Frauenfelder, H., P. G. Wolynes und R. H. Austin. Biological Physics. *Rev. Mod. Phys.* **71**, 419–430 (1999).

[8] Carstens, H. *Konformationsdynamik lichtschaltbarer Peptide: Molekulardynamiksimulationen und datengetriebene Modellbildung*. Dissertation, Ludwig-Maximilians Universität München (2004).

[9] Itoh, H., A. Takahashi, K. Adachi, H. Noji, R. Yasuda, M. Yoshida und K. Kinosita Jr. Mechanically driven ATP synthesis by F1-ATPase. *Nature* **427**, 465–468 (2004).

[10] Wallace, B. A. Gramicidin Channels and Pores. *Ann. Rev. Biophys. Biophys. Chem* **19**, 127–157 (1990).

[11] Frauenfelder, H., S. G. Sligar und P. G. Wolynes. The Energy Landscapes and Motions of Proteins. *Science* **254**, 1598–1603 (1991).

[12] Boyer, P. D. The ATP synthase a splendid molecular machine. *Annu. Rev. Biochem.* **66**, 717–749 (1997).

[13] McCammon, J. A. und S. C. Harvey. *Dynamics of proteins and nucleic acids*. Cambridge University Press, Cambridge (1987).

[14] Geibel, S., J. H. Kaplan, E. Bamberg und T. Friedrich. Conformational dynamics of the Na+ - K+ -ATPase probed by voltage clamp fluorometry. *Proc. Natl. Acad. Sci.* **100**, 964–969 (2003).

[15] Eaton, W. A., V. Munoz, S. J. Hagen, G. S. Jas, L. J. Lapidus, E. R. Henry und J. Hofrichter. Fast kinetics and mechanisms in protein folding. *Annu. Rev. Biophys. Struct.* **29**, 327–359 (2000).

[16] Qiu, L. L., A. Pabit, A. E. Roitberg und S. J. Hage. Smaller and faster: The 20-residue rp-cage protein folds in 4 mu s. *J. Am. Chem. Soc.* **124**, 12952–12953 (2002).

[17] Grubmüller, H. und P. Tavan. Molecular-Dynamics of Conformational Substrates For a Simplified Protein Model. *J. Chem. Phys.* **101**, 5047–5057 (1994).

[18] Hamprecht, F. A., C. Peter, X. Daura, W. Thiel und W. F. van Gunsteren. A strategy for analysis of (molecular) equilibrium simulations: Configuration space density estimation, clustering and visualization. *J. Chem. Phys.* **114**, 2079–2089 (2001).

[19] Blundell, T. L. und L. N. Johnson. *Protein Crystallography*. Academic Press, London (1976).

[20] Canet, D. *NMR – Konzepte und Methoden*. Springer, Berlin (1994).

[21] Neuhaus, D. und M. Williamson. *The nuclear Overhauser effect in structural and conformational analysis*. VCH, Weinheim (1989).

[22] Giacovazzo, C., H. L. Monaco, D. Viterbo, F. Scordari, G. Gilli, G. Zanotti und M. Catti. *Fundamentals of Crystallography*, Kapitel 8. Oxford University Press, Oxford (1992).

[23] Winter, R. und F. Noll. *Methoden der Biophysikalischen Chemie*. Teubner Studienbücher, Stuttgart (1998).

[24] Garman, E. und T. Schneider. Macromolecular Cryocrystallography. *Journal of Applied Crystallography* **30**, 211–237 (1997).

[25] Oesterhelt, D. und W. Stoeckenius. Functions of a new photoreceptor membrane. *Proc. Natl. Acad. Sci.* **70**, 2853–2857 (1973).

[26] Alberts, B., D. Bray, J. Lewis, M. Raff, K. Roberts und J. D. Watson. *Molekularbiologie der Zelle*. VCH, Weinheim (1990).

[27] Zinth, W. und D. Oesterhelt. Primary Photochemical Processes in Bacteriorhodopsin. In E. Riklis (Herausgeber), *Photobiology, the Science and its Applications*, Seiten 531–535. Plenum Press, New York (1990).

[28] Siebert, F., W. Mäntele und W. Kreutz. Evidence for the protonation of two internal carboxylic groups during the photocycle of bacteriorhodopsin. *FEBS Letters* **141**, 82–87 (1982).

[29] Whittal, R. M., H. L. Ball, F. E. Cohen, A. L. Burlingame, S. L. Prusiner und M. A. Baldwin. Copper binding to octarepeat peptides of the prion protein monitored by mass spectrometry. *Protein Sci.* **9**, 332–343 (2000).

[30] Schmidt, M. Protein Kinetics: Structures of Intermediates and reaction mechanism from time-resolved x-ray data. *Proc. Natl. Acad. Sci. USA* **101**, 4799–4804 (2004).

[31] Kaptein, R., R. Boelens, R. M. Scheek und W. F. van Gunsteren. Protein Structure from NMR. *Biochemistry* **27**, 5389–5395 (1988).

[32] Riek, R., S. Hornemann, G. Wider, M. Billeter, R. Glockshuber und K. W<FC>thrich. NMR structure of the mouse prion protein domain PrP(121-231). *Nature* **382**, 180–182 (1996).

[33] Garcia, F. L., R. Zahn, R. Riek und K. W<FC>thrich. NMR structure of the bovine prion protein. *Proc. Natl. Acad. Sci. USA* **97**, 8334–8339 (2000).

[34] Sp<F6>rlein, S., H. Carstens, H. Satzger, C. Renner, R. Behrendt, L. Moroder, P. Tavan, W. Zinth und J. Wachtveitl. Ultrafast spectroscopy reveals subnanosecond peptide con-formational dynamics and validates molecular dynamics simulation. *Proc. Natl. Acad. Sci.* **99**, 7998–8002 (2002).

[35] Nilsson, L., M. Karplus, A. Br<FC>nger, M. Clore und A.Gronenbornt. Determination of 3D molecular structures from NMR/NOE distance data: application to nucleic acids. *J. Mol. Graphics* **4**, 182 (1986).

[36] Kirkpatrick, S., C. D. Gelatt und M. P. Vecchi. Optimization by Simulated Annealing. *Science* **220**, 671–680 (1983).

[37] Berman, H. M. *et al*. The Protein Data Bank. *Nucleic Acids Research* **28**, 235–242 (2000). Http://www.rcsb.org/pdb/.

[38] Demtröder, W. *Laserspektroskopie: Grundlagen und Techniken*. Springer Verlag Berlin, Germany (2007).

[39] Dobler, J., W. Zinth, W. Kaiser und D. Oesterhelt. Excited-state reaction dynamics of Bacteriorhodopsin studied by femtosecond spectroscopy. *Chem. Phys. Lett* **144**, 215–220 (1988).

[40] Gauden, M., S. Yeremenko, W. Laan, I. H. M. van Stokkum, J. A. Ihalainen, R. van Grondelle, K. J. Hellingwerf und J. T. M. Kennis. Photocycle of the flavin-binding photoreceptor AppA, a bacterial transcriptional antirepressor of photosynthesis genes. *Biochemistry* **44**, 3653–3662 (2005).

[41] Gauden, M., I. H. M. van Stokkum, J. M. Key, D. C. Lührs, R. van Grondelle, P. Hegemann und J. T. M. Kennis. Hydrogen-bond switching through a radical pair mechanism in a flavin-binding photoreceptor. *Proc. Natl. Acad. Sci. USA* **103**, 10895–10900 (2006).

[42] Pachler, K. G. R., F. Matlok und H.-U. Gremlich. *Merck FTIR Atlas*. VCH, Weinheim (1988).

[43] Byler, D. M. und H. Susi. Examination of the secondary structure of proteins by deconvolved FTIR spectra. *Biopolymers* **25**, 469–487 (1986).

[44] Susi, H. und D. M. Byler. Resolution-enhanced fourier-transform infrared-spectroscopy of enzymes. *Methods in Enzymology* **130**, 290–311 (1986).

[45] Barth, A. und C. Zscherp. What vibrations tell us about proteins. *Quarterly Reviews of Biophysics* **35**, 369–430 (2002).

[46] Vogel, R., F. Siebert, G. Mathias, P. Tavan, G. Fan und M. Sheves. Deactivation of Rhodopsin in the transition from the signaling state Meta II to Meta III involves a thermal isomerization of the retinal chromophore C=N double bond. *Biochemistry* **42**, 9863–9874 (2003).

[47] Wilson, E. B., J. C. Decius und P. C. Cross. *Molecular Vibrations*. McGraw Hill, London (1955).

[48] Woodward, L. A. *Introduction to the theory of molecular vibrations and vibrational spectroscopy*. At the Clarendon Press, Oxford (1972).

[49] Hauser, K., M. Engelhard, N. Friedman, M. Sheves und F. Siebert. Interpretation of amide I difference bands observed during protein reactions using site-directed isotopically labeled bacteriorhodopsin as a model system. *J. Phys. Chem. A* **106**, 3553–3559 (2002).

[50] Breton, J. *et al*. Binding-sites of quinones in photosynthetic bacterial reaction centers investigated by light-induced FTIR difference spectroscopy. *Biochemistry* **33**, 4953–4965, 12405–12415, 14378–14386 (1994).

[51] Brudler, R., H. J. M., de Groot, W. B. S. van Liemt, W. F. Steggerda, R. Esmeijer, P. Gast, A. J. Hoff, J. J. Lugtenburg und K. Gerwert. Asymmetric binding of the 1-C=O and 4-C=O groups of Q(A) in rhodobacter-sphaeroides-R26 reaction centers monitored by fourier-transform infrared-spectroscopy using site-specific isotopically labeled ubiquinone-10. *EMBO J.* **13**, 5523–5530 (1994).

[52] Nonella, M., G. Mathias und P. Tavan. The infrared spectrum of 1,4-benzoquinone in water obtained from a hybrid molecular dynamics simulation. *J. Phys. Chem. A* **107**, 8638–8647 (2003).

[53] Murgida, D. H., E. Schleicher, A. Bacher, G. Richter und P. Hildebrandt. Resonance Raman spectroscopic study of the neutral flavin radical complex of DNA photolyase from Escherichia Coli. *J. Ram. Spec.* **32**, 551–556 (2001).

[54] Allin, C., M. R. Ahmadian, A. Wittinghofer und K. Gerwert. Monitoring the GAP catalyzed H-Ras GTPase reaction at atomic resolution in real time. *Proc. Natl. Acad. Sci. U.S.A.* **98**, 7754–7759 (2001).

[55] Martin, C. B., M.-L. Tsao, C. M. Hadad und M. S. Platz. The reaction of triplet flavin with indole. A study of the cascade of reactive intermediates using density functional theory and time resolved infrared spectroscopy. *J. Am. Chem. Soc.* **124**, 7226–7234 (2002).

[56] Babizka, G. *Dichtefunktionaltheorie kombiniert mit einem molekülmechanischen Kraftfeld: Die Berechnung der Schwingungsspektren des Retinalchromophores von Bakteriorhodopsin*. Diplomarbeit, Ludwig-Maximilians Universität München, Lehrstuhl für Biomolekulare Optik, AG Theoretische Biophysik (2004).

[57] Masuda, S., K. Hasegawa, H. Ohta und T.-A. Ono. Crucial role in light signal transduction for the conserved Met93 of the BLUF protein PixD/Slr1694. *Plant Cell Physiol.* **49**, 1600–1606 (2008).

[58] Bonetti, C., T. Mathes, I. H. M. van Stokkum, K. M. Mullen, M.-L. Groot, R. van Grondelle, P. Hegemann und J. T. M. Kennis. Hydrogen Bond Switching among Flavin and Amino Acid Side Chains in the BLUF Photoreceptor Observed by Ultrafast Infrared Spectroscopy. *Biochem. J.* **95**, 4790–4802 (2008).

[59] Hohenberg, P. und W. Kohn. Inhomogeneous Electron Gas. *Phys. Rev. B* **136**, 864–870 (1964).

[60] Kohn, W. und L. J. Sham. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **140**, 1133–1138 (1965).

[61] van Gunsteren, W. F. und H. J. C. Berendsen. Computer Simulation of Molecular Dynamics: Methodology, Applications, and Perspectives in Chemistry. *Angew. Chem. Int. Ed. Engl.* **29**, 992–1023 (1990).

[62] Mathias, G. *Elektrostatische Wechselwirkungen in komplexen Flüssigkeiten und ihre Beschreibung mit Molekulardynamiksimulationen*. Dissertation, Dissertation, Ludwig-Maximilians Universität München (2004).

[63] Schropp, B. *Ableitung polarisierbarer Kraftfelder aus der Dichtefunktionaltheorie: Eine Pilotstudie*. Dissertation, Diplomarbeit, Ludwig-Maximilians Universität München (2005).

[64] Eichinger, M., P. Tavan, J. Hutter und M. Parrinello. A hybrid method for solutes in complex solvents: Density functional theory combined with empirical force fields. *J. Chem. Phys.* **110**, 10452–10467 (1999).

[65] Schropp, B. *Entwicklung und Analyse polarisierbarer Potentialfunktionen für Molekulardynamiksimulationen*. Dissertation, Dissertation, Ludwig-Maximilians Universität München (2010).

[66] CPMD V3.9, Copyright IBM Corp 1990-2008, Copyright MPI für Festkörperforschung Stuttgart 1997-2001, see also www.cpmd.org.

[67] Schmitz, M. *Entwicklung, Anwendung und Vergleich von Methoden zur Berechnung von Infrarotspektren einzelner Moleküle in polaren Lösungsmitteln*. Dissertation, Dissertation, Ludwig-Maximilians Universität München (2004).

[68] Brünger, A. *X-PLOR Manual*. The Howard Hughes Medical Institute and Department of Molecular Biophysics and Biochemistry, Yale University, New Haven (1992).

[69] Allen, M. P. und D. Tildesley. *Computer Simulations of Liquids*. Clarendon, Oxford (1987).

[70] McCammon, J. A., B. R. Gelin und M. Karplus. Dynamics of folded proteins. *Nature* **267**, 585–590 (1977).

[71] Grubmüller, H. und P. Tavan. Multiple time step algorithms for molecular dynamics simulations of proteins: How good are they? *J. Comput. Chem.* **19**, 1534–1552 (1998).

[72] Böckmann, R. A. und H. Grubmüller. Conformational dynamics of the F1-ATPase β-subunit: A molecular dynamics study. *Biophys. J.* **85**, 1482–1491 (2003).

[73] de Groot, B. L., T. Frigato, V. Helms und H. Grubmüller. The mechanism of proton exclusion in the aquaporin-1 water channel. *J. Mol. Biol.* **333**, 279–293 (2003).

[74] Stork, M., A. Giese, H. A. Kretzschmar und P. Tavan. Molecular dynamics simulations indicate a possible role of parallel β-helices in seeded aggregation and amyloid formation of poly-Gln. *Biophys. J.* **88**, 2442–2451 (2005).

[75] Mathias, G., B. Egwolf, M. Nonella und P. Tavan. A fast multipole method combined with a reaction field for long-range electrostatics in molecular dynamics simulations: The effects of truncation on the properties of water. *J. Chem. Phys.* **118**, 10847–10860 (2003).

[76] van Gunsteren, W., X. Daura und A. Mark. GROMOS force field. *Encyclopedia of Computational Chemistry* **2**, 1211–1216 (1998).

[77] Pearlman, D., D. Case, J. Caldwell, W. Ross, T. Cheatham III, S. DeBolt, D. Ferguson, G. Seibel und P. Kollman. AMBER, a computer program for applying molecular mechanics, normal mode analysis, molecular dynamics and free energy calculations to elucidate the structures and energies of molecules. *Comp. Phys. Comm.* **91**, 1–41 (1995).

[78] MacKerell, A. D., D. Bashford, M. Bellott, R. L. Dunbrack, J. D. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W. E. Reiher, B. Roux, M. Schlenkrich, J. C. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin und M. Karplus. All-atom empirical potential for molecular modeling and dynamics studies of proteins. *J. Phys. Chem. B* **102**, 3586–3616 (1998).

[79] Ponder, J. W. und D. A. Case. Force Fields for Protein Simulations. *Advances in Protein Chemistry* **66**, 27–85 (2003).

[80] Mathias, G. und P. Tavan. Angular resolution and range of dipole-dipole correlations in water. *J. Chem. Phys.* **120**, 4393–4403 (2004).

[81] Laio, A., J. VandeVondele und U. Rothlisberger. A Hamiltonian electrostatic coupling scheme for hybrid Car-Parrinello molecular dynamics simulations. *J. Chem. Phys.* **116**, 6941–6947 (2002).

[82] Person, B. *Resonanz-Raman-Spektroskopie zur Untersuchung der Photoreaktionen biologischer Blaulichtrezeptoren*. Dissertation, Universität Bielefeld, Germany (2007).

[83] Hellingwerf, K. L. The molecular basis of sensing and responding to light in microorganisms. *Antonie Leeuwenhoek* **81**, 51–59 (2002).

[84] van der Horst, M. A. und K. J. Hellingwerf. Photoreceptor Proteins, Star Actors of Modern Times: A Review of the Functional Dynamics in the Structure of Representative Members of Six Different Photoreceptor Families. *Acc. Chem. Res.* **37**, 13–20 (2004).

[85] Tipler, P. A. *Physik*. Spektrum Akadearticleher Verlag, Heidelberg - Berlin (1998).

[86] Ahmad, M. und A. Cashmore. HY4 gene of A. thaliana encodes a protein with characteristics of a blue-light photoreceptor. *Nature* **366**, 162–166 (1993).

[87] Sancar, A. Structure and function of DNA photolyase cryptochrome blue-light photoreceptors. *Chem. Review.* **103**, 2203–2237 (2003).

[88] Sancar, A. Regulation of the mammalian circadian clock by cryptochrome. *J. Biol. Chem.* **279**, 34079–34082 (2004).

[89] Briggs, W. R. und E. Huala. Blue-light photoreceptors in higher plants. *Annu. Rev. Cell Dev. Biol.* **15**, 33–62 (1999).

[90] Lin, C. Plant blue-light receptors. *Trends Plant Sience* **5**, 337–342 (2000).

[91] Cashmore, A. R. Cryptochroms: enabling plants and animals to determine circadian time. *Cell* **114**, 537–543 (2003).

[92] Menaker, M. Circadian rhythms. Circadian photoreception. *Science* **299**, 213–214 (2003).

[93] Frost, B. J. und H. Mouritsen. The neural mechanisms of long distance animal navigation. *Curr. Opin. Neurobiol.* **16**, 481–488 (2006).

[94] Christie, J. M., T. E. Swartz, R. A. Bogomolni und W. R. Briggs. Phototropin LOV domains exhibit distinct roles in regulating photoreceptor function. *Plant J.* **32**, 205–500 (2002).

[95] Lin, C. Blue-light receptors and signal transduction. *Plant Cell* **14**, 207–225 (2002).

[96] Kinoshita, T., M. Doi, N. Suetsugu, T. Kagawa, M. Wada und K. Shimazaki. Phot1 and Phot2 mediate blue light regulation of stomatal opening. *Nature* **414**, 656–660 (2001).

[97] Sakai, T., T. Kagawa, M. Kasahara, T. E. Swartz, J. M. Christie, W. R. Briggs, M. Wada und K. Okada. Arabidopsis nph1 and npl1: blue light receptors that mediate both phototropism and chloroplast relocation. *Proc. Natl. Acad. Sci. USA* **98**, 6969–6974 (2001).

[98] Whippo, C. W. und R. P. Hangarter. Second positive phototropism results from coordinated co-action of the phototropins and chryptochromes. *Plant Physiol.* **132**, 1499–1507 (2003).

[99] T.Kottke, J. Heberle, D. Hehn, B. Dick und P. Hegemann. Phot-LOV1: Photocycle of a Blue-Light Receptor Domain from the Grenn Alga Chlamydomonas reinhardtii. *Biophys. J.* **414**, 1192–1201 (2003).

[100] Swartz, T., S. B. Corchony, J. M. Christie, J. W. Lewis, I. Szundi, W. R. Briggs und R. A. Bogomolni. The photocycle of a flavin-binding domain of the blue light photoreceptor phototropin. *J. Biol. Chem.* **276**, 36493–36500 (2001).

[101] Salomon, M., J. M. Christie, E. Knieb, U. Lempert und W. R. Briggs. Photochemical and mutational analysis of the FMN-binding domains of the plant blue light photoreceptor phototropin. *Biochemistry* **39**, 9401–9410 (2000).

[102] Gomelsky, M. und G. Klug. BLUF: A novel FAD-binding domain involved in sensory transduction in microorganisms. *Trends Biochem. Sci.* **27**, 497–500 (2002).

[103] Laan, W., T. Bednarz, J. Heberle und K. J. Hellingwerf. Chromophore composition of a heterologously expressed BLUF-domain. *Photochem. Photobiol. Sci.* **3**, 1011–1016 (2004).

[104] Masuda, S. und C. E. Bauer. AppA is a blue light photoreceptor that antirepresses photosynthesis gene expression in *Rhodobacter sphaeroides*. *Cell* **110**, 613–623 (2002).

[105] Jung, A., T. Domratcheva, M. Tarutina, Q. Wu, W. Ko, R. Shoeman, M. Gomelsky, K. Gardner und I. Schlichting. Structure of a Bacterial BLUF Photoreceptor: Insights Into Blue Light-Mediated Signal Transduction. *Proc. Natl. Acad. Sci. USA* **102**, 12350–12355 (2005).

[106] Iseki, M., S. Matsunaga, A. Murakami, K. Ohno, K. Shiga, K. Yoshida, M. Sugai, T. Takahashi, T. Hori und M. Watanabe. A blue light acivated adenylyl cyclase mediates photoavoidance in *Euglena gracilis*. *Nature* **415**, 1047–1051 (2002).

[107] Masuda, S., K. Hasegawa, A. Ishii und T. A. Ono. Light-induced structural changes in a putative blue-light receptor with a noval FAD binding fold sensor of blue light using FAD (BLUF): Slr1694 of *Synechosystis* sp. PCC6803. *Biochemistry* **43**, 5304–5313 (2004).

[108] Okajima, K., S. Yoshihara, Y. Fukushima, X. Geng und M. Katayama. Biochemical and functional characterization of BLUF-type flavin-binding proteins of two species of cyanobacteria. *J. Biochem.* **137**, 741–750 (2005).

[109] Gomelsky, M. und S. Kaplan. AppA, a novel gene encoding a trans-acting factor involved in the regulation of photosynthesis gene expression in Rhodobacter sphaeroides. *J. Biochem.* **177**, 4609–4618 (1995).

[110] Gomelsky, M. und S. Kaplan. Molecular genetic analysis suggesting interactions between AppA and PpsR in regulation of photosynthesis gene expression in Rhodobacter sphaeroides. *J. Biochem.* **179**, 128–134 (1997).

[111] Braatsch, S., M. Gomelsky, S. Kuphal und G. Klug. A single flavoprotein, AppA,integrates both redox and light signals in Rhodobacter sphaeroides. *Mol. Microbiol.* **45**, 827–836 (2002).

[112] Gregor, J. und G. Klug. Regulation of bacterial photosynthesis genes by oxigen and light. *FEMS Microbiol. Letters* **179**, 1–9 (1999).

[113] Masuda, S. und T. A. Ono. Biochemical characterization of the major adenylyl cyclase, Cya1,in the cyanobacterium *Synechocystis* sp. PCP6803. *FEBS Lett.* **577**, 255–258 (2004).

[114] Anderson, S., V. Dragnea, S. Masuda, J. Ybe, K. Moffat und C. Bauer. Structure of a noval photoreceptor, the BLUF domain of AppA from *Rhodobacter sphaeroides*. *Biochemistry* **44**, 7998–8005 (2005).

[115] Jung, A., J. Reinstein, T. Domratcheva, R.-L. Shoeman und I. Schlichting. Crystal Structures of the AppA BLUF Domain Photoreceptor Provide Insights Into Blue Light-Mediated Signal Transduction. *J. Mol. Biol.* **362**, 717–732 (2006).

[116] Yuan, H., S. Anderson, S. Masuda, V. Dragnea, K. Moffat und C. Bauer. Crystal structures of the Synechocystis photoreceptor Slr1694 reveal distinct structural states related to signaling. *Biochemistry* **45**, 12687–12694 (2006).

[117] Grinstead, J., S.-T. Hsu, W. Laan, A. Bonvin, K. Hellingwerf, R. Boelens und R. Kaptein. SOLUTION STRUCTURE OF THE BLUF DOMAIN OF AppA 5-125. *ChemBioChem* **7**, 187–193 (2006).

[118] van der Horst, M. A., W. Laan, S. Yeremenko, A. Wende, P. Palm, D. Oesterhelt und K. J. Hellingwerf. From primary photochemistry to biological function in the blue-light photoreceptors PYP and AppA. *Photochem. Photobiol. Sci.* **4**, 688–693 (2005).

[119] Kraft, B. J., S. Masuda, J. Kikuchi, V. Dragnea, G. Tollin, J. M. Zaleski und C. E. Bauer. Spectroscopic and mutational analysis of the blue-light photoreceptor AppA: A novel photocycle involving flavin stacking with an aromatic amino acid. *Biochemistry* **42**, 6726–6734 (2003).

[120] Toh, K. C., I. H. M. van Stokkum, J. Hendriks, M. T. A. Alexandre, J. C. Arents, M. A. Perez, R. van Grondelle, K. J. Hellingwerf und J. T. M. Kennis. On the signaling mechanism and the absence of photoreversibility in the AppA BLUF domain. *Biochem. J.* **95**, 312–321 (2008).

[121] Masuda, S., K. Hasegawa und T. A. Ono. Tryptophan at Position 104 is involved in Transforming Light Signal into Changes of β-sheet Structure for the Signaling State in the BLUF Domain of AppA. *Plant Cell Physiol.* **46**, 1894–1901 (2005).

[122] Unno, M., R. Sano, S. Masuda und T. A. Ono. Light-induced structural changes in the active side of the BLUF domain in AppA by Resonance Raman spectroscopy. *J. Phys. Chem. B* **109**, 12620–12626 (2005).

[123] V.Dragnea, M. Waegele, S. Balascuta, C. Bauer und B. Dragnea. Time-resolved spectoscopic studies of the AppA blue-light receptor BLUF domain from *Rhodobacter sphaeroides*. *Biochemistry* **44**, 15978–15985 (2005).

[124] Laan, W., M. Gauden, S.Yeremenko, R. van Grondelle, J. T. Kennis und K. J. Hellingwerf. On the mechanism of activation of the BLUF domain of AppA. *Biochemistry* **45**, 51–60 (2006).

[125] Hasegawa, K., S. Masuda und T. Ono. Spectroscopic analysis of the dark relaxation process of a photocycle in a sensor of blue light using FAD (BLUF) protein SLR1694 of the cyanobacterium *Synechocystis sp.* PCC6803. *Plant Cell Physiol.* **46**, 136–146 (2005).

[126] Unno, M., S. Masuda, T. A. Ono und S. Yamauchi. Orientation of a key glutamine residue in the BLUF domain from AppA revealed by mutagenesis, spectroscopy, and quantum chemical calculations. *J. Am. Chem. Soc.* **128**, 5638–5639 (2006).

[127] Masuda, S., Y. Tomida, H. Ohta und K. Takamiya. The critical role of a hydrogen bond between Gln63 and Trp104 in the blue-light sensing BLUF domain that controlls AppA activity. *J. Mol. Biol.* **368**, 1223–1230 (2007).

[128] Stelling, A., K. Ronayne, J. Nappa, P. Tonge und S. Meech. Ultrafast structural dynamics in BLUF domains: Transient infrared spectroscopy of AppA and its mutants. *J. Am. Chem. Soc.* **129**, 15556–15564 (2007).

[129] Takahashi, R., K. Okajima, H. Suzuki, H. Nakamura, M. Ikeuchi und T. Noguchi. FTIR study on the hydrogen bond structure of a key tyrosine residue in the flavin-binding blue light sensor TePixD from *Thermosynechococcus elongatus*. *Biochemistry* **46**, 6459–6467 (2007).

[130] Obanayama, K., H. Kobayashi, K. Fukushima und M. Sakurai. Structures of the chromophore binding sites in BLUF domains as studied by molecular dynamics and quantum chemical calculations. *Photochem. Photobiol.* **84**, 1003–1010 (2008).

[131] Domratcheva, T., B. L. Grigorenko, I. Schlichting und A. V. Nemukhin. Molecular models predict light-induced glutamine tautomerization in BLUF photoreceptors. *Biophys. J.* **94**, 3872–3879 (2008).

[132] Sadeghian, K., M. Bocola und M. Schütz. A Conclusive Mechanism of the Photoinduced Reaction Cascade in Blue Light Using Flavin Photoreceptors. *J. Am. Chem. Soc.* **130**, 12501–12513 (2008).

[133] Sadeghian, K., M. Bocola und M. Schütz. A QM/MM study on the fast photocycle of blue light using flavin photoreceptors in their light-adapted/active form. *Phys. Chem. Chem. Phys* **12**, 8840–8846 (2010).

[134] Schmitz, M. und P. Tavan. On the art of computing the IR spectra of molecules in condensed phase. In S. Tanaka und J. Lewis (Herausgeber), *Modern methods for theoretical physical chemistry of biopolymers*, Kapitel 8, Seiten 157–177. Elsevier, Amsterdam (2006).

[135] Geigeot, M. P. und M. Sprik. Ab initio molecular dynamics computation of the infrared spectrum of aqueous Uracil. *J. Phys. Chem. B* **107**, 10344–10358 (2003).

[136] Ramirez, R., T. Lopez-Ciudad, P. Kumar und D. Marx. Quantum corrections to classical time-correlation functions: Hydrogen bonding and anharmonic floppy modes. *J. Chem. Phys.* **121**, 3973–3983 (2004).

[137] Nonella, M., G. Mathias, M. Eichinger und P. Tavan. Structures and vibrational frequencies of the quinones in *Rb. Sphaeroides* derived by a combined density functional molecular mechanics approach. *J. Phys. Chem. B* **107**, 316–322 (2003).

[138] Baer, M., D. Marx und G. Mathias. Theoretical messenger spectroscopy of microsolvated Hydronium and Zundel cations. *Angew. Chem. Int. Ed.* **49**, 7346–7349 (2010).

[139] Babitzki, G., G. Mathias und P. Tavan. The infra-red spectra of the retinal chromophore in *Bacteriorhodopsin* calculated by a DFT/MM approach. *J. Chem. Phys. B* **113**, 10483–10495 (2009).

[140] Schmitz, M. und P. Tavan. Vibrational spectra from atomic fluctuations in dynamics simulations. I. Theory, limitations, and a sample application. *J. Chem. Phys.* **121**, 12233–12246 (2004).

[141] Schmitz, M. und P. Tavan. Vibrational spectra from atomic fluctuations in dynamics simulations. II. Solvent-induced frequency fluctuations at femtosecond time resolution. *J. Chem. Phys.* **121**, 12247–12258 (2004).

[142] Abe, M., Y. Kyogoku, T. Kitagawa, K. Kawano, N. Ohishi, A. Takai-Suzuki und K. Yagi. Infrared spectra and molecular association of lumiflavin and riboflavin derivatives. *Spectrochim. Acta A* **42**, 1059–1068 (1986).

[143] Birss, V. I., A. S. Hinman, C. E. McGarvey und J. Segal. *In situ* FTIR thin-layer reflectance spectroscopy of flavin adenine dinucleotide at a mercury/gold electrode. *Electrochim. Acta* **39**, 2449–2454 (1994).

[144] Kondo, M., J. Neppa, K. L. Ronayne, A. L. Stelling, P. J. Tonge und S. R. Meech. Ultrafast vibrational Spectroscopy of the Flavin Chromophore. *J. Phys. Chem. B* **110**, 20107–20110 (2006).

[145] Hazekawa, I., Y. Nishina, K. Sato, M. Shichiri, R. Miura und K. Shiga. A Raman Study on the C(4)=O Stretching Mode of Flavins in Flavoenzymes: Hydrogen Bonding at the C(4)=O Moiety. *J. Biochem.* **121**, 1147–1154 (1997).

[146] Nishina, Y., K. Sato, R. Miura, K. Matsui und K. Shiga. Resonance Raman Study on Reduced Flavin in Purple Intermediate of Flavoenzyme: Use of 4-Carbonyl-^{18}O-Enriched Flavin. *J. Biochem.* **124**, 200–208 (1998).

[147] Copeland, R. und T. Spiro. Ultraviolet Resonance Raman Spectroscopy of Flavin Mononucleotide and Flavin Adenine Dinucleotide. *J. Phys. Chem.* **90**, 6654–6657 (1986).

[148] Hellwig, P., D. Scheide, S. Bungert, W. Mäntele und T. Friedrich. FT-IR spectroscopic characterization of NADH:Ubiquinone oxidoreductase (Complex I) from *Escherichia coli*: Oxidation of FeS cluster N2 is coupled with the protonation of an aspertate or glutamine side chain. *Biochemistry* **39**, 10884–10891 (2000).

[149] Wille, G., M. Ritter, R. Friedemann, W. Mäntele und G. Hübner. Redox-triggered FTIR difference spectra of FAD in aqueous solution and bound to flavoproteins. *Biochemistry* **42**, 14814–14821 (2003).

[150] Nishina, Y., K. Sato, C. Setoyama, H. Tamaoki, R. Miura und K. Shiga. Intramolecular and intermolecular perturbation on electronic state of FAD free in solution and bound to flavoproteins: FTIR spectroscopic study by using C=O stretching vibrations as probes. *J. Biochem.* **142**, 265–272 (2007).

[151] Wolf, M. M. N., C. Schumann, R. Gross, T. Domratcheva und R. Diller. Ultrafast Infrared Spectroscopy of Riboflavin: Dynamics, Electronic Structure, and Vibrational Mode Analysis. *J. Phys. Chem. B* **112**, 13424–13432 (2008).

[152] Neugebauer, J. und B. A. Hess. Fundamental vibrational frequencies of small polyatomic molecules from density-functional calculations and vibrational perturbation theory. *J. Chem. Phys.* **118**, 7215–7225 (2003).

[153] Jorgensen, W. L., J. Chandrasekhar, J. D. Madura, R. W. Impey und M. L. Klein. Comparison of simple potential functions for simulating liquid water. *J. Chem. Phys.* **79**, 926–935 (1983).

[154] Perdew, J. P. und W. Yue. Accurate and simple density functional for the electronic exchange energy: Generalized gradient approximation. *Phys. Rev. B* **33**, 8800–8802 (1986).

[155] Becke, A. D. Density-functional exchange-energy approximation with correct asymptotic behavior. *Phys. Rev. A* **38**, 3098–3100 (1988).

[156] Troullier, N. und J. L. Martins. Efficient pseudopotentials for plane-wave calculations. *Phys. Rev. B* **43**, 1993–2005 (1991).

[157] Su, Y. und G. N. R. Tripathi. Time-resolved resonance Raman observation of protein-free riboflavin semiquinone radicals. *J. Am. Chem. Soc.* **116**, 4405–4407 (2010).

[158] Babitzki, G., R. Denschlag und P. Tavan. Polarization effects stabilize *Bacteriorhodopsin's* chromophore binding pocket: A molecular dynamics study. *J. Chem. Phys. B* **113**, 10496–10508 (2009).

[159] Grinstead, J., M.Avila-Perez, K. Hellingwerf, R. Boelens und R. Kaptein. Light-induced flipping of a conserved glutamine sidechain and its orientation in the AppA BLUF domain. *J. Am. Chem. Soc.* **128**, 15066–15067 (2006).

[160] Unno, M., S. Kikuchi und S. Masuda. Structural refinement of a key tryptophan residue in the BLUF photoreceptor AppA by ultraviolet Resonance Raman specroscopy. *Biophys. J.* **98**, 1949–1956 (2010).

[161] Schultheis, V., R. Reichold, B. Schropp und P. Tavan. A Polarizable Force Field for Computing the infrared Spectra of the Polypeptide Backbone. *J. Phys. Chem. B* **112**, 12217–12230 (2008).

[162] Schropp, B., C. Wichmann und P. Tavan. A spectroscopic polarizable force field for amide groups in polypeptides. *J. Phys. Chem. B* **114**, 6740–6750 (2010).

[163] VandeVondele, J., M. Krack, F. Mohamed, M. Parrinello, T. Chassaing und J. Hutter. Quickstep: Fast and accurate density functional calculations using a mixed Gaussian and plane waves approach. *Comp. Phys. Comm.* **167**, 103–128 (2005).

Danksagung

Leider lässt sich eine wahrhafte Dankbarkeit
mit Worten nicht ausdrücken.

(Johann Wolfgang von Goethe)

Zum Schluss meiner Promotion ist es an der Zeit, allen Menschen, die an der erfolgreichen Umsetzung dieser Arbeit beteiligt waren, einmal meinen aufrichtigen Dank zu übermitteln. Unter allen Mitarbeiter des gesamten Lehrstuhls für Biomolekulare Optik herrschte stets eine freundliche und angenehme Arbeitsatmosphäre, welche die Strapazen einer Doktorarbeit deutlich leichter zu bewältigen ermöglichte. Ich hoffe, ich konnte durch die Versorgung mit Kaffee einen angemessenen Beitrag beisteuern. Auch der Deutschen Forschungsgemeinschaft möchte ich danken, die im Rahmen der Forschergruppe 526 und im Rahmen des SFB 749 die Finanzierung meiner Dissertation übernahm. Selbstverständlich gibt es auch einzelne Personen, die sich in ganz besonderem Maße für das Gelingen meiner Promotion hervorgehoben haben. Ihnen möchte ich gerne einzelnt meine Dankbarkeit zum Ausdruck bringen.

Als erstes möchte ich natürlich meinem Doktorvater, Paul Tavan, meinen herzlichen Dank dafür aussprechen, dass er mir das Vertrauen entgegengebracht und dadurch meine Promotion überhaupt erst ermöglicht hat. Auch wenn das Thema meiner Dissertation aus einem gewissen Zwang heraus entstanden ist, so habe ich das Projekt letztendlich sehr lieb gewonnen und mein ganzes Herzblut hineingesteckt. Deshalb bin ich auch sehr froh, dass er diese fazinierende Forschungsarbeit in die Obhut meiner Hände gelegt hat. Seine schier grenzenlose Unterstützung, die zahlreichen Diskussionen, seine Fähigkeit zu motivieren aber auch die ein oder andere Schelte haben maßgeblich zum Gelingen dieser Arbeit beigetragen. Neben dem erfolgreichen Abschluss bin ich auch sehr dankbar für die fachliche sowie besonders für meine persönliche Weiterentwicklung, die mir in Zukunft sehr behilflich sein wird. Alles in Allem ist es mir eigentlich nicht möglich, meine Dankbarkeit und Wertschätzung zu Papier zu bringen, so sage ich einfach: *Danke für alles, „Papa“Paul*!

Auch Gerald Mathias möchte ich meinen Dank aussprechen, der bei all meinen kleinen Problemen stets ein offenes Ohr für mich hatte. Auch wenn er selbst genug zu tun hatte, stand er mir immer sofort mit Rat und Tat zur Seite und war deshalb auf Grund seiner zahlreichen fachlichen und technischen Tipps eine sehr große Hilfe für mich.

Bei den Mitgliedern meiner Arbeitsgruppe, Martina Stork, Verena Schultheis, Galina Babitzki, Christine Lutz, Thomas Hirschberger, Rudolf Reichhold, Martin Lingenheil, Robert Denschlag, Bernhard Schropp, Christoph Wichmann, Konstantin Lorenzen, Philipp Troester und Maximilian Schmidt möchte ich mich fúr das kollegiale Arbeitsklima und das freundschaftliche Miteinander recht herzlich bedanken. Ich werde mich stetsgerne an

diese Zeit erinnern.

Nicht zu vergessen sind die Damen aus dem Sekretariat und der Verwaltung, hier sind besonders Alexandra Michaelis, Marianne Widmann-Diermeier und Sigi Siegert zu nennen, durch die stets ein reibungsloser Ablauf aller organisatorischen Verwatlungstätigkeiten gewährleistet war.

Einen nicht zu vernachlässigen Beitrag haben auch meine Freunde geleistet, hier vor allem meine Kameraden aus der Freiwilligen Feuerwehr Gröbenzell, die mich stets daran erinnert haben, dass es auch eine Welt außerhalb der Universität gab, und mir auch die Stange gehalten haben, wenn ich des öfteren mal keine Zeit für sie hatte. Im selben Atemzug ist natürlich auch meine Familie, meine Mutter Maria-Theresia, mein Vater Rudolf, mein Bruder Telemach und meine Fast-Schwägerin Claudia Kreis zu erwähnen. Euch allen ein herzliches Dankeschön.

Auch meinem Studiengefährten, Bürokollegen, Leidensgenossen und in erster Linie sehr guten Freund Sebastian Bauer möchte ich meinen aufrichtigen Dank aussprechen. All die Höhen und Tiefen des Studiums sowie der Promotion haben wir gemeinsam gemeistert. Die zahlreichen Gespräche privater und wissenschaftlicher Natur werden mir stets in guter Erinnerung bleiben. Ich habe es geschafft, ich habe fertig! Dazu hast auch Du, lieber Sebi, einen ganz entscheidenden Beitrag beigesteuert.

Zu guter Letzt möchte ich mich noch bei meiner Frau Corinna bedanken. Sie hat mir immer, egal was auch kam, den Rücken freigehalten. Danke auch für all die Kleinigkeiten, die einem so selbstverständlich vorkommen. Durch ihre grenzenlose Unterstützung und Liebe hat sie mir die Kraft gegeben, die Strapazen, die Mühen und die zahlreichen Rückschläge während meiner Dissertation zu überwinden. Danke für die unvergesslichen Momente, die wunderschöne gemeinsame Zeit sowie für alles andere natürlich auch. Liebe Nini, ohne Dich hätte ich das hier nie geschafft!