
DIE DIPHOSPHATAKTIVIERUNG IN PROTEINEN
BETRACHTUNGEN AUS VERSCHIEDENEN BLICKWINKELN

DISSERTATION

zur Erlangung des akademischen Grades
doctor rerum naturalium (Dr. rer. nat.)
vorgelegt der

Naturwissenschaftlichen Fakultät II
(Chemie, Physik und Mathematik)
der Martin-Luther-Universität Halle-Wittenberg

von Frau Diplom-Bioinformatikerin Diana Schulze
geb. am 19.05.1982 in Nordhausen

Gutachter:

- | | |
|-------------------------------|--|
| 1. Prof. Dr. L. A. Wessjohann | Martin-Luther-Universität Halle-Wittenberg |
| 2. Dr. T. E. Exner | Universität Konstanz |

Halle (Saale), den 17.06.2011

EINGEREICHT AM: 09.12.2010
VERTEIDIGT AM: 17.06.2011

NUR WENIGE WISSEN,
WIE VIEL MAN WISSEN MUSS,
UM ZU WISSEN,
WIE WENIG MAN WEIß.

WERNER HEISENBERG

FÜR MEINE ELTERN

Die vorliegende Arbeit wurde im Zeitraum von Mai 2006 bis Dezember 2009 in der Abteilung für Natur- und Wirkstoffchemie des Leibniz-Instituts für Pflanzenbiochemie (IPB) in Halle angefertigt.

An dieser Stelle möchte ich **Herrn Prof. Dr. Ludger A. Wessjohann** für die interessante Themenstellung und die freundliche Aufnahme am IPB danken.

Für die stete Diskussionsbereitschaft, die stets kritische Hinterfragung von Ergebnissen, aber auch für den Freiraum zum Experimentieren sei hier Herrn **PD Dr. habil. Wolfgang Brandt** im Besonderen gedankt.

Weiterhin danke ich:

Herrn **Prof. Dr. Horst Bögel** für das kritische Korrekturlesen meiner Arbeit und die damit verbundenen Ratschläge.

Frau **Julia Kufka** für die schöne Zusammenarbeit, die mir viel Freude bereitet hat, und die Einblicke in die Welt der Biochemie.

den Mitarbeitern der Arbeitsgruppe Computerchemie Frau **Dr. Stephanie Tennstedt**, Herrn **Peter-Paul Heym**, Herrn **Robert Klein**, Frau **Eva Schulze**, Frau **Juliane Fischer** und Herrn **Felix Rausch**, sowie Frau **Dr. Susanne Aust** und Herrn **Dr. Frank Broda** für die anregenden Gespräche, das fortwährende Interesse an der Entwicklung meiner Arbeit und die angenehme Arbeitsatmosphäre.

Herrn **Torsten Geissler**, Herrn **Michael Henze**, Frau **Julia Mülbradt** und Herrn **Rainer Preusentanz** für die chemischen Diskussionen und Ratschläge und die schöne, gemeinsame Zeit.

allen Mitarbeitern der Abteilung Natur- und Wirkstoffchemie für die nette Aufnahme.

Meinen guten Freund **Tobias Heintz** möchte ich an dieser Stelle nicht vergessen und mich für die vielen fachlichen und privaten Gespräche bedanken.

Meinen lieben Eltern und meiner Schwester Sabrina danke ich für die Unterstützung während des Studiums und der Promotion, für die vielen ermunternden Worte und den immerwährenden Glauben an mich.

Nicht zuletzt und ganz ausdrücklich möchte ich mich bei meinem lieben Freund **Tobias** bedanken, der mir stets zur Seite stand und ohne dessen Geduld und Unterstützung ich all das hier sicher nicht erreicht hätte.

INHALTSVERZEICHNIS

Abkürzungen	XI
1 Einleitung	1
1.1 Definition der Aktivierung	2
1.2 Diphosphataktivierung an einigen Beispielen	3
1.2.1 Glycosyltransferasen und Glycosidasen	3
1.2.2 Prenyldiphosphatkonvertierende Enzyme	6
1.2.3 Dihydropteroat-Synthase	12
1.2.4 Bedeutung in der chemischen Synthese	13
1.3 Steuerung des Aktivierungsprozesses	14
1.3.1 Hydrolyse allylischer und nichtallylischer Mono- und Diphosphate	14
1.3.2 Enzymabhängige Spaltung aktivierter Diphosphate	18
1.4 Diphosphataktivierung und ATP-Hydrolyse	21
2 Zielstellung der Arbeit	23
3 Methoden	25
3.1 Extraktion von Strukturdaten von Diphosphatbindestellen und daraus abgeleiteter Informationen	25
3.1.1 Extraktion und Aufbereitung der Bindestellen aus der PDB	25
3.2 Clusteringmethoden	30
3.2.1 Distanzmetriken	31
3.2.2 Der Clustering-Algorithmus OPTICS	32
3.3 Die Quantenmechanik	37
3.3.1 Einführung	37
3.3.2 Methoden der Quantenmechanik	38
3.3.3 Basissätze	40
3.3.4 Solvatationsmodelle	41
3.3.5 Übergangszustände von Reaktionen	42
3.3.6 Vergleich mit der Molekülmechanik	43

4	Ergebnisse und Diskussionen.....	45
4.1	Vorstellung des Bindestellen-Datensatzes	45
4.1.1	Liganden und Enzyme	46
4.1.2	Aktivierende Bindestellen.....	48
4.2	Clustering von Diphosphatbindestellen	50
4.2.1	Repräsentation der Bindestellen und Wahl der Parameter.....	50
4.2.2	Vorstellung der Ergebnisse	54
4.2.3	Zusammenfassung und Diskussion.....	58
4.3	Vorstellung der quantenchemischen Vorarbeiten.....	60
4.3.1	Protonierungsformen des Dimethylallyldiphosphats	61
4.3.2	Erzeugung von Ausgangsstrukturen und Etablierung von quantenchemischer Methode und Basissatz.....	62
4.3.3	Vergleich mit Röntgenstrukturen der Cambridge Structural Database.....	65
4.4	Diphosphaterkennung in der Proteinbindetasche.....	67
4.4.1	Die Erkennung des Diphosphats durch Aminosäuren	67
4.4.2	Die Erkennung des Diphosphats durch einfache chemische Gruppen	70
4.4.3	Protonenübertragung betrachteter Brønsted-Säuren	77
4.4.4	Abhängigkeit von Bindungslängen des Diphosphats von Protonierungsgrad und Metallionenkomplexierung.....	80
4.4.5	Zusammenfassung und Diskussion.....	82
4.5	Bedeutung der Ligandenkonformation	84
4.5.1	Torsionswinkel nichtaktivierter Diphosphate.....	84
4.5.2	Torsionswinkel von Prenyldiphosphaten	87
4.5.3	Torsionswinkel von Nukleosiddiphosphat-Zuckern.....	90
4.5.4	Torsionswinkel von Pterindiphosphaten	95
4.5.5	Quantenchemische Untersuchungen von Torsionswinkeln in Prenyldiphos- phaten.....	98
4.5.6	Zusammenfassung und Diskussion.....	104
4.6	Energetische Betrachtungen des Aktivierungsprozesses.....	106
4.6.1	Quantenchemische Beschreibung der Modellreaktion.....	107
4.6.2	Reaktionsenergien des H ₂ O-Angriffes.....	110
4.6.3	Angriff des H ₂ O an das C ¹ -Atom des Dimethylallyldiphosphats.....	113
4.6.4	Angriff des H ₂ O an das Phosphoratom des P _α des Dimethylallyldiphosphats..	117

4.6.5	Zusammenfassung und Diskussion.....	118
5	Zusammenfassung und Ausblick.....	119
6	Summary and Outlook	123
7	Soft- und Hardware, Skripte	127
8	Literaturverzeichnis	129
9	Anhang.....	143
	Curriculum Vitae	161
	Veröffentlichungen.....	162
	Eidesstattliche Erklärung	163

ABKÜRZUNGEN

AO	Atomorbital
aPT	aromatische Prenyltransferase
ATP	Adenosintriphosphat
CoA	Coenzym A
ΔE	Reaktionsenergie
DFT	Dichtefunktionaltheorie
DHPS	Dihydropteroat-Synthase
DMAPP	Dimethylallyldiphosphat
DMSO	Dimethylsulfoxid
E_A	Aktivierungsenergie
EDTA	Ethylendiamintetraacetat (Komplexbildner)
FAD	Flavinadenindinukleotid (oxidierte Form)
FPP	Farnesyldiphosphat
GPP	Geranyldiphosphat
GTP	Guanosintriphosphat
IPP	Isopentenyldiphosphat
IPPS	Isoprenyldiphosphatsynthase
MO	Molekülorbital
NAD(P) ⁺	Nicotinamidadenindinukleotid(-phosphat) (oxidierte Form)
NDP	Nukleosiddiphosphat
NMR	“nuclear magnetic resonance”; Kernspinresonanz
PDB	Protein Data Bank
PFT	Proteinfarnesyltransferase
PGGT I	Proteingeranylgeranyltransferase Typ I
PGGT II	Proteingeranylgeranyltransferase Typ II
PP	Diphosphat
PPT	Proteinprenyltransferase
PÜZ	Pseudoübergangszustand
QM/MM	kombinierte Quanten- und Molekülmechanik (computerchemische Methode)
THF	Tetrahydrofolat
TPP	Thiamindiphosphat
ÜZ	Übergangszustand

1 EINLEITUNG

Phosphate haben eine enorme Bedeutung in biologischen Systemen. So dienen Monophosphate in Zusammenhang mit Proteinkinasen und -phosphatasen der Regulation von Enzymaktivitäten. Eines der wohl bekanntesten Beispiele ist die MAP-Kinase-Kaskade; eine Reihe in Serie geschalteter Kinasen, die an verschiedenen Signaltransduktionswegen beteiligt sind. Triphosphat hingegen ist wesentlicher Bestandteil eines Energielieferanten, der die Grundlage aller Lebensprozesse bildet: Adenosintriphosphat (ATP). Doch auch Diphosphate haben ihren festen Platz in biochemischen Prozessen. Liganden, wie Nicotinamidadenindinukleotid(-phosphat) (NAD(P)⁺), Flavinadenindinukleotid (FAD), Coenzym A (CoA) oder Thiamindiphosphat (TPP) sind Cofaktoren für Oxidations- und Reduktionsprozesse, Überträger von Acylgruppen bzw. unterstützen Decarboxylierungsreaktionen. Trotz ihrer strukturellen Diversität haben diese Moleküle eines gemeinsam: Das Diphosphat bleibt während der Reaktion unberührt. Es dient lediglich der Verankerung in der Proteinbindetasche oder der Verknüpfung weiterer Strukturelemente des Cofaktors. Diese Verbindungen werden daher auch als nichtaktivierte Diphosphatliganden bezeichnet, die zugehörigen Proteine als „nicht-aktivierend“. Jedoch existieren auch Strukturen, bei denen das Diphosphat eine grundlegendere Rolle spielt: Die Spaltung der Bindung zum verbleibenden organischen Rest liefert in diesen Fällen den Antrieb für den Ablauf der Enzymreaktion. Es entsteht ein reaktiver, elektrophiler, organischer Rest, der für die Weiterreaktion zur Verfügung steht. Diphosphatstrukturen, die derartige Prozesse ermöglichen, werden als aktivierte Diphosphate bezeichnet; die beteiligten Enzyme dementsprechend als „aktivierend“. Jedoch ist eine Unterscheidung von aktiviertem oder nichtaktiviertem Liganden nicht in jedem Fall einfach anhand der Struktur möglich, sondern kann auch durch die Proteinumgebung bestimmt werden.

Die vorliegende Arbeit beschäftigt sich mit dem Aktivierungsprozess in Proteinen und insbesondere mit der von der Umgebung abhängigen Aktivierung. Eine Vielzahl von Proteinröntgenstrukturen bildet die Grundlage für die Ermittlung von allgemeinen, weitgehend enzymunabhängigen Kriterien, die für die Steuerung dieses Prozesses verantwortlich sind.

Die folgenden Abschnitte gehen zunächst auf die Definition der Aktivierung und ihre Bedeutung für die chemische Synthese ein. Anschließend werden einige diphosphat-aktivierende Enzymgruppen vorgestellt und literaturbekannte Daten zur Steuerung des Prozesses erläutert. Das Kapitel schließt mit einer Betrachtung der Spaltung der Phosphorsäureanhydridbindung in ATP ab.

1.1 Definition der Aktivierung

Diphosphataktivierung bezeichnet einen Prozess, bei dem aus einem mono- oder disubstituierten Diphosphat, das protoniert oder durch Metalldikationen komplexiert ist, durch Spaltung der R¹O-Bindung, welche organischen und Diphosphatrest miteinander verknüpft, ein reaktives, organisches Elektrophil, sowie ein anorganisches bzw. monosubstituiertes Diphosphat entsteht (Abbildung 1.1). Dieses Elektrophil liegt entweder als Kation vor (S_N1-Mechanismus) oder ist lediglich positiv polarisiert (S_N2-Mechanismus). In beiden Fällen führt (in der schützenden Proteinumgebung) der Angriff eines Nukleophils (Akzeptor) zur Ausbildung der Endprodukte. Auf diese Weise ist es möglich, unter milden Bedingungen sp³-Kohlenstoffzentren zu aktivieren und auf aromatische oder andere elektronenreiche Systeme, wie π-Systeme oder die Atome N, O oder S, zu übertragen. Zusätzlich erlaubt die Enzymumgebung, die eine bestimmte Positionierung der Substrate und Übergangszustände im aktiven Zentrum definiert, eine regioselektive Übertragung auf den Akzeptor.

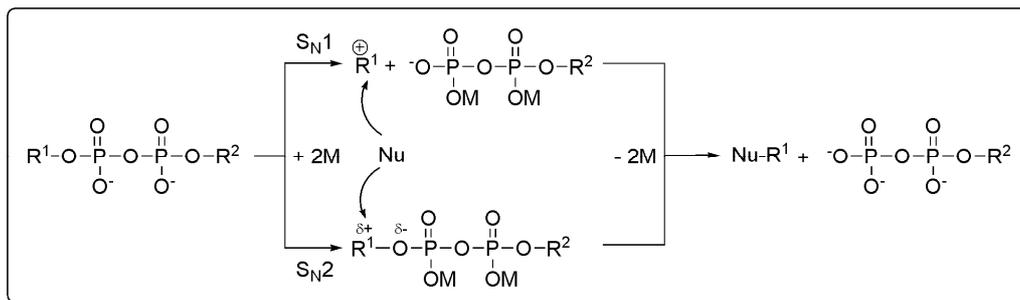


Abbildung 1.1: Schematische Darstellung des Aktivierungsprozesses. R¹ – organischer Rest; R² – organischer Rest (disubstituiertes Diphosphat) oder H (monosubstituiertes Diphosphat); Nu – Nukleophil; M – H oder Lewis-saures Kation (Mg, Mn)

1.2 Diphosphataktivierung an einigen Beispielen

Im Folgenden werden zwei Gruppen aktivierender Enzyme näher beleuchtet. Die Grundstruktur der Diphosphatliganden innerhalb jeder dieser Gruppen ist dabei identisch, jedoch führt eine Variation des Nukleophils zu einer großen Vielfalt organischer Verbindungen. Weiterhin wird ein diphosphataktivierendes Enzym beschrieben, welches keiner der beiden Gruppen zugeordnet werden kann. Besonderes Augenmerk wird auf die Metallionenabhängigkeit der Enzyme, die Erkennung von Metallionen durch das Protein, sowie die Abspaltung des Diphosphates (Aktivierung) gelegt. Der Abschnitt schließt mit der Klärung der Bedeutung der Diphosphataktivierung im Zusammenhang mit der chemischen Synthese ab.

1.2.1 Glycosyltransferasen und Glycosidasen

Glycosyltransferasen^[1] katalysieren die Übertragung von Zuckern von einem aktivierten Donorsubstrat auf einen Akzeptor. Die Zuckerdonoren sind meist Nukleosiddiphosphat-Zucker, wie GDP-Glucose oder UDP-Galactose; jedoch können auch Nukleosidmonophosphat-Zucker (z.B. CMP-N-Acetylneuraminsäure), Lipidphosphat-Zucker (z.B. Dolicholphosphat-Oligosaccharide) oder monosubstituierte Zucker-Diphosphate (z.B. Phosphoribosyldiphosphat) als Glycosyllieferanten dienen. Akzeptoren sind meist andere Zucker, aber auch Lipide, Proteine, Nucleinsäuren, Antibiotika oder andere kleine Moleküle. Dabei wird der Zuckerrest im Allgemeinen auf einen nukleophilen Sauerstoff einer Hydroxylgruppe des Akzeptors übertragen; jedoch ist auch ein Transfer auf Stickstoff-, Schwefel- oder Kohlenstoff-Nukleophile möglich. Die Funktionen solcher Zuckerübertragungen sind vielseitig. Sie dienen beispielsweise der Regulation von Proteinaktivitäten, dem Schutz vor proteolytischem Abbau, der Erhöhung der Löslichkeit oder der Sicherung der korrekten Faltung von Proteinen. Zusätzlich sind Glykosylierungen beteiligt an Zellinteraktionen und Bestandteil von Zellmembranen. Defekte in der Glykosylierung haben daher häufig schwerwiegende Erkrankungen zur Folge.^[2] Im Gegensatz zu Glycosyltransferasen führen Glycosidasen (= Glycosid-Hydrolasen) nicht zur Übertragung von Zuckern, sondern lediglich zu deren Abspaltung von verschiedenen Substraten durch Hydrolyse. Solche Substrate sind oft Di- oder Polysaccharide (z.B. Lactose, Cellulose), aber auch diphosphathaltige Verbindungen (z.B. GDP-Mannose^[3]). Lactoseintoleranz ist beispielsweise auf einen Mangel der Glycosidase Lactase zurückzuführen. Abbildung 1.2 zeigt ein allgemeines Reaktionsschema zur Reaktion von Glycosyltransferasen und Glycosidasen.

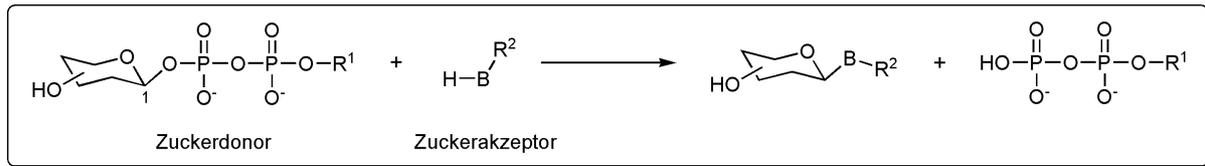


Abbildung 1.2: Allgemeine Darstellung der Reaktion von Glycosyltransferasen und Glycosidasen am Beispiel von Nucleosiddiphosphaten als Zuckerdonoren. B – O, N, S, C_{nukleophil}; R¹ - Nucleosid; R² – H oder organischer Rest (z.B. Zucker, Lipid, Protein, Nucleobase)

Im Gegensatz zu den Glycosidasen, bei denen eine sehr große strukturelle Diversität vorliegt, ist die Zahl der möglichen Faltungen bei Glycosyltransferasen sehr begrenzt.^[1;4;5] Bis heute wurden nur zwei verschiedene Faltungsmuster für alle Strukturen nucleosiddiphosphatabhängiger Glycosyltransferasen beobachtet^[6;7] (GT-A und GT-B); ein dritter Faltungstyp (GT-C) wurde für Enzyme mit Lipidphosphatzucker-Donoren vorhergesagt^[8]. Diese stark begrenzte Anzahl möglicher Faltungen lässt darauf schließen, dass sich die Mehrzahl der Glycosyltransferasen aus einer kleinen Menge von Vorläufersequenzen entwickelt hat, während die Vielzahl der Glycosidase-Strukturen auf eine konvergente Entwicklung hinweist.^[1] Sowohl GT-A- als auch GT-B-Enzyme besitzen Rossmann-Faltungsmuster^[4;7;9] – ein typisches Bindemotiv in mono- und dinucleotidbindenden Proteinen^[10]. Dabei erfolgt in GT-A-Enzymen die Donor-Bindung durch eine einzelne Rossmann-Faltung am N-Terminus des Enzyms, während in GT-B-Enzymen zwei solche Faltungsmuster am C-Terminus gemeinsam die Nucleotidbindetasche ausbilden.^[7] In beiden Fällen handelt es sich im Allgemeinen um lösliche Proteine; es wurden jedoch auch Membrananker für einige Enzyme beobachtet^[7;11]. Nicht alle Proteine mit GT-A- bzw. GT-B-Faltungstyp sind tatsächlich Glycosyltransferasen; beispielsweise besitzen Enzyme der Zucker-1-Phosphat-Pyrophosphorylase/Nucleotidyltransferase-Superfamilie^[12], die für die Biosynthese von Nucleosiddiphosphatzuckern benötigt werden, eine GT-A-Architektur, sowie die UDP-N-Acetylglucosamin-2-Epimerase^[13] eine GT-B-Faltung. GT-A-Enzyme sind stets abhängig von zweiwertigen Metallionen, meist besitzen sie zur Koordinierung ein DXD-Motiv (D – Aspartat, X – beliebige Aminosäure).^[14;15] Glycosyltransferasen vom Faltungstyp GT-B sind dagegen metallionenunabhängig^[1]; eine Beteiligung der Ionen an der Produktfreigabe wurde jedoch postuliert^[16]. Neben diesen beiden Konformationen, deren genaue Struktur durch Röntgenstrukturanalyse nachgewiesen werden konnte, wurde ein dritter Faltungstyp (GT-C) mittels detaillierter Sequenz- und Strukturvergleiche der Glycosyltransferasen der CAZy-Datenbank^[17] vorhergesagt^[8]. Hierbei handelt es sich um integrale Membranproteine mit 8 bis 13 vorhergesagten Transmembranhelices. Weiterhin enthält eine lösliche Loop-Region ein aspartat- bzw. glutamatreiches Motiv (DXD, EXD, DXE, DDX oder DEX); möglicherweise für die Bindung eines Metallions^[18], mechanistische Details der Enzyme sind jedoch unbekannt. Interessanterweise verwenden die meisten Enzyme dieser Gruppe ein Lipid-

phosphat (statt Nucleosiddiphosphat) als aktivierten Donorzucker. Eine erste 3D-Struktur der löslichen C-terminalen Domäne des vorhergesagten GT-C-Enzyms Oligosaccharyltransferase konnte aufgeklärt werden^[18]. Ebenso wurde die Struktur einer peptidoglycansynthetisierenden Glycosyltransferase gelöst^[19;20], die weder GT-A, noch GT-B oder GT-C zugeordnet werden konnte. Das Enzym besitzt eine Faltung, die dem Bacteriophagen-Lysozym ähnlich ist; es kann keine Rossmann-Faltung detektiert werden. Dies ist konsistent mit der Tatsache, dass das Enzym ein Lipidphosphat als Zuckerdonor verwendet und kein Nucleosiddiphosphat. Ebenso gibt es weitere Glycosyltransferase-Familien, die keinem der drei bekannten Faltungsmuster zugeordnet werden können. Sowohl Glycosidasen als auch Glycosyltransferasen lassen sich anhand der Stereochemie der entstehenden Produkte in zwei Gruppen einteilen: Invertierende Enzyme, bei denen die Konfiguration am anomeren Kohlenstoffatom (C¹) während der Reaktion umgekehrt wird und stereochemieerhaltende Enzyme, bei denen die Konfiguration am C¹ beibehalten bleibt. Hierfür sind zwei unterschiedliche Reaktionsmechanismen erforderlich, die jedoch unabhängig von der Faltung des Proteins sind. Insbesondere für Glycosyltransferasen gilt, dass sowohl invertierende GT-A- und GT-B-Enzyme, als auch jeweils die Retentionsvarianten existieren. Der Mechanismus invertierender Glycosidasen und Glycosyltransferasen konnte bereits aufgeklärt werden^[1;5]. In beiden Fällen wird ein Mechanismus mit S_N2-Charakter angenommen, bei dem eine Aminosäureseitenkette des aktiven Zentrums (meist Asp, Glu oder His) als Base wirkt und den Zucker-Akzeptor bzw. das Wassermolekül deprotoniert. Das so entstehende Nucleophil ist dann in der Lage, an das positiv polarisierte anomere Kohlenstoffatom des Donorzuckers anzugreifen. Für eine GT-A-Glycosyltransferase konnte der S_N2-Charakter der Reaktion mittels QM/MM-Berechnungen (kombinierte quanten- und molekülmechanische Methoden) theoretisch gestützt werden, indem gezeigt wurde, dass der nucleophile Angriff des Akzeptors und die Diphosphatabspaltung nahezu simultan erfolgen.^[21] Je nach Faltungstyp der Glycosyltransferasen läuft diese Reaktion mit oder ohne Beteiligung eines zweiwertigen Metallions ab. Für GT-B-Enzyme übernehmen dipolare Helixstrukturen und/oder Wechselwirkungen mit Aminosäureseitenketten die Funktion des Metallions als Stabilisator der entstehenden negativen Ladung am Diphosphat.^[1] Glycosidasen hydrolysieren meist phosphat- und diphosphatfreie Substrate (Di- und Polysaccharide), daher sind sie auch meist unabhängig von Metallionen. Jedoch gibt es auch Ausnahmen. Die GDP-Mannose-Mannosyl-Hydrolase beispielsweise spaltet GDP-Mannose bzw. GDP-Glucose mit Hilfe von Wasser in GDP und den entsprechenden Zucker unter Inversion der Stereochemie am C¹ und ist abhängig von zweiwertigen Metallionen.^[22;23] Diese Inversion erfordert die Spaltung der C¹O-Bindung. Interessanterweise wird das Enzym basierend auf seiner Aminosäuresequenz jedoch der Nudix-Familie zugeordnet^[24] – einer Gruppe von Hydrolasen von Nucleosiddiphosphat-Derivaten, für die ein Angriff an ein Phosphoratom mit resultierender

Spaltung der OP-Bindung typisch ist^[25]. Für stereochemieerhaltende Glycosidasen wird ein „double displacement“-Mechanismus^[26] diskutiert, bei dem der abgespaltene Zucker zunächst auf eine Carboxylgruppe einer Aminosäureseitenkette übertragen (Glycosyl-Enzym-Intermediat) und die so entstehende kovalente Bindung anschließend hydrolysiert wird. Jeder dieser beiden Schritte läuft dabei vermutlich als S_N2 -Mechanismus ab.^[5] Ein S_N1 -Mechanismus ohne Bildung eines Zucker-Enzym-Intermediates wurde durch Phillips postuliert^[27;28], ist aber aufgrund der meist schlechten Abgangsgruppe unwahrscheinlich^[1] – glykosidische Bindungen sind sehr stabil; die Halbwertszeit für die spontane Hydrolyse von Cellulose oder Stärke bei Raumtemperatur und neutralem pH beträgt etwa 5 Mio. Jahre^[29]. Für stereochemieerhaltende Glycosyltransferasen werden hingegen verschiedene Mechanismen in Betracht gezogen. Ein „double displacement“-Mechanismus mit Glycosyl-Enzym-Intermediat wird vorgeschlagen, jedoch sind mögliche zuckerakzeptierende Aminosäureseitenketten nicht eindeutig^[1]. Meist handelt es sich bei geeignet positionierten Gruppen um Rückgrat- oder Seitenkettenamidgruppen. Auch ein S_N1 -Mechanismus wäre für stereochemieerhaltende Glycosyltransferasen aufgrund der guten Hydrogendiphosphat-Abgangsgruppe und unter Abschirmung einer Seite des entstehenden Kations durch Proteinseitenketten möglich.^[1] Jedoch scheint ein S_{Ni} -Mechanismus ($i = \text{„internal“}$) für die meisten dieser Enzyme am wahrscheinlichsten.^[1] Es handelt sich um eine spezialisierte Form des S_N1 -Mechanismus, bei dem nach der Spaltung Zuckerkation und anionische Abgangsgruppe als Ionenpaar vorliegen. Die Abgangsgruppe deprotoniert anschließend den Zuckerakzeptor und positioniert dabei das entstehende Nukleophil in einer Art und Weise zum Zuckerkation, dass der Angriff nur unter Erhalt der Konfiguration am C^1 erfolgen kann.^[1] Die im „double displacement“-Mechanismus vorgeschlagenen zuckerakzeptierenden Amidgruppen dienen dabei vermutlich der Stabilisierung des entstehenden Carbokations ohne aber eine kovalente Bindung auszubilden.

1.2.2 Prenyldiphosphatkonvertierende Enzyme

Prenyldiphosphatkonvertierende Enzyme repräsentieren eine sehr große Gruppe von Proteinen, die beteiligt sind an der Biosynthese der Terpenoide und Meroterpenoide (chimäre Verbindungen mit Terpenoidanteil) – einer Substanzklasse mit insgesamt über 55.000 Verbindungen^[30]. Aufgrund der enormen Diversität und Verbreitung dieser Verbindungen in allen Lebensformen wird ihre Gesamtheit auch als „Terpenom“^[30] bezeichnet. Die biologische als auch pharmazeutische Bedeutung der Substanzen ist vielfältig. So wirken beispielsweise Steroide als Hormone oder Carotinoide als Bestandteil der Photosynthese der Pflanzen und vieler photosynthetisierender Bakterien.^[31] Weiterhin sind Plastochinon^[32] bzw.

Ubichinon^[33-36] am Elektronentransport der Photosynthese bzw. der Atmungskette beteiligt. Viele zyklische Terpeneoide stellen ätherische Öle wie Menthol oder Citronellol.^[37] Eine Übersicht über verschiedene Terpeneoide ist den Literaturstellen ^[38;39] zu entnehmen; die Bedeutung von Pflanzenterpenoiden wird in Referenz ^[40] genauer erläutert. Auch die pharmazeutische Bedeutung dieser Verbindungen ist vielseitig. Insbesondere das Antimalariamedikament Artemisinin^[41] aus den Blättern und Blüten des einjährigen Beifußes *Artemisia annua*, sowie das Chemotherapeutikum Paclitaxel^[42] (Handelsname Taxol[®]) aus der Rinde und den Nadeln verschiedener Eibenarten (insbesondere der pazifischen Eibe *Taxus brevifolia*) sind bekannte Vertreter. Eine Isolierung dieser Terpeneoide aus den genannten Pflanzen ist aufgrund geringer Mengen oder hoher Kosten schwierig^[43]. Auch Totalsynthesen, wie die für Paclitaxel entwickelten (z.B. Literatur ^[44]), sind durch ihre Komplexität und ihre geringen Ausbeuten kommerziell nicht anwendbar^[43]. Daher gibt es immer wieder Ansätze, diese und andere Terpeneoide mittels biotechnologischer Methoden in den entsprechenden Organismen anzureichern oder in anderen, leicht handhabbaren Organismen zu exprimieren.^[43;45-48] Grundlage aller Terpeneoide und Meroterpeneoide bilden die Terpene, die mittels prenyldiphosphatkonvertierender Enzyme biosynthetisiert werden. Heute sind drei Gruppen dieser Enzyme bekannt: Prenyltransferasen, Terpensynthasen und Isomerasen (Abbildung 1.3).

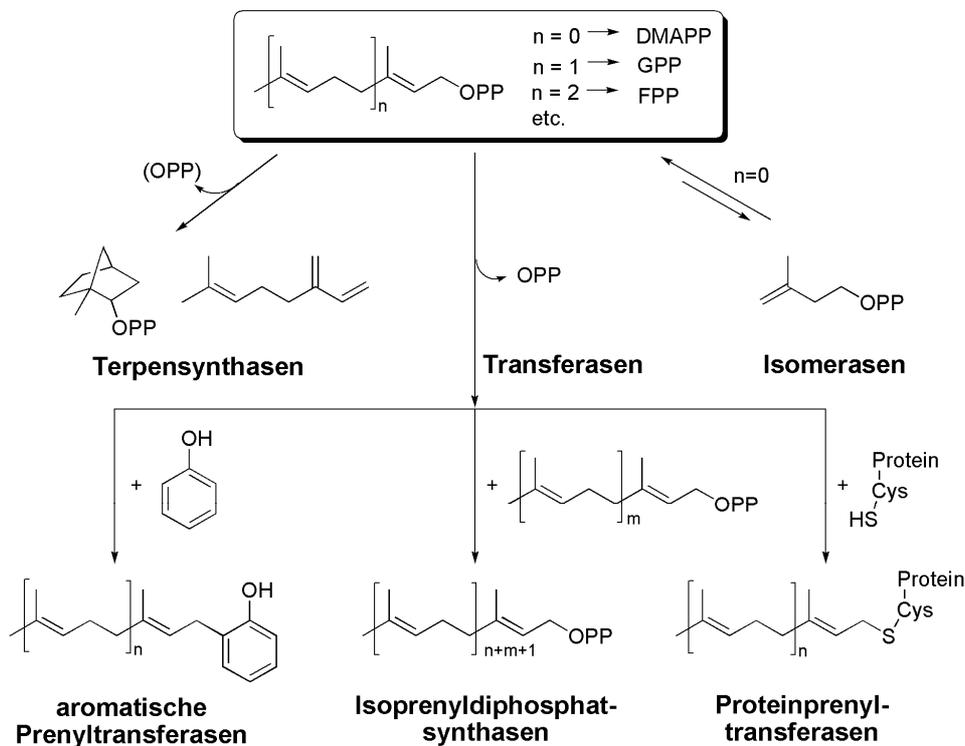


Abbildung 1.3: Darstellung der Funktionsweise prenyldiphosphatkonvertierender Enzyme. Für jede Gruppe ist ein beispielhaftes Produkt und für die Transferasen auch ein Beispiel für ein zweites Substrat angegeben. OPP – Diphosphat

Synthasen und Transferasen scheinen sich von einem gemeinsamen Vorläufer abzuleiten, wie aus den sehr ähnlichen Kristallstrukturen einer Farnesyldiphosphatsynthase^[49] und einer 5-epi-Aristolochensynthase^[50] hergeleitet wurde. Im Gegensatz zur Vielfalt der Verbindungen, die durch die drei Enzymgruppen zugänglich werden, liegt die Grundlage aller Terpenoide in zwei einfachen Vorläufern: Isopentenylidiphosphat (IPP) und Dimethylallyldiphosphat (DMAPP).

Prenyltransferasen überführen Prenylreste, die vom Enzym aus einem Prenyldiphosphat erzeugt werden, auf ein zweites Substrat. Dies kann ein weiteres Prenyldiphosphat sein (Isoprenyldiphosphatsynthasen), ein Aromat (aromatische Prenyltransferasen) oder ein Protein (Proteinprenyltransferasen). Isoprenyldiphosphatsynthasen^[31;51] (IPPS) biosynthetisieren dabei lineare Prenyldiphosphate mit definierter Kettenlänge durch die z.T. mehrstufige Verlängerung von DMAPP mittels IPP. Damit liefern sie die Substrate der Terpensynthasen, der Protein- und aromatischen Prenyltransferasen. Es gibt dabei *cis*- und *trans*-Enzyme – abhängig von der Stereochemie der während der Reaktion gebildeten Doppelbindung. Die *trans*-IPPS sind stets abhängig von Metalldikationen. Für verschiedene Enzyme konnte durch Röntgenkristallographie (siehe Referenz ^[49], sowie PDB^[52]-Eintrag 3ez3) gezeigt werden, dass drei Mg²⁺ komplexiert werden. Dies erfolgt durch insgesamt zwei DDXXD-Motive (D – Aspartat, X – beliebige Aminosäure) auf Helix D und Helix H, die bisher in allen *trans*-IPPSs detektiert werden konnten^[31]. Zusätzlich zeigen die Enzyme untereinander Sequenzähnlichkeiten, die auf einen gemeinsamen Vorläufer schließen lassen^[53;54]. Die *cis*-Enzyme zeigen hingegen keine Sequenzähnlichkeit zu den *trans*-Enzymen.^[55] Es fehlen aspartatreiche Motive, obwohl sie Mg²⁺ für ihre katalytische Aktivität benötigen. Im Gegensatz zu *trans*-IPPS komplexieren sie jedoch nur ein Mg²⁺, das durch ein einzelnes Aspartat gebunden wird, wie für Undecaprenyldiphosphatsynthasen gezeigt werden konnte^[56-59]. Beide Enzymgruppen binden zwei diphosphorylierte Substrate, von denen nur eines während der Reaktion aktiviert wird. Anhand von Röntgenstrukturen (z.B. PDB-Einträge 1yh1^[60] und 1x06^[57]) wird erkennbar, dass stets nur das zu aktivierende Substrat (allylisches Diphosphat) durch das (die) Metallion(en) komplexiert wird. Liegt ein Metallion zwischen beiden Diphosphaten (z.B. Referenz ^[57]), so wird es durch Wassermoleküle, die als Komplexliganden dienen, vom stabilen Diphosphat (IPP) abgeschirmt. Sowohl *cis*- als auch *trans*-IPPS erzeugen carbokationische Intermediate^[31;61]; für beide Enzymgruppen wird demzufolge ein S_N1-Charakter der Reaktion angenommen. Untersuchungen an *trans*-Enzymen mit fluorierten Substratanaloga^[62], sowie die Wirkung kationischer Aza-Analoga als starke Inhibitoren^[63-65] unterstreichen diese Aussage. Sanchez *et al.*^[66] bestätigt die Entstehung eines Übergangszustandes mit carbokationischem Charakter für eine Farnesyldiphosphatsynthase (*trans*-Enzym) mittels kombinierter molekül- und quantenmechanischer Methoden, widerlegt aber die Bildung eines kationischen Intermediates. Die *cis*- und *trans*-Enzyme

führen stets eine Kopf-Schwanz-Verknüpfung („head-to-tail“) durch, wobei „Kopf“ dem Diphosphatrest und „Schwanz“ dem Prenylrest entspricht. Es existieren jedoch auch Enzyme, die in der Lage sind, Kopf-Kopf-Verknüpfungen („head-to-head“) durchzuführen. Beispiele sind an der Bioynthese von Präsqualendiphosphat^[67], einem Intermediat der Sterolbiosynthese, oder Präphytoendiphosphat^[68], einem Intermediat der Carotinoidbiosynthese, beteiligt.

Proteinprenyltransferasen (PPT) katalysieren die S-Prenylierung von Cysteinen in definierten Proteinen. Der Prenylrest ist dabei häufig verantwortlich für die Verankerung des Proteins in Membranen oder dient als Erkennungseinheit.^[69] Ihre immense Bedeutung in lebenswichtigen Prozessen, wie Differenzierung und Apoptose, machen PPT zu einem wichtigen Zielmolekül in der Krebstherapie.^[69] Bisher sind drei Typen von Proteinprenyltransferasen bekannt: Proteinfarnesyltransferase (PFT), Proteingeranylgeranyltransferase Typ I (PGGT I) und Proteingeranylgeranyltransferase Typ II (PGGT II).^[69] PFT und PGGT I überführen eine Farnesylgruppe von Farnesyldiphosphat bzw. eine Geranylgeranylgruppe von Geranylgeranyldiphosphat auf das Cystein eines CXXX-Motivs (C – Cystein; X – beliebige Aminosäure) am C-Terminus des Proteins.^[69] Die drei dem Cystein folgenden Aminosäuren bestimmen, ob das Protein durch eines der beiden Enzyme prenyliert wird und durch welches.^[70;71] Die PGGT II, auch Rab-Geranylgeranyltransferase genannt, prenyliert C-terminale Cysteine in CC- oder CXC-Motiven von Rab-Proteinen zweifach. Dies sind kleine GTPasen, die über den angeknüpften Prenylrest in der Membran verankert werden. Sie sind beteiligt an der Regulation des intrazellulären vesikulären Transports^[72;73]. Proteinprenyltransferasen sind abhängig von Zn^{2+} und meist auch von Mg^{2+} bzw. Mn^{2+} ^[74;75]. Das Zn^{2+} bewirkt eine Aktivierung der Cystein-Thiolgruppe^[75]; die Wechselwirkung mit Zn^{2+} erniedrigt den pK_a -Wert des Thiols und führt zu einem Zn^{2+} -koordinierten Thiolatanion bei physiologischem pH^[76]. Dieses ist somit in der Lage, einen nukleophilen Angriff auf das C¹-Atom des Prenyldiphosphates durchzuführen^[69]. Jedoch findet ohne Mg^{2+} bzw. Mn^{2+} und seine Rolle bei der Aktivierung des Diphosphats meist keine Übertragung des Prenylrestes auf das Protein statt^[74;75]. Für eine PGGT I wurde eine Unabhängigkeit von Mg^{2+} bzw. Mn^{2+} ermittelt (Zugabe von EDTA).^[77] Versuche mit fluoriertem Prenylsubstrat und die Untersuchung des Ausmaßes der Fluorierung auf die Reaktion einer PFT lassen auf eine Reaktion mit S_N2 -Charakter schließen.^[74]

Aromatische Prenyltransferasen^[78] (aPT) katalysieren den Transfer von Prenylresten auf C-, O- oder N-Atome aromatischer Strukturen. Analog zur Proteinprenylierung hat dies oft eine erhöhte Bioaktivität durch verstärkte Affinität der prenylierten Verbindung zu biologischen Membranen zur Folge^[79]. Es sind drei Gruppen aromatischer Prenyltransferasen bekannt, die sich hauptsächlich anhand ihrer unähnlichen Proteinsequenzen unterscheiden. Integrale, membrangebundene aPTs enthalten zwei aspartathaltige Motive (D(R/F)XX(D/G) und

DXXXD), die zur Bindung von Metalldikationen benötigt werden. Sie katalysieren ausschließlich C-Prenylierungen, d.h. der Prenylrest wird an ein Kohlenstoffatom des Aromaten angeknüpft. Bisher existiert keine Röntgenstruktur dieser Gruppe aromatischer Prenyltransferasen, jedoch konnten Modelle erzeugt (UbiA^[80;81], AuaA^[51], LePGT1^[82]) und teilweise mittels Mutationsstudien (UbiA^[81], LePGT1^[82]) validiert werden. Basierend auf theoretischen, semiempirischen Berechnungen wird für das Enzym UbiA dieser Gruppe ein S_N1-Typ der Reaktion angenommen.^[80] Die zweite Gruppe setzt sich ausschließlich aus löslichen Enzymen zusammen.^[83] Sie enthalten keine aspartatreichen Motive, sind teilweise aber Mg²⁺-abhängig^[84]. Die Aufklärung einer Röntgenstruktur^[84] brachte eine neue „Barrel“-Struktur hervor: das Prenyltransferase-Fass (PT-Fass), was – aufgrund seiner außergewöhnlichen Gestalt – den Enzymen dieser Gruppe den Namen ABBA-Prenyltransferasen einbrachte. Aufgrund großer Sequenzähnlichkeit wird diese Faltung für alle Enzyme dieser Gruppe angenommen.^[83] Neben regulären Prenylierungen (Verknüpfung des C¹-Atoms des Prenylrestes mit dem Aromaten), sind diese Enzyme zum Teil auch in der Lage, reverse Prenylierungen (Verknüpfung des C³-Atoms des Prenylrestes mit dem Aromaten) vorzunehmen^[85]. Die Enzyme machen sowohl C- als auch O-Prenylierungen.^[84;85] Die dritte Gruppe aromatischer Prenyltransferasen bilden lösliche Indolprenyltransferasen aus Pilzen.^[86;87] Sie enthalten ebenfalls keine aspartatreichen Motive und keines der bisher charakterisierten Enzyme benötigt Metalldikationen für die katalytische Aktivität. Sie katalysieren C- und N-Prenylierungen, die sowohl regulär als auch revers erfolgen können. Kürzlich konnte eine erste Röntgenstruktur für ein Enzym dieser Gruppe ermittelt werden, die eine dem PT-Fass ähnliche Struktur zeigt.^[88] Ebenso konnte experimentell mittels ¹⁸O-markiertem Prenylsubstrat, für das eine reversible Diphosphatabspaltung im Enzym festgestellt wurde, ein S_N1-Mechanismus für ein Enzym dieser Gruppe nachgewiesen werden.^[89] Zusätzlich gibt es weitere Enzyme, die anhand ihrer Sequenz keiner der oben genannten Gruppen zugeordnet werden können (z.B. LtxC^[90]). Dies lässt auf die Existenz weiterer Klassen aromatischer Prenyltransferasen schließen.

Terpensynthasen bilden neben den Prenyltransferasen die zweite große Gruppe prenyldiphosphatkonvertierender Enzyme. Sie können zum Einen azyklische Produkte synthetisieren, wie die Myrcen- oder Ocimensynthase^[91;92], meist sind aber zyklische Terpene das Resultat der Reaktion. In diesem Fall werden die entsprechenden Enzyme als Terpenzyklen bezeichnet. Sehr häufig synthetisieren Terpensynthasen eine Reihe von Produkten^[91], dennoch sind die Reaktionen oft hochgradig stereospezifisch^[93]. Beispielsweise bildet die Lanosterol-Synthase nur genau eins von 128 möglichen Stereoisomeren. Substrate der Terpenzyklen sind Geranyl- (C₁₀), Farnesyl- (C₁₅), Geranylgeranyl- (C₂₀), Geranyl-farnesyldiphosphat (C₂₅) oder Squalen (C₃₀). Da Squalen diphosphatfrei ist und die squalen-

zyklisierenden Enzyme demnach nicht prenyldiphosphatkonvertierend sind, werden sie in diesem Zusammenhang nicht näher beschrieben. Durch das Fehlen des Diphosphats ist für diese Enzyme kein Metallion notwendig, dennoch besitzen sie ein DXDD-Motiv, wobei das zentrale Aspartat als Protonendonator die Zyklisierung anstößt; Details zum Mechanismus sind den Literaturstellen ^[93;94] zu entnehmen. Terpenzyklen, die diphosphathaltige Prenylverbindungen als Substrate verwenden, komplexieren generell drei Metallionen, die die Abspaltung des Diphosphatrestes vom Substrat unterstützen und somit den Zyklisierungsprozess einleiten.^[93] Die Bindung der Metallionen erfolgt durch zwei aspartathaltige Motive, von denen eines (DDXXD/E-Motiv) topologisch identisch zum DDXXD-Motiv auf Helix D der *trans*-IPPSs ist. Das zweite Motiv auf Helix H jedoch ist – im Gegensatz zum entsprechenden Motiv der *trans*-IPPS – nicht aspartatreich; vereinfacht wird es als NSE/DTE-Motiv bezeichnet^[95]. Kürzliche Untersuchungen^[96] erklären die Tatsache, dass das zentrale Tyrosin bzw. Serin durch Glycin aber nicht durch Alanin ersetzbar ist. Ebenso sind auch die nichtzyklisierenden Terpensynthasen metallionenabhängig und besitzen das stark konservierte DDXXD-Motiv.^[91;92] Die Zyklisierung erfolgt direkt aus dem Substrat, aus einem Isomer^[93;97], bei dem das Diphosphat von C¹ auf C³ übertragen wurde (z.B. bei Monoterpenzyklen oder der Trichodiensynthese) oder anderweitig umgelagerter, aber diphosphathaltiger Substrate^[93] (z.B. bei der Abietadiensynthese) durch einen intramolekularen Angriff von π -Elektronen. Ob diese Zyklisierungsreaktion schrittweise unter Bildung eines allylischen Kations vor Angriff der π -Elektronen (S_N1) oder simultan abläuft (S_N2) konnte bisher nicht eindeutig geklärt werden. Da der Schritt intramolekular ist, führt er jedoch letztendlich generell zur Bildung eines zyklischen Kations.^[93] Weitere Ringschlüsse, Deprotonierung, Hydroxylierung, Diphosphorylierung o.ä. schließen die Reaktion ab.

Die dritte Gruppe prenyldiphosphatkonvertierender Enzyme stellen die Isomerasen dar. Sie katalysieren die reversible Reaktion der Isomerisierung von IPP zum allylischen Diphosphat DMAPP; das Gleichgewicht der Reaktion liegt dabei auf der Seite des DMAPP.^[98] Dabei kommt es nicht zu einer Abspaltung des Diphosphats, d.h. es erfolgt keine Diphosphataktivierung. In höheren Pflanzen, einigen Bakterien und Algen sind sowohl IPP als auch DMAPP Produkte des Methylerythritolphosphat-Reaktionsweges^[99] (MEP). Für diese Organismen sind die Isomerasen keine essentiellen Enzyme. Jedoch gibt es Organismen, die diesen Reaktionsweg nicht besitzen und stattdessen über den Mevalonat-Weg^[99] ausschließlich IPP gewinnen können. Hierzu gehören die meisten Eucaria, Archaea und Pilze. Für diese Organismen sind Isomerasen essentiell, um durch die Synthese von DMAPP zur Vielfalt der Terpeneide gelangen zu können. Es sind zwei nichthomologe Gruppen von IPP-Isomerasen bekannt ^[100;101]. Isomerasen vom Typ I sind in Eucaria und Bacteria zu finden und sind abhängig von Metalldikationen.^[98] Typ II-Isomerasen liegen in Archaea und Bacteria vor

und benötigen neben dikationischen Metallionen auch reduziertes Flavin für ihre katalytische Aktivität.^[98] Das Metallion ist dabei nicht an der Komplexbildung des Diphosphats beteiligt^[98]. Es erhöht durch die Komplexbildung einer Glutaminsäure hingegen deren Azidität, sodass es durch Protonierung des IPP-Substrates zur Einleitung der Isomerisierung kommt.^[98] Eine nachfolgende Deprotonierung des so entstandenen Kations resultiert in der Bildung des Produktes DMAPP.

1.2.3 Dihydropteroat-Synthase

Ein weiteres diphosphataktivierendes Enzym, das insbesondere in der pharmazeutischen Anwendung lange Zeit von großem Interesse war, ist die Dihydropteroat-Synthase (DHPS). Das Enzym katalysiert die Reaktion von 6-Hydroxymethyl-7,8-Dihydropterindiphosphat mit *p*-Aminobenzoat unter Bildung von 7,8-Dihydropteroat und anorganischem Diphosphat (Abbildung 1.4) – ein Schlüsselschritt in der bakteriellen Folat-Biosynthese (Vitamin B9). Es ist Mg²⁺-abhängig und Baca *et al.*^[102], sowie Babaoglu *et al.*^[103] schlagen einen S_N2-Mechanismus vor, jedoch sind Details des Reaktionsmechanismus noch weitgehend ungeklärt^[103;104].

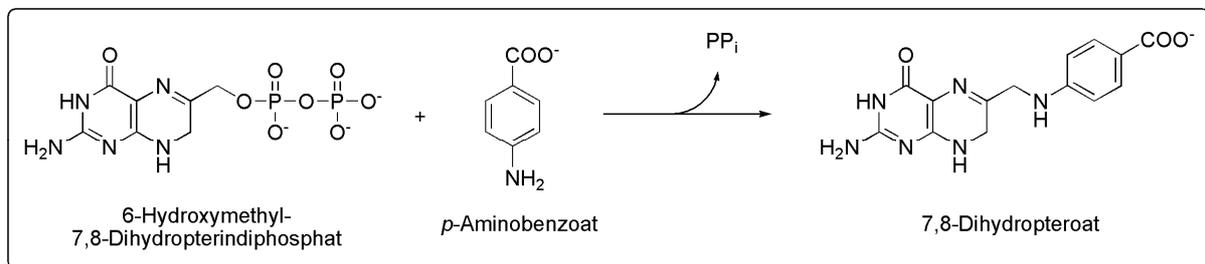


Abbildung 1.4: Reaktionsgleichung der Dihydropteroat-Synthase. PP_i – anorganisches Diphosphat.

Folat dient als Vorstufe des Cofaktors Tetrahydrofolat (THF), der als C₁-Gruppenüberträger am Metabolismus von Aminosäuren und Purinbasen beteiligt und somit essentiell für die Biosynthese von Proteinen und DNA ist. Eine Störung des Folat-Stoffwechsels führt damit zur Inhibierung des Zellwachstums und letztendlich zum Zelltod^[105]. Nur einige Prokaryonten, niedere Eukaryonten, sowie Pflanzen besitzen alle Enzyme für eine vollständige Folat-Biosynthese.^[103] Alle anderen Organismen – auch der Mensch – müssen Folat oder reduzierte Derivate mit der Nahrung aufnehmen und haben dafür hochaffine Transportsysteme entwickelt^[106]. Diese Tatsachen machten die Dihydropteroat-Synthase für fast 70 Jahre zu einem interessanten, pharmazeutischen Zielmolekül.^[103] Mit der Entdeckung der

antibiotischen Wirkung^[107] des Azofarbstoffes Sulfamidochrysoidin (Handelname Prontosil®), der zur Rotfärbung von Textilien verwendet wurde, kam es zur Entwicklung verschiedener Sulfonamidderivate. Nachfolgend konnte gezeigt werden, dass ihre Wirkung auf der Besetzung der Bindestelle des aromatischen Substrats *p*-Aminobenzoat der DHPS beruht^[108-111]. Noch heute werden Sulfonamidantibiotika in Kombination mit dem Dihydrofolat-Reduktase-Inhibitor Trimethoprim bei der Behandlung von Infektionen mit *Pneumocystis carinii* bei HIV- und Krebspatienten eingesetzt^[112], jedoch schränkt die Entwicklung von Resistenzen^[113;114] (verursacht durch Mutationen in der DHPS) ihren derzeitigen klinischen Nutzen ein.

1.2.4 Bedeutung in der chemischen Synthese

Die Erzeugung aktivierter Kohlenstoffzentren und die regioselektive Übertragung auf nukleophile Reaktionspartner ist auch eine Problemstellung der chemischen Synthese.^[115] Jedoch sind effiziente Ansätze bis heute auf Spezialfälle beschränkt.^[115] Diese erfordern oft extreme Reaktionsbedingungen, wie hohe Temperaturen (z.B. Shilov-System^[116]) oder starke Säuren (z.B. Methode nach Bergman und Ellman^[117]). Aufgrund der häufigen Verwendung von Metallkatalysatoren ist auch das Vorhandensein spezieller funktioneller Gruppen zu ihrer Komplexierung notwendig. Zusätzlich ist die Regioselektivität solcher Aktivierungsreaktionen meist nur schwer kontrollierbar, dies gilt insbesondere für intermolekulare Reaktionen.^[115] Häufig sind aufwändige Schützungen und Entschützungen erforderlich. Für Enzymreaktionen ist die Regioselektivität hingegen kein Problem, wie die Ausführungen der vorangehenden drei Abschnitte zeigen. Auch oft gute Stereoselektivitäten und milde Reaktionsbedingungen machen Enzyme in Verbindung mit der organischen Synthese zu interessanten Objekten. Beispielweise wurde in diesem Zusammenhang das relativ breite Substratspektrum der C-C-verknüpfenden Transferase UbiA, die irreversible Alkylierungen vom Friedel-Crafts-Typ katalysiert, detailliert untersucht.^[118;119] Die Synthese von Kuhistanol A, das antibakterielle Aktivität gegen methicillinsensitiven und methicillinresistenten *Staphylococcus aureus* zeigt^[120], konnte ebenfalls durch dieses Enzym erreicht werden^[78;121]. Detaillierte Kenntnisse über den Aktivierungsprozess von Diphosphaten in Enzymen könnten helfen, vereinfachte, enzymähnliche Strukturen zu konstruieren, die auch unter milden Bedingungen eine gezielte C-C-Bindungsknüpfung im Rahmen einer chemischen Synthese vornehmen können.

1.3 Steuerung des Aktivierungsprozesses

Der folgende Abschnitt geht zunächst auf Untersuchungen zur Steuerung des Aktivierungsprozesses in Zusammenhang mit der Hydrolyse allylischer und nichtallylischer Mono- und Diphosphate (enzymunabhängige Systeme) ein. Nachfolgend wird behandelt, inwiefern die so gewonnenen Erkenntnisse auf enzymatische Systeme übertragen werden können.

1.3.1 Hydrolyse allylischer und nichtallylischer Mono- und Diphosphate

Bei der Hydrolyse allylischer und nichtallylischer Mono- und Diphosphate kommt es meist zur Bildung des Alkohols und eines anorganischen Mono- bzw. Diphosphats. Beispielhaft gibt Abbildung 1.5 die Hydrolyse eines allylischen Diphosphats an.

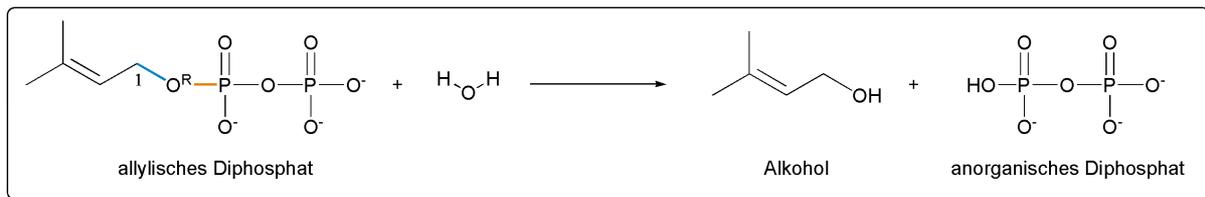


Abbildung 1.5: Reaktionsgleichung der Hydrolyse eines allylischen Diphosphats am Beispiel von Dimethylallyldiphosphat (DMAPP). Denkbar ist – im Zusammenhang mit einer Diphosphat-abspaltung – sowohl die Spaltung der C¹O^R-Bindung (blau) als auch der O^RP-Bindung (orange).

Sowohl für allylische als auch nichtallylische Mono- und Diphosphate konnte eine pH-Abhängigkeit festgestellt werden.^[122-127] Im Allgemeinen gilt: Je niedriger der pH-Wert („saurer“), desto höher ist die Hydrolysegeschwindigkeit. Die verschiedenen Hydrolysesubstrate unterscheiden sich dabei in erster Linie durch die jeweils gespaltene Bindung: Es kommt entweder zur Spaltung der C¹O^R- oder O^RP-Bindung oder bei Diphosphaten in einigen Fällen auch zur Monophosphatabspaltung. Die nichtallylischen Monophosphate Methylphosphat^[122], sowie α -D-Glucose-1-Phosphat^[123] wurden detailliert mittels kinetischer Methoden und Isotopenexperimenten (H₂¹⁸O) untersucht. Dabei wurde festgestellt, dass bestimmte Protonierungsformen des Monophosphats stets durch Spaltung derselben Bindung hydrolysiert werden: Für Monoanionen erfolgt eine Spaltung der O^RP-Bindung, für die neutrale Form hingegen hydrolysiert die C¹O^R-Bindung. Die Reaktivität der zweifach negativ geladenen Form hingegen ist in Bezug auf die Hydrolyse sehr gering, was niedrige Hydrolysegeschwindigkeiten bei basischen pH-Werten erklärt. Weiterhin wurde festgestellt, dass

die Reaktivität des Monoanions relativ unabhängig von der Natur des organischen Substituenten ist. Dies scheint umgekehrt nicht für die Hydrolyseaktivität der neutralen Form zu gelten: Im Falle des Methylphosphats ist sie geringer als die des Monoanions; für α -D-Glucose-1-phosphat dagegen höher. Diese Differenzierung kann möglicherweise durch die Resonanzstabilisierung des durch C^1O^R -Hydrolyse entstehenden Zuckerkations erklärt werden (siehe Abbildung 1.6).

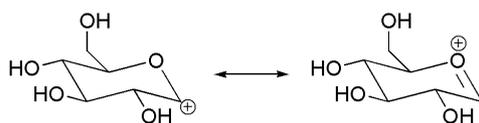


Abbildung 1.6: Mesomere Grenzstrukturen des durch C^1O^R -Spaltung aus α -D-Glucose-1-phosphat entstehenden Kations.

Darauf weist auch die Tatsache hin, dass im Rahmen der Säurekatalyse (Protonierung des O^R) für Methylphosphat sowohl eine Spaltung der $O^R P$ - (zu 35%) als auch der C^1O^R -Bindung (zu 65%) zu beobachten ist; für α -D-Glucose-1-phosphat hingegen findet fast ausschließlich eine Spaltung der C^1O^R -Bindung statt. Auf die verschiedenen Reaktivitäten der unterschiedlichen Protonierungsformen ist auch zurückzuführen, dass der Zusammenhang von pH-Wert und Hydrolysegeschwindigkeit nicht linear sein muss. Für Methylphosphat kommt es zu einem Einbruch der Hydrolysegeschwindigkeit bei ca. pH 0,5; in stärker sauren Umgebungen aufgrund der zunehmenden Säurekatalyse allerdings wieder zu einer Steigerung. Für α -D-Glucose-1-phosphat kommt es bei pH < 5 zu einer deutlichen Steigerung der Raten im Vergleich zur Hydrolyse bei basischeren pH-Werten. Es bleibt anzumerken, dass sowohl bei den Untersuchungen zu Methylphosphat als auch α -D-Glucose-1-phosphat einwertige Metallionen Bestandteil der verwendeten Puffer waren, deren Wirkung auf die Hydrolysegeschwindigkeit nicht auszuschließen ist. Zusätzlich wurden fast ausschließlich Temperaturen verwendet, die weit über der Raumtemperatur liegen. Als Repräsentant für nichtallylische Diphosphate wurde Isopentenylidiphosphat in natriumhaltigem Puffer untersucht.^[125] Dabei wurde festgestellt, dass keine Abspaltung des Diphosphats, sondern eine Hydrolyse der zentralen POP-Bindung mit Abspaltung eines Monophosphats stattfindet. Beim Vergleich der Hydrolysegeschwindigkeit bei 100°C für pH 1,02 gegen pH 4,17 liegt ebenfalls eine Steigerung im sauren Bereich vor. Allylische Monophosphate^[125], wie Dimethylallylphosphat oder Geranylphosphat, zeigen ebenfalls eine klare Abhängigkeit der Hydrolysegeschwindigkeit vom pH-Wert im Bereich 1 bis 8 (Temperatur zwischen 40°C und 70°C). Der Autor schließt anhand von Vergleichen mit der Hydrolyse von Prenylchloriden auf die Spaltung der C^1O^R -Bindung statt der $O^R P$ -Bindung. Für allylische Diphosphate wird durch mehrere Autoren^[125-131] eine spontane Hydrolyse in Säuren beschrieben, wobei die Hydrolysege-

schwindigkeit ebenfalls mit fallendem pH steigt. Bei genauem Studium der entsprechenden Literaturstellen fällt jedoch auf, dass – bis auf eine Ausnahme (Vial *et al.*^[127]) – in allen Fällen Metallionen (vorwiegend Na oder K) beteiligt sind, meist als Bestandteil der verwendeten Pufferlösung. Zusätzlich erfolgten einige der Messungen unter erhöhter Temperatur. Auch Vial *et al.*^[127] zeigen eine spontane Hydrolyse von Geranyldiphosphat bei 40 °C und pH 7 ohne Metallionen – wenn auch mit einer sehr geringen Geschwindigkeitskonstante von $3,4 \cdot 10^{-7} \text{ s}^{-1}$. Jedoch wird durch die Autoren ebenso eine vollständige Einstellung der Hydrolyse bei Komplexierung zugesetzter Metallionen mit EDTA beschrieben. Dies lässt auf Spuren von Metallionen in der vorgeblich metallionenfreien Messung schließen. Unveröffentlichte Arbeiten^[124], bei denen die pH-Abhängigkeit der Hydrolyse der Prenylverbindung Dimethylallyldiphosphat (DMAPP) unter Ausschluss von metallionenhaltigen Puffern untersucht wurde, machen zudem deutlich, dass bei Raumtemperatur (23 °C) bei keinem der eingestellten pH-Werte 1, 3, 5, 7 und 9 eine spontane Spaltung erfolgt (in Tris/HCl, 24 h Inkubationszeit, Einstellung des pH-Wertes mit HCl, jeweils 0,24 M DMAPP). Erst nach Erhitzen (54 °C, 72 h, Tris/HCl) kann eine Hydrolyse beobachtet werden, die auch pH-abhängig ist: pH 1 führt mit 87% zur maximalen Umsetzung des DMAPP, während bei pH 9 mit 58% der geringste Zerfall gemessen werden konnte. Diese pH-Abhängigkeit könnte auf eine verstärkte Säurekatalyse^[129] (Aktivierung des DMAPP durch Protonierung von O^R) oder auf einen durch die erhöhte Protonierung des Diphosphats verursachten Effekt (z.B. leichtere Abspaltung des Diphosphats von allylischen Verbindungen durch vollständige Protonierung^[130]) zurückgeführt werden. Weiterhin unveröffentlicht^[124] sind Messungen zur DMAPP-Hydrolyse in Mg-haltigen Systemen (23 °C, 6 h Inkubationszeit, Einstellung des pH-Wertes mit HCl, jeweils 0,25 M DMAPP und MgCl₂, 10% DMSO zur Verbesserung der Löslichkeit). Eine Ausfällung von Metall-Phosphat-Komplexen war innerhalb des angegebenen Zeitraums visuell nicht erkennbar. Es wurde bewusst auf (Tris/HCl-) Puffer verzichtet, um eine Beeinflussung der Hydrolysegeschwindigkeit durch Komplexierung der Metallionen auszuschließen. Im Gegensatz zu den metallionenfreien Messungen zeigte sich eine Umsetzung zu Produkten bereits bei Raumtemperatur. Bei pH 3 war eine 54%ige Umsetzung feststellbar, die sich für pH 1,5 und pH 1 auf 100% erhöhte. Für die außerdem gemessenen pH-Werte 5 und 7 war jeweils nur eine 1%ige Umsetzung zu beobachten. Die Bestimmung dieser prozentualen Umsetzung erfolgte sowohl bei den Messungen mit als auch ohne Metallionen mittels NMR, wobei jeweils der Anteil an entstehendem anorganischen Diphosphat bestimmt wurde. Eine erneute Messung des pH-Wertes nach jeder NMR-Analyse sicherte die Konstanz des jeweils eingestellten Wertes, was insbesondere bei Ausschluss von Pufferlösungen notwendig wird. Der beobachtete Effekt der Beschleunigung der Hydrolyse allylischer Diphosphate durch Metallionen wird ebenfalls durch andere Autoren^[127;132-134] beschrieben. Dabei konnten sowohl Brems und Rilling^[132], als auch Vial *et al.*^[127] zeigen,

dass Mn^{2+} die Hydrolyse um Faktor 5 stärker beschleunigt als Mg^{2+} . Beide Publikationen beschreiben zudem ein Verhältnis von 2:1 für Metalldikation und Prenylverbindung im Metallion-Prenyldiphosphat-Komplex, um die maximale Hydrolysegeschwindigkeit zu erzielen. Vial *et al.* merken jedoch an, dass vermutlich auch der (1:1)-Komplex einen Beitrag zur Hydrolyse leistet, da bei geringen Metallionenkonzentrationen eine höhere Reaktionsgeschwindigkeit zu beobachten ist, als anhand der Konzentration der dann vorliegenden Prenyldiphosphat-(Mn^{2+})₂-Komplexe zu erwarten wäre. Dies kann mit den oben beschriebenen Analysen von J. Kufka vereinbart werden, durch die gezeigt wurde, dass auch bei (1:1)-Verhältnissen vollständige Umsetzungen erreicht werden können. Weiterhin gibt die Literatur an, dass die Hydrolyse von allylischen Diphosphaten unter Spaltung der C¹O^R-Bindung^[125;127-129;135] nach S_N1-Mechanismus^[126;129] abläuft. Es kommt nicht zu einer O^RP-Spaltung und nicht zu einer Abspaltung von Monophosphat. Monophosphatabspaltungen konnten ebenso durch J. Kufka^[124] weder bei der Anwesenheit noch bei Abwesenheit von Metallionen beobachtet werden. Jedoch erfolgen in den angegebenen Literaturstellen keine

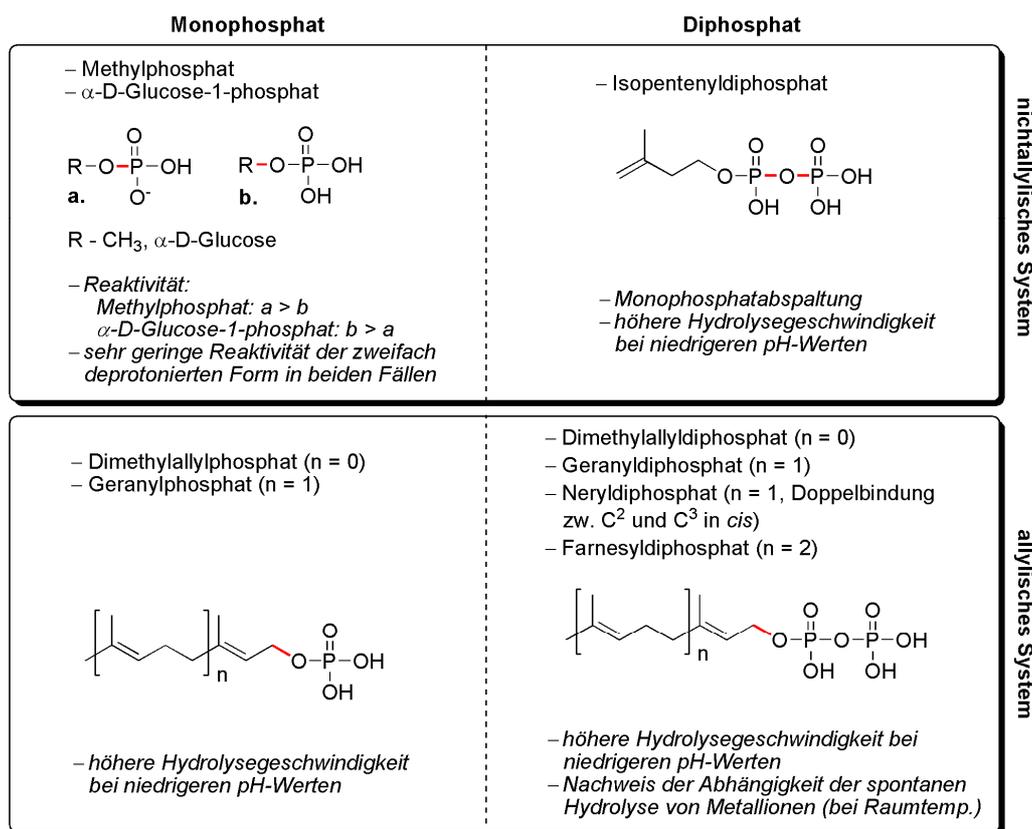


Abbildung 1.7: Zusammenfassende Darstellung der Resultate verschiedener Literaturstellen zur Hydrolyse allylischer und nichtallylischer Mono- und Diphosphatverbindungen. Es sind jeweils die in der Literatur betrachteten Verbindungen namentlich angegeben, sowie ihre Struktur dargestellt. Die gespaltene Bindung ist jeweils in rot gezeigt. In allen Fällen wurde die Hydrolyse bei Präsenz von Metallionen oder unter erhöhten Temperaturen (> Raumtemperatur) beobachtet.

Isotopenexperimente, beispielsweise mit H_2^{18}O , durch die eindeutig zwischen der Spaltung der $\text{C}^1\text{O}^{\text{R}}$ - bzw. $\text{O}^{\text{R}}\text{P}$ -Bindung unterschieden werden könnte. Einen Hinweis darauf, dass für Dimethylallyldiphosphat (DMAPP) bei pH 7 keine $\text{O}^{\text{R}}\text{P}$ -, sondern ausschließlich eine $\text{C}^1\text{O}^{\text{R}}$ -Spaltung stattfindet, liefern Agranoff *et al.*^[136]: Sie konnten nachweisen, dass bei der Hydrolyse neben dem Dimethylallylalkohol (3-Methylbut-2-enol) auch 2-Methylbut-3-en-2-ol zu gleichen Teilen entsteht. Die Ursache hierfür liegt in der Resonanzstabilisierung des durch $\text{C}^1\text{O}^{\text{R}}$ -Spaltung entstehenden Kations. Bei zusätzlicher $\text{O}^{\text{R}}\text{P}$ -Bindungsspaltung wäre ein größerer Anteil des entstehenden Dimethylallylalkohols zu erwarten. Interessant ist, dass Vial *et al.*^[127] bei Vorhandensein höherer Konzentrationen zweiwertiger Metallionen (Mn^{2+}) eine verstärkte Bildung zyklischer Produkte statt linearer Prenylalkohole beobachteten, wenn sie die Werte mit den Hydrolyseprodukten, die bei sehr geringen Konzentrationen von Mn^{2+} entstehen, verglichen. Abbildung 1.7 fasst die wesentlichen Charakteristika der Hydrolyse allylischer und nichtallylischer Mono- und Diphosphate, die in diesem Abschnitt angesprochen wurden, noch einmal zusammen.

1.3.2 Enzymabhängige Spaltung aktivierter Diphosphate

Die während der Hydrolyse allylischer und nichtallylischer Mono- und Diphosphate vorliegende Abhängigkeit vom pH-Wert der Umgebung spiegelt sich auch in diphosphat-aktivierenden Enzymen wider: Die Bindetaschen enthalten im Bereich der Diphosphatbindung häufig Aminosäuren wie Arginin, Lysin, Histidin oder Tyrosin (z.B. ^[1;57;61;66;103;137-140]). Diese können durch Protonierung des Diphosphats eine saure Umgebung simulieren. Ebenso sind auch Metalldikationen (meist Mg^{2+} oder Mn^{2+}) oft in solchen Bindestellen zu finden. Häufig bilden sie vor der Enzymbindung einen Komplex mit dem zu aktivierenden Diphosphatsubstrat und erhöhen damit die Affinität des Substrats zum Enzym^[57;141]. Zum Teil ist die Ausbildung des Enzym-Substrat-Komplexes ohne Metallion ebenfalls möglich, wie King *et al.*^[141] für eine Prenyltransferase zeigen konnten. Die ausschließliche Bindung des Metallions an den Enzym-Substrat- oder den Enzym-Produkt-Komplex weist hier auf einen metallionfreien, nativen Zustand des Enzyms hin und damit auf eine Freigabe des Metallions gemeinsam mit dem Diphosphatprodukt. Guo *et al.*^[57] beschreiben für eine Undecaprenyldiphosphatsynthase (*cis*-Prenyltransferase) die Notwendigkeit des Metallions für die Ausbildung des Enzym-Substrat-Komplexes unter physiologischen Substratkonzentrationen. Abbildung 1.8 zeigt schematisch den Ablauf der Reaktion. Das Enzym katalysiert die aufeinanderfolgende Kondensierung von Farnesyldiphosphat (FPP; allylisches Diphosphat; aktiviertes Substrat) mit 8 Isopentenylidiphosphatmolekülen (IPP; nichtallylisches Diphosphat; nichtaktiviertes Substrat). Dies bedeutet eine schrittweise Verlängerung des FPP bis zu Un-

decaprenyldiphosphat, wobei die entstehenden Prenyldiphosphate jeweils als aktiviertes Substrat wirken. Interessanterweise erfolgt die Bindung des Metallions an das nichtaktivierte Substrat IPP bevor dieser Komplex den Kontakt mit der Enzymbindetasche aufnimmt, während das zu aktivierende Substrat (Prenyldiphosphat) metallionenunabhängig bindet. Erst in der Proteinbindetasche wird das Metallion auf das Prenyldiphosphat übertragen, um die Enzymreaktion einzuleiten. Der Diphosphat-Metallion-Komplex bleibt auch nach Abspaltung des Diphosphats im Zuge der Aktivierung erhalten und verlässt das Enzym. Der nächste Kondensierungsschritt wird dann durch die erneute Bindung eines IPP-Metallion-Komplexes eingeleitet. Das Metallion bindet nicht an den Komplex aus Enzym und allylischem Diphosphat; die geringe Affinität zum substratfreien Enzym kann nur durch sehr hohe Metallionenkonzentrationen überwunden werden.

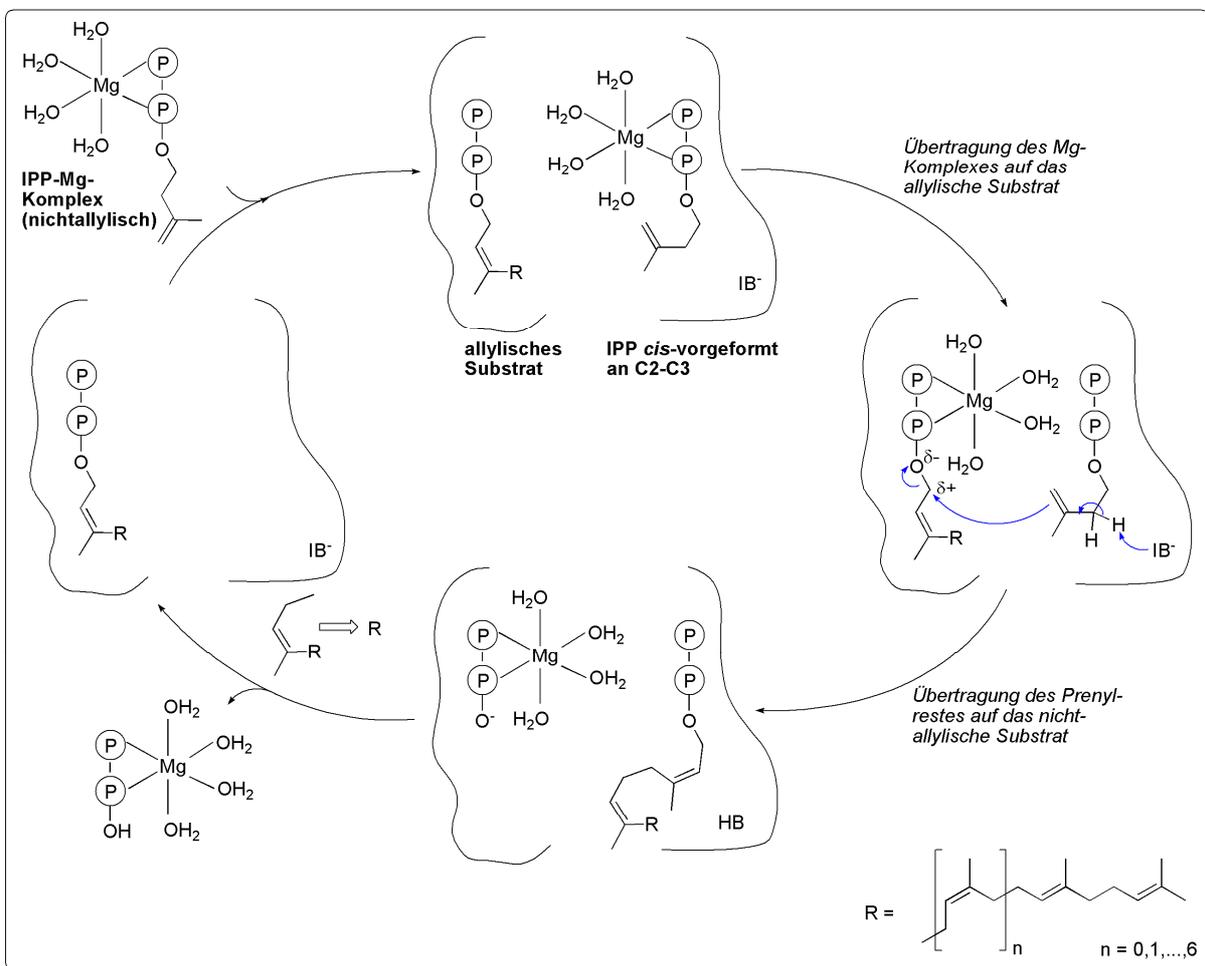


Abbildung 1.8: Vereinfachte, schematische Darstellung des Reaktionsmechanismus der Undecaprenyldiphosphatsynthese ab der zweiten Verlängerung. Auf die Darstellung der ersten Verlängerung (Kondensierung von FPP mit IPP) wurde zugunsten der Verdeutlichung des zyklischen Mechanismus des Enzyms verzichtet. B – Base; OPP – Diphosphat

Es existieren jedoch auch Enzyme, für die ein Einfluss des Metallions auf die Affinität des Substrates zum Enzym nicht detektierbar ist, wie für eine Proteinfarnesyltransferase beschrieben wurde^[142]. Die Funktion des diphosphatkomplexierenden Metallions Mg^{2+} liegt hier vielmehr in der Katalyse. Auch verschiedene andere Autoren beschreiben eine Erleichterung der Aktivierung bzw. eine Stabilisierung der am Diphosphat entstehenden negativen Ladung durch das Metallion^[1;103;132;141]. Neben der Beteiligung des Metallions an der Substratbindung oder Katalyse wird für die eigentlich metallionenunabhängigen Glycosyltransferasen vom Typ GT-B auch die Erleichterung der Produktfreigabe durch Metallionen angegeben^[1;16]. Zusätzlich werden auch Einflüsse der Metallionen auf die korrekte Konformation des Substrats in der Bindetasche beschrieben^[66;141]. Jedoch sind nicht für alle aktivierenden Enzyme Metallionen essentiell (z.B. ^[1;85;143-145]); ihre Zugabe kann, muss aber nicht zur Beschleunigung der Reaktion führen. Die Funktion des Metallions zur Aktivierung der Diphosphatabgangsgruppe bzw. Stabilisierung der während der Aktivierung entstehenden negativen Ladung kann in solchen Fällen durch dipolare Helixstrukturen oder positiv geladene bzw. wasserstoffbrückenbildende Aminosäuren übernommen werden^[1].

Zudem existieren auch Enzyme, die nicht in der Lage sind, eine Diphosphataktivierung durchzuführen, deren diphosphathaltiger Ligand aber trotzdem durch ein Metallion komplexiert werden muss, um die Enzymfunktion zu gewährleisten. Als Beispiel sei Thiamindiphosphat (TPP) genannt. Hier dient das Metallion der Verankerung des Kofaktors im Protein.^[146] Der Nachweis, dass anorganisches Diphosphat in der Lage ist, mit TPP um die Bindestelle im Enzym zu konkurrieren^[147], zeigt, dass die metallionvermittelte Wechselwirkung des Diphosphats mit dem Protein einen wesentlichen Beitrag zur Ausbildung des Kofaktor-Enzym-Komplexes leistet.

1.4 Diphosphataktivierung und ATP-Hydrolyse

Im Zusammenhang mit der zuvor beschriebenen Diphosphataktivierung liegt naturgemäß eine Parallelität zur Hydrolyse des Adenosintriphosphat (ATP; Abbildung 1.9) vor. Der folgende Abschnitt zeigt generelle Gemeinsamkeiten und Unterschiede beider Vorgänge auf.

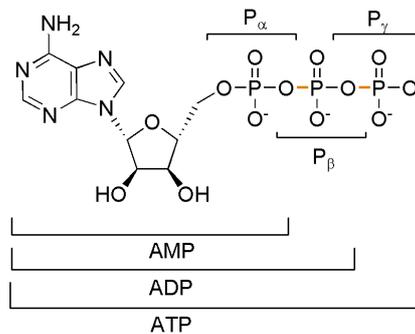


Abbildung 1.9: Struktur von Adenosinmonophosphat (AMP), Adenosindiphosphat (ADP) und Adenosintriphosphat (ATP). P_α, P_β und P_γ bezeichnen die verschiedenen Untereinheiten des Triphosphats. Die während der Hydrolyse spaltbaren Bindungen sind orange markiert.

ATP wurde im Jahr 1929 durch Karl Lohmann entdeckt^[148]; seine Bedeutung als Hauptenergiequelle der Zelle wurde erst 1941 durch Fritz Lipmann dargestellt^[149]. Die ubiquitär vorkommende Verbindung liefert Energie für die Ausführung mechanischer Arbeit (z.B. bei der Muskelkontraktion), für den aktiven Transport von Molekülen und Ionen, sowie die Synthese von Biomolekülen aus einfachen Vorstufen. Durch Kopplung der stark exothermen ATP-Hydrolyse mit endothermen Reaktionen werden diese thermodynamisch möglich. Dabei kann es zur Bildung der Produkte ADP und P_i (anorganisches Phosphat) oder AMP und PP_i (anorganisches Diphosphat) kommen. Unter Standardbedingungen (298 K, 101,3 kPa) und pH 7 werden dabei $-30,5 \text{ kJ}\cdot\text{mol}^{-1}$ bzw. $-32,2 \text{ kJ}\cdot\text{mol}^{-1}$ frei.^[150] Beispielsweise wird die Phosphorylierung der Glucose zu Glucose-6-phosphat im Rahmen der Glykolyse durch die ATP-Hydrolyse zu ADP und P_i angetrieben. Bei der Spaltung des ATP in AMP und PP_i kommt es stets durch die anorganische Pyrophosphatase zu einer weiteren Zerlegung des PP_i in 2 P_i, was wiederum $-33,5 \text{ kJ}\cdot\text{mol}^{-1}$ unter Standardbedingungen bei pH 7 liefert.^[150] Damit können auch stärker endotherme Reaktionen, wie der Aktivierungsschritt des Fettsäureabbaus durch β -Oxidation, möglich gemacht werden. Die Grundlage der starken thermodynamischen Begünstigung der ATP-Hydrolyse liefern dabei die starke elektrostatische Abstoßung der negativen Ladungen des Triphosphats unter physiologischen Bedingungen, die erhöhte Resonanzstabilisierung sowie entropische Vorteile der entstehenden Produkte. Es existieren auch Biomoleküle mit größerem Phosphorylgruppenübertragungspotential als ATP – wie beispielsweise Creatinphosphat –, die zur Regeneration des ATP aus ADP genutzt werden.

Im Gegensatz zur Diphosphataktivierung erfolgt die ATP-Hydrolyse nicht durch Spaltung einer organischen Phosphoresterbindung (C-OP), sondern durch Spaltung einer Phosphoranhydridbindung (P-O-P; siehe Abbildung 1.9). Die meisten Anhydride hydrolysieren in wässriger Lösung schnell, die Spaltung von Phosphoranhydridbindungen erfordert jedoch die Überwindung ungewöhnlich hoher Aktivierungsbarrieren, was zu einer erhöhten Stabilität des ATP unter physiologischen Bedingungen führt und eine enzymkatalysierte Reaktion notwendig macht.^[150] Die Abhängigkeit der ATP-Hydrolyse zu ADP und anorganischem Phosphat von der Konzentration von Metallkationen und dem pH-Wert ist unter enzymfreien Bedingungen gut untersucht. Dabei wurde festgestellt, dass – im Gegensatz zur pH-Abhängigkeit der Diphosphataktivierung – mit steigendem (basischerem) pH eine verstärkte Umsetzung erfolgt (erhöhte Gleichgewichtskonstante).^[151] Untersuchungen zur Reaktionsenthalpie zeigten einen höheren thermodynamischen Energiegewinn unter basischen Bedingungen.^[152] Quantenmechanische Berechnungen zur Hydrolyse von Methyltriphosphat unterstützen diese experimentellen Ergebnisse und zeigen außerdem einen Zusammenhang von Reaktionsenergie und Protonierungszustand des Triphosphats auf: Je geringer die Protonierung desto negativer ist die Reaktionsenergie.^[153] Auch der Einfluss von zweiwertigen Metallionen entspricht nicht dem bei der Diphosphataktivierung beobachteten. Es konnte festgestellt werden, dass eine Metallkomplexierung eher zu einer verringerten Hydrolyse des ATP zu ADP und P_i führt.^[154] Die Ursache hierfür liegt in der verringerten Affinität der Metallionen zu den Produkten im Vergleich zur Substrataffinität. Bei Diphosphataktivierungen steigt die Affinität des Metallions zum Diphosphatprodukt hingegen eher. Erst sehr hohe Konzentrationen von Metallkationen, die auch eine Komplexierung der entstehenden Produkte ermöglichen, können unter basischen Bedingungen für ATP-Hydrolysen Umsätze ähnlich den metallionenfreien Reaktionen wiederherstellen. Auch Untersuchungen zur Reaktionsenthalpie zeigen einen verringerten thermodynamischen Energiegewinn bei Metallionenkomplexierung.^[155] Trotz Allem sind ATP-hydrolysierende Enzyme i.A. abhängig von Metallkationen.^[156] Nemukhin *et al.*^[157] stellt die Betrachtung von Reaktionsmechanismen einiger adenosin- und guanosintriphosphathydrolysierender Enzyme durch kombinierte quanten- und molekülmechanische Methoden (QM/MM) zusammen. Weiterhin konnte ermittelt werden, dass die Reaktionsenthalpie der ATP-Hydrolyse zu ADP und P_i unter physiologischen Bedingungen und bei Betrachtung physiologischer Konzentrationen von ATP, ADP und anorganischem Phosphat etwa $-50 \text{ kJ}\cdot\text{mol}^{-1}$ beträgt.^[150;156]

2 ZIELSTELLUNG DER ARBEIT

Diphosphate und ihre Aktivierung spielen eine grundlegende Rolle in biologischen Systemen. Ohne sie wäre keine Photosynthese oder Atmung in uns heute bekannter Form denkbar. Auch die Regulation von Proteinaktivitäten oder der Löslichkeit von Molekülen wäre in der bewährten Art und Weise nicht oder nur eingeschränkt möglich. Daher haben Defekte entsprechender Proteine oft schwerwiegende, existenzbedrohende Folgen. Gerade diese Eigenschaft macht die Proteine auch interessant für die pharmazeutische Industrie. Beispiele sind das Antimalariamedikament Artemisinin, das Chemotherapeutikum Paclitaxel oder die lange Zeit eingesetzten Sulfonamidantibiotika. Nicht zuletzt sind organische Diphosphatverbindungen interessante Zielmoleküle der chemischen Synthese, insbesondere im Zusammenhang mit der Erzeugung von auf herkömmlichem Wege nur schwer zugänglichen aktivierten sp^3 -Kohlenstoffzentren.

Detaillierte Informationen über die Aktivierung von Diphosphaten in Proteinen sind nur beispielhaft bekannt. Eine generelle Untersuchung, die den Prozess unabhängig von Details der Enzymbindetasche erklärt, wurde nach Wissen des Autors bisher nicht durchgeführt. Ziel der vorliegenden Arbeit ist es, die Diphosphataktivierung aus verschiedenen Blickwinkeln zu betrachten und zu ermitteln, welche Faktoren den Prozess beeinflussen können, in welcher Art und Weise sie dies tun und wie sich diese Faktoren evolutionär entwickelt haben. Dafür wurde zunächst ein Datensatz aus allen zum Zeitpunkt des 28. August 2008 vorliegenden 3D-Röntgenstrukturen diphosphatbindender Proteine der Protein Data Bank erzeugt und aufbereitet, um redundante Informationen auszuschließen. Eine Unterscheidung in Bindestellen, die eine Aktivierung durchführen und Bindestellen, für die es nicht zur Diphosphatabspaltung kommt, ermöglicht dabei einen Vergleich der Diphosphaterkennung beider Fälle. Zunächst soll ein Clustering Aufschluss über die Evolution dieser Bindestellen geben. Dabei ist insbesondere von Interesse, ob aktivierende und nichtaktivierende Bindestellen mit sehr ähnlichen Liganden einen gemeinsamen Vorläufer besitzen oder ob eher ein Vorläufer für alle aktivierenden und ein weiterer für alle nichtaktivierenden Bindestellen existieren. Zusätzlich sollen Bindemotive aktivierender Proteine mit diesem Ansatz detektiert werden. Einfache Statistiken sollen nachfolgend einen unmittelbaren Vergleich aktivierender und nichtaktivierender Bindestellen ermöglichen und Aufschluss über die Bedeutung bestimmter funktioneller Gruppen für die Aktivierung geben. Analysen der Torsionswinkel der Diphosphate in den Bindetaschen sollen die Bedeutung der Diphosphatkonformation aufdecken. Alle Untersuchungen werden dabei durch quantenchemische Berechnungen unterstützt, um die geringe Zahl verfügbarer 3D-Strukturen aktivierender Bindestellen auszugleichen. Neben der Erkennung und Konformation des Diphosphats in der Bindetasche ist jedoch auch der

Aktivierungsprozess als solches von Interesse. Mit Hilfe einer Modellreaktion und der Unterstützung durch experimentelle Daten sollen Reaktionsenergien und Aktivierungsbarrieren für verschiedene Aktivierungsreaktionen bestimmt werden. Dabei wird insbesondere auf die Bedeutung des Protonierungsgrades und der Metallionenkomplexierung des Diphosphats eingegangen.

All diese Ansätze sollen – im Zusammenhang betrachtet – helfen, allgemeine Kriterien für eine Aktivierung bzw. Nichtaktivierung zu finden, um somit Reaktionsmechanismen von Enzymen besser verstehen zu können oder Grundlagen für die Konstruktion einfacher Katalysatoren für den Einsatz in der chemischen Synthese zu liefern.

3 METHODEN

3.1 Extraktion von Strukturdaten von Diphosphatbindestellen und daraus abgeleiteter Informationen

Den zugrundeliegenden Datenpool bildete die PDB^[52] mit Stand vom 28. August 2008. Alle Programme zur Erkennung von Diphosphaten und ihrer Mimikri in den Proteinen, sowie ihrer Wechselwirkungen innerhalb der Bindetasche, als auch aller unmittelbar daraus abgeleiteten Informationen, wie z.B. Torsionswinkel des Diphosphats, die Zahl und Art der Wechselwirkungen und die verschiedenen Bindestellenrepräsentationen für das Clustering, wurden mit Hilfe der Skriptsprache MOETM^[158] SVL erzeugt. Im Folgenden soll detailliert auf die verschiedenen Schritte der Informationsgewinnung und dabei auftretender Probleme eingegangen werden.

3.1.1 Extraktion und Aufbereitung der Bindestellen aus der PDB

Zur Speicherung der Bindestellen und zugehöriger Informationen wurde eine MOETM-Datenbank erzeugt, um eine einfache Verarbeitung für nachfolgende Skripte zu ermöglichen. Zunächst wurde jede Struktur der PDB nach Diphosphaten durchsucht. Eine sehr mächtige Methode hierfür bieten SMARTS[®], was sich von „SMiles ARbitrary Target Specification“ herleitet, wobei der Begriff SMILESTM wiederum „Simplified Molecular Input Line Entry System“ bezeichnet. Sowohl SMILES[®], als auch SMARTSTM wurden von Daylight Chemical Information Systems, Inc.^[159] entwickelt. Es handelt sich hierbei um eine Möglichkeit, Strukturen oder Strukturfragmente als Zeichenketten zu definieren, die, durch bereits in MOETM vorimplementierte Algorithmen, sehr einfach für Struktursuchen angewendet werden können.

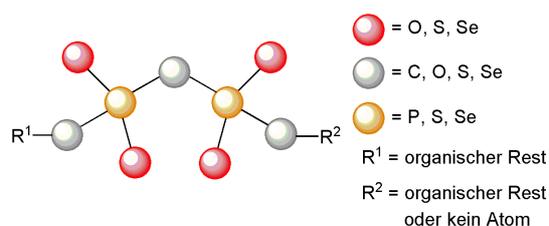


Abbildung 3.1: Darstellung der Diphosphate und einfacher Mimikri für die Definition der SMARTS.

Die Diphosphate, sowie einige definierte Mimikri wurden dabei, wie in Abbildung 3.1 abgebildet, dargestellt. Es wird klar, dass anorganische Diphosphate bei der Suche ausgeschlossen wurden. Sie binden oft unspezifisch und sind daher für eine Analyse von Diphosphatbindemotiven nicht interessant. Als Mimikri werden nur solche Strukturen betrachtet, bei denen lediglich einzelne Atome ausgetauscht sind und damit die Grundstruktur des Diphosphats erhalten bleibt. Um eine ausreichende Qualität der verwendeten Strukturen zu gewährleisten, gleichzeitig aber nicht zu viele Strukturen von der Analyse auszuschließen, wurden nur Röntgenstrukturen mit einer Mindestauflösung von 2,0 Å für die Weiterverarbeitung verwendet. Jede Struktur, die diese Kriterien erfüllt, wurde nun mit Hilfe der Methode Protonate3D^[160] überarbeitet. Dieser Algorithmus ist u.a. in der Lage, den Imidazolrest von Histidin bzw. den Carbonsäureamidrest der Seitenketten von Asparagin und Glutamin zu drehen, sodass optimale Wechselwirkungen mit der Umgebung erfolgen können. Dieser Ansatz ist deshalb interessant und wurde hier angewendet, da es für die genannten Aminosäurefragmente oft schwierig ist, während der Strukturaufklärung die korrekten Atomidentitäten in der Elektronendichtekarte zuzuordnen. Für diese so vorverarbeiteten Strukturen wurden nun die mit jedem Diphosphat direkt in Wechselwirkung tretenden Gruppen ermittelt. Hierbei kann es sich um Aminosäuren, aber auch um Liganden, Metallionen oder Wassermoleküle handeln. Dabei wurden alle Wechselwirkungen berücksichtigt, die festgelegte, nachfolgend erläuterte Kriterien erfüllen, auch, wenn ein Atom des Diphosphats Wechselwirkungen mit mehreren umgebenden Gruppen ausbildet. Aufgrund der Hydrophilie des Diphosphats und damit auch seiner Wechselwirkungen werden lediglich Wasserstoffbrücken und ionische Wechselwirkungen in Betracht gezogen; Van-der-Waals- sowie hydrophobe Interaktionen spielen bei der Diphosphaterkennung eine untergeordnete bzw. keine Rolle. Ionische Beziehungen wurden sehr einfach definiert durch eine Distanz von maximal 4,5 Å zwischen einem Anion und einem Kation. Dieser geringe Abstand sichert, dass zwischen beiden geladenen Gruppen kaum abschirmende Atome liegen können, bezieht aber gleichzeitig die größere Reichweite der ionischen Wechselwirkungen im Vergleich zu Wasserstoffbrücken ein, die bereits bei einer Distanz von über 3 Å zwischen H-Atom und seinem Akzeptor vernachlässigt wurden. Die Erkennung der Wasserstoffbrücken gestaltete sich als etwas schwieriger, da hier neben der Distanz auch der Winkel zwischen H-Donor, H-Atom und H-Akzeptor bedeutsam ist. Hierfür wurde ein bereits in MOETM implementierter Ansatz^[161] verwendet, der unabhängig ist von der Addition von Wasserstoffen an die Struktur. Dadurch wird das Vernachlässigen von Wasserstoffbrücken aufgrund falscher Positionierung der Wasserstoffe verhindert. Eine mögliche Protonierung des Diphosphats, die in den Röntgenstrukturen nicht erkennbar ist, wurde ebenfalls während der Berechnung der Wechselwirkungen berücksichtigt (z.B. für Interaktionen mit der Carbonylgruppe eines Aminosäurerückgrates oder einer Carboxylgruppe). Jedes Diphosphat

wurde anschließend gemeinsam mit den jeweils bindenden Gruppen (= Diphosphatbindetasche), der Auflösung der Struktur in Å, dem für jede Ligandenstruktur spezifischen 3-Letter-Code und einer Grobklassifizierung des Proteins (aus dem „Header“-Eintrag der PDB-Datei) einzeln in einer MOETM-Datenbank mit letztlich insgesamt 2253 Einträgen abgelegt. Außerdem wurde jede Bindestelle anhand der Grobklassifizierung des Proteins einer der sechs obersten Ebenen der EC-Klassifizierung (Oxidoreduktase, Transferase, Hydrolase, Lyase, Isomerase, Ligase) zugeordnet. Einige, wenige Strukturen konnten keiner der genannten Gruppen zugeordnet werden. Es handelt sich hierbei um Strukturen, die am Elektronentransport (z.B. in der Atmungskette) beteiligt sind, um Signalproteine oder Strukturen, deren Funktion unbekannt ist. Da die auf dem resultierenden Datensatz auszuführenden Analysen eine Unterscheidung von aktivierten und nichtaktivierten Liganden notwendig machen, wird jeder Bindestelle zusätzlich das Schlüsselwort „aktiviert“, „nichtaktiviert“ oder „Produkt“ zugeordnet. Bindestellen, die Diphosphatliganden als Produkte enthalten, werden nachfolgend vernachlässigt.

Die Erzeugung einer aussagekräftigen Statistik, die aus diesen Daten gewonnen werden soll, erfordert einen repräsentativen Datensatz, der möglichst frei von Redundanzen ist. Die mehrfache Kristallisierung desselben Proteins bei verschiedenen Auflösungen, mit unterschiedlichen Liganden oder Mutationen, sowie das Vorliegen von Homomultimeren macht die PDB jedoch zu einem Datenpool von stark ausgeprägter Redundanz. Es muss berücksichtigt werden, dass in solchen Strukturen häufig verschiedene Aminosäurekonformationen und unterschiedliche Muster von Wassermolekülen innerhalb der Ligandenbindungstasche vorliegen und damit redundante Bindestellen nicht zwangsläufig in ihrer dreidimensionalen Struktur identisch sind. Diese Beobachtung wird durch Boström *et al.*^[162] bestätigt, die feststellten, dass bei der Bindung ähnlicher Liganden an identische Proteine häufig eine Rotation von Aminosäureseitenketten oder eine Veränderung der Architektur der Wassermoleküle auftritt. Es ist daher nicht möglich, durch das Suchen dreidimensional identischer Strukturen alle redundanten Bindestellen zu erkennen. Auch die Verwendung von EC-Nummern als Basis ist schwierig, da diese für viele Einträge der PDB nicht angegeben sind und dies auch keine Lösung für heteromultimere Proteine mit mehreren, verschiedenen Diphosphatbindestellen darstellen würde. Daher wurde ein Ansatz unter Verwendung des Programms CD-HIT^[163] entwickelt. CD-HIT klassifiziert gegebene Proteinsequenzen anhand der Identität (= Übereinstimmung) ihrer Aminosäuren. Jede resultierende Klasse enthält dabei diejenigen Sequenzen, die untereinander eine vorgegebene Mindestidentität haben. Erfüllen zwei Sequenzen dieses Kriterium nicht, werden sie verschiedenen Klassen zugeordnet. Um Klassen redundanter Bindestellen zu erhalten, wurden zunächst alle Sequenzen der Proteine des erzeugten Datensatzes mit CD-HIT klassifiziert. Dabei wurde eine Mindestidentität von

90% verwendet, um identische Sequenzen unterschiedlicher Organismen in einer Klasse zu gruppieren. Anschließend wurden für alle Bindestellen die Proteinketten der mit dem Diphosphat wechselwirkenden Gruppen bestimmt, was es in einem nachfolgenden Schritt ermöglichte, jede Bindestelle einer oder einer Kombination von CD-HIT-Klassen zuzuordnen. Jedoch sind nicht zwangsläufig alle Bindestellen mit gleichen CD-HIT-Klassen als redundant zu bezeichnen. In dem hier vorliegenden Zusammenhang wird auch eine Berücksichtigung des gebundenen Liganden benötigt. Beispielsweise gibt es Proteine, die in einer großen gemeinsamen Bindetasche zwei Diphosphatliganden binden (z.B. Isoprenyldiphosphatsynthasen). Auch darf nicht vernachlässigt werden, dass zwei völlig unterschiedliche Liganden in sehr ähnlichen Bindetaschen gebunden sein können. Aus diesem Grund wurde jeder Diphosphatligand auf Basis seines eindeutigen 3-Letter-Codes einer Ligandengruppe (z.B. NAD⁺, CoA, UDP-Zucker, Geranyldiphosphat) zugeordnet. Eine direkte Verwendung des 3-Letter-Codes ist nicht geeignet, da dies beispielsweise NAD⁺ von NAD⁺-Inhibitoren trennen würde, die zum Zweck der Kristallisation des Enzym-Ligand-Komplexes zugesetzt wurden. Die erhaltenen CD-HIT-Klassen der Bindestellen wurden nachfolgend anhand der erhaltenen Ligandengruppen aufgespalten. Somit lagen redundante Bindestellen in dem hier benötigten Zusammenhang in einer Klasse vor. Um nun einen nichtredundanten Datensatz zu erhalten, war es notwendig aus jeder Klasse einen Repräsentanten auszuwählen. Hier wurde jeweils die Bindestelle mit den meisten Interaktionen des Diphosphats und der besten Auflösung gewählt. Das Resultat war hier eine Reduktion der Zahl der Bindestellen auf 506.

Diese Zahl ist gering genug, um eine manuelle Bewertung der Strukturen vorzunehmen. In einigen Fällen war es den Kristallographen aufgrund großer Beweglichkeit von Gruppen nicht möglich, eine exakte Positionierung festzustellen. Häufig werden dann mehrere mögliche Konformationen in der Röntgenstruktur abgebildet. Für eine Analyse von Wechselwirkungen, wie sie hier durchgeführt werden soll, sind solche Strukturen ungeeignet und wurden daher entfernt. Weiterhin wurden Bindestellen entfernt, bei denen eine unspezifische Bindung des Diphosphats (oft aufgrund von Kristallpackungseffekten) auf der Oberfläche des Proteins offensichtlich war. Diese Nachbearbeitung reduzierte den Datensatz auf letztendlich 429 nichtredundante Bindestellen.

Der Prozess der Extraktion der diphosphathaltigen Bindestellen, sowie die Aufbereitung des Datensatzes sind in Abbildung 3.2 schematisch zusammengefasst.

Um die Erzeugung der Datenbank mit der Gewinnung der benötigten Informationen zu beschleunigen, wurden zeitintensive Prozesse des Programms mit MOETM/smp implementiert, das eine parallele Verarbeitung ermöglicht.

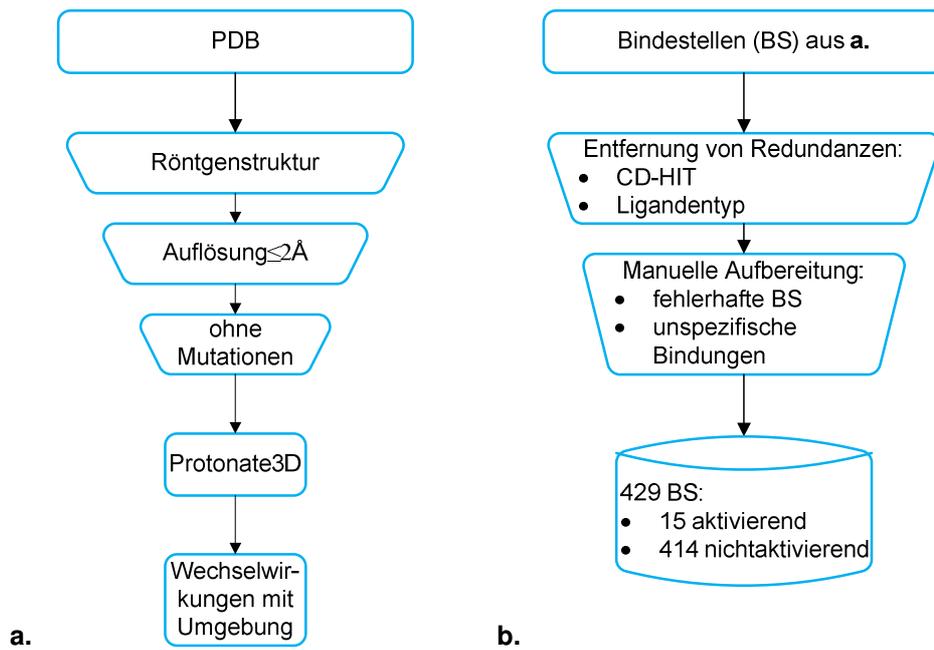


Abbildung 3.2: Darstellung der Extraktion (a.) und Aufbereitung (b.) der Bindestellen in einem Ablaufschema.

3.2 Clusteringmethoden

Clustering^[164;165] ist eine häufig eingesetzte Methode der Datenanalyse, um wiederkehrende Strukturen in großen Datenmengen zu erkennen. Hierbei liegt oft keine gesicherte Information über die exakte Gestalt der Cluster vor – Ziel der Analyse ist es, diese erst zu detektieren. Auch die Zugehörigkeit der Datenpunkte zu den einzelnen, unbekanntem Clustern ist nicht gegeben. Hier findet sich auch der Unterschied zur Klassifizierung, bei der bereits eine Menge von Clustern (Klassen) und zugeordneter Daten bekannt ist und neue Daten mit unbekannter Klassenzugehörigkeit zugewiesen werden. Die Voraussetzungen und Ziele beider Methoden sind also verschieden und doch sind sie eng verknüpft: Auf ein Clustering folgt häufig eine Klassifizierung. Die Problematik der meisten heute angewendeten Clustering-Methoden (hierarchische oder partitionierende Methoden) liegt in ihrer oft starken Parameterabhängigkeit. Neben algorithmenspezifischen Parametern müssen meist die Klassenzahl, sowie Informationen über die Form der Klassen berücksichtigt werden. Beide können durch den Nutzer sehr häufig nur erahnt werden. Während die Klassenzahl direkt angegeben wird, fließen Restriktionen über die Gestalt der Cluster lediglich implizit über die Wahl der verwendeten Methode ein. Jeder Algorithmus forciert dabei eine ganz bestimmte Clusterform – unter Umständen auch, wenn sie den zu clusternden Daten nicht zugrunde liegt. Teilweise wird somit eine Klassenstruktur erzeugt, die nicht der natürlichen Struktur der Daten entspricht. Die genaue Kenntnis über die jeweils bevorzugte Klassenform ist also vor Anwendung der Algorithmen von großer Wichtigkeit. Um diese Probleme einzuschränken, kann eine Reihe möglicher Klassenzahlen und Algorithmen auf den Datensatz angewendet werden. Die entstehenden Resultate („Clusterings“) können dann durch verschiedene Validierungsmethoden auf ihre Zuverlässigkeit getestet werden. Einen guten Überblick über diesen Ansatz bieten Handl et al.^[166] und Yeung et al.^[167]. Problematisch bleibt allerdings der Zeitaufwand, der für das vielfache Clustering investiert werden muss. Hinzu kommt, dass auch durch die Validierungsmethoden die Wahl der geeigneten Kombination aus Klassenzahl und Algorithmus nicht immer eindeutig ist. Geeigneter wäre ein Algorithmus, für den die (implizite) Wahl einer Klassenzahl oder -form nicht notwendig ist. Einen Ansatz stellen hier die dichtebasierten Clusteringmethoden dar. Diese Gruppe von Methoden ist in der Lage, ohne Vorgabe einer Klassenzahl, beliebig geformte Cluster zu erkennen. Dies sind hier solche Bereiche des Datenraums, in denen die Datenpunkte dicht liegen. Sie sind durch weniger dichte Regionen voneinander abgegrenzt. Für einige dieser Algorithmen muss jedoch durch den Nutzer eine bestimmte Dichte vorgegeben werden, die ein Cluster definiert. Für den Algorithmus OPTICS^[168;169], der auch in der vorliegenden Arbeit genutzt wird, ist dies jedoch nicht der Fall. Dieser Ansatz

produziert allerdings auch kein festes Clustering (keine klar definierte Klassenstruktur), sondern gibt vielmehr Aufschluss über die (Dichte-) Struktur des Datensatzes.

Nachfolgend werden kurz einige Distanzmetriken vorgestellt, sowie detailliert auf den OPTICS-Clustering-Algorithmus eingegangen.

3.2.1 Distanzmetriken

Jeder Clustering-Algorithmus benötigt die Definition einer Distanzmetrik, die die Abstände (und damit implizit auch die Ähnlichkeit) von Datenpunkten definiert. Ziel jedes Algorithmus ist es dann, ähnliche Datenpunkte ein- und demselben Cluster zuzuordnen. Damit hat die Wahl des Distanzmaßes einen großen Einfluss auf das letztendlich entstehende Ergebnis. Sie muss stark abhängig gemacht werden von der genauen Datenstruktur (Dimensionalität, Variablentyp usw.), aber auch von der exakten Zielstellung des Clusterings. Seien p und q nachfolgend zwei reelle, n -dimensionale Vektoren.ⁱ Formel 3.1 definiert einige Distanzmaße für diese Vektoren. Eine der bekanntesten und am häufigsten eingesetzten Metriken ist die euklidische Distanz. In Zusammenhang mit ganzzahligen Daten wird auch häufig die Manhattan-Metrik (auch City-Block-Metrik genannt) verwendet. Zusätzlich gibt es auch skalierte Distanzmaße, die Quotienten enthalten. Ein Beispiel für eine skalierte Manhattan-Distanz ist die Bray-Curtis-Distanz (auch Sørensen-Distanz). Durch die Skalierung wird erreicht, dass sich, falls alle Koordinaten i der Datenpunkte p und q nichtnegativ sind, für die Distanz ein Wert zwischen 0 und 1 ergibt.ⁱⁱ Auch die Canberra-Distanz ist eine skalierte Metrik. Durch die vorgezogene Summe kommt es zur Auswertung jeder Koordinate i unabhängig von den verbleibenden Koordinaten $j \neq i$.ⁱⁱⁱ Dabei liegt die Teildistanz d_C^i einer Koordinate i stets zwischen 0 und 1.^{iv} Gilt $p_i = 0$ oder $q_i = 0$ so ergibt sich stets der Wert 1 für die entsprechende Teildistanz; unabhängig vom anderen Wert.^v Dies kann auch als „Boole’sche Eigenschaft“ der Canberra-Distanz bezeichnet werden, da sie genau dann 1 zurückliefert, wenn die betrachtete Eigenschaft (Koordinate) in einem der Vektoren nicht vor-

ⁱ $p, q \in \mathfrak{R}^n$

ⁱⁱ $\forall_{i=1}^n : p_i, q_i \geq 0 \Rightarrow d_{BC} \in [0,1]$

ⁱⁱⁱ $d_C = \sum_{i=1}^n d_C^i$

^{iv} $\forall_{i=1}^n : d_C^i \in [0,1]$

^v $\exists i : p_i = 0, q_i \neq 0 \Rightarrow d_C^i = 1$

handen ist, jedoch im anderen Vektor ausgeprägt ist. Es sei an dieser Stelle auch zu beachten, dass dieser Wert die Maximaldistanz für eine Koordinate darstellt. Dies führt zu einer starken „Bestrafung“ solcher binärer Differenzen zwischen den zu vergleichenden Vektoren. Vektoren mit gleicher Koordinatenausprägung (d.h. eine Eigenschaft ist stets gleichzeitig für beide Vektoren gleich 0 oder ungleich 0, aber niemals für einen Vektor 0 und für den anderen ungleich 0) sind damit stets ähnlicher – unabhängig von der Differenz der Koordinaten – als Vektoren mit unterschiedlicher Koordinatenausprägung.^{vi}

$$\begin{array}{ll}
 \text{a.} & d_E = d(p, q) = \sqrt{\sum_{i=1}^n (p_i - q_i)^2} \\
 \text{b.} & d_M = d(p, q) = \sum_{i=1}^n |p_i - q_i| \\
 \text{c.} & d_{BC} = d(p, q) = \frac{\sum_{i=1}^n |p_i - q_i|}{\sum_{i=1}^n (p_i + q_i)} \\
 \text{d.} & d_C = d(p, q) = \sum_{i=1}^n \frac{|p_i - q_i|}{|p_i| + |q_i|}
 \end{array}$$

Formel 3.1: Definition verschiedener Distanzmetriken: a. euklidische Distanz, b. Manhattan-Distanz, c. Bray-Curtis-Distanz, d. Canberra-Distanz. Die Vektoren p und q seien reell und n -dimensional.

3.2.2 Der Clustering-Algorithmus OPTICS

Der OPTICS-Algorithmus („Ordering Points To Identify the Clustering Structure“) wurde erstmals 1999 vorgestellt und ist ein Beispiel für einen dichte-basierte Clustering-Ansatz.^[168;169] Er ist daher in der Lage, Häufungen von Datenpunkten zu detektieren (dichte Regionen). Als Resultat ergibt sich dabei keine feste Klassenzuordnung der Datenpunkte, sondern eine Neuordnung der Datenpunkte („ordering“), sowie eine Menge von Zahlenwerten, die – aufgetragen in einem Diagramm in der in „ordering“ vorgegeben Reihenfolge – ein Maß für die Datendichte der Umgebung jedes Punktes p angibt. Diese Erreichbarkeitsdistanz („reachability distance“) gibt die minimale Entfernung von p zu allen bisher betrachteten Punkten, die sich in einer gewissen Mindestdichte befinden, an. Punkte, die in einer Umgebung mit solch einer Mindestdichte liegen, werden als Kernobjekte („core objects“) bezeichnet. Jedes Kernobjekt hat eine zugehörige Kerndistanz („core distance“), die genaue Aussagen über die Dichte der Umgebung macht. Der Algorithmus startet mit einem beliebigen Punkt p_1 des zu clusternden Datensatzes. Da dieser Punkt von keinem anderen, bisher betrachteten Punkt erreichbar ist (er ist der erste untersuchte Punkt), erhält er eine

^{vi} $\exists i \in [1, n]: p_i, q_i \neq 0, \exists j \in [1, n]: p_j \neq q_j = 0 \Rightarrow d_C^i < d_C^j$

unendlich große Erreichbarkeitsdistanz. Außerdem wird er zu „ordering“ hinzugefügt. Ist p_1 kein Kernobjekt, d.h. befindet sich nicht in einer Umgebung mit einer gewissen Mindestdichte, so wird der nächste Punkt des Datensatzes analog analysiert. Erfüllt p_1 hingegen die Bedingungen eines Kernobjektes, so breitet sich der Algorithmus auch über seine Nachbarpunkte aus: Ihre Erreichbarkeitsdistanz in Anhängigkeit von p_1 wird bestimmt und in einer Liste („orderseeds“) gespeichert. Die Reihenfolge der Nachbarn in dieser Liste ist festgelegt: Sie sind nach ihrer Erreichbarkeitsdistanz sortiert, beginnend mit der kleinsten Distanz. Schritt für Schritt wird nun diese Liste beginnend mit dem Punkt p_2 mit der kleinsten Erreichbarkeitsdistanz abgearbeitet. Der Punkt wird zunächst zu „ordering“ hinzugefügt. Anschließend wird bestimmt, ob es sich um ein Kernobjekt handelt; falls dies der Fall ist, werden seine Nachbarn ermittelt, ihre Erreichbarkeitsdistanz in Abhängigkeit von p_2 bestimmt und die Punkte in die Liste „orderseeds“ einsortiert. Dabei kann es auch zu einer Aktualisierung der Distanz eines Punktes kommen, falls dieser besser von p_2 als p_1 erreichbar ist. Ist p_2 hingegen kein Kernobjekt, so werden seine Nachbarn nicht untersucht. Diese Vorgehensweise führt zu einer immer stärkeren Ausbreitung über den Datensatz, wobei immer der Punkt als nächstes untersucht – und gegebenenfalls ausgedehnt – wird, der am besten von einem der bereits betrachteten Kernobjekte aus erreichbar ist. Manchmal sind Datenpunkte von keinem Kernobjekt aus erreichbar, da sie in einer Region mit sehr geringer Dichte liegen und damit nie als Nachbar eines Kernobjektes erkannt werden. Für diese Ausreißer ist die Erreichbarkeitsdistanz stets unendlich groß. Allgemein gibt die letztendlich resultierende Erreichbarkeitsdistanz jedes Datenpunktes die beste Erreichbarkeit zu allen der bis zu diesem Zeitpunkt der Ausbreitung über den Datensatz betrachteten Kernobjekte an. Es werden nachfolgend einige Definitionen eingeführt.

Definition 1: p , ε , $MinPts$.

p sei ein M -dimensionaler, reeller Vektor ($p \in \mathfrak{R}^M$). ε sei eine positive, reelle Zahl ($\varepsilon \in \mathfrak{R}_+^*$). $MinPts$ sei eine positive, natürliche Zahl ($MinPts \in \mathfrak{N}^*$).

Definition 2: ε -Nachbarschaft („ ε -neighborhood“).

Die ε -Nachbarschaft $N_\varepsilon(p)$ eines Datenpunktes p ist die Menge aller Datenpunkte $q \in \mathfrak{R}^M$, die sich in einer (Hyper)-Sphäre mit Radius ε und Mittelpunkt p befinden.

Definition 3: Kernobjekt („core object“).

Ein Datenpunkt p ist ein Kernobjekt, wenn gilt: $|N_\varepsilon(p)| \geq MinPts$, d.h. es befinden sich mindestens $MinPts$ Nachbarn in der ε -Nachbarschaft von p .

Definition 4: Kerndistanz („core distance“).

Sei $dist(p, N_{MinPts})$ die durch eine beliebige Metrik definierte Distanz eines Datenpunktes p und seines $MinPts$ -nächsten Nachbarn. Dann gibt

$$coredist_{\epsilon, MinPts}(p) = \begin{cases} \infty, & \text{wenn } |N_{\epsilon}(p)| < MinPts \\ dist(p, N_{MinPts}), & \text{sonst} \end{cases}$$

die Kerndistanz von p an.

Definition 5: Erreichbarkeitsdistanz („reachability distance“).

Sei $dist(q, p)$ die durch eine beliebige Metrik definierte Distanz der Datenpunkte $q, p \in \mathfrak{R}^M$. Dann gibt

$$reachdist_{\epsilon, MinPts}(q, p) = \begin{cases} \infty, & \text{wenn } |N_{\epsilon}(p)| < MinPts \\ \max(coredist(p), dist(q, p)), & \text{sonst} \end{cases}$$

die Erreichbarkeitsdistanz von q in Bezug auf p an.

Abbildung 3.3 zeigt den Pseudocode des OPTICS-Algorithmus. Neben ϵ und $MinPts$ werden vier D-dimensionale Vektoren $v = (v_1, \dots, v_D)$ definiert, wobei D die Anzahl der Datenpunkte des Datensatzes darstellt. Damit steht jedes Element $v(i)$ ($i \in \mathfrak{R}^D$) einem Datenpunkt zur Verfügung:

- *ordering* gibt die Abarbeitungsreihenfolge der Datenpunkte während des Algorithmus an. Zuerst abgearbeitete Datenpunkte stehen vorn; es handelt sich um eine fortlaufende Liste. Im Raum benachbarte Punkte folgen in *ordering* aufeinander. Bei Auftragung der Erreichbarkeitsdistanzen im Diagramm wird die in *ordering* festgelegte Reihenfolge der Datenpunkte verwendet.
- *processed(i)* gibt für Datenpunkt i des Datensatzes an, ob er bereits abgearbeitet wurde, d.h. ob er bereits in *ordering* eingetragen wurde. Die Erreichbarkeitsdistanz des Punktes ist damit nicht mehr veränderbar.
- *reachdist* gibt die Erreichbarkeitsdistanz jedes Datenpunktes an. Dabei entspricht *reachdist(i)* jeweils dem Minimum aller Erreichbarkeitsdistanzen zu Kernobjekten, die zum Zeitpunkt des Anfügens des Datenpunktes i bereits besucht waren (und damit bereits selbst an *ordering* angefügt waren).
- *coredist(i)* gibt die Kerndistanz des Datenpunktes i an.

```

initialize  $\epsilon$ , MinPts by user-defined values
initialize ordering with null set
initialize all elements of processed with FALSE
initialize all elements of reachdist with INFINITE
initialize all elements of coredist with INFINITE

-----
main function
-----
function optics()
begin
  foreach p in dataset do
    if processed(p)==FALSE then
      expandClusterOrder(p)
    fi
  od
end

-----
- sets vectors processed, ordering, reachdist
  and coredist
- expands the neighbourhood of core objects
-----
function expandClusterOrder(p)
begin
  initialize orderseeds by empty set
  set processed(p) to TRUE
  set coredist(p) by definition
  append p to ordering
  if  $|N_{\epsilon}(p)| \geq \text{MinPts}$  then // p is core object
    updateOrderSeeds(p,  $N_{\epsilon}(p)$ )
    while orderseeds not empty do
      get first element of orderseeds and store in cur
      set processed(cur) to TRUE
      set coredist(cur) by definition
      append cur to ordering
      if  $|N_{\epsilon}(\text{cur})| \geq \text{MinPts}$  then // cur is core object
        updateOrderSeeds(cur,  $N_{\epsilon}(\text{cur})$ )
      fi
    od
  fi
end

-----
- determines the reachability distance of all
  neighbors  $N_{\epsilon}(p)$  of p
- sorts them into orderseeds
-----
function updateOrderSeeds(p,  $N_{\epsilon}(p)$ )
begin
  forall n in  $N_{\epsilon}(p)$  do // for all neighbours of p
    if processed(n)==FALSE then
      get reachdist(n,p) by definition and store in new_rd
      if reachdist(n)==INFINITE then // n has not been visited yet
        set reachdist(n) to new_rd
        insert n into orderseeds
      elseif reachdist(n)>reachdist(n,p) then
        set reachdist(n) to new_rd
        update position of n in orderseeds
      fi
    fi
  od
end

```

Abbildung 3.3: Pseudocode des OPTICS-Algorithmus.

Bei Betrachtung des Pseudocodes wird deutlich, dass die Parameter ε und *MinPts* nutzerdefinierte Werte sind. Der Vorteil der OPTICS-Methode ist jedoch, dass sie weitgehend robust gegenüber der exakten Wahl der Werte ist: Es existiert ein breites Spektrum möglicher Werte, die zu denselben Ergebnissen führen^[168]. Dies ist ein Vorteil gegenüber vielen anderen Clustering-Ansätzen, die oft sehr stark abhängig von der Wahl der Eingabeparameter sind. Es ist offensichtlich, dass mit fallenden Werten für ε weniger Datenpunkte die geforderten Eigenschaften eines Kernobjektes erfüllen und auch weniger Datenpunkte in die ε -Nachbarschaft dieser Kernobjekte fallen. Damit sind weniger Objekte von solchen Kernobjekten aus erreichbar. Dies hat zur Folge, dass Datenpunkte in weniger dichten Regionen des Datensatzes eine unendlich große Erreichbarkeitsdistanz erhalten und damit ihre Clusterstruktur innerhalb des Diagramms nicht mehr erkennbar ist. Je strenger die Kriterien gewählt sind, desto stärker ist dieser Effekt. Im Allgemeinen wird jedoch die Sichtbarkeit aller Clusterebenen (dichter und weniger dichter Regionen) angestrebt. Um dies zu erreichen, muss ε mindestens so groß gewählt werden, dass für keinen der Datenpunkte eine unendlich große Erreichbarkeitsdistanz resultiert, d.h. alle Datenpunkte in der Nachbarschaft eines Kernobjektes liegen. Werte für *MinPts* haben eher einen Einfluss auf die Glätte der im Diagramm dargestellten Erreichbarkeitsdistanz: Kleinere Werte führen allgemein zu einer raueren, „zackigeren“ Oberfläche. Dies ist in erster Linie auf eine häufigere Abweichung der Erreichbarkeitsdistanz von Datenpunkten von der Kerndistanz des zugehörigen Kernobjektes (siehe Definition 5) zurückzuführen. Größere Werte für *MinPts* glätten hingegen den Verlauf der Erreichbarkeitsdistanz im Diagramm: Die Erreichbarkeitsdistanz weicht für Punkte in der ε -Umgebung eines Kernobjektes seltener von der Kerndistanz ab. Zusätzlich führen größere Werte für *MinPts* – ähnlich wie kleine Werte für ε – zu einer Reduzierung der Zahl der Kernobjekte und damit zu mehr Datenpunkten mit unendlich großer Erreichbarkeitsdistanz. Ankerst *et al.*^[168] geben Werte zwischen 10 und 20 als optimale Werte für *MinPts* an.

Neben den bereits zu Beginn des Kapitels genannten allgemeinen Vorteilen dichtebasierter Algorithmen, hebt sich der OPTICS-Algorithmus durch eine weitere Eigenschaft hervor: Die Eigenschaft, keine Cluster im engeren Sinne zu produzieren. Es wird lediglich eine Reihenfolge der Datenpunkte festgelegt, die Aufschluss über die (Dichte-) Struktur des Datensatzes gibt. Dies ermöglicht es dem Nutzer, auch verschieden dichte Regionen als Cluster auszuwählen. Sehr viele andere Clustermethoden bieten diesen Vorteil nicht.

Der beschriebene Algorithmus wurde im Rahmen dieser Arbeit mit Gnu R^[170] implementiert.

3.3 Die Quantenmechanik

Die Quantenmechanik bietet einen zuverlässigen Ansatz zur mathematischen Modellierung von Molekülen und ihrer Eigenschaften, wie Molekülstrukturen, -energien oder Partialladungen, die durch andere Ansätze nur mit wesentlich geringerer Genauigkeit oder auch gar nicht verfügbar sind. Ebenso ermöglicht sie – durch die Berücksichtigung von Elektronen und Orbitalwechselwirkungen – die Betrachtung chemischer Reaktionen und somit die Ermittlung von Übergangszuständen, sowie die Berechnung von Reaktionsenergien. Aufgrund der benötigten Rechenkapazitäten bleibt die Quantenmechanik jedoch heute noch auf kleinere Molekülsysteme beschränkt. Alle in der vorliegenden Arbeit durchgeführten quantenchemischen Berechnungen erfolgten mit Gaussian® 03^[171].

Die nachfolgenden Abschnitte beschreiben zunächst die Grundlagen der Quantenmechanik und gehen nachfolgend auf Methoden, Basissätze und Solvatationsmodelle ein. Anschließend erfolgt ein kurzer Abriss über die Berechnung von Reaktionsübergangszuständen. Ein Vergleich mit den Methoden der Molekülmechanik schließt das Kapitel ab.

3.3.1 Einführung

Die Grundlage aller heutigen quantenmechanischen Methoden bildet eine Theorie, die in den Jahren 1925/26 unabhängig durch Erwin Schrödinger^[172], sowie Werner Heisenberg, Max Born und Pascual Jordan^[173-175] entwickelt wurde. Beide Gruppen verwenden verschiedene Ansätze, es konnte jedoch bereits 1926 gezeigt werden, dass sie mathematisch absolut äquivalent sind.^[176] Schrödinger ging vom Eigenwertproblem der schwingenden Saite aus und – mit dem Hintergrund der Darstellung von Elektronen als Welle basierend auf dem 1924 von Louis de Broglie beschriebenen Welle-Teilchen-Dualismus – gelang es ihm, die Wellenfunktion Ψ und die Gesamtenergie (totale Energie) E von Atomen oder Atomsystemen durch die heute als Schrödinger-Gleichung bezeichnete Formulierung in Zusammenhang zu bringen. Die exakte Bedeutung von Ψ war zunächst weitgehend unklar; jedoch schlug Max Born im Jahr 1926 eine statistische Interpretation^[177] vor (Kopenhagener Deutung), indem er das Quadrat der Wellenfunktion Ψ^2 als eine Wahrscheinlichkeitsdichte für den Aufenthalt der Elektronen des Systems an bestimmten Orten darstellte.

Die Schrödinger-Gleichung hat die Form $\hat{H}\Psi = E\Psi$, wobei \hat{H} als Hamilton-Operator bezeichnet wird und einen Differentialoperator darstellt. Die Wellenfunktionen Ψ , die durch

Lösung des Eigenwertproblems (Schrödingergleichung) erhalten werden, werden in Zusammenhang mit der Berechnung von Molekülen häufig auch „Molekülorbitale“ genannt. Um ihre Berechnung möglich zu machen, werden sie oft mittels LCAO-Näherung („linear combination of atomic orbitals“) durch Linearkombination aus Atomorbitalen (repräsentiert durch Basisfunktionen) zusammengesetzt. Für einfache Systeme, wie das Wasserstoffatom oder die wasserstoffähnlichen Atome He^+ , Li^{2+} , Be^{3+} , bei denen sich ein einziges Elektron im Feld einer positiven Punktladung (=Repräsentation des Atomkerns) bewegt, ist die Schrödinger-Gleichung analytisch geschlossen lösbar. Aber bereits für das He-Atom, das zwei Elektronen besitzt, ist dies aufgrund der entstehenden Elektronenwechselwirkung nicht mehr möglich. Daher wurden Methoden entwickelt, die aufgrund ihrer vereinfachenden Annahmen zumindest näherungsweise Lösungen ermöglichen. Auf diese wird im nachfolgenden Abschnitt näher eingegangen. Eine analytisch exakte Lösung der Schrödinger-Gleichung ist ebenfalls nicht möglich, wenn das System mehr als einen Atomkern besitzt. In solchen Fällen wird die Born-Oppenheimer-Näherung^[178] eingesetzt, die auf der Annahme beruht, dass die Bewegung der Atomkerne von der Bewegung der Elektronen getrennt betrachtet werden kann. Grundlage hierfür bildet der Masseunterschied zwischen Kern und Elektronen, der zur sofortigen Anpassung der Elektronen an die jeweilige Kernanordnung führt. Dieses Verhältnis ermöglicht die Betrachtung der Wellengleichung des molekularen Gesamtsystems als ein Produkt der Wellengleichung der Elektronen und der Wellengleichung der Kerne.

3.3.2 Methoden der Quantenmechanik

Systeme mit mehr als einem Elektron benötigen Vereinfachungen, um zumindest eine näherungsweise Lösung der Schrödinger-Gleichung zu ermöglichen. Die einfachste und grundlegendste Methode, die in diesem Zusammenhang zu nennen ist, ist die SCF-Näherung^[179;180] („self consistent field“). Hierbei wird die Wechselwirkung eines Elektrons i mit den verbleibenden Elektronen $j \neq i$ des Systems nicht als exakte Wechselwirkung zwischen Elektronen i und j betrachtet, sondern als Wechselwirkung des Elektrons i mit dem „gemittelten Feld“ aller Elektronen j . Auf jedes Elektron wirkt dann das Potential der (ortsfesten) Kerne und zusätzlich ein Potential, das von den verbleibenden Elektronen verursacht wird. Damit wird die Berechnung der Wellenfunktion und Energie auf ein Eielektronensystem reduziert. Eine vielfache Iteration über alle Elektronen des Systems führt meist zur Konvergenz dieser Energie, die somit als Näherung der Energie des gesamten Systems angenommen werden kann. Häufig wird das SCF-Verfahren auch als Hartree-Fock-Methode

bezeichnet. Ein Nachteil dieser strikten Vereinfachung ist jedoch, dass die Elektronenkorrelation – also die Reaktion der Elektronen auf die Bewegung des aktuell betrachteten Elektrons – vernachlässigt wird. Daher wurde eine Reihe von post-SCF-Methoden entwickelt, die durch Einführung von Korrekturtermen versuchen, den Fehler, der durch die vereinfachende Annahme der SCF-Näherung entsteht, auszugleichen. Neben verschiedenen anderen Ansätzen stellt die Störungstheorie hierfür eines der wichtigsten und am häufigsten verwendeten Prinzipien dar. Vorgeschlagen wurde diese Methode erstmals von Møller und Plesset^[181] (MP-Theorie) für die Anwendung auf Atome; eine Erweiterung für Moleküle erfolgte durch Pople *et al.*^[182]. In der Quantenchemie der Moleküle werden heute insbesondere Rechnungen MP2 (MP-Theorie zweiter Ordnung) für die Optimierung von Molekülstrukturen und MP4 (MP-Theorie vierter Ordnung) für die Verfeinerung von berechneten Energien eingesetzt.^[183]

Ein großer Nachteil dieser aus dem Hartree-Fock-Ansatz abgeleiteten Methoden (*ab-initio*-Molekülorbitaltheorie) besteht jedoch darin, dass sie auch bei kleinen Molekülen schnell sehr rechenaufwändig werden, insbesondere bei Einbeziehung der oben beschriebenen Elektronenkorrelation. Zusätzlich zeigen die Methoden in vielen Fällen Schwächen bei Berücksichtigung von Übergangsmetallen.^[183] Derartige Probleme treten bei Methoden der Dichtefunktionaltheorie (DFT) nicht auf. Hierbei werden nicht – wie bei den oben beschriebenen Ansätzen – Molekülorbitale (Wellenfunktionen) berechnet, sondern Elektronendichten. Ihren Ursprung finden diese Ansätze bereits 1927, ein Jahr nach der Veröffentlichung der Schrödinger-Gleichung, als Thomas^[184] und Fermi^[185] erstmals zeigten, dass die kinetische Energie eines chemischen Systems direkt durch seine Elektronendichte dargestellt werden kann. Im Jahr 1964 bewiesen Hohenberg und Kohn^[186], dass auch die Gesamtenergie (totale Energie) eines Systems ausschließlich durch seine Elektronendichte – also völlig unabhängig von der Wellenfunktion – beschrieben werden kann. Die Ansätze waren jedoch sehr komplex. Daher entstanden im Laufe der Zeit verschiedene Näherungen, um einzelne Energiebeiträge, wie die Elektronenkorrelationsenergie, aus der Elektronendichte zu erhalten. Jedoch gelang es nicht, auf diesem Weg eine ausreichend genaue Beschreibung der kinetischen Energie zu finden.^[187] Eine Lösung wurde durch Verwendung des kinetischen Energieterms aus der *ab-initio*-Molekülorbitaltheorie erreicht. Dieser Ansatz wird heute als Kohn-Sham-Dichtefunktionaltheorie^[186;188;189] bezeichnet; ein Beispiel für eine Methode ist LSDA^[190;191] („local spin density approximation“). Nach diesem ersten Schritt kam es – basierend auf Methoden der Kohn-Sham-Theorie – zu weiteren Einführungen von Termen aus der Molekülorbitaltheorie in die Dichtefunktionaltheorie: Der Hybrid-Ansatz entstand. Hierbei werden sogar einzelne Energieterme teils aus Termen der Molekülorbitaltheorie und teils aus Termen der Dichtefunktionaltheorie zusammengesetzt. Der relative Anteil der

jeweiligen Theorie wird durch Koeffizienten festgelegt, die an experimentellen Daten optimiert werden. Damit handelt es sich bei den Hybridmethoden nicht mehr um reine *ab-initio*-Methoden. Eines der bekanntesten und am stärksten genutzten Funktionale des Hybrid-Ansatzes ist B3LYP^[192] (Becke, 3 Parameter, Lee-Yang-Parr). Moderne Dichtefunktionalmethoden – auch B3LYP – berücksichtigen die Elektronenkorrelation. Insgesamt stellt die Dichtefunktionaltheorie einen guten Kompromiss zwischen Genauigkeit und Rechengeschwindigkeit dar. Jedoch hat auch diese Methodenklasse Schwächen, wie die häufige Unterschätzung von Aktivierungsenergien oder die oft eingeschränkte Reproduktion von schwachen Wasserstoffbrücken oder Van-der-Waals-Wechselwirkungen.^[183]

Neben *ab-initio*-Molekülorbital- und Dichtefunktionaltheorie liefert die Semiempirik die dritte Gruppe quantenchemischer Methoden. Wie der Name bereits andeutet, kommt es hierbei zur Verschmelzung von empirischen und *ab-initio*-Ansätzen. Grundsätzlich ähneln semiempirische Methoden der allgemeinen Struktur des Hartree-Fock-Ansatzes. Bestandteile der Gleichung werden dabei vereinfacht oder vollständig vernachlässigt. Zusätzlich werden oft nur Valenzelektronen betrachtet und nur minimale Basissätze (siehe Abschnitt 3.3.3) verwendet. Um den entstehenden Fehler auszugleichen, wird die Methode parametrisiert – dies macht den empirischen Teil der Methodik aus. Die Parameter werden dabei durch Angleichung der Resultate an experimentelle Werte oder *ab-initio*-Rechnungen erhalten. Der Vorteil dieses Ansatzes liegt in der Geschwindigkeit, die meist wesentlich höher ist, als bei *ab-initio*-Molekülorbital- oder DFT-Methoden. Der dafür zu zahlende Preis ist die Güte der berechneten Strukturen: Weicht das zu berechnende Molekül vom Parametrisierungsdatensatz ab, so können die Ergebnisse unbefriedigend ausfallen. Jedoch ist diese Abhängigkeit vom Parametrisierungsdatensatz hier nicht so stark ausgeprägt, wie bei Anwendung molekülmechanischer Methoden (siehe auch Abschnitt 3.3.6). Bekannte semiempirische Methoden sind beispielsweise PM6^[193] („parameterization method 6“) oder AM1^{*[194-199]} („Austin model 1*“). Eine Übersicht über weitere semiempirische Methoden kann den Literaturquellen^[183] und^[200] entnommen werden.

3.3.3 Basissätze

Neben einer Vielzahl von Methoden zur Beschreibung der Energie des Systems existieren auch verschiedenste Ansätze zur Repräsentation der Atomorbitale. Bereits 1930 schlug Slater^[201] Funktionen vor, die eine gute Repräsentation von Atomorbitalen darstellen. Die Gestalt dieser „Slater-Type Orbitals“ (STO) macht die quantenchemischen Rechnungen jedoch sehr schwierig und damit zeitaufwändig. Daher werden sie meist nicht direkt genutzt,

sondern es wird versucht, sie durch Linearkombinationen von Gauss-Funktionen zu approximieren, was eine einfache und damit schnelle Berechnung ermöglicht. Einen frühen Ansatz, der heute kaum noch genutzt wird, stellt der STO-3G-Basissatz^[202] („Slater-Type Orbitals approximated by 3 Gaussian functions“) dar. Modernere Basissätze lassen mehr Spielraum zur besseren Modellierung von Valenzorbitalen; ein typisches Beispiel ist der Basissatz 6-31G^[203], bei dem die inneren Orbitale durch sechs linearkombinierte Gauss-Funktionen und die Valenzorbitale durch drei linearkombinierte und eine separate Gauss-Funktion repräsentiert werden. Die separate Gauss-Funktion ermöglicht es, einen Aufenthalt der Valenzelektronen mit größerem Abstand vom Kern darzustellen, was der Bindung von Atomen in Molekülen Rechnung trägt. Jedoch setzt dieser Basissatz eine Gleichverteilung der Elektronen um den Kern voraus. Bei Betrachtung von Atomen in Molekülen ist diese Annahme jedoch nicht korrekt. Durch die Ausbildung von Bindungen oder Wechselwirkungen mit anderen Atomen kommt es zur Verschiebung von Elektronen zu einer Seite des Atoms. Eine Verwendung von Polarisierungsfunktionen als zusätzliche Basisfunktionen kann diesen Effekt modellieren. Dabei werden stets Basisfunktionen benötigt, die eine geringere Symmetrie als die zu polarisierenden Orbitale besitzen. Somit werden s-Orbitale mit p-Funktionen und p-Orbitale mit d-Funktionen polarisiert. Basissätze werden mit einem „*“ gekennzeichnet, wenn sie Polarisierungsfunktionen beinhalten, z.B. berücksichtigt 6-31G* (6-31G(d)) d-Funktionen für Nicht-Wasserstoffatome und 6-31G** (6-31G(d,p)) zusätzlich p-Funktionen für Wasserstoff.^[204;205] Bei Betrachtung von anionischen Systemen reichen die bisher beschriebenen Basissätze nicht aus. Es müssen zusätzliche Basisfunktionen hinzugefügt werden, die die höhere Aufenthaltswahrscheinlichkeit von Elektronen bei größeren Abständen vom Kern beschreiben können. Diese Basisfunktionen werden als diffuse Funktionen bezeichnet und mit einem „+“ gekennzeichnet (z.B. 6-31+G mit diffusen Funktionen für Nicht-Wasserstoffatome).^[206]

3.3.4 Solvationsmodelle

Viele der in der Literatur veröffentlichten quantenchemischen Rechnungen werden in Gasphase, d.h. im Vakuum, durchgeführt; meist um Rechenzeit und/oder -kapazität einzusparen. Dies entspricht jedoch in den allermeisten Fällen nicht den experimentell vorliegenden Bedingungen, bei denen im Allgemeinen Lösungsmittel beteiligt sind. Die natürlichste Methode zur Berücksichtigung von Lösungsmittelmolekülen in quantenchemischen Rechnungen ist ihre unmittelbare Einbeziehung in die Berechnung, die durch eine geeignete Platzierung des Lösungsmittels um das zu berechnende Molekül erreicht wird („Supermolekül“-Näherung). Dieses explizite Solvationsmodell hat jedoch verschiedenste Schwächen, die es in der

Praxis kaum einsetzbar machen. Zum Einen können nur wenige Lösungsmittelmoleküle platziert werden, um die benötigte Rechenzeit und -kapazität nicht zu stark zu erhöhen. Das weitaus größere Problem liegt jedoch in der „geeigneten Platzierung“ der Lösungsmittelmoleküle. Da einfache Strukturoptimierungen nur lokale Energieminima finden und damit die erhaltenen Ergebnisse von der Ausgangspositionierung der Moleküle abhängen, ist es notwendig, mehrere solcher Platzierungen zu erzeugen und zu rechnen. Dies führt zu einer zusätzlichen, erheblichen Steigerung der benötigten Rechenleistung. Um dennoch den Einfluss von Lösungsmittel in quantenmechanische Betrachtungen einbeziehen zu können, wurden implizite Solvationsmodelle entwickelt. Die bekanntesten Methoden sind hierbei solche, bei denen die Lösungsmittelmoleküle als Kontinuum mit zugehöriger Dielektrizitätskonstante modelliert werden. Dieser Ansatz schafft einen guten Kompromiss zwischen benötigter Rechenleistung und realitätsnaher Beschreibung der Molekülumgebung. Die verschiedenen Ansätze unterscheiden sich dabei im Wesentlichen nur in der Modellierung des Moleküls (z.B. Oberfläche, Polarität) und des Lösungsmittels, sowie in der Darstellung verschiedener Energiebeiträge. Eine der genauesten Methoden ist die SCRF-Methode („self consistent reaction field“), die mit ihrer Variation PCM^[207;208] („polarized continuum model“) auch in der vorliegenden Arbeit verwendet wird. Eine Einführung in die verschiedenen Solvationsmodelle, sowie weiterführende Literatur, kann Referenz^[209] entnommen werden.

3.3.5 Übergangszustände von Reaktionen

Im Allgemeinen werden quantenmechanische Methoden zur Optimierung einzelner Molekülstrukturen verwendet, es ist jedoch auch die Berechnung chemischer Reaktionen möglich. Diese werden dabei als ein „Pfad“ auf einer Energiehyperfläche modelliert („Reaktionskoordinate“). Sehr häufig ist man dabei auch interessiert an der Gestalt des Übergangszustandes (energiereicher, instabiler Zustand; ÜZ; quantenchemisch definiert als stationärer Punkt mit mindestens einem negativen Eigenwert der Matrix der partiellen zweiten Ableitungen der Energie) der Reaktion. Ein moderner Ansatz zur Berechnung solcher ÜZ-Strukturen ist die STQN-Methode^[210;211] („Synchronous Transit-guided Quasi-Newton method“). Hierbei werden zwei Molekülstrukturen vorgegeben: Es handelt sich um den Vorgänger V und Nachfolger N der energiereichsten berechneten Konformation M entlang der Reaktionskoordinate. Der Übergangszustand wird nun als eine Struktur entlang eines Pfades zwischen beiden Strukturen ermittelt. Bei den LST-Methoden („linear synchronous transit“) entspricht dieser Pfad einer linearen Interpolation zwischen V und N , während QST-Methoden („quadratic synchronous transit“) einen kurvigen Pfad nutzen. Zur Bestimmung der Startstruktur des ÜZ gibt es verschiedene Methoden. Sie kann als eine Durchschnittsstruktur von

V und N (QST2) berechnet werden oder aus den drei Konformationen V , N und M (QST3). Die Berechnung eines Sattelpunktes erster Ordnung – also eines Maximums in genau einer Richtung der Potentialenergiehyperfläche, die alle Pfade zwischen V und N beinhaltet – führt dann zum Übergangszustand.

3.3.6 Vergleich mit der Molekülmechanik

In der Literatur ist neben der Quantenmechanik eine weitere große Klasse von Methoden vertreten, die zur Molekülmechanik (Kraftfeldmethoden) zusammengefasst werden. Im Folgenden soll sehr kurz und nur einführend ein Vergleich beider Methodenklassen vorgenommen werden. Eine detaillierte Einführung in die Methoden der Molekülmechanik kann beispielsweise Literatur ^[212] entnommen werden.

Wie die Quantenmechanik erlaubt auch die Molekülmechanik Struktur- und Energieoptimierungen von Molekülen oder Molekülmengen. Sie berechnet dabei ebenfalls nur lokale Minima. Der wesentliche Unterschied lässt sich sofort bei Anwendung in der Geschwindigkeit beider Methodenklassen erkennen: Molekülmechanische Methoden sind meist sehr viel schneller als quantenmechanische Methoden. Die Ursache hierfür liegt im angewendeten Grundprinzip der Methoden. Während die Quantenmechanik während der Optimierungen Elektronen berücksichtigt, kommt es in der Molekülmechanik lediglich zur Betrachtung von Massepunkten der Atome. Dabei wird eine vorgegebene Kraftfeldfunktion („potential energy function“) minimiert, die Bindungslängen, Bindungswinkel, Torsionswinkel und nichtbindende Wechselwirkungen optimiert. Referenzwerte für die Parameter solcher Kraftfeldfunktionen werden aus experimentellen Datensätzen gewonnen. Damit sind molekülmechanische Methoden rein empirisch, im Gegensatz zu quantenmechanischen Methoden, die zumindest *ab-initio*-Anteile besitzen. Die Tatsache, dass es bei der Molekülmechanik nicht zur Berücksichtigung von Elektronen kommt, führt dazu, dass es durch diese Methoden nicht möglich ist, Reaktionen zu beschreiben. Ebenso ist eine exakte Molekülstruktur, die durch spezielle Orbitalüberlappungen nicht benachbarter Atome verursacht wird, mit Kraftfeldern nicht zu ermitteln. Auch die durch die Molekülmechanik gelieferte Energie kann nicht mit der totalen Energie der Quantenmechanik gleichgesetzt werden. Es handelt sich vielmehr um ein Maß dafür, wie stark die Ähnlichkeit der Bindungslängen, -winkel, Torsionswinkel und nichtbindenden Wechselwirkungen der optimierten Struktur zu den entsprechenden Parametern der experimentellen Referenzwerte ausgeprägt ist. Man erhält also relative Energien, die zwar einen Vergleich verschiedener Konformationen desselben Moleküls, aber nicht verschiedener Moleküle ermöglichen. Damit sind keine Rückschlüsse auf Reaktionsenergien

oder gar Aktivierungsbarrieren mittels der Molekülmechanik möglich. Ein häufig verwendetes Kraftfeld ist beispielsweise MMFF94^[213-217] („Merck Molecular Force Field“), das insbesondere für kleine, organische Moleküle parametrisiert wurde. Bei Betrachtung von Proteinen und Nukleinsäuren (ohne Liganden) findet beispielsweise das AMBER-Kraftfeld^[218-220] („Assisted Model Building with Energy Refinement“) Anwendung.

Die hohe Geschwindigkeit der Struktur- und Energieberechnung durch molekülmechanische Methoden, macht sie interessant für erste Optimierungsschritte oder die Berechnung großer Strukturdatenbanken. Für exakte Strukturen, Energien oder für die Beschreibung von Reaktionen muss jedoch auf die Quantenmechanik zurückgegriffen werden.

4 ERGEBNISSE UND DISKUSSIONEN

4.1 Vorstellung des Bindestellen-Datensatzes

Der in dieser Arbeit erzeugte und für nachfolgend beschriebene Analysen verwendete Datensatz beinhaltet insgesamt 429 nichtredundante Diphosphatbindestellen, von denen 15 aktivierend sind. Diese Zahl aktivierender Bindestellen erscheint in diesem Kontext zunächst gering, lässt sich aber unter Berücksichtigung der folgenden Punkte erklären. Dies sind:

- *oft erschwerte Kristallisierungsbedingungen und*
- *die natürliche Verbreitung aktivierender Proteine.*

Zum Einen sind aktivierte Substrate in vielen Fällen nicht kristallisierbar. Es kommt nach Aufnahme in der Proteinbindetasche meist sehr schnell zu enzymkatalysierten Umsetzungen. Um trotzdem die Bindung des Substrats beobachten zu können, ist zunächst die chemische Synthese möglichst substratähnlicher Inhibitoren notwendig, bevor mit dem Kristallisierungsprozess begonnen werden kann. In vielen Fällen kommen Thioester zum Einsatz, bei denen das an den organischen Rest gebundene Sauerstoffatom des Diphosphats durch ein Schwefelatom ersetzt ist. Dies führt durch verringerte Elektronegativitätsdifferenzen zur Herabsetzung der Reaktivität der aktivierten Bindung. Zum Anderen sind diphosphataktivierende Enzyme generell seltener als viele nichtaktivierende Proteine. So haben beispielsweise die nichtaktivierten NAD^+ - oder FAD -Liganden durch ihre Bedeutung als Kofaktoren eine sehr starke Verbreitung in verschiedensten Oxidoreduktasen. Aufgrund dieser Zusammenhänge ist auch in den folgenden Jahren nicht mit einer wesentlichen Verbesserung des Verhältnisses der Anzahl der Röntgenstrukturen aktivierender im Vergleich zu nichtaktivierenden Bindestellen zu rechnen.

In den folgenden Abschnitten soll der Datensatz zunächst als Ganzes und anschließend mit besonderem Fokus auf die enthaltenen aktivierenden Bindestellen beschrieben werden.

4.1.1 Liganden und Enzyme

Abbildung 4.1 zeigt die Verteilung von Enzymklassen und Ligandentypen abhängig von der Anzahl der durch das Diphosphat mit der umgebenden Bindetasche ausgebildeten Wechselwirkungen. Für einige wenige Bindestellen war die Zuordnung zu einer Enzymklasse nicht eindeutig möglich – beispielsweise aufgrund ungeklärter Enzymfunktion. Solche Bindestellen sind in Abbildung 4.1a nicht inbegriffen. Bei der Gruppe der NDP-Liganden in Abbildung 4.1b handelt es sich nicht um Produktstrukturen – diese wurden aus dem Datensatz ausgeschlossen – sondern vielmehr um Substratinhibitoren (z.B. bei Nukleosiddiphosphat-Zucker-Dehydratasen) oder um echte Substrate (z.B. bei Nukleosiddiphosphat-Kinasen). Besonders auffällig ist die starke Streuung der Zahl der Interaktionen, die pro Diphosphat ausgebildet werden. Dies hängt insbesondere mit der verwendeten Mindestauflösung der Röntgenstrukturen und der Nutzung von Schwellenwerten während der Berechnung der Wechselwirkungen zusammen. Schlechtere Auflösungen haben zur Folge, dass eine gewisse Unsicherheit bei der Positionierung von Gruppen in den Elektronendichtekarten vorliegt. Die darauf angewendeten „harten“ Schwellenwerte (siehe Abschnitt 3.1) können dann zur Nicht-Berücksichtigung von Gruppen führen, die nur geringfügig von einer für eine Wechselwirkung geeigneten Position abweichen. Geringere Zahlen im Bereich von zwei bis vier Interaktionen kommen nicht durch unspezifische Bindung der Liganden in der Bindetasche zustande – solche Bindestellen wurden bei Erzeugung des Datensatzes ausgeschlossen, sie treten vielmehr bei nichtaktivierten Liganden auf, die in erster Linie durch Wechselwirkungen ihrer organischen Reste in der Bindetasche verankert sind und weniger durch das Diphosphat selbst. Als maximal mögliche Anzahl an Interaktionen wird 16 ermittelt, obwohl ein Diphosphat nur sieben Sauerstoffe enthält. Die Ursache hierfür ist, dass alle Wechselwirkungen eines Sauerstoffatoms mit seiner Umgebung berücksichtigt wurden, die die im Abschnitt 3.1 beschriebenen Kriterien erfüllen. Dies können dann auch mehr als zwei Wechselwirkungen pro Sauerstoffatom sein. Die meisten Diphosphate bilden jedoch acht bis neun Wechselwirkungen aus. Bereits bei einer Analyse von nur 38 phosphatbindenden Proteinen im Jahr 1994 wurde eine starke Streuung in der Anzahl der durch das Phosphat ausgebildeten Wechselwirkungen festgestellt.^[221]

Weiterhin wird eine Ungleichverteilung in den Häufigkeiten sowohl der Ligandentypen, als auch der Enzymklassen deutlich. Besonders stark äußert sich dies in einem häufigen Auftreten von Diphosphatbindestellen in Oxidoreduktasen, was sich auch in der großen Anzahl von darin enthaltenen Kofaktoren NAD(P)H und FADH₂ widerspiegelt. In diesem Zusammenhang soll nochmals darauf hingewiesen werden, dass redundante Bindestellen weitestgehend aus dem Datensatz entfernt wurden (siehe dazu Abschnitt 3.1) und somit in den gezeigten Statistiken nicht dargestellt sind.

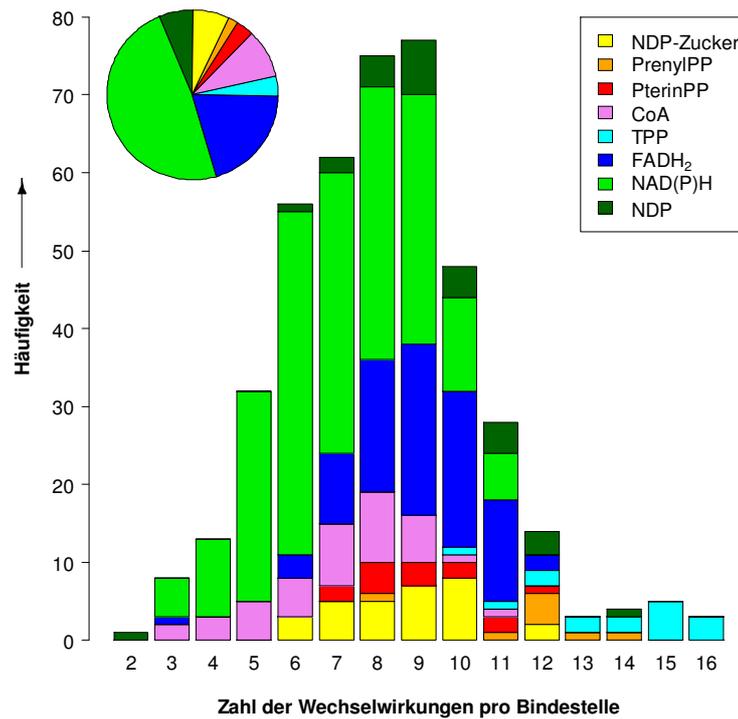
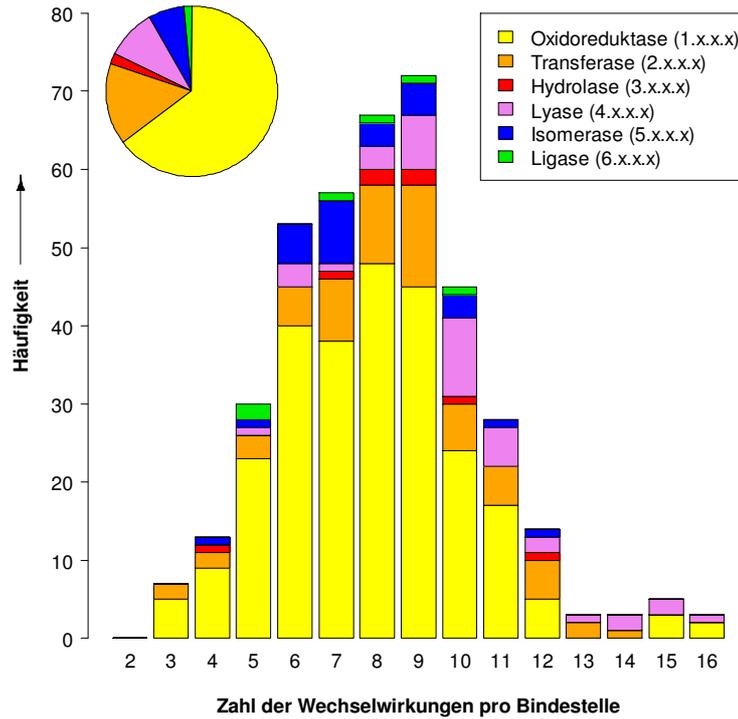


Abbildung 4.1: Darstellung der Verteilung von Enzymklassen (a; mit Angabe der EC-Nummer) und Ligandentypen (b) in Abhängigkeit von der Anzahl der durch das Diphosphat ausgebildeten Wechselwirkungen. Die Tortendiagramme geben jeweils die Gesamtverteilung der verschiedenen Enzymklassen bzw. Ligandentypen unabhängig von der Zahl der ausgebildeten Wechselwirkungen an. NDP-Zucker – Nucleosiddiphosphat-Zucker; PrenylPP – Prenyldiphosphat; PterinPP – Pterindiphosphat; CoA – Coenzym A; TPP – Thiamindiphosphat; FADH₂ – Flavinadenindinukleotid; NAD(P)H – Nikotinamadenindinukleotid(phosphat); NDP – Nucleosiddiphosphat

4.1.2 Aktivierende Bindestellen

Die im Datensatz enthaltenen aktivierenden Bindestellen können anhand des gebundenen Liganden in drei Gruppen eingeteilt werden. Dies sind:

- Glycosyltransferasen und Glycosidasen (mit Nukleosiddiphosphat-Zucker-Substrat)
- prenyldiphosphatkonvertierende Enzyme (mit Prenyldiphosphat-Substrat)
- Dihydropteroat-Synthase (mit Pterindiphosphat-Substrat)

Die nachfolgende Abbildung 4.2 zeigt allgemeine Schemata der Reaktionen dieser Enzyme. Tabelle 4.1 listet alle aktivierenden Bindestellen des Datensatzes, sowie einige wesentliche Eigenschaften auf.

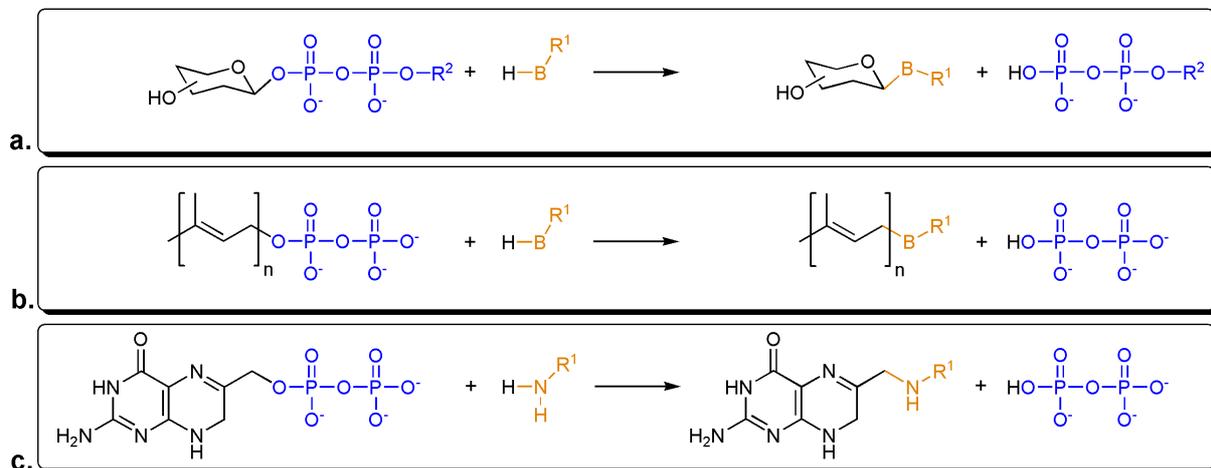


Abbildung 4.2: Allgemeine Reaktionsschemata der Glycosyltransferasen und Glycosidasen (a), der prenyldiphosphatkonvertierenden Enzyme (b), sowie der Dihydropteroat-Synthase (c). B – O, S, N, C_{nukleophil}; R¹ – H oder organischer Rest; R² – Nukleosid.

Tabelle 4.1: Aktivierende Bindestellen des Datensatzes sowie einige Eigenschaften. blau – Glycosyltransferasen und Glycosidasen; grün – prenyldiphosphatkonvertierende Enzyme; rot – pterindiphosphatspaltende Enzyme

PDB-Eintrag	Enzymname	EC-Nummer	Reaktionschema	notwendiges Metallion	natürlicher, aktivierter Diphosphatligand	Kokristall/Soaking	Ref.
1FOA	N-Acetylglucosaminyltransferase I	2.4.1.101	a	Mn ²⁺	UDP-N-Acetyl- α -D-Glucosamin	Kokristall	[222]
1LL2	Glycogenin	2.4.1.186	a	Mn ²⁺	UDP- α -D-Glucose	Soaking	[223]
2P72	Putative Glycosyltransferase	-	a	Mn ²⁺	unbekannt, vermutlich UDP- α -D-Glucose	Soaking	[224]
2RJ7	Galactosyltransferase	2.4.1.37	a	Mn ²⁺	UDP- α -D-Galactose	Soaking	[225]
2I8T	GDP-Mannose-Mannosyl-Hydrolase	3.2.1.-	a	Mg ²⁺	GDP- α -D-Mannose	Kokristall	[3]
1J39	DNA- β -Glycosyltransferase	2.4.1.27	a	-	UDP- α -D-Glucose	Soaking	[16]
1UQU	α , α -1,1-Trehalose-6-Phosphat-Synthase	2.4.1.15	a	-	UDP- α -D-Glucose	Kokristall	[226]
2C1Z	Flavonoid 3-O-Glycosyltransferase	2.4.1.91	a	-	UDP- α -D-Glucose	Kokristall	[227]
2IW1	α -1,3-Glycosyltransferase	2.4.1.-	a	-	UDP- α -D-Glucose	Kokristall	[228]
2VCE	Hydrochinon-Glycosyltransferase	2.4.1.218	a	-	UDP- α -D-Glucose	Kokristall	[229]
1X06	Undecaprenyldiphosphat-Synthase	2.5.1.31	b	Mg ²⁺	Oligoprenyldiphosphat (C ₁₀ bis C ₅₀)	Soaking	[57]
1JCR	Protein-Farnesyltransferase	2.5.1.58	b	Mg ²⁺ , Zn ²⁺	FPP (C ₁₅)	Soaking	[230]
1ZB6	Aromatische Prenyltransferase NphB	-	b	Mg ²⁺	GPP (C ₁₀)	Soaking	[84]
2VG0	E,Z-Farnesylidiphosphat-Synthase	2.5.1.68	b	Mg ²⁺	GPP (C ₁₀)	Kokristall	[231]
1AJ2	Dihydropteroat-Synthase	2.5.1.15	c	Mg ²⁺	6-Hydroxymethyl-7,8-Dihydropterindiphosphat	unbekannt	[137]

4.2 Clustering von Diphosphatbindestellen

Zur Analyse der Evolution aktivierender und nichtaktivierender Bindestellen, sowie als Mittel zur Detektion von Diphosphatbindemotiven wurde ein Clustering-Ansatz gewählt. Der hierfür verwendete Algorithmus ist eine dichtebasierte Methode (siehe Abschnitt 3.2.2). Hierbei werden ähnliche Bindestellen (bezüglich eines gewählten Distanzmaßes) gefunden und in einem Diagramm benachbart dargestellt. Bereiche größerer Dichte (Cluster) sind dabei durch Bereiche geringerer Dichte voneinander getrennt. Eine Anordnung von Bindestellen in einem gemeinsamen Cluster weist damit auf ähnliche Bindemotive hin, eine Anordnung in verschiedenen Clustern auf verschiedene Bindemotive. Damit würde eine gemeinsame Anordnung von Bindestellen in einem Cluster Hinweise auf einen evolutionären Zusammenhang liefern. Wären sie über verschiedene Cluster verteilt und damit die Ähnlichkeiten zu Bindestellen anderer Liganden größer, so wäre ein gemeinsamer Vorläufer eher unwahrscheinlich.

Der nachfolgende Abschnitt beschreibt zunächst die Repräsentation der Bindestellen während des Clusterings, sowie die verwendeten Parameter. Anschließend werden die Ergebnisse vorgestellt.

4.2.1 Repräsentation der Bindestellen und Wahl der Parameter

Um die Bindestellen des in Abschnitt 4.1 beschriebenen Datensatzes clustern zu können, ist die Wahl einer geeigneten Repräsentation notwendig. Hierfür bietet sich zunächst ein dreidimensionales Abbild der Bindestellen an. Dabei ist es wichtig, die Variabilität der absoluten Positionierung der bindenden Gruppen, die auch in sehr ähnlichen Enzymen auftreten kann (siehe Abschnitt 3.1.1), zu berücksichtigen. Einen möglichen Ansatz hierfür bietet eine Darstellung jeder Bindestelle durch ein Pharmakophor. Dies ist eine Menge von Kugeln mit definiertem Mittelpunkt, Radius, sowie einer Eigenschaft (z.B. H-Donor, Anion), um die wesentlichen Charakteristika eines Moleküls zu beschreiben. Durch eine Überlagerung jeder Bindestelle mit jedem berechneten Pharmakophor können paarweise Ähnlichkeiten erhalten werden, die als Basis für eine Einteilung des Datensatzes in Gruppen („Cluster“) dienen können. Die Verwendung von Pharmakophoren ermöglicht auch die Berücksichtigung der Austauschbarkeit bindender Gruppen mit ähnlichen chemischen Eigenschaften. Jedoch zeigten bereits erste Versuche mit einer Teilmenge des Datensatzes nur sehr unbefriedigende Ergebnisse bezüglich der Überlagerung von Pharmakophor und Bindestelle; es

war kaum möglich, eine eindeutige Positionierung der Bindestelle zum Pharmakophor zu finden. Als Ursache hierfür wird die starke Varianz der absoluten Position der diphosphatbindenden Gruppen in der Bindetasche vermutet. Um dennoch ein Clustering zu ermöglichen, erfolgte die Anwendung eines vereinfachten Systems, das weitgehend unabhängig ist von der exakten Positionierung der bindenden Gruppen, ähnlich wie „Fingerprints“. Dies sind binäre Vektordarstellungen von Molekülen, bei denen jedes Vektorelement eine bestimmte Eigenschaft des Moleküls betrachtet. Ist diese im Molekül vorhanden, so wird der Wert des entsprechenden Vektorelements auf 1 gesetzt; ansonsten erhält es den Wert 0. Analog hierzu kann auch jede Bindestelle in Form eines Vektors $v \in \mathbb{N}^M$ dargestellt werden, wobei jedes Vektorelement eine für die Diphosphatwechselwirkung geeignete funktionelle Gruppe repräsentiert (z.B. Hydroxylgruppe, Aminogruppe, Metallion); M sei dabei die Gesamtzahl der berücksichtigten funktionellen Gruppen (Wechselwirkungspartner). Der Wert jedes Vektorelements v_i ($i \in [1, M]$) entspreche der absoluten Häufigkeit eines Wechselwirkungspartners in der betrachteten Bindestelle. Es wurden die in Tabelle 4.2 angegebenen Gruppen berücksichtigt. Die Betrachtung von funktionellen Gruppen – statt vollständiger Aminosäuren – ermöglicht auch die Einbeziehung von Liganden, Metallionen oder Wassermolekülen als Wechselwirkungspartner, die funktionelle Gruppen von Aminosäuren substituieren können.

Tabelle 4.2: Wechselwirkungspartner. Es erfolgt eine Einteilung in vier Klassen: 1. Potentielle Protonendonoren (Brønsted-Säuren) und Metallionen (Lewis-Säuren), 2. Wasserstoffbrückendonoren, 3. Wasserstoffbrückenakzeptoren, 4. Wasser.

Bezeichnung	Beschreibung	Beispiele
Guanidinium	Guanidinogruppe	Arginin
Ammonium	nichtaromatisches N ⁺ -H, <i>nicht</i> Carbonsäureamid oder Guanidinogruppe	Lysin, <i>nicht</i> Arginin
Imidazolinium	protoniertes Imidazol	Histidin
O_Phenol	OH des Phenols	Tyrosin
Metall	Metallion	Mn ²⁺ , Mg ²⁺
N_Aromat	ungeladenes, aromatisches NH, aber <i>nicht</i> Imidazol	Tryptophan, <i>nicht</i> Histidin
N_Amid	NH der Carbonsäureamide	Amidbindung, Glutamin, FAD
O_Hydroxy	OH der Hydroxylgruppe, aber <i>nicht</i> Phenol	Threonin, Serin, <i>nicht</i> Tyrosin
S_Thiol	SH der Thiolgruppe	Cystein
Carboxyl	Carboxylgruppe oder Gruppen, bei denen C gegen S, N oder P ausgetauscht ist	Aspartat, Sulfonat
O_Carbonyl	O der Carbonylgruppe, aber <i>nicht</i> Carboxylgruppe	Amidbindung, Glutamin, FAD
Ether	O, S des Ethers	Zucker, Methionin
Wasser	Wassermolekül	-

Die Tabelle zeigt eine Zusammenfassung der Wechselwirkungspartner zu vier größeren Klassen: die Gruppe der potentiellen Protonendonoren und Metallionen (nachfolgend als „sauer“ bezeichnet), der Wasserstoffbrückendonoren (H-Donor), der Wasserstoffbrücken-

akzeptoren (H-Akzeptor) und die Gruppe der Wassermoleküle. Für die Brønstedsäuren der sauren Gruppe kann gezeigt werden, dass sie tatsächlich in der Lage sind, Diphosphat zu protonieren (siehe Abschnitt 4.4.3). Eine eindeutige Klassenzuordnung ist jedoch für einige der chemischen Gruppen nicht unproblematisch, insbesondere wenn sie stark abhängig vom Protonierungsgrad der betrachteten Gruppe ist, Wasserstoffatome jedoch in Röntgenstrukturen der Proteine aufgrund der unzureichenden Auflösung nicht sichtbar sind. Für Imidazole (Histidin) treten beispielsweise positiv geladene (protonierte) und neutrale Formen auf. Aufgrund der Nähe zum (oft negativ geladenen) Diphosphat wird die Gruppe in der vorliegenden Arbeit stets als protoniert angenommen. Imidazole können jedoch auch nah am während der Aktivierung entstehenden Kation positioniert (z.B. PDB-Eintrag 2vg2^[231]) oder an der Komplexbildung von Metallionen (z.B. PDB-Eintrag 1gx1^[232]) beteiligt sein. In solchen (Einzel-)Fällen ist die protonierte Form nicht zu erwarten. Eine weitere schwierig zuzuordnende Gruppe ist die Carboxylgruppe. Es erfolgte daher eine manuelle Beurteilung des Protonierungszustands. Diese zeigte, dass die Gruppen primär deprotoniert, also negativ geladen und damit nicht als Brønsted-Säure aktiv sind. Lediglich für 7 der insgesamt 39 Carboxylgruppenhaltigen Bindestellen des Datensatzes wäre eine Protonierung des Diphosphats durch die Gruppe denkbar. Meist sind die Carboxylgruppen an der Komplexbildung von Metallionen beteiligt. In wenigen Fällen ist die Gruppe so orientiert, dass eine Wechselwirkung mit anderen Gruppen des aktiven Zentrums (nicht mit dem Diphosphat) am wahrscheinlichsten ist. Bemerkenswert ist außerdem, dass sich die sieben genannten Fälle der möglichen Protonierung des Diphosphats durch die Carboxylgruppe ausschließlich auf nichtaktivierende Bindestellen beziehen. Aufgrund dieser Ergebnisse werden Carboxylgruppen in dieser Arbeit stets als deprotoniert – also H-Akzeptoren – betrachtet.

Abbildung 4.3 zeigt beispielhaft die Vektorrepräsentation v einer Bindestelle unter Berücksichtigung der in Tabelle 4.2 genannten Wechselwirkungspartner. Die Erzeugung dieser Vektoren für alle Bindestellen bildet die Basis für die Berechnung von paarweisen Distanzen, die nachfolgend als Grundlage für ein Clustering dienen. Um eine Berücksichtigung der Austauschbarkeit von bindenden Gruppen ähnlicher chemischer Eigenschaften zu garantieren, ist es notwendig, die entsprechenden Vektorelemente zusammenzufassen. Dies geschieht durch Addition der Werte. Somit erfolgt die Reduktion von v ausgehend von 13 Dimensionen zu drei: die saure Gruppe, Wasserstoffbrückendonoren und Wasserstoffbrückenakzeptoren; Wechselwirkungen mit Wassermolekülen werden aufgrund der Unsicherheit ihrer Positionierung in der Bindetasche (siehe auch Abschnitt 3.1.1) vernachlässigt. Die Zusammenfassung der Vektorelemente kann dabei durch Verschmelzung von Clustern zu einer Vergrößerung einzelner Klassen führen. Zusätzlich werden auch alle Wechselwirkungen mit Carboxylgruppen vernachlässigt, da ihre Funktion meist die Kom-

plexierung von Metallionen ist und sie damit keine direkten Wechselwirkungen mit dem Diphosphat ausbilden (Abschnitt 4.4.2). Damit wird die Vektorrepräsentation der in Abbildung 4.3 gezeigten Bindestelle zu $v = (2,2,0)$.

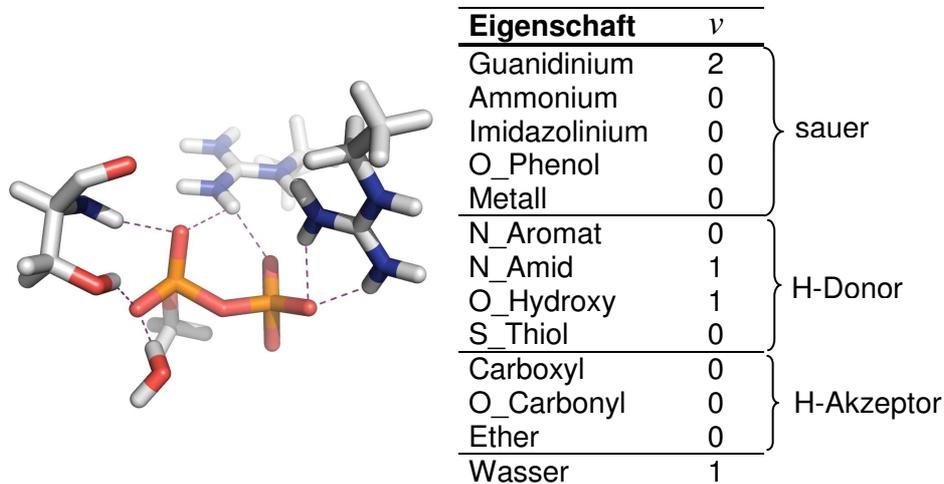


Abbildung 4.3: Beispielhafte Darstellung einer Diphosphatbindestelle mit Ligand, sowie die zugehörige Ausprägung der Vektorrepräsentation v . Die Wechselwirkungspartner des Diphosphats sind in Gruppen eingeteilt: potentielle Protonendonoren und Metallionen (sauer), Wasserstoffbrückendonoren (H-Donor), Wasserstoffbrückenakzeptoren (H-Akzeptor) und Wassermoleküle. Der Diphosphatligand ist durch ein Methyldiphosphat repräsentiert, zwei Argininreste durch 1-Ethylguanidinium. Weiterhin sind ein Serin und ein Wassermolekül gezeigt. Wechselwirkungen mit dem Liganden sind in violett angedeutet.

Zur Berechnung paarweiser Distanzen aus diesen vereinfachten Vektorrepräsentationen der Bindestellen, ist die Wahl eines geeigneten Distanzmaßes notwendig. Aus Abbildung 4.1, Abschnitt 4.1.1 ist bekannt, dass die Gesamtzahl der diphosphatbindenden Gruppen pro Bindestelle stark variieren kann. Eine mögliche Lösung wäre die Aufteilung des Datensatzes in Teildatensätze gleicher Zahl bindender Gruppen^{vii} und ein sich anschließendes getrenntes Clustering. Dies würde jedoch – ohne eine aufwändige Nachbearbeitung und manuelle Zusammenfassung ähnlicher bzw. identischer Cluster – nicht berücksichtigen, dass sehr ähnliche Bindestellen auch geringfügig in der Zahl der bindenden Gruppen variieren können (dies gilt insbesondere für Bindestellen mit einer großen Anzahl bindender Gruppen). Eine mögliche Lösung bietet hier die Skalierung von Distanzen unter Einbeziehung der Ausprägung der zu vergleichenden Vektoren. Hierzu wurden die Bray-Curtis- und die Canberra-Distanz in Abschnitt 3.2.1 vorgestellt. Diese Metriken berücksichtigen durch die Skalierung zusätzlich, dass ein Austausch oder Verlust einer bestimmten Anzahl k von Wechselwirkungspartnern in Bindestellen mit generell vielen Diphosphatwechselwirkungen weniger ins

^{vii} entspricht gleicher Vektorsumme $s(v) = \sum_{i=1}^M v_i$

Gewicht fällt, als von k Wechselwirkungspartnern in Bindestellen, in der die Zahl der Wechselwirkungen des Diphosphats ohnehin gering ist.

Die mittels dieser Metriken erhaltenen paarweisen Distanzen dienen als Grundlage für das Clustering der Bindestellen. Verwendet wird der in Abschnitt 3.2.2 vorgestellte Algorithmus OPTICS. Hierzu wird auch die Angabe der Parameter ε (definiert einen Radius) und $MinPts$ (definiert eine Zahl an Datenpunkten in einer Hypersphäre mit Radius ε um einen Datenpunkt) benötigt. Mit Hilfe der Vorgaben aus Abschnitt 3.2.2 werden geeignete Werte gewählt: Nachfolgend gilt $\varepsilon = 1,25$ und $MinPts = 15$, um die Darstellung aller Clusterebenen zu gewährleisten.

4.2.2 Vorstellung der Ergebnisse

Abbildung 4.4 zeigt die Ergebnisse des Clusterings des Datensatzes mit den Distanzmetriken Bray-Curtis und Canberra. Die Werte der Bray-Curtis-Distanz liegen *p.d.* stets zwischen 0 und 1 (siehe Abschnitt 3.2.1). Das Clustering mittels dieses Distanzmaßes (siehe Abbildung 4.4a) führt zur Ausbildung eines einzelnen Clusters; die Erreichbarkeitsdistanz ist i.A. nicht größer als 0,2. Durch die sehr starke Vereinfachung der Bindestellenrepräsentation zu dreidimensionalen Vektoren kann es – wie bereits angemerkt – zur Verschmelzung von Clustern kommen. Um auszuschließen, dass dies die Ausbildung des einzigen beobachteten Clusters verursacht, wurde eine weitere Zusammenfassung gewählt, bei der neben der Variablen „sauer“ auch „N_Aromat“, „N_Amid“, „O_Hydroxy“, „S_Thiol“, „O_Carbonyl“ und „Ether“ verwendet wurden. D.h. die einzelnen Vertreter der Klassen der Wasserstoffbrückendonoren und -akzeptoren wurden getrennt betrachtet. Eine Aufspaltung der Gruppe der potentiellen Protonendonoren und Metallionen erfolgte nicht, um ein Trennung der Brønsted- und Lewissäuren im resultierenden Clustering zu verhindern. Es entstand ebenfalls nur ein Cluster mit sehr hoher Dichte der Datenpunkte (Daten nicht gezeigt). Damit ist die Ausbildung eines einzelnen Clusters nicht auf die starke Zusammenfassung der Variablen zurückzuführen. Zusätzlich zeigte auch die Verteilung der Bindestellen der aktivierten und nichtaktivierten NDP-Zucker, Prenyldiphosphate, sowie der Pterindiphosphats ein Abbildung 4.4a sehr ähnliches Muster. Um eine durch die Skalierung verursachte Verschmierung von Clustern auszuschließen, erfolgte die Anwendung von OPTICS auch unter Verwendung der Manhattan-Distanz (Bray-Curtis ist eine skalierte Manhattan-Distanz, siehe Abschnitt 3.2.1). Auch in diesem Fall ist die Ausbildung lediglich eines dichten Clusters erkennbar, bei dem sich die einzelnen Datenpunkte um maximal den Wert 1 unterscheiden. Dies entspricht

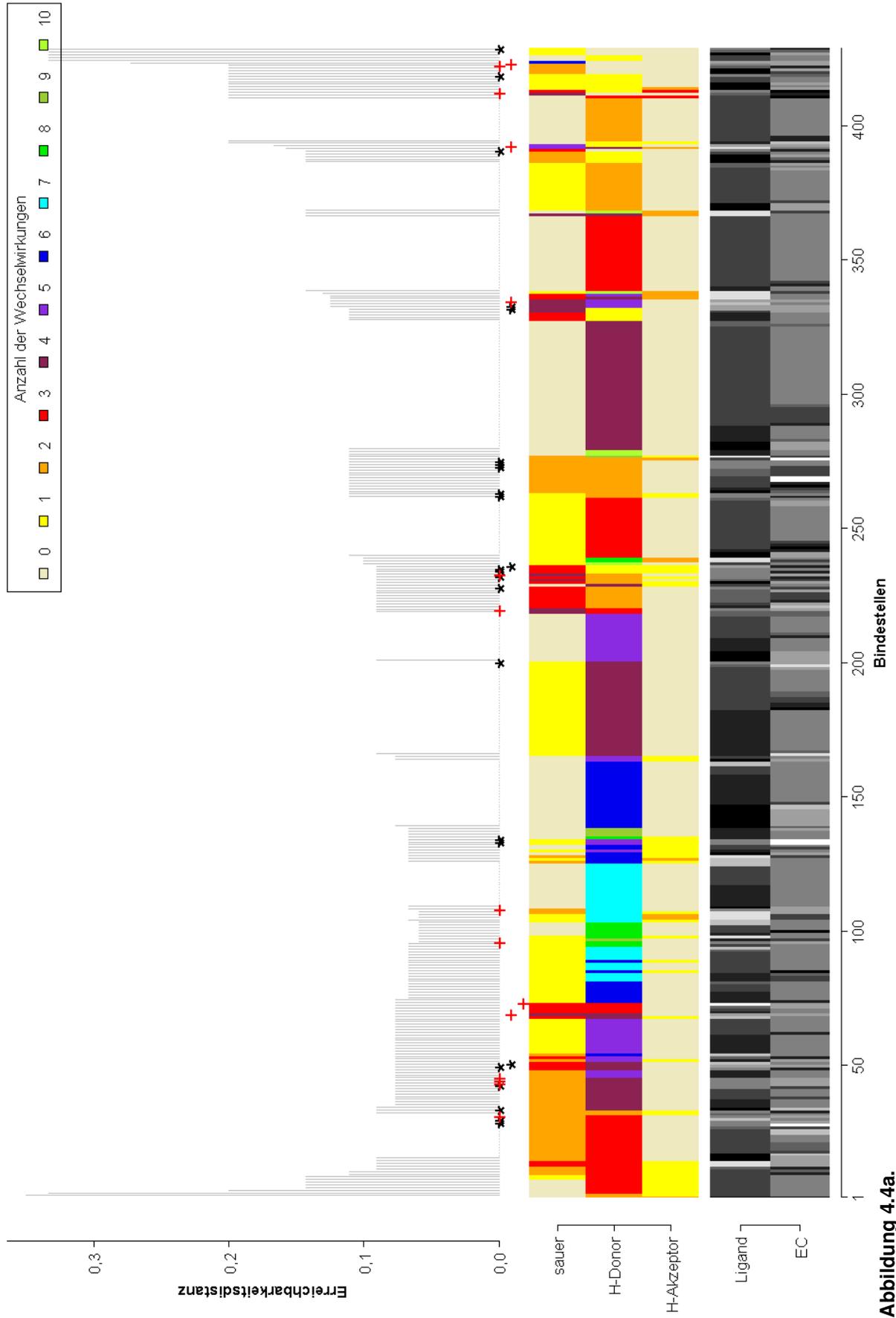


Abbildung 4.4a.

Abbildung 4.4b.

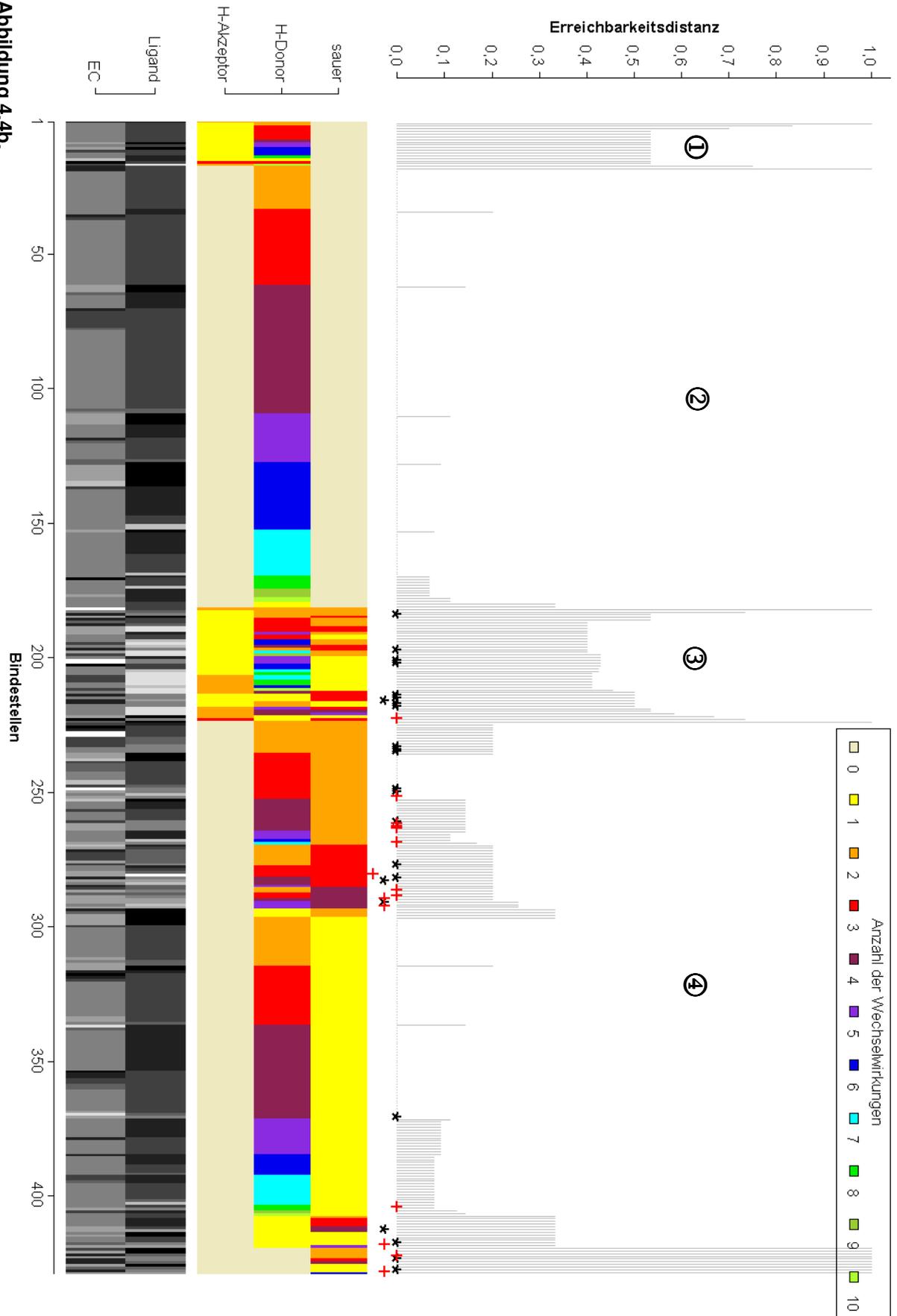


Abbildung 4.4 (Seiten 55 und 56): Darstellung der Ergebnisse des Clusterings mit OPTICS unter Verwendung der Bray-Curtis-Distanz (a.) bzw. der Canberra-Distanz (b.). Die Auftragung der Erreichbarkeitsdistanz für jede Bindestelle ist ein Maß für die Datendichte der unmittelbaren Umgebung des Datenpunktes. Direkt unter der Darstellung der Erreichbarkeitsdistanz ist in drei Reihen die Positionierung der Diphosphatbindestellen aktivierter (rot, +) und nicht-aktivierter (schwarz, *) NDP-Zucker (Reihe 1), Prenyldiphosphate (Reihe 2) und Pterindiphosphate (Reihe 3) innerhalb des Histogramms markiert. Die unter dem Histogramm gezeigte Farbkarte gibt die Werte der für das Clustering verwendeten Variablen (sauer, H-Donor, H-Akzeptor) farbkodiert an; die in der Darstellung oben rechts befindliche Legende zeigt die zu jeder Farbe zugehörigen Zahlenwerte. Die Art des gebundenen Liganden und EC-Nummer jeder einzelnen Bindestellen sind in Grautönen abgebildet, um ihre Verteilung innerhalb der mit OPTICS berechneten Clustering-Reihenfolge darzustellen. Die unter Verwendung der Canberra-Distanz entstehenden Cluster sind mit den Zahlen eins bis vier markiert.

beispielsweise einer Unterscheidung der Anzahl der H-Donoren um 1, wobei die verbleibenden Variablen identische Werte annehmen.

Bei Betrachtung der Auftragung der Erreichbarkeitsdistanz unter Verwendung der Canberra-Distanz (Abbildung 4.4b) zeigt sich klar die Ausbildung von zwei Clustern sehr hoher (Cluster 2) bzw. hoher (Cluster 4) Dichte (geringe Erreichbarkeitsdistanz) und zwei Clustern geringerer Dichte (Cluster 1 und 3; hohe Erreichbarkeitsdistanz). Ursache für die Entstehung der Cluster, die bei Verwendung der Bray-Curtis-Distanz nicht vorliegen, ist der Boole'sche Charakter der Canberra-Distanz (siehe Abschnitt 3.2.1). Die Bindestellen beider Cluster hoher Dichte (Cluster 2 und 4) beziehen keine H-Akzeptoren in die Bindung des Diphosphats ein; sie unterscheiden sich lediglich durch die Beteiligung bzw. den Ausschluss saurer Gruppen. Dies gilt auch für beide Cluster geringerer Dichte (Cluster 1 und 3). Allerdings beziehen diese stets H-Akzeptoren in die Diphosphaterkennung ein. Es wird deutlich, dass aktivierende Bindestellen stets in Clustern vorliegen, die durch die Beteiligung potentieller Protonendonoren oder Metallionen an der Diphosphatbindung gekennzeichnet sind. Interessanterweise sind auch die nichtaktivierenden Bindestellen mit den in aktivierenden Bindestellen vorkommenden Ligandentypen (siehe Abschnitt 4.1) ausschließlich in solchen Bereichen zu finden. Gleichzeitig wird aber auch offenbar, dass bei einer großen Zahl nicht-aktivierender Bindestellen keine Beteiligung von Brønsted- und Lewis-Säuren vorliegt (Cluster 1 und 2). Dies lässt den Schluss zu, dass die Beteiligung der sauren Gruppe an der Diphosphaterkennung ein für die Aktivierung notwendiges, aber – als Ganzes betrachtet – kein hinreichendes Kriterium darstellt. Möglicherweise wird jedoch durch die gezielte Betrachtung einzelner Vertreter dieser Gruppe eine Unterscheidung zwischen aktivierenden und nichtaktivierenden Bindestellen möglich, was in den nachfolgenden Abschnitten näher beleuchtet wird. Die Notwendigkeit saurer Gruppen deckt sich auch mit den Ergebnissen zur Hydrolyse von allylischen Diphosphaten, bei der ebenfalls eine Diphosphataktivierung stattfindet und die stark pH- und metallionenabhängig ist (siehe Abschnitt 1.3.1). In beiden Diagrammen nehmen größere Bereiche der Erreichbarkeitsdistanzen den Wert 0 an. Ursache hierfür ist die starke Reduzierung der Bindestellenrepräsentation auf lediglich die

drei Variablen „sauer“, „H-Donor“ und „H-Akzeptor“, wodurch identische Vektorrepräsentationen entstehen können. Es wird deutlich, dass in beiden Diagrammen – also auch bei Wahl verschiedener Distanzmetriken – eine starke Streuung der aktivierenden Bindestellen und ihrer nichtaktivierenden Gegenspieler vorliegt. Rückschlüsse zur gemeinsamen Entwicklung von aktivierenden und nichtaktivierenden NDP-Zucker- bzw. Prenyldiphosphatbindestellen sind daher nicht möglich. Ebenso kann nicht auf eine getrennte Entwicklung aktivierender und nichtaktivierender Bindestellen gleichen Ligandentyps geschlossen werden, da keine klare Trennung dieser Bindestellen innerhalb der Erreichbarkeitsdistanzen vorliegt. Unmittelbare Rückschlüsse auf die Evolution aktivierender Bindestellen sind damit nicht möglich. Auch ist keine klare Clusterbildung aktivierender Bindestellen gleichen Ligandentyps erkennbar; Bindemotive können daher aus den vorliegenden Daten nicht ermittelt werden. Bei Betrachtung der Verteilung von Ligandentyp (siehe Abbildung 4.1 zu den im Datensatz enthaltenen Ligandentypen) bzw. EC-Nummer ist ebenfalls eine starke Streuung erkennbar. Dies zeigt die starke Varianz der die Bindestellen charakterisierenden Vektoren und weist auf eine mangelnde Eignung dieser Darstellung für die Detektion evolutionärer Zusammenhänge oder Bindemotive hin.

4.2.3 Zusammenfassung und Diskussion

Das Kapitel beschreibt das Clustering von Diphosphatbindestellen mit dem Ziel der Detektion von Diphosphatbindemotiven, sowie evolutionärer Zusammenhänge aktivierender und nichtaktivierender Bindestellen. Ein 3D-Ansatz auf Basis von Pharmakophoren führte bereits in ersten Versuchen zu unbefriedigenden Ergebnissen. Daher erfolgte eine Reduzierung der Bindestellenrepräsentation auf „Fingerprint“-ähnliche Vektoren, die mittels des Algorithmus OPTICS unter Verwendung der Distanzmetriken Bray-Curtis und Canberra geclustert wurden. Dabei konnten Hinweise geliefert werden, dass die Beteiligung von Brønstedtsäuren und lewis-sauren Metallionen an der Diphosphatbindung ein notwendiges, aber kein hinreichendes Kriterium für die Aktivierung darstellt: Alle aktivierenden Bindestellen des dieser Arbeit zugrundeliegenden Datensatzes enthalten Vertreter der sauren Gruppe, doch gilt dies ebenfalls für viele nichtaktivierende Bindestellen, insbesondere wenn Prenyldiphosphat- oder NDP-Zucker-Liganden gebunden werden. Es zeigte sich weiterhin, dass die für das Clustering verwendete Vektordarstellung der Bindetasche zu wenig detailliert ist, um klare Bindemotive aktivierender Bindestellen zu detektieren oder Rückschlüsse auf evolutionäre Zusammenhänge ziehen zu können. Die Ursache hierfür liegt aber nicht ausschließlich in der Art und Weise der Repräsentation der Bindestellen, sondern vor Allem in Mängeln der zu-

grundlegenden Röntgenstrukturen (Ungenauigkeit) und in der stark variierenden Zahl von Wechselwirkungspartnern des Diphosphats (siehe Abbildung 4.1).

Wo die Vektorrepräsentation der Bindestellen zu grob ist, scheiterte das dreidimensionale Clustering (Pharmakophoransatz) der Bindestellen an einer zu detaillierten Beschreibung der Bindestellen, die eine automatisierte Überlagerung verhinderte. Es lässt sich vermuten, dass die Diphosphatbindung – und insbesondere die Aktivierung von Diphosphaten – nicht nur durch die Art der bindenden Gruppen, sondern auch durch ihre räumliche Anordnung und somit auch die Konformation des Diphosphats in der Bindetasche gesteuert wird. Daher wird in den nachfolgenden Kapiteln die Aktivierung mit einfachen Statistiken in Kombination mit quantenchemischen Methoden, die zur Prüfung aufgestellter Hypothesen dienen sollen, detailliert untersucht – auch im Hinblick auf die Diphosphatkonformation. Dabei wird insbesondere der unmittelbare Vergleich möglichst ähnlicher aktivierender und nichtaktivierender Bindestellen angestrebt. Das nachfolgende Kapitel stellt dabei zunächst den quantenchemischen Ansatz dar, der den verbleibenden Kapiteln als Grundlage dient.

4.3 Vorstellung der quantenchemischen Vorarbeiten

Alle in dieser Arbeit durchgeführten quantenchemischen Berechnungen basieren auf Dimethylallyldiphosphat (DMAPP, siehe Abbildung 4.5).

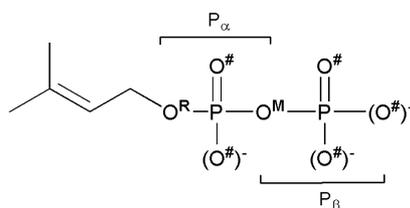


Abbildung 4.5: Struktur der deprotonierten Form des Dimethylallyldiphosphats (3-Methyl-2-butenyldiphosphat), sowie Kennzeichnung von P_α und P_β und der Sauerstoffatome $O^\#$, O^R und O^M .

Dieses Molekül wurde aufgrund zweier seiner Eigenschaften ausgewählt. Es handelt sich um:

- einen zu aktivierenden Liganden und
- ein kleines Molekül.

DMAPP ist der zu aktivierende Ligand in vielen prenyldiphosphatkonvertierenden Enzymen – einer Gruppe von Proteinen, die auch Bestandteil des in dieser Arbeit betrachteten Datensatzes ist. Diese Enzyme aktivieren neben DMAPP auch strukturell ähnliche Moleküle, wie Geranyl- oder Farnesyldiphosphat (Abbildung 4.6). Damit ist DMAPP Bestandteil aller zu aktivierenden Diphosphate dieser Enzymgruppe und kann somit als Repräsentant gewählt werden. Zum Anderen ist DMAPP ein kleines Molekül, das bei deprotoniertem Diphosphat lediglich aus 23 Atomen besteht. Dies macht es für die sehr zeit- und speicheraufwändigen quantenchemischen Berechnungen interessant und ermöglicht auch die Berechnung von Reaktionen, wie sie in Abschnitt 4.6 gezeigt sind.

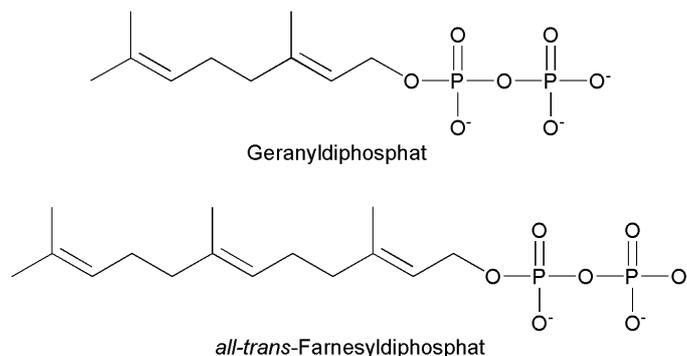


Abbildung 4.6: Strukturen des deprotonierten Geranyl- und all-trans-Farnesyldiphosphats.

In den folgenden Abschnitten werden die verschiedenen Formen des DMAPP, die für die Rechnungen eingesetzt wurden, sowie der verwendete Basissatz und die quantenchemische Methode vorgestellt. Abschließend werden die erzeugten Strukturen mit Röntgenstrukturen verglichen.

4.3.1 Protonierungsformen des Dimethylallyldiphosphats

Aufgrund der Ununterscheidbarkeit von Sauerstoffen des Diphosphats, die lediglich an ein einziges Nichtwasserstoffatom gebunden sind (nachfolgend als O[#] bezeichnet, siehe auch Abbildung 4.5), sowie verschiedener möglicher Protonierungszustände sind grundsätzlich sechs Formen des DMAPP denkbar. Dies sind eine Form mit dreifach (vollständig) protoniertem Diphosphat, jeweils zwei Formen mit zwei- bzw. einfach protoniertem Diphosphat und eine Form mit vollständig deprotoniertem Diphosphat. Dies sind jedoch lediglich theoretische Betrachtungen – anhand der pK_a-Werte des Diphosphats des DMAPP können die natürlich auftretenden Protonierungsformen ermittelt werden. Da nach bestem Kenntnis des Autors keine experimentellen Daten verfügbar sind, wurden die pK_a-Werte für DMAPP (Tabelle 4.3), sowie der Protonierungsreihenfolge mit Hilfe von ACD/pK_a DB[®] 10.01^[233] berechnet.

Tabelle 4.3: Mittels ACD/pK_a DB[®] berechnete pK_a-Werte des DMAPP.

$$pK_{a1} = 6,90 \pm 0,53$$

$$pK_{a2} = 1,93 \pm 0,10$$

$$pK_{a3} = 1,08 \pm 0,50$$

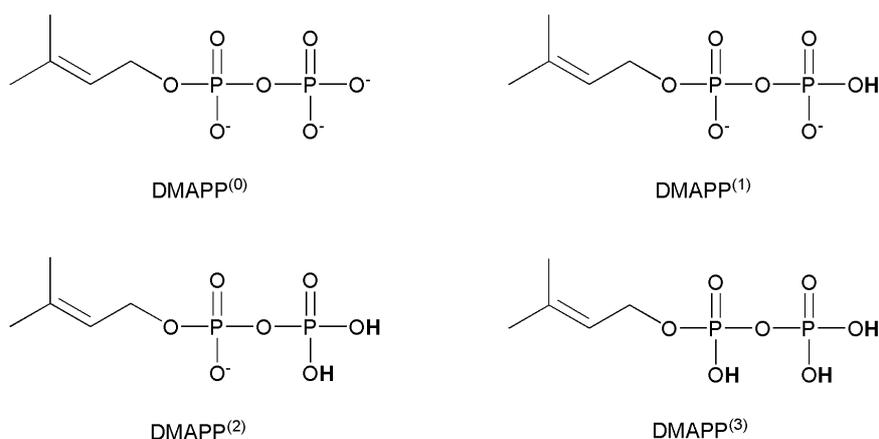


Abbildung 4.7: Berücksichtigte Protonierungsformen des Dimethylallyldiphosphats.

Anhand dieser Ergebnisse ergeben sich vier Protonierungsformen für DMAPP, die in Abbildung 4.7 schematisch dargestellt sind und die Grundlage aller durchgeführten quantenchemischen Rechnungen dieser Arbeit bilden. Die in der Abbildung verwendeten Bezeichnungen DMAPP⁽⁰⁾ bis DMAPP⁽³⁾ referenzieren im Folgenden auf die entsprechenden Strukturen.

4.3.2 Erzeugung von Ausgangsstrukturen und Etablierung von quantenchemischer Methode und Basissatz

Die Tatsache, dass Strukturoptimierungen mit quantenchemischen Methoden lediglich lokale Energieminima detektieren können, die nicht zwangsläufig mit dem (bzw. den) globalen Minimum (bzw. Minima) übereinstimmen müssen, macht es notwendig, die Optimierungen mit geeigneten Ausgangskonformationen zu starten. Hierzu wurde die Methode der systematischen Suche gewählt, bei der rotierbare Bindungen eines Moleküls um einen vorgegebenen, konstanten Winkel im Intervall ($0^\circ, 360^\circ$] gedreht werden. Dies erzeugt eine Menge von Strukturen mit fixierten Winkelkombinationen aller rotierbaren Bindungen. Eine nachfolgende Energieoptimierung dieser Strukturen liefert eine Reihe von lokalen Energieminima, aus denen das globale Minimum durch Wahl der energieärmsten (stabilsten) Konformation gewonnen werden kann. Häufig ist die Zahl der erzeugten Konformationen so groß, dass eine quantenchemische Energieminimierung unpraktikabel wird. In solchen Fällen werden molekülmechanische Methoden (Kraftfelder) genutzt, um die Energien der Konformationen abschätzen zu können. In diesem Zusammenhang muss jedoch berücksichtigt werden, dass die durch Kraftfelder erhaltenen Energien nicht mit den totalen Energien der quantenchemischen Methoden vergleichbar sind. Es handelt sich vielmehr um relative Energien, die zwar einen Vergleich verschiedener Konformationen desselben Moleküls, aber nicht verschiedener Moleküle ermöglichen (siehe auch Abschnitt 3.3.6).

In der vorliegenden Arbeit wurde die systematische Suche unter Rotation aller rotierbaren Bindungen anhand des vollständig protonierten DMAPP⁽³⁾ durchgeführt. Die Rotation zur Erzeugung der Konformationen erfolgte in 60° -Schritten. Zur Energieminimierung der Strukturen wurde das Kraftfeld MMFF94x^[213-217] verwendet, das besonders gut für kleine, organische Moleküle geeignet ist. Die Berechnungen erfolgten zunächst in Gasphase, d.h. unter Vernachlässigung eines Lösungsmittels. Die Konformationen der verbleibenden Protonierungsstufen des DMAPP wurden aus der Struktur mit minimaler Kraftfeldenergie durch Entfernen von Protonen erzeugt. Wie bereits angesprochen (siehe Abschnitt 3.3.6), sind molekülmechanische Methoden – im Gegensatz zu quantenmechanischen Methoden –

nicht in der Lage, Orbitalwechselwirkungen zu berücksichtigen. Daher wurden die MO-Methoden HF und MP2, sowie die Hybrid-Methode B3LYP (siehe Abschnitt 3.3.2) auf die mit dem Kraftfeld voroptimierten Strukturen angewendet. Wie bereits bei der systematischen Konformationssuche erfolgte hierbei zunächst eine Berechnung in Gasphase. Die Vernachlässigung des Lösungsmittels zwingt das Molekül zur ausschließlichen Ausbildung intramolekularer Wechselwirkungen. Da diese in ihrer Zahl beschränkt sind, kann die Qualität der entstehenden Struktur und damit die Eignung der quantenchemischen Methode leicht durch einen Betrachter eingeschätzt werden. Insbesondere die Wiederfindung von Wasserstoffbrücken innerhalb des Diphosphats wurde dabei berücksichtigt. Ausgehend von den in Gasphase optimierten Strukturen erfolgte nun in einem weiteren Optimierungsschritt die Einbeziehung des Solvatationsmodells PCM (siehe auch Abschnitt 3.3.4); als Lösungsmittel wurde Wasser verwendet. Hierbei zeigten sich für alle drei gewählten quantenchemischen Methoden sehr ähnliche Konformationen. Die jeweils berechneten Energien können Tabelle 9.3 (Anhang) entnommen werden. Eine anschließende Energieberechnung für alle Strukturen mittels MP4-Methode zeigte außerdem eine energetische Äquivalenz der mit den verschiedenen quantenchemischen Methoden HF, MP2 und B3LYP und unter Verwendung des Solvatationsmodells optimierten Strukturen. Diese Reproduzierbarkeit der Konformationen und MP4-Energien lässt auf eine Korrektheit der erzeugten Strukturen schließen. Abbildung 4.8 zeigt die mit B3LYP optimierten Konformationen.

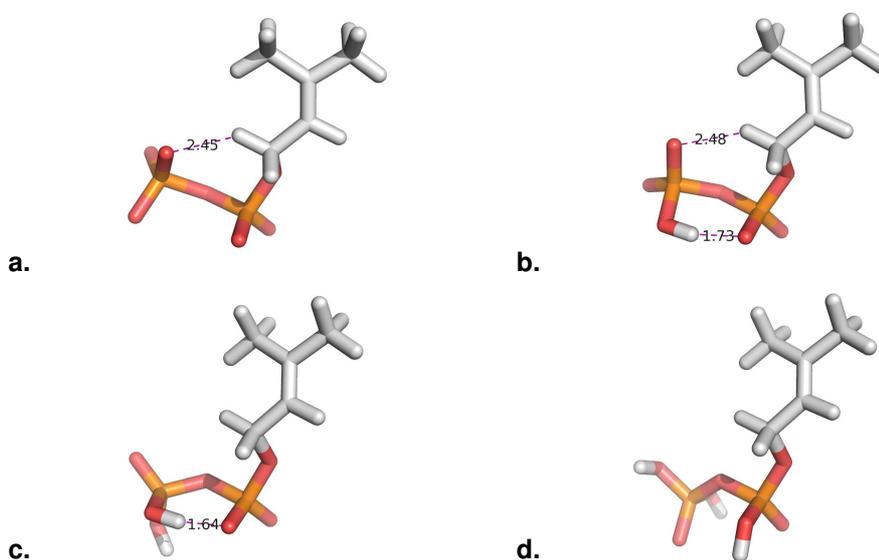


Abbildung 4.8: Konformationen verschiedener DMAPP-Protonierungsgrade nach Struktur- und Energieoptimierung mit B3LYP, 6-31+G** und PCM-Solvatationsmodell: a. DMAPP⁽⁰⁾, b. DMAPP⁽¹⁾, c. DMAPP⁽²⁾, d. DMAPP⁽³⁾. Intramolekulare Wasserstoffbrücken sind in violett gekennzeichnet; die Distanzen sind in Å angegeben.

Für alle quantenchemischen Berechnungen dieser Arbeit wurde B3LYP in Kombination mit dem PCM-Solvatationsmodell ausgewählt, sofern keine andere Methode genannt ist. Eine Ursache für die Wahl der Hybridmethode liegt in der Tatsache, dass die MP2-Methode nach Meinung des Autors zu zeit- und speicherintensiv ist, um beispielsweise Reaktionen zu berechnen und auch für die Durchführung einer großen Anzahl von Berechnungen eher ungeeignet ist. Die HF-Methode liefert zwar schnell Ergebnisse, zeigt aber bereits bei den in diesem Abschnitt beschriebenen Optimierungen Energieoszillationen, d.h. es kommt nicht zu einer Konvergenz der SCF-Methode. Weiterhin wären inkorrekte Ergebnisse aufgrund der Einfachheit der HF-Methode (vor Allem die Vernachlässigung der Elektronenkorrelation) für größere Systeme, wie sie nachfolgend betrachtet werden, zu befürchten (siehe auch Abschnitt 3.3.2). Angegebene Energien entsprechen nachfolgend stets B3LYP-Energien; auf eine Berechnung von MP2- oder sogar MP4-Energien aus den erzeugten Strukturen wird verzichtet. Als Basissatz wurde in diesem und allen nachfolgenden Abschnitten 6-31+G** verwendet, der aufgrund der Berücksichtigung von diffusen Funktionen für Nichtwasserstoffatome und Polarisierungsfunktionen für alle Atome durch den Autor als geeignet eingeschätzt wurde (siehe auch Abschnitt 3.3.3).

Neben verschiedenen Protonierungsstufen des DMAPP wurden auch Magnesiumkomplexe betrachtet. Hierzu wurden die oben beschriebenen, mit B3LYP und Solvatationsmodell PCM optimierten Strukturen von DMAPP⁽⁰⁾ (vollständig deprotoniert) und DMAPP⁽²⁾ (zweifach protoniert) als Basis verwendet. Für beide Moleküle wurden Komplexierungen mit einem einzelnen Metallion, für DMAPP⁽⁰⁾ zusätzlich mit drei Metallionen berücksichtigt. Als Grundlage für die Positionierung der Metallionen in DMAPP⁽⁰⁾ dienten die Röntgenstrukturen mit den PDB-Kodierungen 1x06^[57] (ein Magnesiumion) und 1rqi^[234] (drei Magnesiumionen). Für DMAPP⁽²⁾ wurde eine Komplexierung des Magnesiumions mit einem Sauerstoffatom des P_α erzeugt – beide Protonen sitzen am P_β. Um eine oktaedrische Komplexierung jedes Magnesiumions, wie sie auch in den Röntgenstrukturen vorkommt, zu sichern, wurden noch freie, nicht durch das Diphosphat eingenommene Koordinationsstellen, durch Acetat (Modellierung von Aspartat bzw. Glutamat) und H₂O besetzt. Für DMAPP⁽⁰⁾ wurde die Positionierung der Liganden aus den genannten Röntgenstrukturen übernommen. Für DMAPP⁽²⁾ erfolgte ebenfalls eine Anordnung entsprechend der Röntgenstruktur 1x06; die verbleibende freie Koordinationsstelle, die in 1x06 durch ein weiteres Sauerstoffatom des Diphosphats belegt ist, wurde mit H₂O besetzt. Komplexierungen mit zwei Magnesiumionen wurden vernachlässigt, da dem Autor die Existenz aktivierender Proteine mit dieser Anzahl an Metallbindestellen nicht bekannt ist. Abbildung 4.9 zeigt die optimierten Metallkomplexe. Es wird deutlich, dass auch nach der quantenchemischen Optimierung die oktaedrische Kom-

plexierung der Metallionen erhalten bleibt. Die zugehörigen Energien sind in Tabelle 9.4 (Anhang) dargestellt.

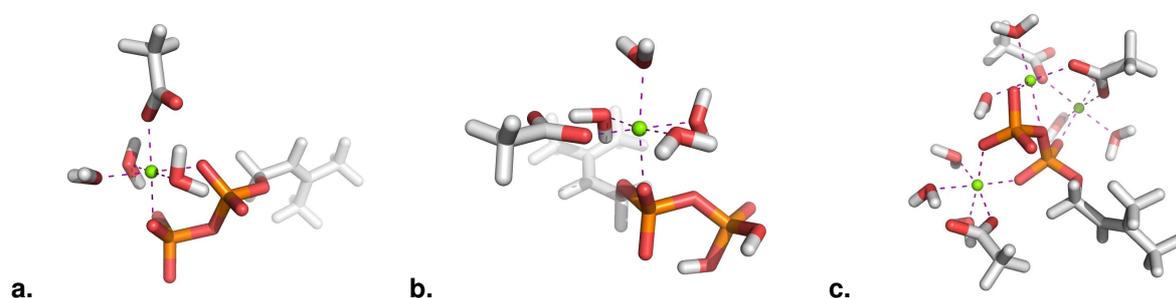


Abbildung 4.9: Konformationen der Komplexierung verschiedener DMAPP-Protonierungsgrade mit Magnesiumionen (grün) nach Struktur- und Energieoptimierung mit B3LYP, 6-31+G** und PCM-Solvatationsmodell: a. DMAPP⁽⁰⁾-Mg-Komplex; b. DMAPP⁽²⁾-Mg-Komplex; c. DMAPP⁽⁰⁾-Mg₃⁺-Komplex. Die Wechselwirkungen des Magnesiumions mit den umgebenden Liganden sind in violett angedeutet.

4.3.3 Vergleich mit Röntgenstrukturen der Cambridge Structural Database

Zur Überprüfung der Qualität der mittels B3LYP, Basissatz 6-31+G** und Solvationsmodell erzeugten Strukturen wurde ein Vergleich mit Röntgenstrukturen von organischen Verbindungen aus der CCDC Cambridge Structural Database (CSD) Version 5.30^[235] durchgeführt. Hierbei wurde insbesondere der Winkel O^RPO^M aufgrund seiner speziellen Ausprägung, sowie PO^M aufgrund seiner unzureichend genauen Berechnung mittels bestimmter theoretischer Methoden berücksichtigt (siehe auch Abbildung 4.5). Um eine Vergleichbarkeit mit DMAPP zu gewährleisten, wurden nur monosubstituierte Diphosphatstrukturen (Diphosphatmonoester; terminale Diphosphate) aus der CSD ausgewählt. Es konnten insgesamt sieben metallionenfreie und sieben metallionenhaltige Röntgenstrukturen extrahiert werden, bei denen es sich um Thiamin-, Adenosin- und Uridindiphosphat handelt.^[236-244] Das Diphosphat der metallionenhaltigen Strukturen war stets vollständig deprotoniert. Abbildung 4.10 zeigt den Vergleich der Winkel O^RPO^M und PO^M abhängig vom Protonierungsgrad und der Metallkomplexierung des Diphosphats. Der Winkel O^RPO^M zeigt sowohl für die Röntgen- als auch für die optimierten Strukturen einen leicht vom Tetraederwinkel ($109,47^\circ$) abweichenden Wert von 100° bis 106° , was durch den Einfluss der d-Orbitale des Phosphoratoms verursacht wird. Der PO^M -Winkel kann durch einige semiempirische Methoden nicht oder nur sehr mangelhaft reproduziert werden: AM1 und PM3 erzeugen lineare bzw. sehr stumpfe Winkel^[245]. Ebenso kommt es für Berechnungen mit B3LYP und 6-31+G**, die das Solvationsmodell nicht berücksichtigen (Gasphaserechnungen), zu einer starken Ab-

weichung von den experimentellen Daten (Daten nicht gezeigt). Die hier gewählte Kombination aus quantenchemischer Methode, Basissatz und Solvationsmodell ist hingegen in der Lage, die Werte der in Röntgenstrukturen beobachteten Winkel zu reproduzieren. Es bleibt jedoch anzumerken, dass die in Abbildung 4.8 gezeigten intramolekularen Wasserstoffbrücken der optimierten Strukturen nicht oder nur schwach in den Röntgenstrukturen zu beobachten sind. Diese Wechselwirkungen sind möglicherweise Artefakte; ihr Ursprung liegt in der Optimierung in Gasphase. Da jedoch keine Auflösung dieser Interaktionen durch das Solvationsmodell erfolgte und auch Röntgenstrukturen nur eine Momentaufnahme des Verhaltens der Strukturen in Lösung darstellen, dienen diese mit B3LYP, 6-31+G** und PCM-Solvationsmodell berechneten Strukturen als Grundlage aller in dieser Arbeit folgenden quantenchemischen Berechnungen.

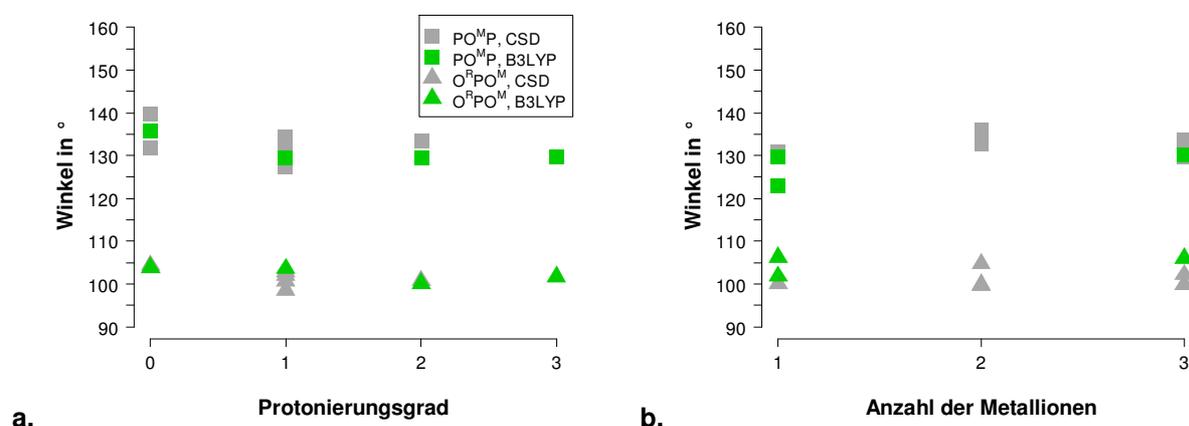


Abbildung 4.10: Vergleich der Winkel O^RPO^M und PO^MP des Diphosphats für metallionenfrie (a) und metallionenhaltige (b) Röntgen- und berechnete Strukturen. Dabei korrespondieren Quadrate zum Winkel PO^MP und Dreiecke zum Winkel O^RPO^M ; in grau sind jeweils die Strukturen der Cambridge Structural Database und in grün die mit B3LYP optimierten Strukturen dargestellt.

4.4 Diphosphaterkennung in der Proteinbindetasche

Das vorliegende Kapitel beschäftigt sich mit einer vergleichenden Analyse aktivierender und nichtaktivierender Diphosphatbindestellen, um die für eine Erkennung aktivierter Diphosphate relevanten Faktoren zu ermitteln. Dabei werden insbesondere die bereits in Kapitel 4.2 als bedeutsam identifizierten potentiellen Protonendonoren (Brønsted-Säuren) und Metallionen (Lewis-Säuren) hervorgehoben. Die Grundlage der Untersuchungen bildet der in den Kapiteln 3.1 und 4.1 beschriebene Datensatz. Ziel ist es, aus diesen Daten Hypothesen zu gewinnen und diese anschließend mittels quantenchemischer Berechnungen zu stützen (siehe Kapitel 3.3 und 4.3). Die Analysen beginnen mit der Erkennung des Diphosphats durch Aminosäuren der Proteinbindetasche und einem Vergleich mit der Phosphaterkennung im Allgemeinen, die in der Literatur^[246] bereits untersucht wurde. Eine detailliertere Untersuchung, bei der auch benachbarte Liganden, Metallionen und Wassermoleküle einbezogen werden, schließt sich an. Nach quantenchemischen Betrachtungen zum Zusammenhang von Bindungslängen und Aktivierung schließt das Kapitel mit Zusammenfassung und Diskussion ab.

4.4.1 Die Erkennung des Diphosphats durch Aminosäuren

Es soll zunächst die Bindung der Diphosphate durch Aminosäuren der Proteinbindetasche dargestellt werden. In diesem Zusammenhang ist zu erwähnen, dass bereits im Jahr 2007 durch Diederich *et al.*^[246] eine Analyse zur Phosphaterkennung in Proteinen veröffentlicht wurde. Sie untersuchten die Beteiligung von Aminosäuren, wobei nur Wasserstoffbrücken der Aminosäureseitenketten und der NH-Gruppe des Rückgrats berücksichtigt wurden; eine Betrachtung der Wasserstoffbrückenakzeptorfunktion der Carbonylgruppe des Rückgrats erfolgte nicht. Der verwendete Datensatz enthielt neben Diphosphaten auch Mono-, Tri- und Pentaphosphate. Es wurde durch Diederich *et al.* keine Unterscheidung zwischen aktivierenden und nichtaktivierenden Bindestellen vorgenommen. Abbildung 4.11 zeigt die Verteilung der Wechselwirkungen der Aminosäuren mit Diphosphat basierend auf dem in Abschnitt 4.1 beschriebenen Datensatz zum Vergleich. Analog zu Diederich *et al.* gingen in die dargestellten Zahlenwerte alle Wasserstoffbrücken zwischen Diphosphat und Aminosäureseitenkette bzw. der NH-Gruppe des Rückgrats ein.

Wie bereits durch Diederich *et al.* beschrieben, erfolgt auch die Bindung des Diphosphats häufig durch NH-Gruppen des Peptidrückgrats von Glycin (G), das schon seit Langem für seine Beteiligung an der Erkennung von Phosphaten bekannt ist (siehe z.B. Referenzen

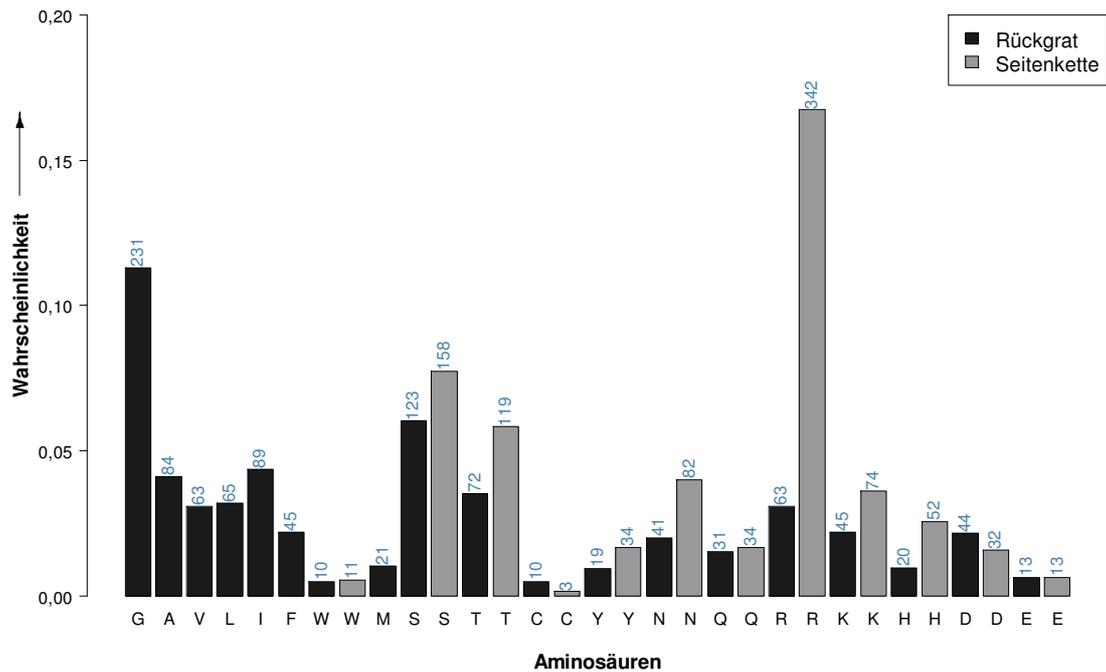


Abbildung 4.11: Verteilung der Zahl der Wechselwirkungen von Aminosäuren (Einbuchstabenkodierung) mit Diphosphat. Aminosäuren, deren Seitenketten Wechselwirkungen mit dem Diphosphat ausbilden können, sind zweifach gezeigt, wobei sich der erste Balken stets auf die NH-Gruppe des Rückgrates und der zweite Balken stets auf die Aminosäureseitenkette bezieht. Die Zahlen oberhalb jedes Balkens geben die absolute Häufigkeit einer Interaktion der jeweiligen Gruppe (Aminosäureseitenkette oder NH-Gruppe des Rückgrats) mit Diphosphat im betrachteten Datensatz an.

[10;221;247-250]). Auch das Rückgrat von Serin (S) tritt gehäuft in Wechselwirkung mit dem Diphosphat; die NH-Gruppe von Threonin hingegen tritt im Vergleich mit der NH-Gruppe hydrophober Aminosäuren wie Alanin (A) oder Isoleucin (I) beinahe in den Hintergrund – im Gegensatz zu den Ergebnissen von Diederich *et al.*, wo Wechselwirkungen mit Serin- und Threonin-NH etwa gleich häufig auftreten und neben dem Glycinrückgrat die wichtigste Rolle spielen. Die vermehrte Erkennung der in Diederich *et al.* beschriebenen Phosphate durch Peptidseitenketten von Serin, Threonin, Arginin (R) und Lysin (K) kann zumindest für die drei erstgenannten auch bei Diphosphaten wiedergefunden werden. Besonders auffällig ist dabei die stärkere Bedeutung der Argininseitenkette im Vergleich zum Glycinrückgrat während der Diphosphaterkennung, die in Diederich *et al.* nicht beobachtet werden konnte. Die deutliche Bevorzugung von Arginin gegenüber Lysin liegt hingegen auch bei der Phosphaterkennung im Allgemeinen vor. Die geringe Bedeutung von Tyrosin (Y), die in Diederich *et al.* beobachtet wird, kann auch für die Diphosphaterkennung festgestellt werden. Als Fazit lässt sich anmerken, dass sich die Erkennung von Phosphaten im Allgemeinen und Diphosphaten im Speziellen durch Aminosäuren nur in Details unterscheidet. Lediglich die enorm große Häufigkeit der Wechselwirkung mit Argininseitenketten während der Diphosphatbindung kann – insbesondere im Vergleich mit der Bedeutung des Glycin-NH – als markanter Unter-

schied herausgestellt werden. Eine Unterscheidung in aktivierende und nichtaktivierende Bindestellen ist durch Diederich *et al.* nicht erfolgt. Jedoch soll für die Diphosphaterkennung die Verteilung der Wechselwirkungen mit Aminosäuren für beide Gruppen verglichen werden. Da sich der in der vorliegenden Arbeit verwendete Datensatz hauptsächlich aus nichtaktivierenden Bindestellen zusammensetzt (siehe Abschnitt 4.1), entspricht eine Auftragung der Verteilung der Aminosäurewechselwirkungen für diese Bindestellengruppe im Wesentlichen der in Abbildung 4.11 gezeigten. Abbildung 4.12 zeigt zum Vergleich die Verteilung für die 15 aktivierenden Bindestellen.

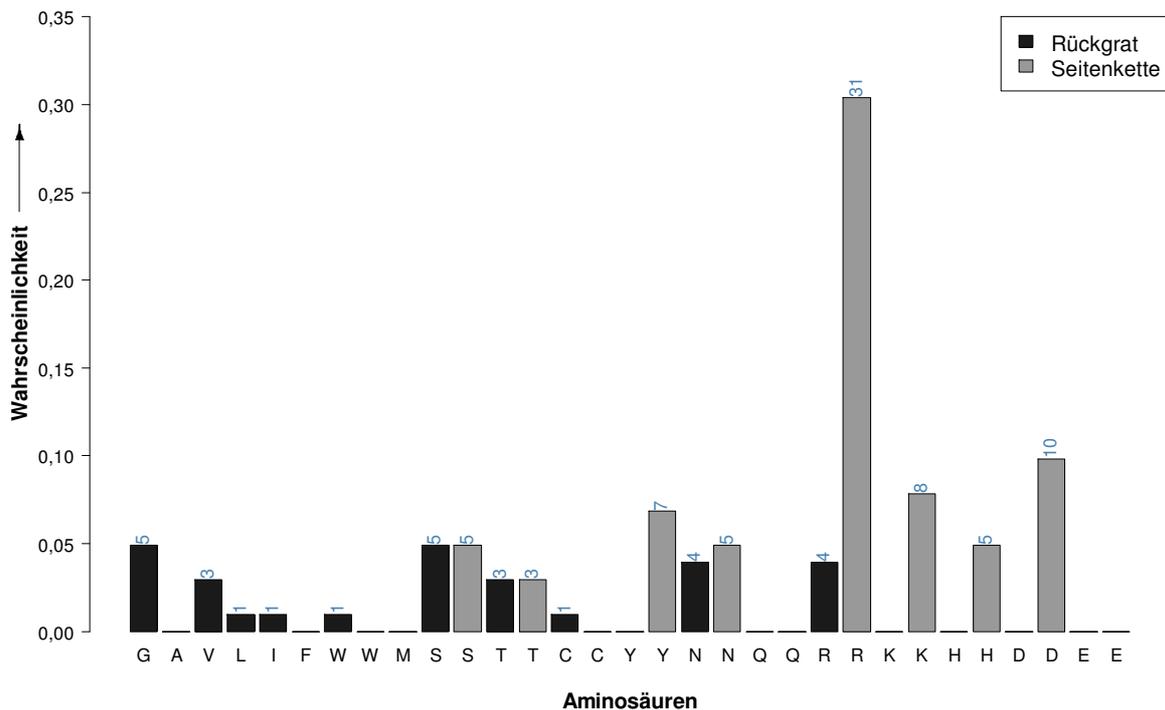


Abbildung 4.12: Verteilung der Zahl der Wechselwirkungen von Aminosäuren (Einbuchstabenkodierung) mit Diphosphat bei ausschließlicher Berücksichtigung aktivierender Bindestellen. Die Zahlen oberhalb jedes Balkens geben die absolute Häufigkeit einer Interaktion der jeweiligen Gruppe (Aminosäureseitenkette oder NH-Gruppe des Rückgrats) mit Diphosphat an.

Ein Vergleich der Abbildung mit Abbildung 4.11 zeigt eine geringere Bedeutung der NH-Gruppe des Glycinerückgrats bei der Erkennung aktivierter Diphosphate: Die Wechselwirkung mit der NH-Gruppe scheint beispielsweise mit der des Rückgrats von Serin (S), Asparagin (N) oder Arginin (R) gleichwertig zu sein. Zusätzlich nehmen auch Wechselwirkungen mit Tyrosinseitenketten (Y) im Vergleich zu Serin- oder Threoninseitenketten (T) zu, was bei Berücksichtigung des stärker sauren Charakters der Hydroxylgruppe des Tyrosins verständlich wird. Interessant ist außerdem die Tatsache, dass Glutamin (Q) nicht als Binder aktivierter Diphosphate beobachtet werden kann – im Gegensatz zu Asparagin (N). Analoges kann für Glutamat (E) und Aspartat (D) festgestellt werden. In diesem Zusammen-

hang muss jedoch angemerkt werden, dass die Aspartatseitenketten stets an der Komplexierung von Metallionen beteiligt sind, die wiederum das Diphosphat komplexieren. Die komplexierenden Sauerstoffatome des Aspartats und Diphosphats sind demzufolge nicht protoniert. Eine direkte Wechselwirkung zwischen diesen beiden Gruppen – ausgenommen der in der vorliegenden Arbeit nicht explizit berücksichtigten Van-der-Waals-Wechselwirkungen – kann also nicht stattfinden. Unbedingt zu beachten ist an dieser Stelle außerdem, dass aus dem Nichtauftreten einer Aminosäure in Abbildung 4.12 aufgrund der geringen Zahl verfügbarer aktivierender Bindestellen kein genereller Ausschluss dieser Aminosäure von der Erkennung geschlussfolgert werden darf.

4.4.2 Die Erkennung des Diphosphats durch einfache chemische Gruppen

Im vorangehenden Abschnitt wurde die Wechselwirkung des Diphosphats mit Aminosäuren der Bindetasche berücksichtigt. Jedoch werden die spezifischen Bindungseigenschaften von Aminosäuren meist nicht durch die Struktur als Ganzes bestimmt, sondern durch einfache, chemische Gruppen, wie beispielsweise Hydroxyl- oder Aminogruppen. Dies ermöglicht ein Zusammenfassen von Aminosäuren in Gruppen ähnlicher chemischer Eigenschaften. Zusätzlich können auch Liganden, Metallionen oder Wassermoleküle an der Bindung des Diphosphats beteiligt sein und funktionelle Gruppen von Aminosäuren substituieren. Eine Beschreibung der Bindetasche nicht durch komplette Aminosäurestrukturen, sondern durch einfache, chemische Gruppen erlaubt außerdem einen einfacheren und vielseitigeren Zugang von Verbindungen während des Designs synthetischer Diphosphatrezeptoren (siehe Abschnitt 1.2.4). Die nachfolgend gezeigten Diagramme orientieren sich daher an der in Tabelle 4.2 gezeigten Menge einfacher chemischer Gruppen zur Diphosphatbindung. Dabei wird auch auf die dort vorgestellte Einteilung der Gruppen in die Gruppe der potentiellen Protonendonoren und Metallionen (Brønsted- und Lewis-Säuren; nachfolgend als „saure“ Gruppe bezeichnet), der Wasserstoffbrückendonoren, der Wasserstoffbrückenakzeptoren und Wassermoleküle zurückgegriffen. Zusätzlich wurden in Abschnitt 4.4.1 – in Anlehnung an die Betrachtungen durch Diederich *et al.* – einzelne Wechselwirkungen ermittelt und gingen in die gezeigten Zahlenwerte ein. Dies führt für größere Molekülstrukturen, wie Arginine, die gleichzeitig mit mehreren Sauerstoffatomen des Diphosphats Wechselwirkungen ausbilden können, zu erhöhten Absolutwerten. In diesem Abschnitt korrespondiert daher die angegebene Häufigkeit zur Anzahl des Auftretens der entsprechenden funktionellen Gruppen in der Bindetasche und nicht zur Anzahl der ausgebildeten Wechselwirkungen (siehe Abbildung 4.3 für ein Beispiel). Abbildung 4.13 vergleicht die Wechselwir-

kungen aktivierter und nichtaktivierter Diphosphate unter Berücksichtigung dieser vier Gruppen. Es ist erneut zu beachten, dass Carboxylgruppen nicht als Protonendonoren klassifiziert wurden, da nach manueller Beurteilung der Bindestellen primär die deprotonierte Form zu erwarten ist (siehe Abschnitt 4.2.1). Hier zeigt sich wie schon in Kapitel 4.2 die Bedeutung der potentiellen Protonendonoren und Metallionen für aktivierende Bindestellen: Sie sind etwa gleich häufig an Wechselwirkungen mit dem Diphosphat beteiligt wie Wasserstoffbrückendonoren. Bei nichtaktivierenden Bindestellen ist dieses Verhältnis stark verschoben: Wasserstoffbrückendonoren sind hier etwa fünfmal so häufig wie Vertreter der sauren Gruppen.

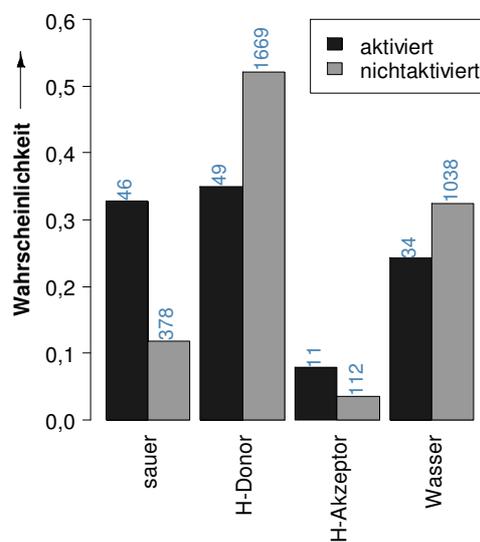


Abbildung 4.13: Verteilung der Häufigkeiten potentieller Protonendonoren und Metallionen (sauer), Wasserstoffbrückendonoren (H-Donor), -akzeptoren (H-Akzeptor) und Wassermolekülen (Wasser) in Bindestellen aktivierter und nichtaktivierter Diphosphate. Die Zahlen oberhalb jedes Balkens geben die absolute Häufigkeit der jeweiligen Gruppe an.

Um eventuelle Bevorzugen bestimmter funktioneller Gruppen in aktivierenden bzw. nicht-aktivierenden Bindetaschen deutlich zu machen, zeigt Abbildung 4.14 die Verteilung der Häufigkeiten unter Vernachlässigung der Zusammenfassung der Gruppen. Es wird deutlich, dass alle berücksichtigten Brønsted- und Lewis-Säuren in aktivierenden Bindestellen häufiger auftreten. Es soll hervorgehoben werden, dass auch die lewissäuren Metallionen bei der Bindung aktivierter Diphosphate eine größere Bedeutung haben als bei der Bindung nichtaktivierter Diphosphate. Dies deckt sich zwar mit den Beobachtungen verschiedener Autoren, dass Metallionen eine Aktivierung fördern (Abschnitt 1.3.1), ist hier aber zunächst unter Vorbehalt zu betrachten, da Kristallographen in einigen Fällen – im Hinblick auf die Unterbindung einer Enzymreaktion – dem Kristallisationsansatz essentielle Metallionen nicht zufügen. Analog treten alle berücksichtigten Wasserstoffbrückendonoren in nichtaktivieren-

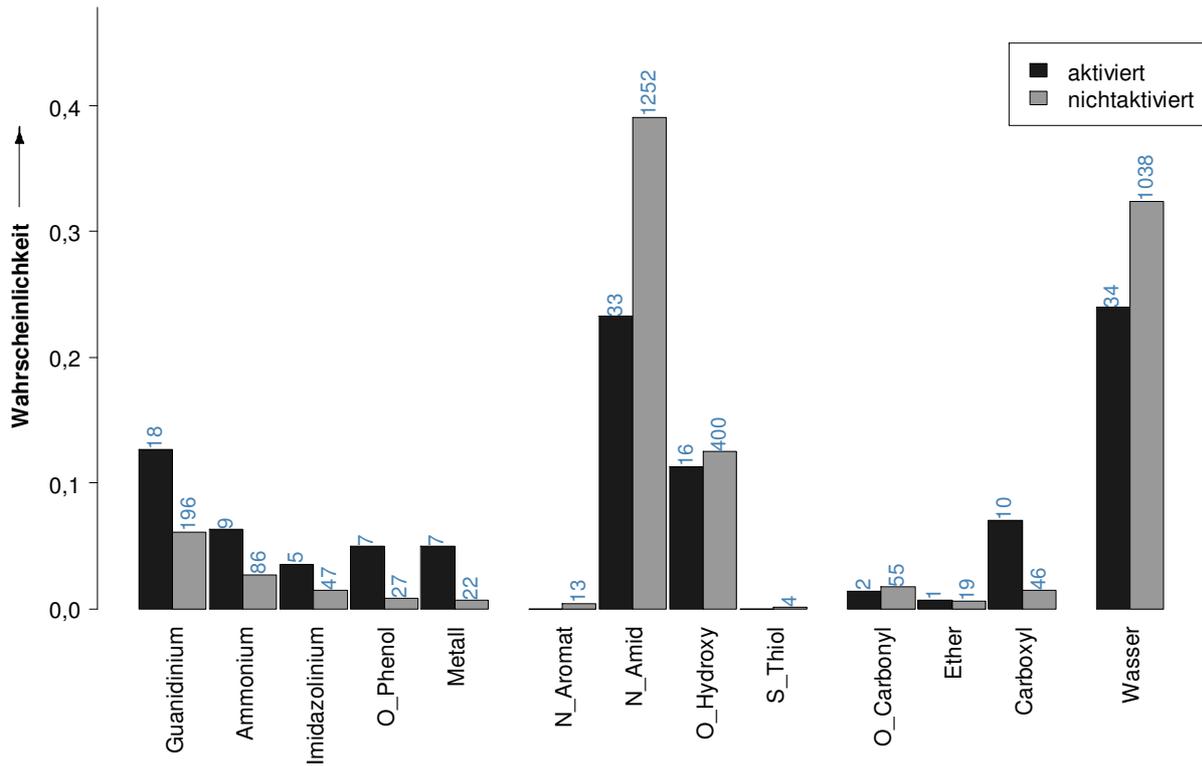


Abbildung 4.14: Verteilung der Häufigkeiten einfacher, chemischer Gruppen in Bindestellen aktivierter und nichtaktivierter Diphosphate. Die Zahlen oberhalb jedes Balkens geben die absolute Häufigkeit der jeweiligen Gruppe an.

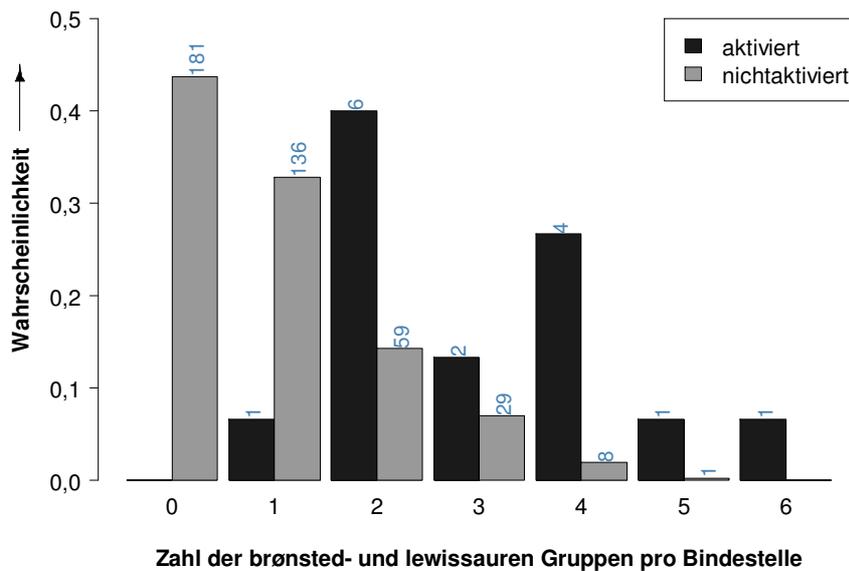


Abbildung 4.15: Darstellung der Häufigkeit des Auftretens potentieller Protonendonoren (Brønsted-Säuren) und Metallionen (Lewis-Säuren) pro aktivierender bzw. nichtaktivierender Diphosphatbindestelle. Die Zahlen oberhalb jedes Balkens geben die absolute Häufigkeit an.

den Bindestellen häufiger auf als in aktivierenden Bindestellen. Dieser Effekt wird für die Wechselwirkung mit „N_Amid“ besonders deutlich: Während die Erkennung des Diphosphats durch diese Gruppe für aktivierende Bindestellen nur gut 20% der Gesamtzahl ausmacht, ist der Wert für nichtaktivierende Bindestellen beinahe doppelt so hoch. Das für aktivierende Bindestellen häufige Auftreten von Carboxyl- und ähnlichen Gruppen („Carboxyl“) ist zurückzuführen auf ihre Funktion als Metallliganden; eine direkte Wechselwirkung einer protonierten Carboxylgruppe mit Diphosphat ist für die aktivierenden Bindestellen des vorliegenden Datensatzes nicht zu beobachten. Eine Interpretation der Häufigkeiten der Wassermoleküle in den Diphosphatbindestellen im Hinblick auf verschiedene Bedeutungen in aktivierenden und nichtaktivierenden Bindestellen ist schwierig, da selbst für homomultimere Proteine oft unterschiedliche Anordnungen der Wassermoleküle im Kristall bzw. verschiedene Anzahlen gefunden werden. Aus diesem Grund werden sie nachfolgend nicht weiter berücksichtigt.

Um die Bedeutung der Gruppe der potentiellen Protonendonoren und Metallionen (Brønsted- und Lewis-Säuren) für die Diphosphaterkennung näher zu beleuchten, wurde die Gesamtzahl der Strukturen dieser Gruppen pro Bindestelle ermittelt (Abbildung 4.15). Dabei wurden alle in Tabelle 4.2 der sauren Gruppe zugeordneten einfachen, chemischen Gruppen einbezogen – auch Wechselwirkungen mit Metallionen. Es ist für nichtaktivierende Bindestellen deutlich ein Trend zu weniger Brønsted- und Lewis-Säuren pro Bindestelle erkennbar. Für die nichtaktivierenden Bindestellen mit drei bis fünf sauren Gruppen pro Bindestelle ist kein Zusammenhang zu Enzymklasse oder gebundenem Liganden vorhanden. Eine Interpretation für aktivierende Bindestellen ist aufgrund der geringen Zahl verfügbarer Beispiele schwierig, es wird jedoch eine Tendenz zu einer höheren Anzahl an Brønsted- und Lewis-Säuren pro Bindestelle deutlich. Bemerkenswert ist, dass aktivierte Diphosphate stets durch mindestens eine dieser Gruppen gebunden werden. Eine vertiefende Analyse zur Abhängigkeit der Aktivierung vom Protonierungsgrad und Metallkomplexierung des Diphosphats ist den Abschnitten 4.4.4 und 4.6 zu entnehmen.

An dieser Stelle kann die Frage gestellt werden, ob die beobachtete verstärkte Bindung aktivierter Diphosphate durch brønsted- und lewissaure Gruppen das gesamte Diphosphat oder nur Bereiche des Diphosphats nahe der zu spaltenden Bindung betrifft. Die Struktur wurde daher in zwei Bestandteile P_{α} und P_{β} zerlegt, die in Abbildung 4.16 zu sehen sind. Abbildung 4.17 zeigt die Häufigkeiten des Auftretens potentieller Protonendonoren und Metallionen, Wasserstoffbrückendonoren, -akzeptoren und Wassermolekülen in den entsprechenden Bindestellen von P_{α} und P_{β} aktivierter Diphosphate.

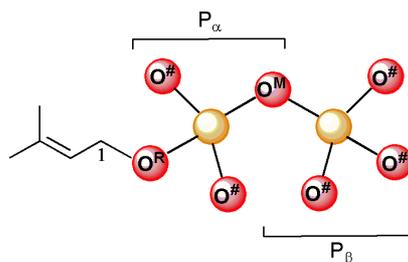


Abbildung 4.16: Darstellung von P_α und P_β des Diphosphats, sowie der Bezeichnung der Sauerstoff-Atomtypen am Beispiel des Dimethylallyldiphosphats (DMAPP).

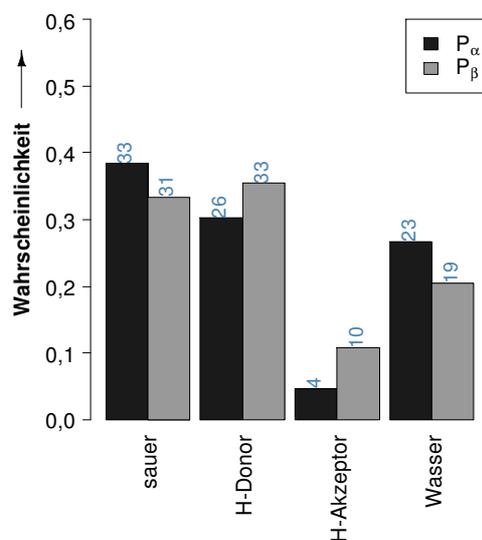


Abbildung 4.17: Verteilung der Häufigkeiten der Beteiligung potentieller Protonendonoren und Metallionen (sauer), Wasserstoffbrückendonoren (H-Donor), -akzeptoren (H-Akzeptor) und Wassermolekülen (Wasser) an der Erkennung von P_α und P_β aktivierter Diphosphate. Die Zahlen oberhalb jedes Balkens geben die absolute Häufigkeit an.

Es wird deutlich, dass die berücksichtigten potentiellen Protonendonoren und Metallionen sowohl das P_α als auch das P_β aktivierter Diphosphate binden; ersteres wird nicht signifikant besser gebunden. Auch bei den anderen Wechselwirkungsgruppen (Wasserstoffbrückendonoren, -akzeptoren, Wassermoleküle) gibt es kaum Unterschiede. Beim Vergleich der Absolutwerte der Abbildung mit denen in Abbildung 4.13 zeigt sich eine Abweichung. Diese wird verursacht durch Wechselwirkungen mit O^M – dieses Atom ist sowohl Bestandteil von P_α als auch von P_β (siehe Abbildung 4.16); Wechselwirkungen mit diesem Atom werden daher zweifach gezählt. Auch sind Interaktionen – insbesondere größerer Gruppen, wie der Guanidinogruppe – gleichzeitig mit P_α und P_β möglich.

Man hat in Abbildung 4.13, Abbildung 4.14 und Abbildung 4.15, sowie in Kapitel 4.2 gesehen, dass ein direkter Zusammenhang zwischen Aktivierung und der Zahl der potentiellen Protonendonoren und Metallionen im aktiven Zentrum des Enzyms besteht. Allerdings muss hier auch berücksichtigt werden, dass die nichtaktivierten Diphosphate, die dabei als Referenz verwendet wurden, sehr divers zusammengesetzt sind – ein Einfluss des organischen Restes auf den Aktivierungsprozess kann *per se* nicht ausgeschlossen werden. Abbildung 4.18 zeigt daher die Häufigkeiten der Gruppen ausschließlich für Prenyldiphosphat- bzw. Nukleosiddiphosphatzuckerbindestellen (NDP-Zucker). Für beide Ligandentypen existieren sowohl aktivierte als auch nichtaktivierte Formen. Die aktivierten Formen sind in Abschnitt 4.1.2 beschrieben. Nichtaktivierte Formen der NDP-Zucker liegen beispielsweise in Isomerasen (z.B. PDB-Eintrag 3beo^[251]) oder Ligasen (z.B. PDB-Eintrag 4uag^[252]) vor, während nichtaktivierte Prenyldiphosphate beispielsweise in Isoprenyldiphosphatsynthasen (z.B. PDB-Eintrag 1rqj^[234]) oder IPP- Δ -Isomerasen (z.B. PDB-Eintrag 2ick^[98]) gefunden werden können. Da für Pterindiphosphate keine nichtaktivierten Formen im betrachteten Datensatz enthalten sind, ist für diese Ligandengruppe keine Abbildung möglich.

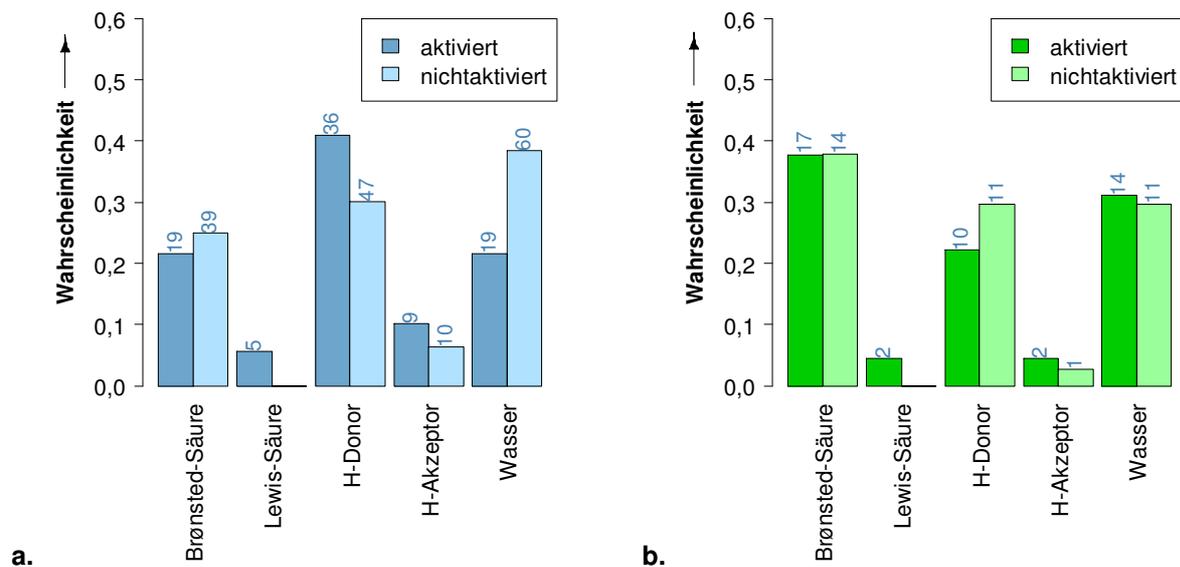


Abbildung 4.18: Verteilung der Häufigkeiten potentieller Protonendonoren (Brønsted-Säure) und Metallionen (Lewis-Säure), Wasserstoffbrückendonoren (H-Donor), -akzeptoren (H-Akzeptor) und Wassermolekülen (Wasser) in Bindestellen aktivierter und nichtaktivierter NDP-Zucker (a) und Prenyldiphosphate (b). Die Zahlen oberhalb jedes Balkens geben die absolute Häufigkeit an.

Sowohl für NDP-Zucker als auch für Prenyldiphosphate ist kein Zusammenhang zwischen vermehrter Anzahl an Brønsted-Säuren in der Diphosphatbindetasche und Aktivierung erkennbar. Es wird jedoch deutlich, dass für beide Ligandengruppen nichtaktivierte Diphosphate niemals durch Metallionen komplexiert sind. Für die im Datensatz berücksichtigten

aktivierten Prenyldiphosphate gilt, dass – auch, wenn die Kristallisierung nicht immer mit Metallion erfolgt ist – stets ein Metallion für die Aktivierung benötigt wird (siehe auch Tabelle 4.1). Es sind jedoch auch prenyldiphosphataktivierende Enzyme bekannt, die ihre Funktion vermutlich auch ohne Metallionen ausführen können (z.B. aromatische Prenyltransferasen aus Pilzen^[86;87], siehe Abschnitt 1.2.2). Ähnlich verhält es sich für die NDP-Zucker: Auch hier gibt es Enzyme, die eine metallionenunabhängige Aktivierung durchführen können (z.B. Glycosyltransferasen vom GT-B-Faltungstyp^[1], siehe Abschnitt 1.2.1). Durch Abbildung 4.18 wird deutlich, dass brønstedsaure Gruppen allein für eine Unterscheidung aktivierender und nichtaktivierender Bindestellen nicht ausreichend sind; Metallionen beeinflussen die Aktivierung zusätzlich. Bei Auftragung der Häufigkeiten der Gruppen für aktivierende NDP-Zucker-Bindestellen der metallionenunabhängigen GT-B-Glycosyltransferasen im Vergleich zu nichtaktivierenden NDP-Zucker-Bindestellen analog zu Abbildung 4.18 zeigt sich ebenfalls keine Häufung brønstedsaurer Gruppen in den aktivierenden Bindestellen (siehe Abbildung 9.1 im Anhang). Demzufolge müssen für diese Gruppe von Enzymen, die ohne die Komplexbildung von Metallionen in der Lage sind, eine Diphosphataktivierung durchzuführen, weitere Faktoren existieren (siehe auch Abschnitt 4.5).

Zusätzlich zur Untersuchung der Zusammensetzung aktivierender und nichtaktivierender Bindestellen, kann auch die Anzahl der Wechselwirkungen der Sauerstoffatome des Diphosphats ohne Berücksichtigung des Wechselwirkungspartners analysiert werden (Tabelle 4.4). Die Bezeichnung der verschiedenen im Diphosphat vorliegenden Sauerstoff-Atomtypen ist in Abbildung 4.16 erläutert.

Tabelle 4.4: Zahl der Wechselwirkungen, die unabhängig von der Art der Wechselwirkung durch verschiedene Sauerstoff-Atomtypen aktivierter und nichtaktivierter Diphosphate im Durchschnitt ausgebildet werden. In Klammern ist die Gesamtzahl des jeweiligen Atomtyps im Datensatz angegeben.

	O [#]	O ^M	O ^R
aktiviert	2,6 (65)	1,4 (15)	0,5 (25)
nichtaktiviert	2,0 (1697)	0,5 (414)	0,6 (787)

Besonders auffällig ist, dass der verbrückende Sauerstoff O^M in aktivierten Diphosphaten im Durchschnitt fast dreifach so viele Wechselwirkungen ausbildet, wie in nichtaktivierten Diphosphaten. Die Wechselwirkungspartner für O^M können sowohl Brønsted-Säuren, Wasserstoffbrückendonoren oder Wasser sein. Eine genauere Untersuchung der aktivierenden Bindestellen ergab, dass sich der Effekt der verstärkten Wechselwirkung von O^M insbesondere für prenyldiphosphatkonvertierende Enzyme und Glycosyltransferasen vom Faltungstyp

GT-B (metallionenunabhängige Gruppe) zeigt. In Kapitel 4.5 wird auf diesen Sachverhalt erneut Bezug genommen.

4.4.3 Protonenübertragung betrachteter Brønsted-Säuren

Im vorangehenden Abschnitt erfolgte eine Auswahl brønstedsaurer Gruppen unter der Voraussetzung, dass sie in der Lage sind, Diphosphat zu protonieren. Um zu zeigen, dass tatsächlich eine Protonenübertragung auf das Diphosphat stattfinden kann, wurden Reaktionsenergien ΔE dieses Prozesses am Beispiel des Dimethylallyldiphosphats (DMAPP) berechnet (Abbildung 4.19). Aktivierungsbarrieren wurden jedoch in diesem Zusammenhang nicht ermittelt.

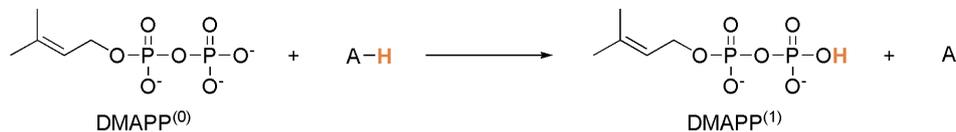


Abbildung 4.19: Darstellung der Protonierung von Diphosphat durch eine Säure A-H am Beispiel von vollständig deprotoniertem DMAPP.

Die Berechnungen der Energien erfolgten quantenchemisch. Dafür wurden zunächst die Parameter analog Abschnitt 4.3.2 gewählt, die ein implizites Wassermodeill beinhalten. Dieser Ansatz wurde verwendet, obwohl die in aktiven Zentren vorliegenden Bedingungen eher einem Vakuum gleichen, das oft gut durch Berechnungen in Gasphase modelliert werden kann. Die Ursache für die Verwendung des implizites Solvatationsmodells liegt in der Tatsache, dass das Gasphasemodell – im Gegensatz zur Berechnung mit Solvatationsmodell – zur Ausbildung intramolekularer Wechselwirkungen des Diphosphats führt (siehe Abschnitt 4.3.2), die nicht mit den in Proteinen zu beobachtenden Wechselwirkungen zwischen Ligand und Aminosäuren zu vereinbaren sind. Eine nachfolgende Berechnung der Energien (ohne Strukturoptimierung) der mit Solvatationsmodell erhaltenen Strukturen unter Gasphasebedingungen führt zu einer angenäherten Simulation der vakuumähnlichen Eigenschaften einer Proteinumgebung.

Folgende Formel diente für die anschließende Berechnung der Reaktionsenergien als Grundlage:

$$\Delta E = (E_{A^-} + E_{DMAPP-H}) - (E_{AH} + E_{DMAPP})$$

wobei E_{AH} und $E_{DMAPP-H}$ den totalen Energien der protonierten Formen der Säure bzw. des DMAPP und E_{A^-} bzw. E_{DMAPP} den totalen Energien der zugehörigen deprotonierten Formen (korrespondierende Basen) entsprechen. Für DMAPP-Metallionen-Komplexe – deren Reaktionsenergie entsprechend Abbildung 4.19 hier auch ermittelt wurde – erfolgte die Protonierung analog zu den metallionfreien Verbindungen (siehe Abbildung 4.7), z.B. an P_{β} für die deprotonierte Form des Diphosphats. Im Falle mehrerer möglicher Protonierungsstellen bei Metallkomplexen wurde stets das Sauerstoffatom mit der negativsten Ladung für die Protonierung ausgewählt (höchste Protonenaffinität). Die protonierte Form des DMAPP⁽¹⁾-Mg-Komplexes wurde durch zweifache Protonierung des DMAPP⁽⁰⁾-Mg-Komplexes (Abbildung 4.9) konstruiert. Für den Komplex mit drei Metallionen wurde nur die Reaktionsenergie für die einfache Protonierung berechnet; weitere Protonierungen des Komplexes sind nicht möglich, da bereits bei der einfachen Protonierung alle O[#]-Atome durch ein Metallion oder Proton besetzt sind. Die brønstedsauren Gruppen mit den Bezeichnungen „Guanidinium“, „Imidazolinium“, „Ammonium“ und „O_Phenol“ wurden dabei durch 1-Ethylguanidinium, 4-Ethylimidazolinium, Ethylammonium bzw. *p*-Ethylphenol dargestellt. Zusätzlich wurde auch die Protonenübertragung durch Propansäure (als Repräsentant für „Carboxyl“) ermittelt. Es bleibt an dieser Stelle jedoch zu berücksichtigen, dass die Carbonsäure im vorangehenden Abschnitt nicht als Protonenüberträger betrachtet wurde, da für die meisten der betrachteten Bindestellen aufgrund der Wechselwirkung beispielsweise mit Metallionen der deprotonierte Zustand angenommen wird. Tabelle 4.5 zeigt die ermittelten Reaktionsenergien. Die zugehörigen totalen Energien, die die Grundlage für die Berechnungen der Werte bilden, können Tabelle 9.5 (Anhang) entnommen werden. Generell zeigt sich, dass die Protonenübertragung mit steigender Ladung des DMAPP energetisch stärker begünstigt ist. Erklärt werden kann dies durch die höhere Protonenaffinität des Diphosphats. Ethylammonium, 4-Ethylimidazolinium und 1-Ethylguanidinium sind in der Lage Diphosphat vollständig zu protonieren, während Propansäure und *p*-Ethylphenol lediglich die zweite Protonierungsstufe erzeugen können. Dies erklärt sich durch die Tatsache, dass die korrespondierenden Basen der beiden zuletzt genannten Säuren geladen sind, was für Gasphasebetrachtungen aufgrund der mangelnden Abschirmung von Ladungen energetisch ungünstig ist. Im Enzym kann für diese Gruppen die Säurestärke durch sekundäre, die entstehenden Ladungen stabilisierende Wechselwirkungen daher durchaus anders sein. Auch für Komplexe mit Metallionen wird deutlich, dass die Protonenübertragung mit steigendem Protonierungsgrad des Diphosphats weniger stark begünstigt wird. Es zeigt sich, dass stets zumindest die einfache Protonierung des vollständig deprotonierten DMAPP⁽⁰⁾-Mg-Komplexes durch die betrachteten Säuren möglich ist. Die höhere Reaktionsenergie des DMAPP⁽²⁾-Mg-Komplexes im Vergleich zum DMAPP⁽⁰⁾-Mg₃-Komplexes wird verursacht durch

Tabelle 4.5: Reaktionsenergien ΔE in $\text{kJ}\cdot\text{mol}^{-1}$ für die Übertragung von Protonen auf verschiedene Protonierungsstufen und Metallionenkomplexierungen des DMAPP durch brønstedsaure Aminosäuremimetika.

Protonierung bzw. Metallkomplexierung	Ethylammonium	4-Ethylimidazolium	1-Ethylguanidinium
DMAPP ⁽²⁾	-344,95	-293,68	-245,65
DMAPP ⁽¹⁾	-764,47	-713,20	-665,18
DMAPP ⁽⁰⁾	-1256,46	-1205,19	-1157,16
DMAPP ⁽²⁾ -Mg	16,87	68,13	116,16
DMAPP ⁽⁰⁾ -Mg ₃	-181,62	-130,36	-82,33
DMAPP ⁽¹⁾ -Mg	-322,99	-271,72	-223,69
DMAPP ⁽⁰⁾ -Mg	-785,51	-734,25	-686,22

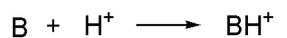
Protonierung bzw. Metallkomplexierung	Propansäure	<i>p</i> -Ethylphenol
DMAPP ⁽²⁾	186,60	193,32
DMAPP ⁽¹⁾	-232,93	-226,20
DMAPP ⁽⁰⁾	-724,92	-718,19
DMAPP ⁽²⁾ -Mg	548,41	555,13
DMAPP ⁽⁰⁾ -Mg ₃	349,92	356,64
DMAPP ⁽¹⁾ -Mg	208,56	215,28
DMAPP ⁽⁰⁾ -Mg	-253,97	-247,25

eine Destabilisierung der Metallwechselwirkung durch die Protonierung, was sich in einer Verlängerung der Diphosphat-Magnesium-Bindung äußert. Die höhere Reaktionsenergie des DMAPP⁽⁰⁾-Mg₃-Komplexes im Vergleich zum DMAPP⁽¹⁾-Mg-Komplex hingegen hängt mit einer weniger stark negativen Ladung des zu protonierenden O[#] zusammen. Beim Vergleich der metallionenhaltigen mit den metallionenfremen Verbindungen wird außerdem deutlich, dass die Reaktionsenergie für den DMAPP⁽⁰⁾-Mg-Komplex annähernd der des einfach protonierten DMAPP⁽¹⁾ und die Reaktionsenergie für den DMAPP⁽¹⁾-Mg-Komplex annähernd der des zweifach protonierten DMAPP⁽²⁾ entspricht; die geringfügigen Unterschiede in den Reaktionsenergien sind in erster Linie durch geringfügige Unterschiede in den Ladungen der O[#]-Sauerstoffatome zu erklären.

Aufgrund der Daten aus Tabelle 4.5 lässt sich daher annehmen, dass in metallionenfremen Proteinbindetaschen vollständig protonierte Formen des Diphosphats vorliegen können. Bei Vorliegen eines Diphosphat-Metallion-Komplexes sind ebenfalls Protonierungen des Diphosphats möglich; insbesondere können bei Komplexen mit einem Metallion bis zu zweifache Protonierungen schon durch schwach saure Aminosäureseitenketten erzeugt werden.

Um die Korrektheit der in Tabelle 4.5 implizit angegebenen Stärken der betrachteten Säuren nachzuweisen, wurde ein Vergleich von berechneten und experimentellen Protonenaffinitäten der korrespondierenden Basen durchgeführt. Die Berechnung der Protonenaffinitäten erfolgte anhand der auch Tabelle 4.5 zugrundeliegenden Energiewerte der Säuren

und ihrer korrespondierenden Basen (siehe Anhang, Tabelle 9.5). Die Protonenaffinität E_{PA} ist dabei definiert durch das Negative der Enthalpieänderung für die Gasphasereaktion



Hohe Werte korrespondieren daher zu einer hohen Protonenaffinität, kleine Werte zu einer geringen. Die quantenchemisch bestimmte Protonenaffinität nähert E_{PA} durch Berechnung von

$$\Delta E = -(E_{BH^+} - E_B)$$

an.

Tabelle 4.6 zeigt die resultierenden Daten. Es wird deutlich, dass die in der vorliegenden Arbeit berechneten Werte die auf experimenteller Basis ermittelten annähern und insbesondere die Reihenfolge der Protonenaffinitäten durch die Berechnung wiedergegeben werden kann.

Tabelle 4.6: Vergleich der quantenchemisch berechneten und auf experimenteller Basis bestimmten Protonenaffinitäten verschiedener Basen in $\text{kJ}\cdot\text{mol}^{-1}$. Die Quelle der experimentellen Werte ist jeweils als Referenz angegeben.

Berechnung		Experiment	
B	ΔE in $\text{kJ}\cdot\text{mol}^{-1}$	B	E_{PA} in $\text{kJ}\cdot\text{mol}^{-1}$
Ethylamin	948,98	Ethylamin	912,0 ^[253]
4-Ethylimidazol	1000,25	Imidazol	942,8 ^[253]
1-Ethylguanidin	1048,27	Guanidin	986,3 ^[253]
Propionat	1480,52	Acetat	1458 ^[254]
<i>p</i> -Ethylphenolat	1487,24	Phenolat	1470 ^[254]

4.4.4 Abhängigkeit von Bindungslängen des Diphosphats von Protonierungsgrad und Metallionenkomplexierung

Die vorangehenden Abschnitte zeigen, dass brønstedsaure Gruppen und Metallionen häufig im katalytischen Zentrum aktivierender Bindestellen vorliegen. Es stellt sich in diesem Zusammenhang die Frage nach der Bedeutung dieser Gruppen für den Aktivierungsprozess. Allgemein ergeben sich für die Abspaltung des Diphosphats von einem organischen Rest – beispielsweise während einer Hydrolyse – zwei Möglichkeiten: Es kann zur Spaltung der Kohlenstoff-Sauerstoff-Bindung (C^1O^R) oder aber zur Spaltung der Bindung zwischen Sauerstoffatom und dem Phosphoratom des P_α ($O^R P$) kommen (siehe auch Abbildung 4.16).

Abbildung 4.20 stellt die Längen dieser Bindungen für die in Abschnitt 4.3 beschriebenen, mit quantenchemischen Methoden berechneten DMAPP-Strukturen dar.

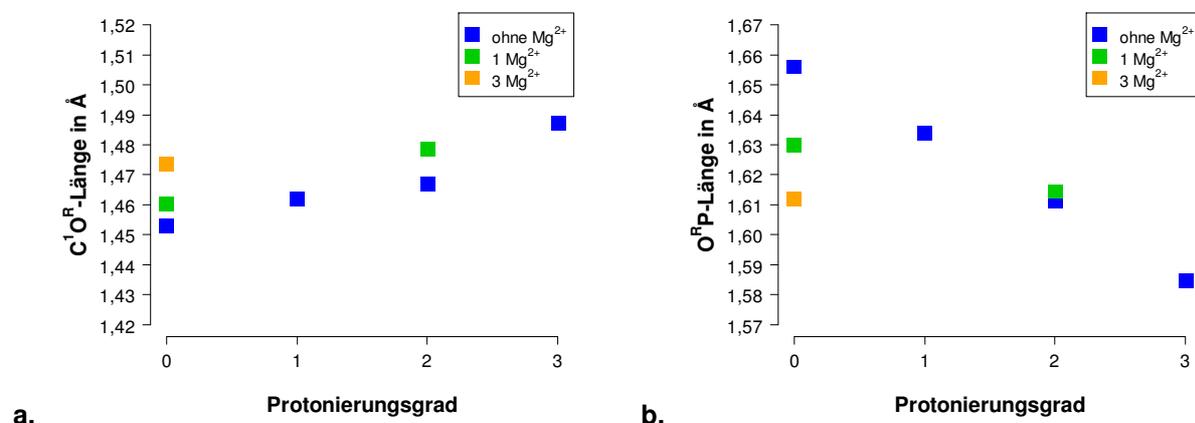


Abbildung 4.20: Darstellung der C¹O^R- (a) sowie der O^RP-Bindungslänge (b) verschiedener Protonierungsgrade sowie der Metallkomplexierungen des DMAPP.

Es wird deutlich sichtbar, dass die Länge der C¹O^R-Bindung mit steigendem Protonierungsgrad zunimmt, jedoch gleichzeitig die Länge der O^RP-Bindung abnimmt. Um auszuschließen, dass diese Abhängigkeit zufällig und aufgrund der gewählten Ausgangskonformationen der Strukturen vorliegt, wurde die optimierte Struktur des dreifach negativ geladenen DMAPP (DMAPP⁽⁰⁾) vollständig protoniert und optimiert (Parameter analog Abschnitt 4.3.2). Die entstehenden Längen der C¹O^R-, sowie der O^RP-Bindung variieren von den entsprechenden Werten des in Abbildung 4.20 gezeigten, vollständig protonierten DMAPP⁽³⁾ lediglich sehr geringfügig in der dritten Nachkommastelle. Metallionen verstärken die Effekte der Protonierung auf die Bindungslängen. Durch Komplexierung des vollständig deprotonierten DMAPP⁽⁰⁾ mit einem Metallion wird ein Zustand der C¹O^R- bzw. O^RP-Bindungslänge erzeugt, der einer einfachen Protonierung ähnelt. Für die Komplexierung von DMAPP⁽²⁾ mit einem Magnesiumion findet ebenfalls eine Verlängerung der C¹O^R-Bindung statt; die Länge der O^RP-Bindung hingegen wird im Vergleich zum DMAPP⁽²⁾ ohne komplexierendes Metallion kaum verändert. Eine Komplexierung des DMAPP⁽⁰⁾ mit drei Metallionen entspricht – bezogen auf die betrachteten Bindungslängen – etwa einer zweifachen Protonierung des Diphosphats. Die Daten weisen auf eine erleichterte Spaltung der C¹O^R-Bindung in brønsted-saurem Milieu oder bei Komplexierung mit Metallionen hin, während die Spaltung der O^RP-Bindung unter diesen Bedingungen erschwert scheint.

4.4.5 Zusammenfassung und Diskussion

Aus experimentellen Ergebnissen (siehe Abschnitt 1.3.1) in enzymfreier Umgebung ist bekannt, dass die Geschwindigkeit der Aktivierung mit steigender Konzentration von Säuren oder Metallionen ebenfalls steigt. Die Experimente zeigten außerdem, dass nur die Kombination von Säure und Metallion eine Aktivierung herbeiführen kann; keiner der beiden Faktoren kann die Aktivierung allein bewirken. Die Ursache für die Abhängigkeit der Aktivierungsgeschwindigkeit von Säuren und Metallionen liegt zum Einen in einer verstärkten Säurekatalyse (elektrophiler Angriff an O^R). Zum Anderen haben aber auch Protonierungsgrad und Metallionenkomplexierung des Diphosphats direkten Einfluß – vermutlich vor allem durch die Überführung des Diphosphats in eine bessere Abgangsgruppe durch Kompensation der negativen Ladung, die während der Heterolyse entsteht. Erste Hinweise für eine Bedeutung von Protonierungsgrad und Komplexierung durch Metallionen auch in Enzymen liefern die Ergebnisse von Kapitel 4.2, in dem dargestellt wurde, dass stets eine Beteiligung von Brønsted- bzw. Lewis-Säuren an der Bindung aktivierter Diphosphate vorliegt. Das vorliegende Kapitel zeigt, dass das Auftreten dieser Gruppen in aktivierenden Bindestellen tatsächlich erhöht ist: Die betrachteten Brønsted- bzw. Lewis-Säuren sind etwa gleich häufig wie Wasserstoffbrückendonoren, während in nichtaktivierenden Bindestellen lediglich ein (1:5)-Verhältnis im Vergleich zu Wasserstoffbrückendonoren vorliegt. Dabei konnte keine Bevorzugung aktivierter Diphosphate für bestimmte Brønsted-Säuren festgestellt werden: Alle berücksichtigten Brønstedsäuren, sowie auch die Metallionen treten in aktivierenden Bindestellen gehäuft auf. Zusätzlich wurde gezeigt, dass die Summe der Protonendonoren und Metallionen pro Bindestelle tendenziell in aktivierenden Bindestellen größer ist als in nichtaktivierenden Bindestellen. Es wurde dargelegt, dass nichtaktivierte Diphosphate, die in erster Linie durch Wasserstoffbrückendonoren gebunden werden, eine Bevorzugung für „N_Amid“ besitzen, was auch auf das erhöhte Auftreten von Glycin in diesen Bindestellen zurückzuführen ist. Bei einem Vergleich der Anzahl der Wechselwirkungen, die durch jedes einzelne Sauerstoffatom des Diphosphats im Durchschnitt ausgebildet werden, zeigte sich in aktivierenden Bindestellen interessanterweise eine verstärkte Interaktion mit dem zentralen Sauerstoffatom (O^M). Mit Hilfe quantenchemischer Untersuchungen konnte weiterhin gezeigt werden, dass die organischen Säuren innerhalb der Proteinbindetaschen in der Lage sind, höhere Protonierungsgrade des Diphosphats und z.T. sogar vollständige Protonierungen zu erzeugen. Ebenso wurde dargestellt, dass deprotoniertes Diphosphat, das mit einem oder drei Metallionen komplexiert ist, zusätzlich durch die Brønsted-Säuren protoniert werden kann. Jedoch ist eine weitere Protonierung nicht begünstigt, wenn das metallkomplexierte Diphosphat bereits zweifach protoniert ist. Weitere quantenchemische Untersuchungen belegen, dass eine starke gegenläufige Abhängigkeit der C^1O^R - und der $O^R P$ -Bindungs-

längen von Protonierungsgrad und Metallionenkomplexierung besteht. Erwartungsgemäß scheint die Metallionenkomplexierung (Lewis-Säuren) Zustände höherer Protonierung zu imitieren. Die präsentierten Daten lassen die Schlussfolgerung zu, dass die mit erhöhter Konzentration starker Säuren oder Metallionen verstärkt stattfindende, im Experiment beobachtete Aktivierung nicht ausschließlich auf eine zunehmende Säurekatalyse, sondern tatsächlich auch auf strukturelle Veränderungen innerhalb der Moleküle (verursacht durch einen erhöhten Protonierungsgrad des Diphosphats bzw. eine Metallkomplexierung) zurückzuführen ist. Jedoch zeigt die Beteiligung einer ebenfalls erhöhten Anzahl brønstedsaurer Gruppen an der Bindung nichtaktivierter NDP-Zucker bzw. Prenyldiphosphate, dass – zumindest für metallionenunabhängige, aktivierende Bindestellen – ein weiterer, die Aktivierung beeinflussender Faktor existieren muss.

4.5 Bedeutung der Ligandenkonformation

Die in Kapitel 4.4 durchgeführten Analysen zur Diphosphaterkennung in Proteinbindetaschen zeigen, dass der Protonierungsgrad des Diphosphats, der durch die Nachbarschaft von brønstedsauren Gruppen bestimmt wird, sowie die Metallionenkomplexierung den Aktivierungsprozess beeinflussen. Es wird jedoch auch deutlich, dass ein weiterer Faktor existieren muss, insbesondere da die in metallionenunabhängigen, aktivierenden Bindestellen vorliegenden erhöhten Protonierungsgrade ebenso in nichtaktivierenden Bindestellen wahrscheinlich sind (siehe Abbildung 4.18 und zugehörige Diskussion). Neben Protonierungsgrad und Metallionenkomplexierung des Diphosphats können Enzyme auch die Konformation des gebundenen Liganden beeinflussen. Dies ist eine häufig von Enzymen genutzte Strategie, um Übergangszustände mit geringeren Aktivierungsbarrieren zu begünstigen und damit die Reaktion voranzutreiben. Das vorliegende Kapitel geht daher auf die Konformation des Liganden in der Proteinbindetasche ein. Dabei sind insbesondere Bereiche um die zu aktivierende Bindung interessant, die durch verschiedene Torsionswinkel repräsentiert werden (Abbildung 4.21).

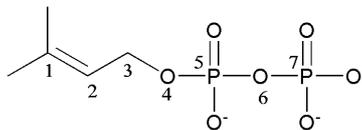


Abbildung 4.21: Darstellung der berücksichtigten Torsionswinkel am Beispiel von Dimethylallyldiphosphat (DMAPP). Die Nummerierung entspricht folgenden Bezeichnungen der Winkel: 1-2-3-4 (CCCCO), 2-3-4-5 (CCOP), 3-4-5-6 (COPO), 4-5-6-7 (OPOP).

Die folgenden Abschnitte behandeln zunächst die Konformation des Liganden in der Proteinbindetasche. Grundlage hierfür bildet der in den Abschnitten 3.1 und 4.1 beschriebene Datensatz. Es werden ausgewählte Torsionswinkel analysiert und ihre Bedeutung in aktivierenden und nichtaktivierenden Bindetaschen verglichen. Quantenchemische Betrachtungen dieser Torsionswinkel anhand des Modellmoleküls Dimethylallyldiphosphat (DMAPP) folgen.

4.5.1 Torsionswinkel nichtaktivierter Diphosphate

Abbildung 4.22 zeigt die paarweise Auftragung der Werte der Torsionswinkel CCCO, CCOP, COPO und OPOP für nichtaktivierte Diphosphate.

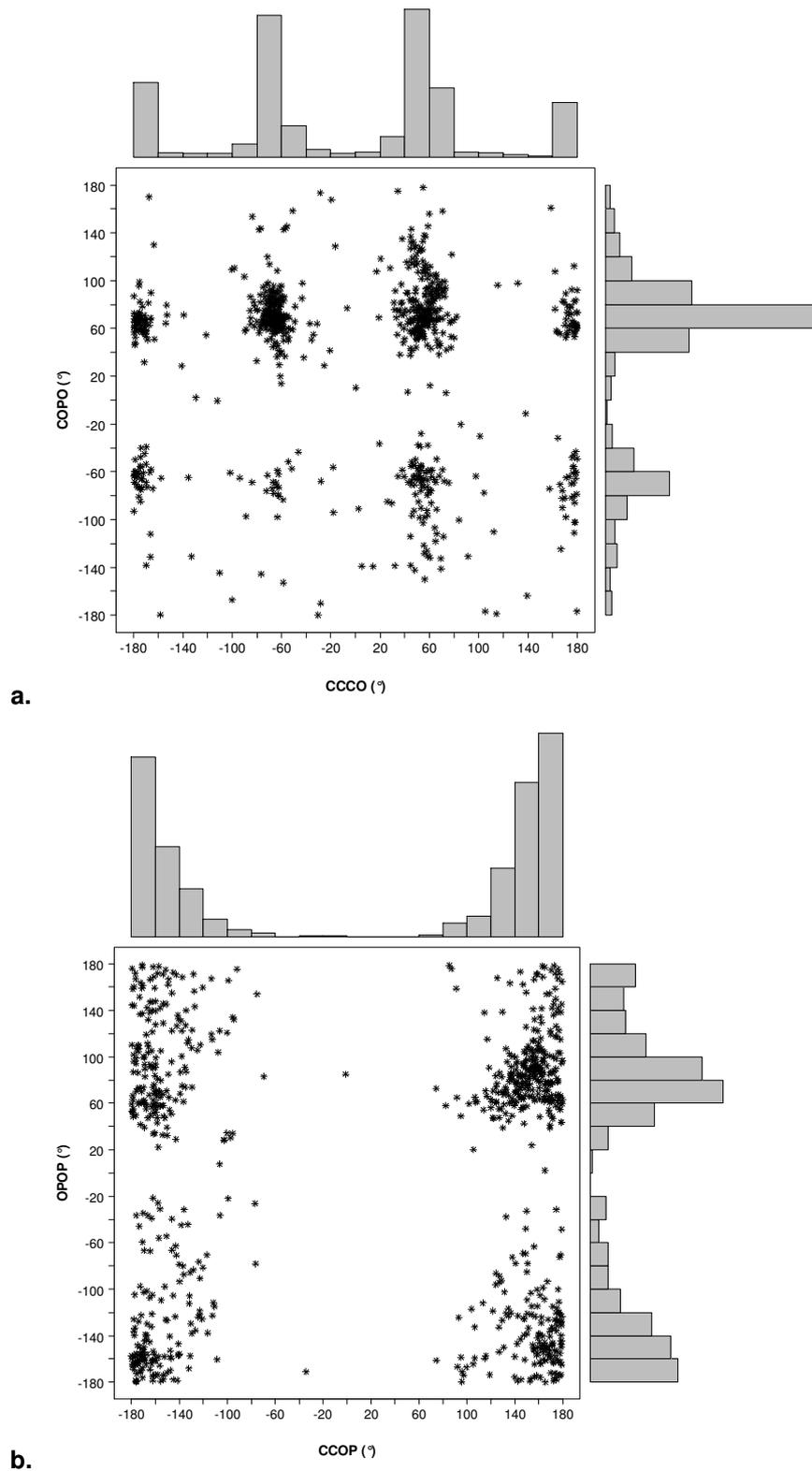


Abbildung 4.22: Paarweise Auftragung der Torsionswinkel CCCO und COPO (a) sowie CCOP und OPOP (b) für nichtaktivierte Diphosphate. Histogramme an den Achsen spiegeln die Häufigkeiten der jeweiligen Werte wieder. Jeder Balken der Histogramme fasst Werte in einem Intervall von 20° zusammen.

Allgemein führen Torsionswinkel A-B-C-D (A, B, C, D – tetraedrisch gebundene Atome) von $\pm 60^\circ$ oder $\pm 180^\circ$ zu einer gestaffelten Anordnung von A und D und sind somit oft energetisch besonders günstig. Insbesondere für die Torsionswinkel CCCO und COPO treten Wertehäufungen in diesen Bereichen auf. Die geringe Häufigkeit von COPO = $\pm 180^\circ$ kann durch die dadurch verursachte ekliptische Anordnung der Wasserstoffatome des C¹-Atoms des organischen Restes und der O[#]-Sauerstoffatome des P_α des Diphosphat erklärt werden. Durch COPO = $\pm 60^\circ$ kann dies nicht auftreten; hier kommt es jedoch zur ekliptischen Anordnung mit einem O[#] des P_α und mit O^M (Abbildung 4.23). Eine quantenchemische Energieberechnung (ohne Energieminimierung) der in Abbildung 4.23 gezeigten Propyldiphosphatkonformationen mit den in Abschnitt 4.2 vorgestellten Parametern (B3LYP, 6-31+G**, PCM) zeigt eine Stabilisierung der Struktur mit COPO = -60° gegenüber der Struktur mit COPO = $\pm 180^\circ$ um etwa 7,54 kJ (Energien beider Strukturen siehe Tabelle 9.6 im Anhang).

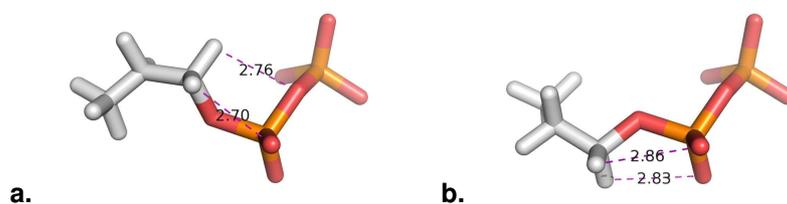


Abbildung 4.23: Darstellung der Auswirkung des Torsionswinkels COPO auf intramolekulare Wechselwirkungen. Es ist *n*-Propyldiphosphat mit COPO = -60° (a; stabiler) und COPO = $\pm 180^\circ$ (b; weniger stabil) gezeigt. Die Distanzen sind in Å angegeben.

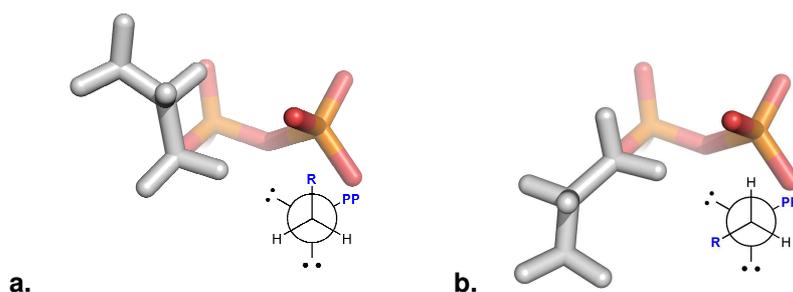


Abbildung 4.24: Darstellung der Auswirkung des Torsionswinkels CCOP am Beispiel von *n*-Propyldiphosphat mit CCOP = 60° (a) und CCOP = $\pm 180^\circ$ (b) auf mögliche sterische Hindernisse zwischen Diphosphat und organischem Rest. Zusätzlich zu den Molekülstrukturen sind jeweils auch die Newman-Projektionen angegeben. R – Propylrest, PP – Diphosphat

Für den Torsionswinkel CCOP wird eine starke Häufung im Bereich von $\pm 180^\circ$ deutlich; die Streuung der Werte ist jedoch wesentlich stärker als bei CCCO. Die Ursache für die Begünstigung von CCOP = $\pm 180^\circ$ kann in der Tatsache vermutet werden, dass dieser Wert die geringste sterische Hinderung des organischen Restes und des Diphosphats zur Folge hat. Werte um $\pm 60^\circ$ hingegen bringen das Diphosphat und den organischen Rest dichter zu-

sammen (Abbildung 4.24). Die Werte von OPOP zeigen die stärkste Streuung aller betrachteten Torsionswinkel; es lassen sich jedoch auch hier schwache Häufungen in den Bereichen um 60° und unterhalb von -180° erkennen, deren Ursache vermutlich ebenfalls sterisch günstigere Anordnungen sind.

Es bleibt anzumerken, dass ligandenspezifische Häufungen bestimmter Werte (beispielsweise für COPO = 60° im Vergleich zu -60°) nicht ausgeschlossen sind. Die Ursache hierfür liegt in der Diversität der an das Diphosphat gebundenen organischen Reste, sowie ihrer zum Teil stark differierenden Häufigkeit im zugrundeliegenden Datensatz. Aus diesem Grund werden nachfolgend die Torsionswinkel der in Abschnitt 4.1.2 beschriebenen Ligandentypen separat betrachtet. Dabei werden jeweils die Torsionswinkel aktivierter und nichtaktivierter Liganden miteinander verglichen, um Hinweise auf die Bedeutung bestimmter Ausprägungen für den Aktivierungsprozess zu erhalten.

4.5.2 Torsionswinkel von Prenyldiphosphaten

Abbildung 4.25a zeigt einen Vergleich der Torsionswinkel CCCO und COPO für aktivierte und nichtaktivierte Prenyldiphosphate; alle weiteren nichtaktivierten Diphosphate des zugrundeliegenden Datensatzes dienen als Referenz (in grau dargestellt).

Bei Betrachtung der Werteverteilung der aktivierten Prenyldiphosphate für CCCO fällt auf, dass diese außerhalb der in grau gezeigten Gruppierungen nichtaktivierter Diphosphate liegen. Insbesondere treten die Werte in Intervallen um CCCO = $\pm 120^\circ$ auf. Möglicherweise unterstützt diese Winkelausprägung also den Aktivierungsprozess. Dies wird gestützt durch die Tatsache, dass sich nichtaktivierte Prenyldiphosphate bis auf eine Ausnahme innerhalb der grauen Referenzbereiche bewegen: Es handelt es sich um den Liganden einer Isopentenylidiphosphat- Δ -Isomerase (PDB-Eintrag 2ick^[98]). Das Enzym führt in einem, in der Natur seltenen Protonierungs-Deprotonierungs-Mechanismus eine Umwandlung von Isopentenylidiphosphat (IPP) in Dimethylallyldiphosphat (DMAPP) durch (Abbildung 4.26). Der Vorgang ist reversibel; das Gleichgewicht liegt jedoch auf der Seite des DMAPP. Die Protonierung des IPP erfolgt in 2ick durch Glu149 (protonierte Form) am C⁴. Dadurch entsteht ein Carbokation (am C³), das durch Cys87 an C² deprotoniert wird, was zur Bildung des Produktes DMAPP führt. Die Rückreaktion führt ebenfalls über das angesprochene Carbokation. Eine Abspaltung des Diphosphats von diesem instabilen Intermediat würde zur Ausbildung einer zusätzlichen positiven Ladung am C¹ führen und ist dadurch nicht begünstigt. Die Aktivierung des Intermediats ist damit auch bei Ausbildung eines den Prozess begünstigenden Winkels für CCCO nicht denkbar.

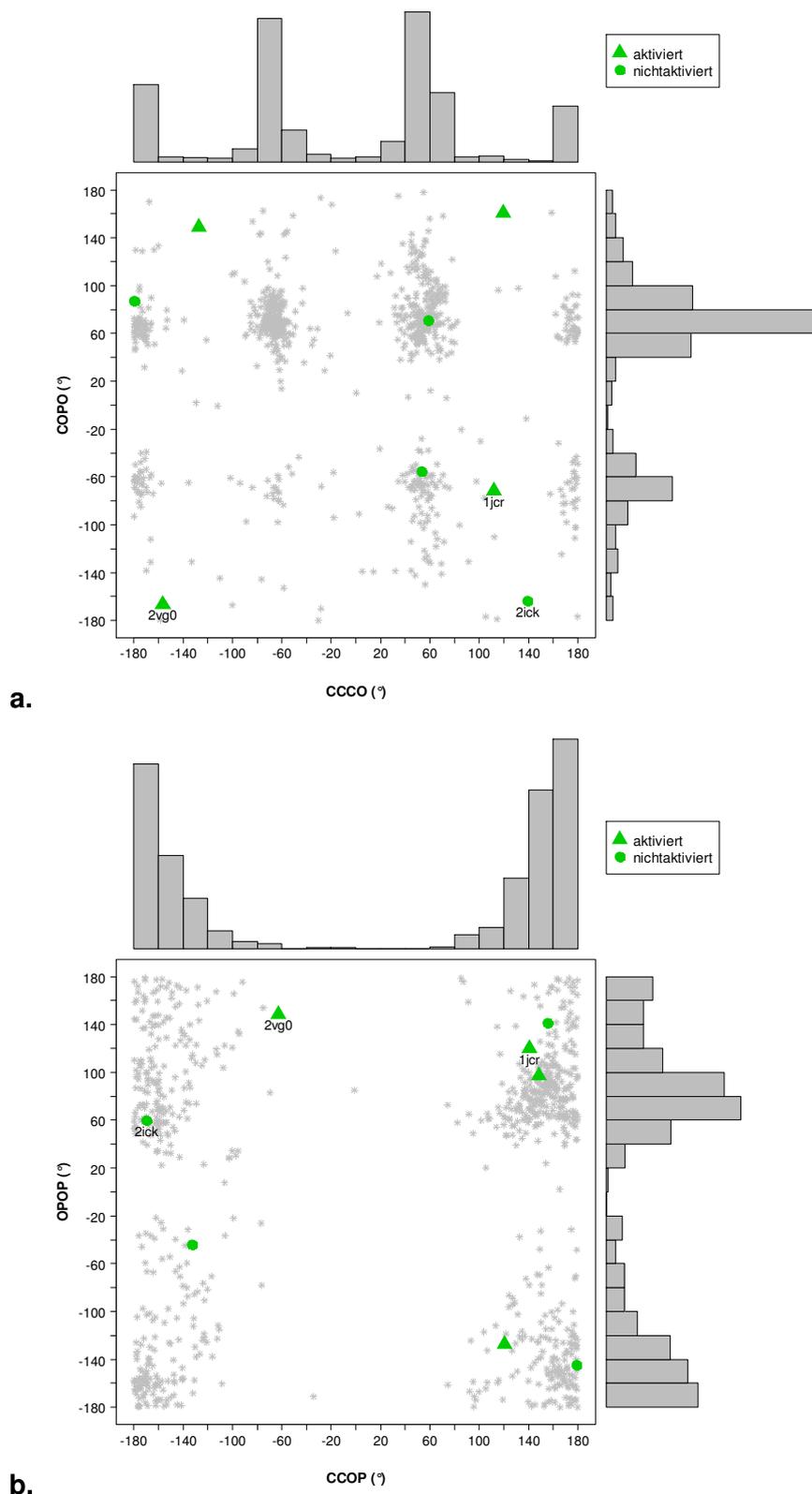


Abbildung 4.25: Paarweise Auftragung der Torsionswinkel CCCO und COPO (a), sowie CCOP und OPOP (b) aktivierter und nichtaktivierter Prenyldiphosphate (grün), sowie nichtaktivierter Diphosphate, die keine Prenyldiphosphate sind, als Referenz (graue Sterne). Histogramme an den Achsen spiegeln die Häufigkeiten der jeweiligen Winkel wieder. Jeder Balken der Histogramme fasst Werte in einem Intervall von 20° zusammen.

Die Entstehung des Carbokations an C³ des DMAPP-Substrates bzw. die Freigabe des DMAPP-Produktes laufen dabei vermutlich so schnell ab, dass eine Abspaltung des Diphosphats vom Substrat bzw. Produkt nicht stattfindet. Zusätzlich ist das Enzym zwar abhängig von zweiwertigen Metallionen (Mn²⁺ oder Mg²⁺), jedoch komplexiert dieses nicht das Diphosphat, sondern erhöht durch Wechselwirkung mit Glu149 die Azidität dieser Aminosäure.

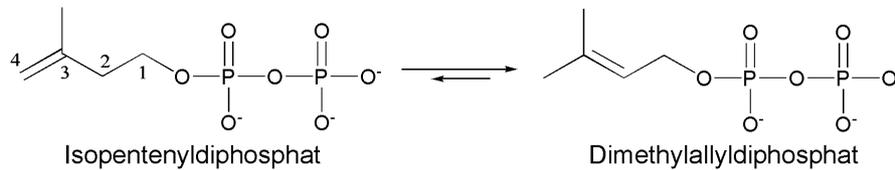


Abbildung 4.26: Schematische Darstellung der durch Isopentenyl diphosphat- Δ -Isomerase durchgeführten Isomerisierung.

Das aktivierte Diphosphat der E,Z-Farnesyl diphosphatsynthase mit PDB-Eintrag 2vg0^[231] liegt mit einem Wert für CCCO von -160° nahe am häufig für nichtaktivierte Diphosphate auftretenden Wert von CCCO = $\pm 180^\circ$. Dies kann leicht geklärt werden, wenn der in der Binde-tasche kristallisierte Ligand genauer betrachtet wird: Es handelt sich um Citronellyl diphosphat, das dem Enzym anstelle des natürlichen Substrats Geranyl diphosphat (GPP) zuge-setzt wurde und das im Gegensatz zu GPP zwischen C² und C³ gesättigt ist, also keine Doppelbindung besitzt (Abbildung 4.27). Dies hat zur Folge, dass in diesem Bereich keine Planarität des organischen Restes vorliegt, was sich auch auf den Wert des CCCO-Torsions-winkels auswirken kann.

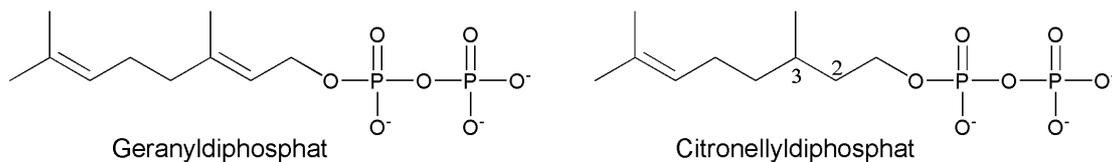


Abbildung 4.27: Schematische Darstellung von Geranyl- und Citronellyl diphosphat.

Ähnlich wie für CCCO fällt anhand von Abbildung 4.25 für COPO mit Werten im Bereich von $\pm 180^\circ$ ebenfalls eine Häufung der aktivierten Prenyl diphosphate außerhalb der Haupt-gruppierungen nichtaktivierter Diphosphate auf und könnte somit auf eine Bedeutung dieses Torsionswinkels für den Aktivierungsprozess hindeuten. Beim „Ausreißer“ mit PDB-Eintrag 1jcr^[230] handelt es sich um eine Proteinfarnesyltransferase. Die Kristallstruktur enthält das natürliche Substrat Farnesyl diphosphat (FPP), sowie einen Peptidinhibitor, der aufgrund seiner Konformation in der Proteinbindetasche eine Reaktion verhindert. Ein Einfluss dieser unnatürlichen Konformation auf die in der Röntgenstruktur ausgeprägte Konformation des FPP, insbesondere des COPO-Torsionswinkels, kann *per se* nicht ausgeschlossen werden.

Für die Torsionswinkel CCOP und OPOP (Abbildung 4.25b) konnte keine klare Abgrenzung der aktivierten von nichtaktivierten (Prenyl-) Diphosphaten detektiert werden. Eine Bedeutung dieser Winkel für den Aktivierungsprozess wird daher – zumindest für die betrachteten aktivierenden Enzyme – vom Autor als unwahrscheinlich angesehen. Für die im Datensatz nicht enthaltenen Terpenzyklen könnten CCOP und OPOP jedoch durchaus eine Rolle spielen: Bei einigen dieser Enzyme findet vor der Zyklisierungsreaktion zunächst eine Übertragung des Diphosphats von der C¹- auf die C³-Position des Prenylrestes statt. Eine Ausprägung der Torsionswinkel CCOP und OPOP derart, dass Sauerstoffe des P_β (siehe Abbildung 4.16) in räumliche Nähe zur C³-Position gebracht werden, könnte diesen Umlagerungsprozess entscheidend fördern. Für die im Datensatz berücksichtigten prenylkonvertierenden Enzyme scheinen lediglich die Winkel CCCO und COPO einen möglichen (stereo-elektronischen) Einfluss auf die Aktivierung zu haben; der Zusammenhang wird daher in Abschnitt 4.5.5 mittels quantenchemischer Methoden detaillierter untersucht.

4.5.3 Torsionswinkel von Nucleosiddiphosphat-Zuckern

Im Gegensatz zu Prenyldiphosphaten ist die Bestimmung einiger Torsionswinkel in Nucleosiddiphosphat-Zuckern (NDP-Zuckern) aufgrund der unmittelbaren Anknüpfung der Zucker-Ringstruktur an das Diphosphat zunächst nicht eindeutig (siehe Abbildung 4.28). Insbesondere kann neben CCCO auch COCO bestimmt werden. Analog gilt dies auch für CCOP und OCOP – hier geben beide Winkel sogar eine Rotation um dieselbe Bindung an.

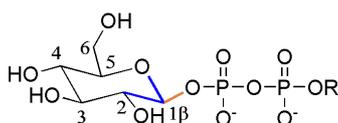


Abbildung 4.28: Darstellung einer NDP- β -D-Glucose (Sesselkonformation) zur Veranschaulichung der Torsionswinkel CCCO, COCO (zentrale Bindung blau), CCOP und OCOP (zentrale Bindung orange). R – Nucleosidrest

Es kann jedoch gezeigt werden, dass – bei Voraussetzung der energetisch stabilen Sesselkonformation des aktivierten Zuckers – eine Betrachtung von CCCO und CCOP ausreichend ist, da sich die Winkel für COCO und OCOP in diesem Fall aus den Werten von CCCO und CCOP ableiten lassen. Aufgrund der Sesselkonformation kann der CCCO-Torsionswinkel nur Werte nahe -60° (α - bzw. axiale Konformation) oder $\pm 180^\circ$ (β - bzw. äquatoriale Konformation) annehmen. Der korrespondierende Winkel COCO ergibt sich dann zu Werten nahe $\pm 180^\circ$ bzw. $+60^\circ$. Betrachtet man CCOP und OCOP, so stellt man fest, dass sich der Betrag der Differenz ihrer Werte stets zu 120° ergibt. Aufgrund dieser Zusammenhänge werden von

diesen vier Torsionswinkeln nachfolgend nur die Winkel CCCO und CCOP berücksichtigt. Zusätzlich werden – wie auch bei den Prenyldiphosphaten – die eindeutigen Winkel COPO und OPOP in Betracht gezogen.

Abbildung 4.30 zeigt die Ausprägung der betrachteten Torsionswinkel CCCO, CCOP, COPO und OPOP in aktivierten und nichtaktivierten NDP-Zuckern im Vergleich. Für CCCO zeigen sich die erwarteten Winkel von -60° (α - bzw. axiale Konformation) und $\pm 180^\circ$ (β - bzw. äquatoriale Konformation), wobei die aktivierten NDP-Zucker ausschließlich die axiale Konformation ausprägen. Das deutlich häufiger beobachtete Auftreten der axialen (α) gegenüber der sterisch weniger gehinderten äquatorialen (β) Konformation macht den Einfluss des anomeren Effektes deutlich (Abbildung 4.29). Die Stabilisierung der axialen Anordnung bewegt sich dabei in einer Größenordnung von etwa $1,4 \text{ kcal} \cdot \text{mol}^{-1}$ – $2,8 \text{ kcal} \cdot \text{mol}^{-1}$ gegenüber der äquatorialen Konformation.^[255] Die axiale Anordnung führt – im Gegensatz zur äquatorialen – zu einer Überlappung eines sp^3 -Hybridorbitals des Brückensauerstoffatoms O^{Br} des Zuckers mit einem antibindenden Orbital am anomeren Kohlenstoff (Hyperkonjugation). Die Folgen sind eine Verlängerung (Schwächung) der $\text{C}^1\text{O}^{\text{R}}$ -Bindung (Abnahme der Elektronendichte), sowie eine Verkürzung (Stabilisierung) der $\text{O}^{\text{Br}}\text{C}^1$ -Bindung (Zunahme der Elektronendichte). Somit aktiviert eine axiale Anordnung zusätzlich die während der Aktivierung gespaltene $\text{C}^1\text{O}^{\text{R}}$ -Bindung. Dennoch ist der anomere Effekt nicht zur Unterscheidung aktivierter und nichtaktivierter NDP-Zucker geeignet.

Der Torsionswinkel CCOP tritt in Abbildung 4.30 a gestreut auf, nimmt aber bis auf eine Ausnahme stets negative Werte an. Die Ursache hierfür liegt in sehr engen Kontakten von Atomen des Diphosphats mit Atomen des aktivierten Zuckers für positive Winkel. Der Ausreißer (PDB-Eintrag 2rj7^[225]) liegt mit einem Wert von 5° noch sehr nah am sterisch nicht gehinderten Bereich für CCOP.

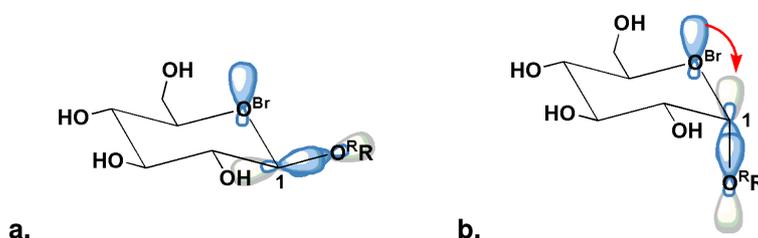
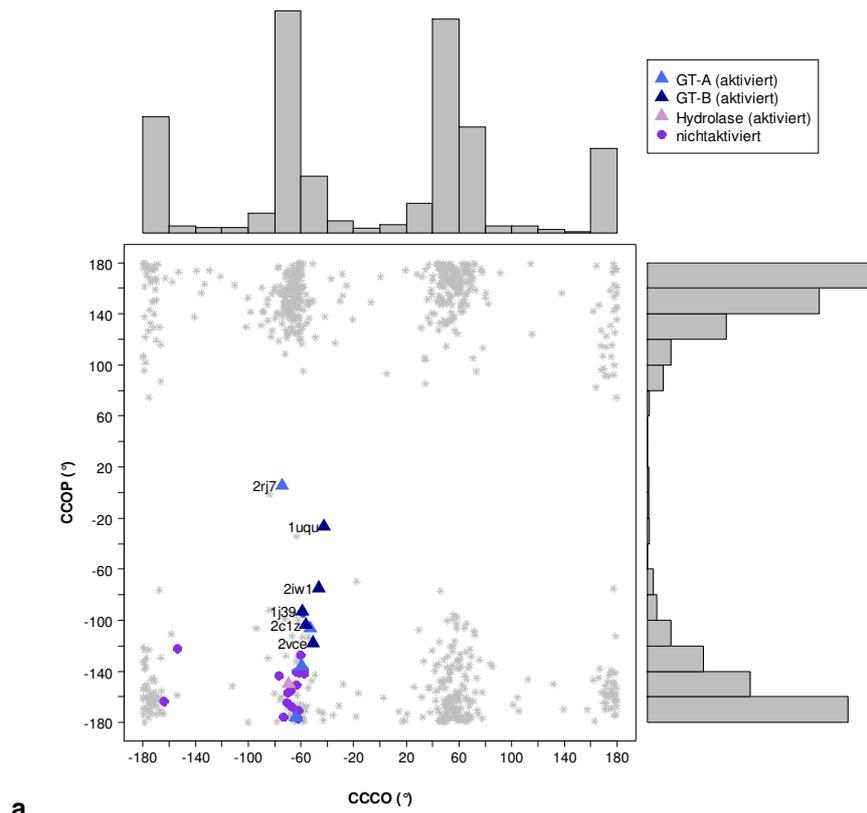
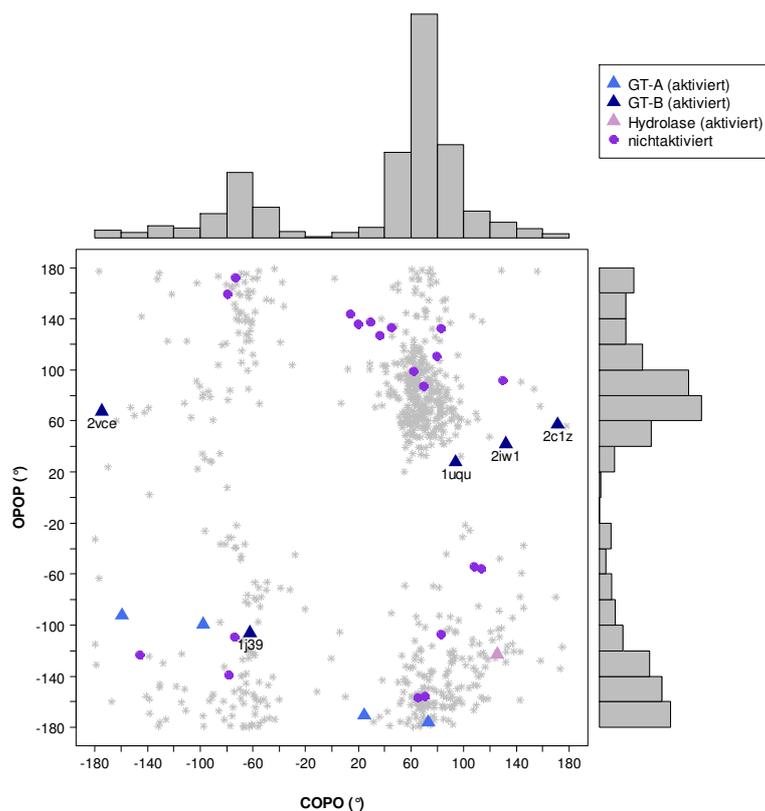


Abbildung 4.29: Darstellung der äquatorialen (a.) und axialen (b.) Anordnung des Substituenten am C^1 -Atom am Beispiel von Glucose-1-Phosphat ($\text{R} = \text{PO}_3^{2-}$). Für das Brückensauerstoffatom des Zuckerrings ist nur eines der beiden nichtbindenden sp^3 -Orbitale gezeigt. Das antibindende Molekülorbital der $\text{C}^1\text{O}^{\text{R}}$ -Bindung ist in grau dargestellt.



a.



b.

Abbildung 4.30: Paarweise Auftragung der Torsionswinkel CCCO und CCOP (a), sowie COPO und OPOP (b) aktivierter und nichtaktivierter NDP-Zucker (blau), sowie nichtaktivierter Diphosphate, die keine NDP-Zucker sind, als Referenz (graue Sterne). Histogramme an den Achsen spiegeln die Häufigkeiten der jeweiligen Winkel wieder. Jeder Balken der Histogramme fasst Werte in einem Intervall von 20° zusammen.

Die Werte für COPO und OPOP sind stark gestreut (Abbildung 4.30 b). Es ist keine klare Trennung der Torsionswinkelausprägungen aktivierter Formen gegenüber den nichtaktivierten Formen erkennbar. Die in der vorliegenden Arbeit betrachteten aktivierten NDP-Zucker können in drei Gruppen unterschieden werden: Glycosidasen, Glycosyltransferasen vom GT-A-Faltungstyp und Glycosyltransferasen vom Faltungstyp GT-B (siehe Abschnitt 1.2.1). Weiterhin hat man in Abbildung 4.18 gesehen, dass die betrachteten nichtaktivierten NDP-Zucker in keinem Fall durch Metallionen komplexiert sind. Da für die Glycosidase des Datensatzes (PDB-Eintrag 2i8t^[3]) und für Glycosyltransferasen vom Typ GT-A im Allgemeinen eine Abhängigkeit der Aktivität von Metallionen vorliegt, könnte dies für diese Enzyme den für die Aktivierung notwendigen Faktor darstellen. Die nicht vorhandene klare Abgrenzung der berücksichtigten Torsionswinkel der Liganden dieser Enzyme von nichtaktivierten NDP-Zuckern lässt die Bedeutung dieser Winkel für die Aktivierung in den Hintergrund treten. Da GT-B-Glycosyltransferasen jedoch unabhängig von Metallionen sind und für sie – im Vergleich zu nichtaktivierenden NDP-Zucker-Bindestellen – auch keine signifikant höhere Zahl von Brønsted-Säuren in der Bindetasche vorliegt (siehe Abbildung 9.1), die eine Protonierung des Diphosphats herbeiführen könnten, muss es für diese Gruppe aktivierender Enzyme andere Mechanismen geben, die Aktivierung zu steuern. Bei Betrachtung von Abbildung 4.30 fällt auf, dass für vier der fünf GT-B-Enzyme eine Gruppierung der Torsionswinkel vorliegt – sie liegen auf einer Geraden, wenn man $+180^\circ = -180^\circ$ berücksichtigt. Für den Ausreißer mit PDB-Eintrag 1j39^[16] erfolgte – im Gegensatz zu den verbleibenden vier GT-B-Typ-Enzymen – eine Kristallisierung durch „Soaking“ des Liganden UDP- α -D-Glucose in die bereits auskristallisierte Proteinstruktur. Konformationsänderungen des Proteins, die durch Ligandenbindung verursacht werden und die Ausbildung bestimmter Ligandenkonformationen erwirken, sind damit stark eingeschränkt. Es konnte jedoch gezeigt werden, dass die Kristallstruktur in der geschlossenen Konformation, also der Konformation mit vollständig ausgebildetem aktivem Zentrum, vorliegt.^[16] Zusätzlich geben Moréra *et al.*^[256] an, dass bei einer Kokristallisierung von UDP- α -D-Glucose (ohne Vorliegen eines Zucker-Akzeptors, wie auch in 1j39 der Fall) eine Aktivierung stattfindet. Demnach wird eine Unterbindung der Aktivierung durch Fehlpositionierung oder eine veränderte Konformation des Substrates in 1j39 wahrscheinlich, was auch eine Veränderung der in Abbildung 4.30 berücksichtigten Torsionswinkel einbeziehen könnte.

Nun stellt sich die Frage nach der Bedeutung der in Abbildung 4.30 b beobachteten Ausprägungen von COPO und OPOP für GT-B-Liganden. Betrachtet man die Konformationen dieser Strukturen, so wird eine „U-Form“ sichtbar (Abbildung 4.31). Diese kann geschlossen – bei den Retentionsvarianten der Glycosyltransferasen mit PDB-Kodierungen 1uqu^[226], 2iw1^[228] – oder etwas geöffnet – bei den invertierenden Enzymen 2c1z^[227] und 2vce^[229] – sein.

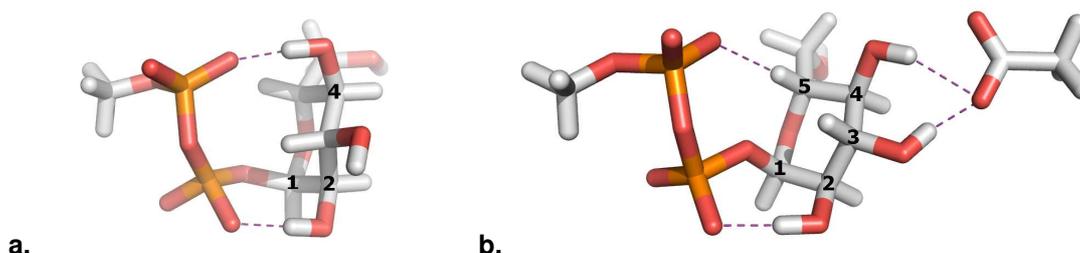


Abbildung 4.31: Darstellung der geschlossenen (a) und geöffneten (b) U-Form von NDP-Zucker-Liganden. Die Diposphate sind in ihrer deprotonierten Form gezeigt. Die Nucleosidreste sind durch Methylgruppen ersetzt, Aspartat bzw. Glutamat durch Acetat vereinfacht dargestellt, z.T. ist die Nummerierung der Kohlenstoffatome des Zuckers angegeben. Die ausgebildeten Wechselwirkungen sind in violett angedeutet. Als Grundlage für die Darstellung der geschlossenen Form diente der Ligand von 1uqu^[226], für die offene Form der Ligand von 2c1z^[227].

Für die aktivierten NDP-Zucker-Liganden der GT-A-Typ-Enzyme, die Glycosidase oder die nichtaktivierten NDP-Zucker-Liganden konnte solch eine U-Form nicht beobachtet werden. Die Gruppierung der GT-B-Glycosyltransferasen für die Torsionswinkel COPO und OPOP in Abbildung 4.30 spiegelt diese Konformation wieder. Sowohl die geschlossene als auch die geöffnete Form bilden eine Wasserstoffbrücke zwischen C²-OH des aktivierten Zuckers und Diphosphat aus. Diese, sowie die Bedeutung aller weiteren Hydroxylgruppen des aktivierten Zuckers wurden bereits in Glycosidasen (i.A. ohne diphosphathaltiges Substrat) mittels kinetischer Messungen untersucht – ein Beitrag von 22 kJ*mol⁻¹^[257] bis 33 kJ*mol⁻¹^[258] zur Stabilisierung des bei der Aktivierung entstehenden kationischen Übergangszustandes (Verringerung der Aktivierungsenergie) wurde dabei für C²-OH ermittelt. Auch der „Ausreißer“ 1j39 bildet diese Wasserstoffbrücke aus. Ebenso ist sie bei zwei nichtaktivierten NDP-Zuckern (PDB-Kodierungen 1p4n^[259], 2q3e^[52]) zu finden. Die U-Form ermöglicht die Ausbildung einer weiteren intramolekularen Wechselwirkung zwischen C⁴-OH und Diphosphat bei der geschlossenen und C⁵-H und Diphosphat bei der geöffneten Form (Abbildung 4.31). Die freie Hydroxylgruppe des C⁴-Atoms, sowie des C³-Atoms treten in der geöffneten Form intermolekular mit einer Aspartat- oder Glutamatseitenkette des Proteins in Wechselwirkung. Im Zusammenhang mit Glycosidasen wurde festgestellt, dass das Fehlen des C⁴-OH zur Verringerung der Umsatzgeschwindigkeit des Enzyms führt.^[257] McCarter *et al.* schätzen die durch die Wechselwirkung verursachte Stabilisierung des Übergangszustandes auf mindestens 16,7 kJ*mol⁻¹.^[258] Die C³-OH-Wechselwirkung senkt die Aktivierungsenergie um 7 kJ*mol⁻¹^[257] bis 16,7 kJ*mol⁻¹^[258]. Zusätzlich wurde beobachtet, dass eine Mutation der Aminosäure D374 der invertierenden Glycosyltransferase 2c1z, die an der Wechselwirkung mit der C⁴-Hydroxylgruppe beteiligt ist, zu Alanin zu einem völligen Verlust der Enzymaktivität führte.^[227] Während die beschriebene C⁴-OH-Diphosphat-Wechselwirkung in keinem der betrachteten nichtaktivierten NDP-Zucker beobachtet werden kann, liegt die C⁴-OH-Carboxyl-Wechselwirkung (teils in Kombination mit der C³-OH-Wechselwirkung) in vier

nichtaktivierten Liganden (PDB-Kodierungen 3beo^[251], 1muu^[260], 1uae^[261], 2yvw^[52]) vor. Für diese Bindestellen ist aber weder eine C²-OH, noch eine C⁵-H-Diphosphat-Interaktion festzustellen.

Es lässt sich schlussfolgern, dass der Aktivierungsprozess von NDP-Zuckern im Fall von GT-A-Enzymen und der vorliegenden Hydrolase – Glycosidasen sind nicht generell metallionenabhängig – durch Metallionen oder im Fall der GT-B-Enzyme durch die Summe mehrerer Wasserstoffbrücken beeinflusst wird. Für detaillierte Aussagen – insbesondere zu exakten Energiebeträgen – wären jedoch reine quantenmechanische oder QM/MM-Rechnungen sinnvoll, die in der vorliegenden Arbeit nicht durchgeführt wurden.

4.5.4 Torsionswinkel von Pterindiphosphaten

Ähnlich wie bei NDP-Zuckern ist auch bei dem betrachteten Pterindiphosphatliganden (Abbildung 4.32) die Bestimmung der berücksichtigten Torsionswinkel nicht in jedem Fall eindeutig: Neben CCCO kann aufgrund der an das Diphosphat angeknüpften Ringstruktur auch NCCO ermittelt werden. Beide Torsionswinkel beschreiben eine Rotation um dieselbe Bindung.

Da jedoch nur ein einziger Ligand der in Abbildung 4.32 dargestellten Struktur im dieser Arbeit zugrundeliegenden Datensatz vorliegt und sich NCCO für diese Struktur durch $NCCO = CCCO + 179,08^\circ$ ergibt, wird nachfolgend nur der Torsionswinkel CCCO angegeben.

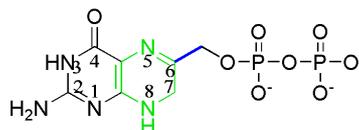


Abbildung 4.32: Struktur des 6-Hydroxymethyl-7,8-Dihydropterindiphosphats – des aktivierten Liganden in der Dihydropteroatsynthese 1aj2. Die zentrale Bindung der Torsionswinkel CCCO bzw. NCCO ist blau markiert, der planare Dihydropyrazinring ist in grün dargestellt.

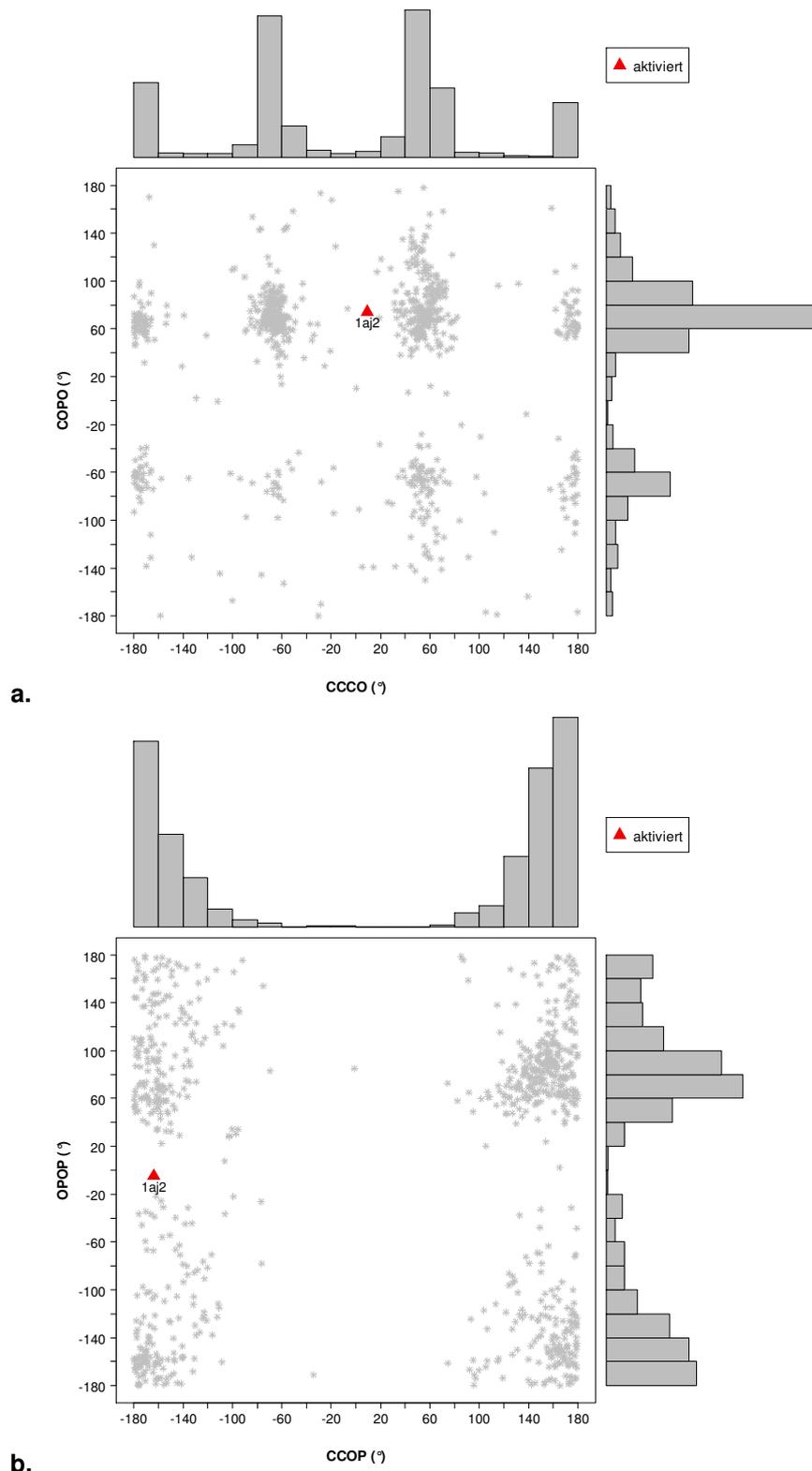


Abbildung 4.33: Paarweise Auftragung der Torsionswinkel CCCO und COPO (a) sowie CCOP und OPOP (b) des aktivierten Pterindiphosphatliganden aus der PDB-Struktur 1aj2 (rot), sowie nichtaktivierter Diphosphate (graue Sterne) als Referenz. Histogramme an den Achsen spiegeln die Häufigkeiten der jeweiligen Werte wieder. Jeder Balken der Histogramme fasst Werte in einem Intervall von 20° zusammen.

Aufgrund des Vorliegens lediglich einer einzigen aktivierten Pterindiphosphatstruktur (PDB-Eintrag 1aj2^[137]), sowie des Fehlens nichtaktivierter Pterindiphosphate der in Abbildung 4.32 gezeigten Struktur als Referenz, ist eine Interpretation der Torsionswinkelausprägungen schwierig. Der Datensatz beinhaltet zwar 13 nichtaktivierte Pterindiphosphate, diese sind aber durch ihre strukturelle Abweichung von der in Abbildung 4.32 gezeigten Struktur nicht mit dieser vergleichbar: Es handelt sich um molybdopterinhaltige Strukturen (siehe Abbildung 4.34) – ein Kofaktor in den meisten Mo- oder W-abhängigen Enzymen. Für alle 13 im Datensatz enthaltenen Strukturen ist die Methylphosphatgruppe des Molybdopterins durch ein Nucleosiddiphosphat ersetzt.

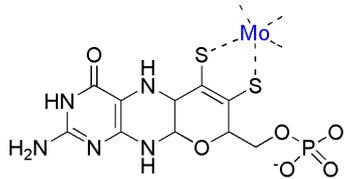


Abbildung 4.34: Schematische Darstellung von Molybdopterin mit komplexiertem Mo. Koordinative Bindungen zum Molybdän sind als unterbrochene Linie angedeutet.

Aufgrund der Doppelbindung zwischen Atom 5 und 6 des Pterinring-Systems lässt sich mutmaßen, dass – ähnlich wie bei den Prenyldiphosphaten – der Torsionswinkel CCCO einen stereoelektronischen Einfluss auf die Aktivierung haben könnte (siehe vergleichsweise Abschnitt 4.5.5). Aus Abbildung 4.33 kann dies jedoch nicht abgeleitet werden. Mögliche Ursachen hierfür sind das Fehlen des Akzeptors, sowie des für die Katalyse notwendigen Metallions in der Röntgenstruktur, was möglicherweise zu einer nichtnativen Konformation des Pterindiphosphatliganden führt. Bisher existiert keine Röntgenstruktur einer Dihydropteroat-Synthase, die sowohl Donor- als auch Akzeptorliganden enthält. Der Einfluss des Akzeptors auf die Donorkonformation – insbesondere auf den Torsionswinkel CCCO – kann damit nicht bestimmt werden. Es liegt zwar mit der Röntgenstruktur 1aj0^[137] eine Struktur vor, die ein Donoranalogen und einen Akzeptorinhibitor enthält; die Übereinstimmung der Inhibitorpositionierung mit der eines echten Akzeptors wird jedoch in verschiedenen Quellen angezweifelt^[102;103]. Zusätzlich wird für Dihydropteroat-Synthasen ein S_N2-Mechanismus angenommen^[102;103]. Damit ist die stereoelektronische Bedeutung von CCCO für den Aktivierungsprozess dieser Enzyme als eher gering einzuschätzen, da der Torsionswinkel für eine erleichterte Ausbildung einer Mesomeriestabilisierung eines während der Aktivierung entstehenden Kations verantwortlich gemacht wird (siehe Abschnitt 4.5.5). Da solch ein Kation in Zusammenhang mit einem S_N2-Mechanismus nicht oder nur partiell vorliegt, wird der Effekt der Resonanzstabilisierung geringer sein. Es ist in der Ligandenkonformation außerdem keine intramolekulare Wechselwirkung ähnlich zu NDP-Zuckern erkennbar.

Ebenso ist dem Autor keine Quelle bekannt, die einen Hinweis auf die Bedeutung intramolekularer stereoelektronischer Kontrolle oder Wechselwirkungen, wie Wasserstoffbrücken, innerhalb des Liganden auf den Aktivierungsprozess bei Pterindiphosphaten gibt. Eine starke Konservierung der Aminosäuren der Pterindiphosphatbindetaschen von Dihydropteroatsynthasen in verschiedenen Organismen^[137] weist jedoch auf einen starken Selektionsdruck hin. Eventuell ist auch die Sicherung einer fest definierten Ligandenkonformation und damit festgelegter Torsionswinkel hier ausschlaggebend. Genauere Aussagen einer Torsionswinkelabhängigkeit der Aktivierung von Pterindiphosphaten der in Abbildung 4.32 gezeigten Struktur erfordern einen größeren Pool von proteingebundenen Pterinliganden bzw. quantenchemische Berechnungen, die in der vorliegenden Arbeit nicht durchgeführt wurden. Abschließend soll darauf hingewiesen werden, dass das Enzym Mg^{2+} -abhängig ist; ob die Funktion dieses Metallion jedoch in der Komplexbildung des Diphosphats liegt, ist bisher nicht abschließend geklärt^[102;103].

4.5.5 Quantenchemische Untersuchungen von Torsionswinkeln in Prenyldiphosphaten

In Abschnitt 4.5.2 wurden in Röntgenstrukturen Torsionswinkel beobachtet, die möglicherweise im Zusammenhang mit dem Aktivierungsprozess stehen, da sie sich für aktivierte und nichtaktivierte Prenyldiphosphate unterscheiden. Dies bezieht sich insbesondere auf die Torsionswinkel COPO und CCCO für die in aktivierenden Bindestellen Werte um $\pm 180^\circ$ für COPO bzw. $\pm 120^\circ$ für CCCO gesammelt auftreten. Für die weiterhin in Abschnitt 4.5.2 berücksichtigten Torsionswinkel CCOP und OPOP konnte hingegen kein Unterschied zwischen aktivierten und nichtaktivierten Diphosphaten festgestellt werden: OPOP tritt in aktivierten wie nichtaktivierten Prenyldiphosphaten gestreut auf, während CCOP sich in beiden Fällen grob um $\pm 180^\circ$ zeigt, was vermutlich auf die damit verbundene langgestreckte Konformation mit wenigen sterischen Hinderungen zurückzuführen ist. Zudem halten sich sowohl für aktivierte als auch für nichtaktivierte Prenyldiphosphate beide Torsionswinkel CCOP und OPOP im Feld der Referenzwerte der nichtaktivierten, nichtprenylierten Diphosphate auf. Basierend auf diesen Ergebnissen sollen im vorliegenden Abschnitt die Torsionswinkel COPO und CCCO detaillierter untersucht, sowie ihr (möglicher) Zusammenhang mit der Aktivierung geklärt werden. Die Torsionswinkel CCOP und OPOP werden hingegen aus oben genannten Gründen nicht weiter betrachtet.

Die Verwendung eines einfachen Modellmoleküls auf Grundlage von DMAPP und ohne Einbeziehung einer Enzymumgebung soll die Rechenprozesse verkürzen und allgemeine

Aussagen über die Effekte der Torsionswinkel zulassen. Es wurde durch schrittweise Verlängerung der C¹O^R-Bindung des DMAPP eine Struktur erzeugt, die dem reaktiven Intermediat der Aktivierungsreaktion, welches aus Dimethylallylkation (DMA-Kation) und anionischem Diphosphat besteht, entspricht. Da sich CCCO und COPO über beide Teilstrukturen (DMA-Kation und Diphosphatanion) erstrecken, ist es notwendig, aus der Menge der durch das schrittweise Vorgehen entstehenden Strukturen genau den Zustand zu ermitteln, für den das DMA-Kation vollständig ausgebildet ist, aber gleichzeitig der C¹O^R-Abstand gerade noch gering genug ist, sodass ein Einfluss des Kations und Anions auf ihre gegenseitige Konformation vorliegt. Ist diese Distanz zu groß, so findet eine unabhängige Konformationsbildung beider Strukturen statt. Durch Rotation der Torsionswinkel COPO und CCCO dieser so ermittelten Struktur können die optimalen (energetisch günstigsten) Ausprägungen der Torsionswinkel bestimmt werden. Unter der Voraussetzung, dass die im Intermediat beobachteten, energetisch optimalen Werte durch das Enzym im Übergangszustand eingestellt werden, können die in Abschnitt 4.5.2 in den Röntgenstrukturen beobachteten Winkelwerte interpretiert werden.

Ausgangspunkt der Erzeugung der Intermediatstruktur bildet das vollständig protonierte DMAPP⁽³⁾ (Grundzustand). Die Bindungsverlängerung erfolgt schrittweise um je 0,2 Å ausgehend von 1,49 Å für den Grundzustand von DMAPP⁽³⁾ bis 3,69 Å (12 Schritte). Nach jedem Schritt erfolgt eine Energie- und Strukturoptimierung, wobei die Länge der C¹O^R-Bindung dabei fixiert bleibt. Die optimierte Struktur des vorhergehenden Schrittes dient jeweils als Ausgangskonformation des aktuellen Schrittes. Die verwendete quantenchemische Methode und Basissatz, sowie die Solvatationsbedingungen sind in Abschnitt 4.3 beschrieben. Es entsteht eine Reihe von Strukturen, deren CCCO- und COPO-Torsionswinkel in Abbildung 4.35 dargestellt sind. Die den Strukturen zugehörigen Energien sind in Tabelle 9.7 (Anhang) gezeigt. Es zeigt sich für COPO zunächst ein konstanter Wert bei ca. 60° (gestaffelte Anordnung), der sich bei 2,49 Å steigert und sich schließlich auf ca. 85° einstellt. Der Torsionswinkel CCCO fällt von 111° auf etwa 90°, bleibt dort zunächst konstant, bevor er ab 3,09 Å wieder steigt. Die Wahl der reaktiven Intermediatstruktur aus der Menge der erzeugten Strukturen erfolgt nach obigen Vorgaben: Insbesondere eine Veränderung des Monotonieverhaltens des in Abbildung 4.35 dargestellten Kurvenverlaufes deutet auf eine unabhängige Konformationsbildung hin. Solch ein Verhalten ist für den Torsionswinkel CCCO im Bereich ab einem C¹O^R-Abstand von etwa 3,09 Å zu beobachten. Ein weiteres Kriterium für die Wahl der Struktur ergibt sich aus der Orbitalhybridisierung am C¹: Während der Aktivierung geht das Atom von einem sp³- in einen sp²-hybridisierten Zustand über. Dies ist gleichbedeutend mit einem vollständigen Übergang der im Grundzustand vorliegenden tetraedrischen Konformation am C¹ in eine planare Konformation des entstehenden Dimethylallylkations (DMA-Kation). Diese Planarität äußert sich in Abbildung 4.35 durch CCCH = 180°. Die Struktur mit

einer C^1O^R -Distanz von 2,69 Å nähert sich der Planarität stark an. Aufgrund dieses Kriteriums und der Tatsache des veränderten Monotonieverhaltens im CCCO-Torsionswinkelverlauf wurde diese Struktur für die nachfolgenden Torsionswinkelanalysen ausgewählt. Die Konformation ist in Abbildung 4.36b dargestellt.

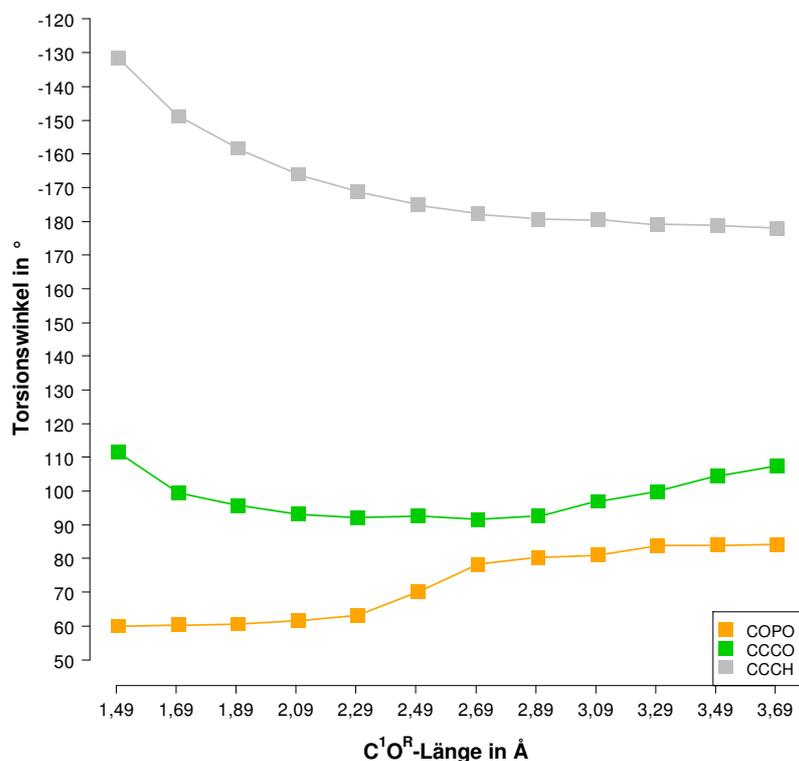


Abbildung 4.35: Verlauf der Torsionswinkel CCCO, COPO und CCCH ($C^3C^2C^1H^{C1}$) während der Verlängerung der C^1O^R -Bindung in DMAPP⁽³⁾.

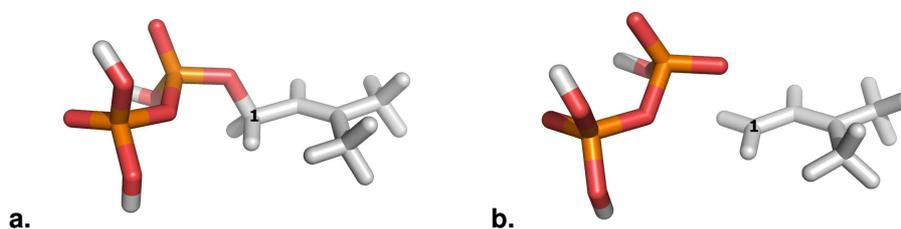


Abbildung 4.36: Vergleich der Konformationen des Grundzustandes (a; C^1O^R -Distanz = 1,49 Å) und des reaktiven Intermediates (b; C^1O^R -Distanz = 2,69 Å) der C^1O^R -Bindungsverlängerung. Insbesondere zu beachten ist der Übergang von einer sp^3 - in eine sp^2 -Hybridisierung am C^1 -Atom.

Ausgehend von der erzeugten reaktiven Struktur des DMAPP⁽³⁾ erfolgte die aufeinanderfolgende Rotation der Torsionswinkel COPO und CCCO in 10°-Schritten; eine kombinierte Rotation beider Torsionswinkel wäre durch die damit verbundene Erzeugung sehr vieler

Strukturen zu zeitaufwändig. Für CCCO können die berücksichtigten Winkel zusätzlich auf den Bereich von 0° bis 180° reduziert werden. Dies ist aufgrund der Planarität des DMA-Kations und der Rotation um eine Bindung in der Ebene des Kations ohne großen Informationsverlust möglich. Die Energiewerte für den verbleibenden Winkelbereich können durch die Werte des berücksichtigten Intervalls approximiert werden. Zunächst wurde COPO variiert. Dies erfolgte unter Fixierung der C^1O^R -Bindungslänge, um die Trennung von DMA-Kation und Diphosphatanion zu gewährleisten; CCCO unterlag hingegen während dieser Berechnung keinerlei Fixierung. Aus der Menge der so erzeugten Konformationen diente nachfolgend die stabilste (energieärmste) Struktur als Grundlage für die Variierung des Torsionswinkels CCCO. Während dieser Optimierungen blieb COPO frei drehbar, während die C^1O^R -Bindungslänge wie zuvor fixiert wurde. Die jeweils verwendete quantenchemische Methode und Basissatz, sowie die Solvatationsbedingungen sind Abschnitt 4.3 zu entnehmen. Diese Vorgehensweise erlaubt einen unmittelbaren Vergleich des Einflusses verschiedener Ausprägungen für COPO und CCCO auf die Gesamtenergie der in Abbildung 4.36b gezeigten Struktur. Durch das Zulassen der freien Bewegung aller weiteren Torsions- und Bindungswinkel, sowie Bindungslängen des Moleküls (außer der C^1O^R -Bindungslänge) wird sichergestellt, dass sich für diese (zumindest lokale) Minima einstellen.

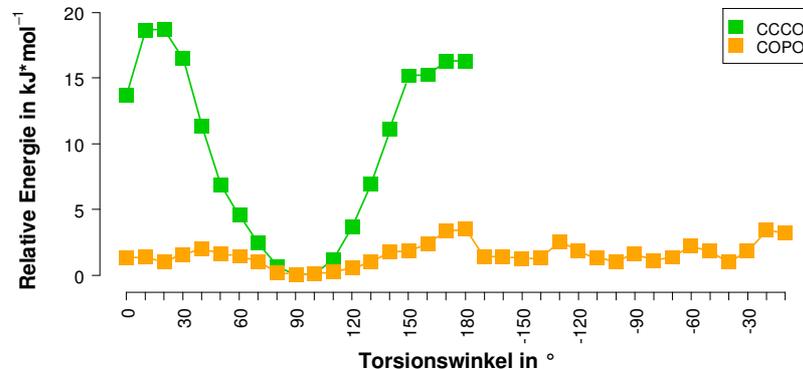


Abbildung 4.37: Darstellung der relativen Energie des reaktiven Zwischenzustandes in Abhängigkeit von den Torsionswinkeln COPO und CCCO. Um eine Vergleichbarkeit beider Kurven zu gewährleisten, wurde jeweils der minimale Energiewert gleich null gesetzt.

Abbildung 4.37 zeigt die relativen Energien der erzeugten Strukturen abhängig vom Torsionswinkel COPO und CCCO. Die totalen Energien der Strukturen können Tabelle 9.8 (Anhang) entnommen werden. Anders als anhand der Analyse des Torsionswinkels in Röntgenstrukturen (siehe Abbildung 4.25) zu erwarten, ergibt sich das Minimum für COPO nicht bei 180° , sondern bei 90° ; 180° stellt sogar das Energiemaximum dar (das bei -20° vorliegende Maximum ist um $0,07 \text{ kJ*mol}^{-1}$ stabiler). Die Energiedifferenz der beiden Extrema beträgt $3,47 \text{ kJ*mol}^{-1}$ ($0,84 \text{ kcal*mol}^{-1}$). Die Strukturen unterscheiden sich dabei in der Positionierung

des positiven Ladungszentrums im Bereich von C¹, C² und C³ (Delokalisation) des DMA-Kations zum negativen Ladungszentrum am P_α-Phosphat, die für COPO = 90° geringfügig günstiger ist, als für COPO = 180°. Zusätzlich kann eine Annäherung eines C¹H an ein Sauerstoffatom des Diphosphats beobachtet werden, die für das Energieminimum COPO = 90° stärker ausgeprägt ist und möglicherweise einen (geringen) Beitrag zur Stabilisierung leistet (siehe Abbildung 4.38).

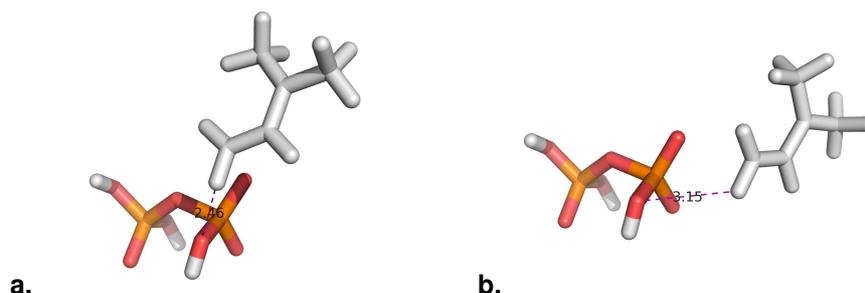


Abbildung 4.38: Darstellung des Energieminimums (a; COPO = 90°) und des Energiemaximums (b; COPO = 180°) der Variation des COPO-Torsionswinkels. Die Wasserstoffbrücke zwischen C¹H und Diphosphat ist in violett dargestellt. Distanzen sind in Å angegeben.

Trotz der geringen Energiedifferenz zwischen den Konformationen stellt sich die Frage nach Ursache der Ausprägung des energetischen Maximums (COPO = 180°) im Enzym. Bei näherer Betrachtung von Abbildung 4.38 fällt die bessere räumliche Zugänglichkeit des O^M-Sauerstoffatoms des Diphosphats im Vergleich zur 90°-Konformation auf. Unter Berücksichtigung von Tabelle 4.4, die für dieses Atom eine verstärkte Wechselwirkungszahl für aktivierte Diphosphate zeigt, kann angenommen werden, dass der durch COPO = 180° verursachte Stabilitätsverlust durch erhöhte Interaktion des O^M mit dem Protein ausgeglichen werden kann. Zusätzlich können auch Entropieeffekte, die hier nicht betrachtet wurden, eine Rolle spielen: Langgestreckte Konformationen sind aufgrund der geringeren Einschränkung der Beweglichkeit entropisch günstiger.

Bemerkenswert ist auch die Stabilität von CCCO während der Variation von COPO. Der Winkel bleibt in einem Bereich von 83° bis 93° konstant, wobei die stärksten Abweichungen von 90° im Bereich um die COPO-Energiemaxima 180° und -20° vorliegen.

Bezogen auf den Torsionswinkel CCCO zeigt Abbildung 4.37, dass – anders als bei den Berechnungen zu COPO – die Energiedifferenz zwischen Minimum (CCCO = 90°) und Maximum (CCCO = 20°) hier wesentlich größer ist. Sie beträgt 18,7 kJ·mol⁻¹ (4,5 kcal·mol⁻¹). Das Minimum von 90° (und entsprechend auch -90°) liegt damit – unter Berücksichtigung des durch die mangelnde Auflösung der Röntgenstrukturen verursachten Fehlers, sowie die

Kristallisierung von eher substrat- statt intermediatähnlichen Konformationen – in der Nähe des in Abbildung 4.25 beobachteten Bereiches. Es entspricht genau dem Wert, der auch während der gesamten Rotation von COPO relativ konstant blieb. Diese Tatsachen weisen auf eine Bedeutung des Torsionswinkels CCCO für den Aktivierungsprozess hin. Abbildung 4.39 zeigt die Struktur des Energieminimums, sowie eine schematische Darstellung, die Aufschluss über die Grundlagen der Stabilität der Konformation geben soll.

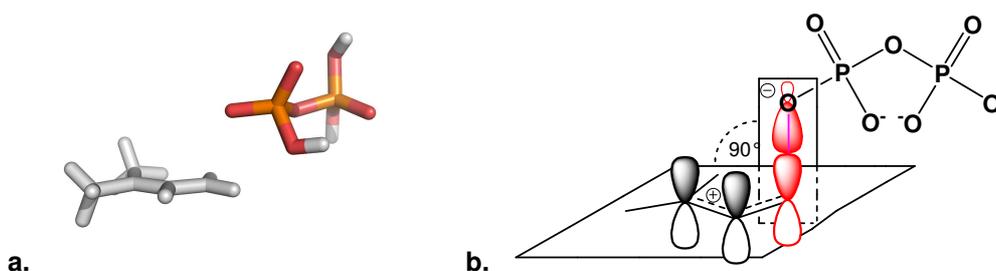


Abbildung 4.39: Darstellung der Minimumsstruktur für CCCO als Stabmodell (a) und schematisch (b).

Während der Aktivierung entstehen aus DMAPP ein DMA-Kation, sowie ein Diphosphatanion. Im Zuge dieses Prozesses werden aus vier besetzten sp³-Hybridorbitalen am C¹-Atom drei besetzte sp²-Hybridorbitale und ein p-Orbital, das mit den π-Elektronen der Doppelbindung in Konjugation tritt und damit zu einem essentiellen Energiegewinn führt (Delokalisierung der π-Elektronen). Ein CCCO-Winkel von 90° fördert den Übergang des an der C¹O^R-Bindung beteiligten sp³-Orbitals in dieses p-Orbital und erleichtert somit die Ausbildung des stabilen mesomeren Zustands des DMA-Kations (Allylkation). Zusammenhänge dieser Art werden allgemein als stereoelektronische Kontrolle bezeichnet.

Der Energiesprung im Bereich von 0° bis 30° für CCCO (Differenz von 5 kJ (=1,2 kcal) zwischen 0° und 20°) kommt vermutlich durch die Interaktion der Ladungszentren am DMA-Kation und Diphosphatanion zustande, die für CCCO = 0° günstiger als für CCCO = 20° sind, wohingegen ab CCCO = 30° der Effekt des in Abbildung 4.39 beschriebenen Zusammenhangs überwiegt, sodass die Stabilität wieder zunimmt. Zusätzlich führt eine Annäherung eines C¹H des DMA-Kations an das O^R des Diphosphats zu einer weiteren Stabilisierung der 0° gegenüber der 20°-Konformation des CCCO-Torsionswinkels. Der geringere Stabilitätsgewinn zwischen verschiedenen COPO-Konformationen im Vergleich zu den CCCO-Konformationen (siehe Minimum-Maximum-Differenzen) äußert sich auch im Verhalten des Winkels während der Variierung von CCCO: Er schwankt in einem Bereich von 73° bis 101° und damit wesentlich stärker als CCCO während der COPO-Variierung.

Abschließend soll auf die Abhängigkeit des Torsionswinkels CCCO von Protonierungsgrad und Metallionenkomplexierung hingewiesen werden (Abbildung 4.40).

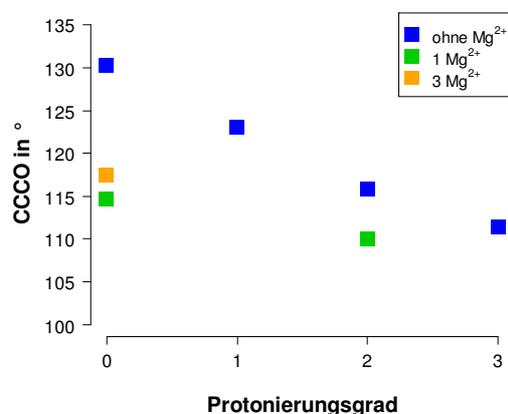


Abbildung 4.40: Quantenchemisch berechnete Werte des Torsionswinkels CCCO bei verschiedenen Protonierungsgraden und Metallionenkomplexierungen von Dimethylallyldiphosphat (Strukturen aus Abschnitt 4.3.2).

Die Metallkomplexierung bewirkt – ähnlich wie es bereits für die C¹O^R- und O^RP-Bindungslängen gezeigt wurde (siehe Abbildung 4.20) – im Allgemeinen eine Verschiebung der Torsionswinkel zu Werten, die höheren Protonierungsgraden entsprechen. Allerdings ist hier – im Gegensatz zu Abbildung 4.20 – nicht zu beobachten, dass eine höhere Zahl an komplexierenden Magnesiumionen proportional zur Stärke der Verschiebung ist. Interessanterweise kommt es mit zunehmendem Protonierungsgrad bzw. mit Metallionenkomplexierung am Diphosphat zu einer Annäherung an das beschriebene Optimum von CCCO = 90°. Damit bewirkt allein eine Metallionenkomplexierung bzw. Protonierung des Diphosphats eine Vorjustierung dieses für die Aktivierung wichtigen Winkels.

4.5.6 Zusammenfassung und Diskussion

Es konnte gezeigt werden, dass aktivierte Diphosphate in Proteinen häufig bestimmte Konformationen bevorzugen. Dabei ist die Ausprägung und Ursache der ausgebildeten Konformationen stark abhängig von der Struktur des Liganden. Für alle drei im Datensatz verfügbaren aktivierten Liganden erfolgte eine Analyse verschiedener Torsionswinkel. Dabei zeigte sich für einige aktivierte Nukleosiddiphosphat-Zucker eine spezifische Konformation (U-Form). Diese ermöglicht bei metallionenunabhängigen Enzymen die Ausbildung intra- und intermolekularer Wechselwirkungen, die die während der Aktivierung am Diphosphat entstehende negative Ladung stabilisieren. Für Pterindiphosphate konnten aufgrund der ge-

ringen Zahl verfügbarer aktivierter Liganden nur Vermutungen bezüglich einer Bedeutung des Torsionswinkels CCCO angestellt werden. Sowohl für NDP-Zucker als auch für Pterindiphosphate sind detaillierte quantenchemische oder QM/MM-Berechnungen notwendig, um endgültige Aussagen über die Bedeutung der Ligandenkonformation während des Aktivierungsprozesses treffen zu können.

Für die Prenylverbindung Dimethylallyldiphosphat wurden im Rahmen dieser Arbeit quantenchemische Berechnungen als Ergänzung zu den Torsionswinkelanalysen der verfügbaren Röntgenstrukturen durchgeführt. Es zeigte sich, dass der in den Röntgenstrukturen aktivierender Enzyme vermehrt beobachtete Wert von $\text{COPO} = 180^\circ$ nicht wie erwartet dem Energieminimum des berechneten reaktiven Intermediates entspricht. Der Winkel wird daher vielmehr mit einer durch ihn verursachten, erhöhten Zugänglichkeit des O^{M} des Diphosphats für Wechselwirkungen mit dem Protein, sowie mit entropischen Vorteilen in Verbindung gebracht. Es konnte gezeigt werden, dass die Variation des Winkels lediglich zu minimalen (vernachlässigbaren) Veränderungen der Gesamtenergie des Moleküls führt. Für den Torsionswinkel CCCO zeigte sich in den Röntgenstrukturen aktivierender Prenyldiphosphatbindestellen eine vermehrte Ausprägung von 120° . Quantenchemische Berechnungen ergaben eine starke Stabilisierung der betrachteten Intermediatstruktur durch $\text{CCCO} = 90^\circ$, wie aus der Orbitaltheorie erwartet wurde. Es wurde gezeigt, dass die Stabilität dieser Konformation durch die erleichterte Ausbildung eines resonanzstabilisierten Allylkations verursacht wird. Zusätzlich konnte dargelegt werden, dass bereits höhere Protonierungsgrade bzw. Komplexierungen mit Metallionen eine Vorjustierung des CCCO-Winkels in Richtung des optimalen Wertes von 90° erwirken. Die in den Röntgenstrukturen beobachtete Ausprägung von $\text{CCCO} = 120^\circ$ ist vermutlich auf diesen Fakt zurückzuführen.

Alle bisherigen Kapitel der vorliegenden Arbeit beschäftigten sich mit der Art und Weise der Bindung des Diphosphatsubstrates innerhalb der Proteinbindetasche. Es konnten aber keine Aussagen bezüglich der energetischen Abhängigkeit der Aktivierungsreaktion von Protonierungsgrad und Metallionenkomplexierung des Diphosphats getroffen werden. Dies soll im nachfolgenden Kapitel betrachtet werden.

4.6 Energetische Betrachtungen des Aktivierungsprozesses

In den vorangehenden Abschnitten 4.4 und 4.5 wurden Bindung und Konformation des Diphosphats innerhalb der Proteinbindetasche behandelt. Das vorliegende Kapitel beschäftigt sich mit dem Prozess der Aktivierung. Ziel der durchgeführten Analysen ist es dabei, den Einfluss von Protonierungsgrad und Metallionenkomplexierung des Diphosphats auf die Aktivierungsreaktion zu untersuchen. In diesem Zusammenhang wurde die Reaktion von H₂O mit Dimethylallyldiphosphat (DMAPP) als Modellreaktion gewählt und mittels quantenchemischer Rechnungen die Energetik dieses Hydrolyseprozesses nachvollzogen. Diese Vorgehensweise hat verschiedene Vorteile gegenüber der direkten Betrachtung einer einzelnen Enzymreaktion beispielsweise mit kombinierten quantenchemischen und molekulmechanischen (QM/MM) Methoden. Dies sind vor Allem:

- *Einfachheit des Szenarios*
- *Möglichkeit systematischer Untersuchungen*
- *Unabhängigkeit von einer Proteinumgebung*
- *Verfügbarkeit experimenteller Referenzdaten*

Die Vernachlässigung von Aminosäuren, die das Diphosphat umgeben und auch beeinflussen, erzeugt ein stark reduziertes, aber einfaches und damit weniger rechenzeit- und -speicheraufwändiges chemisches System. Dies ermöglicht eine systematische Analyse des Einflusses von Protonierungsgrad und Metallkomplexierung des Diphosphats auf die Aktivierung. Anhand der Betrachtung konkreter Enzymbeispiele wäre eine analoge Analyse nur unter sehr großem Aufwand realisierbar. Des Weiteren sind durch die Unabhängigkeit von Aminosäuren allgemeine Aussagen möglich, die nicht auf konkrete Enzyme und damit Umgebungen des Diphosphats beschränkt sind. Nicht zuletzt war das Vorhandensein experimenteller Vergleichsdaten (siehe Abschnitt 1.3.1) ausschlaggebend für die Wahl dieser Vorgehensweise. Damit können erste systematische Einblicke in die energetische Steuerung des Aktivierungsprozesses geliefert werden. Es soll an dieser Stelle jedoch auch darauf hingewiesen werden, dass für das Treffen detaillierter, enzymespezifischer Aussagen QM/MM-Rechnungen unter Einbeziehung der entsprechenden Enzymumgebung unumgänglich sind. Für einige Proteine wurden derartige Rechnungen bereits durchgeführt.^[21;66;245]

Das Kapitel beginnt mit einer kurzen Darstellung der quantenchemischen Beschreibung der Modellreaktion. Es folgt die Auswertung der berechneten Reaktionsenergien und anschließend der ermittelten Aktivierungsbarrieren.

4.6.1 Quantenchemische Beschreibung der Modellreaktion

Als Modellreaktion wurde die Reaktion von DMAPP mit H_2O aus den zuvor genannten Gründen gewählt. H_2O kann dabei als allgemeiner Repräsentant für ein einfaches, neutrales Nukleophil betrachtet werden, das den Akzeptor des während der Aktivierung entstehenden Prenylkations darstellt. Damit ersetzt H_2O beispielsweise die in Enzymreaktionen vorliegenden Zucker- oder Prenylakzeptoren. Für eine Abspaltung des Diphosphats vom organischen Rest des DMAPP kann sowohl eine Spaltung der $\text{C}^1\text{O}^{\text{R}}$ - als auch der $\text{O}^{\text{R}}\text{P}$ -Bindung erfolgen; beides führt im Zusammenhang mit einem H_2O -Angriff zu gleichen Produkten: einem Dimethylallylalkohol und Diphosphat (Abbildung 4.41). Die Abspaltung eines Monophosphates – beispielsweise durch nukleophilen Angriff an das Phosphoratom des P_β -Phosphats – wird im Rahmen einer Diphosphataktivierung durch Enzyme, sowie im enzymfreien Hydrolyseexperiment nicht beobachtet (siehe Abschnitt 1.3.1) und wird daher nachfolgend nicht berücksichtigt. Um eine Spaltung der $\text{C}^1\text{O}^{\text{R}}$ -Bindung zu erzielen, ist ein nukleophiler Angriff an C^1 des organischen Restes notwendig; die Spaltung der $\text{O}^{\text{R}}\text{P}$ -Bindung hingegen wird durch den nukleophilen Angriff an das Phosphoratom des P_α ermöglicht.

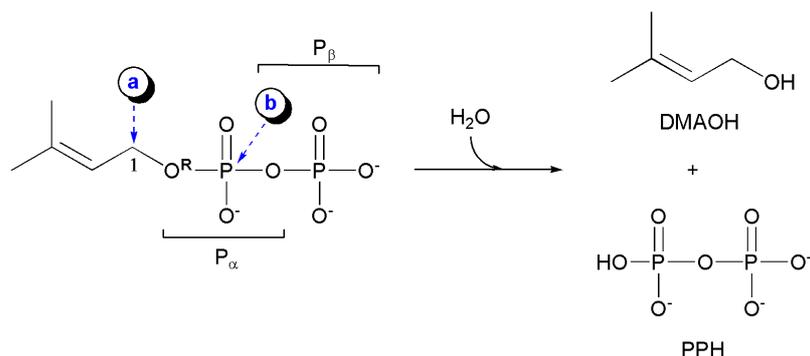


Abbildung 4.41: Mögliche Angriffe eines Wassermoleküls an DMAPP (am Beispiel der vollständig deprotonierten Form), die zur Abspaltung des Diphosphats führen können: nukleophil an C^1 (a) oder das Phosphoratom des P_α -Phosphats (b).

Eine Protonierung des O^{R} -Atoms des entstehenden Diphosphats wird hier aufgrund der approximierten pK_a -Werte von DMAPP (siehe Tabelle 4.3) und anorganischem Diphosphat angenommen. Um die Abhängigkeit der Aktivierung von verschiedenen Protonierungsgraden des Diphosphats, als auch Metallionenkomplexierungen studieren zu können, wurden die Angriffe des Wassermoleküls mit allen in Abschnitt 4.2 beschriebenen Protonierungsformen, sowie Metallionenkomplexierungen (außer dem Komplex mit drei Metallionen) gerechnet und miteinander verglichen. Abbildung 4.42 zeigt die betrachteten Reaktionen im Überblick.

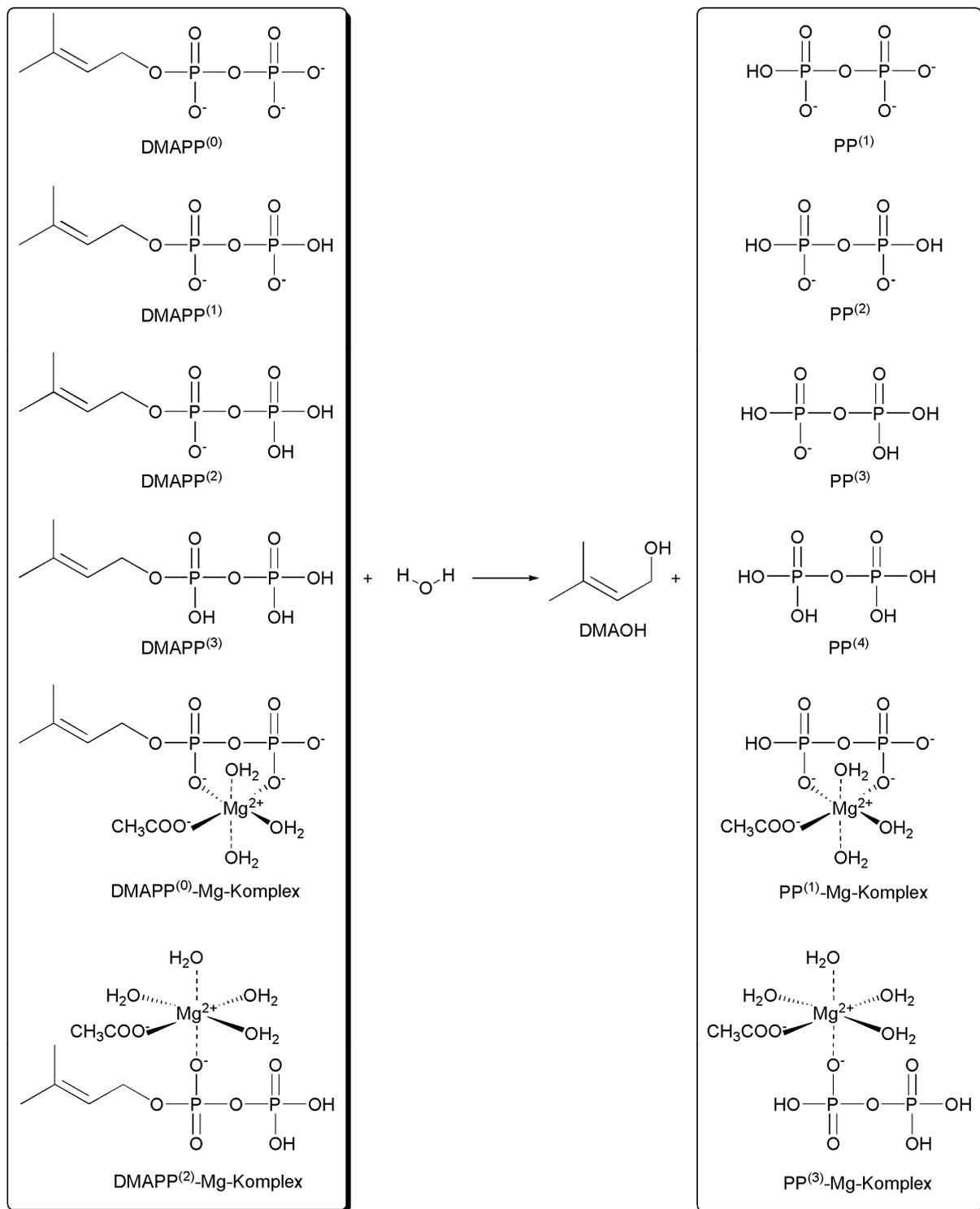


Abbildung 4.42: Schematische Darstellung der Hydrolyseaktion verschiedener Protonierungsgrade und Metallionenkomplexierungen des DMAPP.

Tabelle 4.5 zeigt, dass alle betrachteten Protonierungsstufen tatsächlich innerhalb von Proteinbindetaschen durch Protonenübertragung von Aminosäuren erzeugt werden können. Die Ausrichtung des H_2O während dieser Angriffe wurde dabei stets so gewählt, dass eine Protonierung des O^{R} nach der Spaltung möglich wird. Um die Vergleichbarkeit der Berech-

nungen zu gewährleisten, wurden für alle Angriffe an C¹, sowie an P möglichst gleiche relative Positionierungen des Wassermoleküls zum DMAPP gewählt. Beispiele für solche Konformationen sind in den nachfolgenden Abschnitten gezeigt. Es bleibt an dieser Stelle anzumerken, dass die hier durch einen Wasserangriff modellierte Hydrolyse nicht unmittelbar mit der im Experiment (siehe Abschnitt 1.3.1) beobachteten Hydrolyse verglichen werden kann. Ursache hierfür ist die Vernachlässigung von Angriffen durch Hydronium- (H₃O⁺) bzw. Hydroxidionen (OH⁻), die bei unterschiedlichen pH-Werten in verschiedenen Verhältnissen vorliegen. Hydroniumionen sind beispielsweise in der Lage, durch Protonierung des O^R eine aktivierte Form des DMAPP zu erzeugen, die dann für eine Hydrolyse zur Verfügung steht. Im Gegensatz zu H₃O⁺ ist die Säurestärke von H₂O zu gering, um eine Protonierung des O^R erzielen zu können. Dies konnte durch die quantenchemische Berechnung des Angriffs eines Wasserprotons an O^R des vollständig protonierten DMAPP (DMAPP⁽³⁾) gezeigt werden (Daten nicht gezeigt): Nur ein enormer Energieaufwand führte zur Übertragung des Protons auf O^R, zusätzlich bildeten sich bei einer Nachoptimierung der Produktstruktur die Substrate Wasser und DMAPP⁽³⁾ zurück. Hydroxidionen, die insbesondere im basischen Milieu zunehmende Bedeutung erlangen, sind stärkere Nucleophile als H₂O, was sich ebenfalls auf die Energetik des Prozesses auswirken kann. Sowohl Hydronium- als auch Hydroxidionen führen aber – analog zu H₂O – entweder zur Spaltung der C¹O^R- oder der O^RP-Bindung.

Für die quantenchemische Berechnung der Reaktionsenergien werden die Substrate und Produkte aus Abbildung 4.42 unabhängig voneinander optimiert, um den Einfluss intermolekularer Wechselwirkungen zwischen den einzelnen Substrat- bzw. Produktmolekülen auf die berechneten Energien auszuschließen. Die verwendete quantenchemische Methode und Basissatz, sowie die Solvatationsbedingungen sind in Abschnitt 4.2 beschrieben. Die Reaktionsenergie ΔE für die betrachteten Reaktionen ergibt sich dann folgendermaßen:

$$\Delta E = (E_{DMAOH} + E_{PPH}) - (E_{DMAPP} + E_{H_2O}),$$

wobei E_{PPH} und E_{DMAPP} die totalen Energien der verschiedenen Protonierungsstufen und Metallionenkomplexierungen des Diphosphats repräsentieren.

Die quantenchemische Berechnung der Aktivierungsenergien hingegen erfordert das Nachvollziehen des Angriffs des H₂O an DMAPP. Dies wird durch Annäherung der entsprechenden Atome modelliert. So wurde hier beginnend mit einem Atomabstand von 3 Å die Distanz zwischen O^{H₂O} und C¹ bzw. dem Phosphoratom des P_α des DMAPP schrittweise um jeweils 0,2 Å verringert (siehe auch Abbildung 4.41). In jedem Schritt erfolgt eine Struktur- und Ener-

gieoptimierung, wobei die optimierte Struktur des vorhergehenden Schrittes jeweils als Ausgangskonformation des aktuellen Schrittes dient. Der C¹O^{H₂O}- bzw. PO^{H₂O}-Abstand wird dabei in jedem Schritt fixiert. Die verwendete quantenchemische Methode und Basissatz, sowie die Solvatationsbedingungen sind in Abschnitt 4.2 beschrieben. Es entsteht eine Reihe optimierter Strukturen, deren Energien ein Reaktionsprofil („Reaktionskoordinate“) ergeben. Für die Ermittlung von Aktivierungsbarrieren müssen die genannten Abstandsfixierungen in den Substraten jedoch beseitigt werden. Dies geschieht durch Nachoptimierung der Startstrukturen (DMAPP und Wasser im Komplex) ohne die C¹O^{H₂O}- bzw. PO^{H₂O}-Distanz (auf 3 Å) zu fixieren. Als Übergangszustand wird die instabilste (energiereichste) Konformation der berechneten Reaktionskoordinate angenommen (Pseudoübergangszustand; PÜZ). Es soll an dieser Stelle angemerkt werden, dass der Pseudoübergangszustand nicht dem tatsächlichen Übergangszustand entsprechen muss, sondern lediglich eine Annäherung darstellt. Es gibt verschiedenste quantenchemische Methoden, die eine exakte Berechnung von Übergangszuständen ermöglichen (siehe Abschnitt 3.3.5). Diese wurden auch in der vorliegenden Arbeit angewendet, führten jedoch nicht zum Erfolg. Die Aktivierungsbarrieren wurden dann wie folgt approximiert:

$$E_A = E_{PÜZ} - E_{DMAPP-H_2O-Komplex}$$

Dabei beziehen sich $E_{PÜZ}$ und $E_{DMAPP-H_2O-Komplex}$ jeweils auf die Energien der berücksichtigten Protonierungsgrade und Metallionenkomplexe.

4.6.2 Reaktionsenergien des H₂O-Angriffes

Tabelle 4.7 zeigt die Reaktionsenergien des H₂O-Angriffes bei den verschiedenen betrachteten Protonierungsgraden bzw. in Komplex mit Magnesiumionen (siehe Abbildung 4.42). Die zur Berechnung benötigten Energien können Tabelle 9.9 (Anhang) entnommen werden. Es zeigt sich, dass mit steigendem Protonierungsgrad auch die Reaktionsenergie ΔE ansteigt. Unter der Voraussetzung, dass ein zweiwertiges Metallion ein oder zwei Protonen ersetzen kann, wird deutlich, dass die Reaktion im DMAPP⁽⁰⁾-Mg-Komplex stärker begünstigt ist, als für DMAPP⁽¹⁾ oder DMAPP⁽²⁾. Hingegen ist beim Vergleich der Reaktionsenergien des DMAPP⁽²⁾-Mg-Komplexes mit DMAPP⁽³⁾ eine geringere Begünstigung für den Metallionenkomplex zu beobachten. Es ist zu sehen, dass die Werte für alle betrachteten Protonierungsgrade und Metallionenkomplexierungen negativ sind und damit alle berücksichtigten Reaktionsvarianten energetisch begünstigt sind.

Tabelle 4.7: Reaktionsenergien des H₂O-Angriffs an DMAPP in Abhängigkeit von Protonierungsgrad des DMAPP bzw. der Metallkomplexierung.

Protonierung bzw. Metallkomplexierung	ΔE in kJ*mol ⁻¹
DMAPP ⁽⁰⁾	-36,88
DMAPP ⁽¹⁾	-12,94
DMAPP ⁽²⁾	-10,57
DMAPP ⁽³⁾	-8,92
DMAPP ⁽⁰⁾ -Mg-Komplex	-22,90
DMAPP ⁽²⁾ -Mg-Komplex	-3,83

Die Energien dieser Reaktionen werden durch folgende Faktoren beeinflusst und unterscheidbar: Die Stabilität des Substrates DMAPP, sowie die Stabilität des Diphosphatproduktes. Je instabiler dabei das Substrat und je stabiler das Produkt ist, umso stärker ist die Reaktion begünstigt. Nachfolgend werden daher die Stabilitäten der verschiedenen stark protonierten DMAPP-Substrate und der zugehörigen Diphosphatprodukte beschrieben, um die Reihenfolge der in Tabelle 4.7 dargestellten Reaktionsenergien zu stützen. Ein Vergleich mit der Stabilität der betrachteten Metallionenkomplexe, sowie der Metallionenkomplexe untereinander, ist aufgrund der stark voneinander abweichenden Strukturen (siehe Abbildung 4.8, Abbildung 4.9 und Abbildung 4.42) nicht aussagekräftig und erfolgt daher nicht. Tabelle 4.8 vergleicht die Stabilitäten der verschiedenen Protonierungsgrade des Substrates DMAPP. Die zur Berechnung dieser Werte benötigten Energien können Tabelle 9.9 (Anhang) entnommen werden.

Tabelle 4.8: Protonenaffinitäten verschieden stark protonierter Formen des DMAPP, sowie Energien relativ zur stabilsten Form DMAPP⁽³⁾; x entspricht der Zahl der Protonen.

DMAPP ^(x)	Energiedifferenzen in kJ*mol ⁻¹	
	$E_{\text{DMAPP}(x)} - E_{\text{DMAPP}(x+1)}$	$E_{\text{DMAPP}(x)} - E_{\text{DMAPP}(3)}$
DMAPP ⁽⁰⁾	1267,17	3622,84
DMAPP ⁽¹⁾	1195,33	2355,67
DMAPP ⁽²⁾	1160,34	1160,34
DMAPP ⁽³⁾	-	0

Es wird deutlich, dass die Stabilität der DMAPP-Substrate mit steigendem Protonierungsgrad zunimmt. Gleichzeitig ist zu erkennen, dass die Protonenaffinität mit steigendem Protonierungsgrad abnimmt. Beides ist auf die mit steigendem Protonierungsgrad am Diphosphat abnehmende Ladungsdichte zurückzuführen.

Tabelle 4.9 vergleicht die Stabilitäten der bei der Hydrolyse entstehenden, metallionenfreien Diphosphatprodukte. Die zur Berechnung dieser Werte benötigten Energien können wieder Tabelle 9.9 (Anhang) entnommen werden. Auch hier wird deutlich, dass die Stabilität mit steigendem Protonierungsgrad zunimmt und die Protonenaffinität gegenläufig abnimmt.

Tabelle 4.9: Protonenaffinitäten verschiedener Protonierungsgrade von Diphosphat $PP^{(x)}$, sowie Energien relativ zur stabilsten Form $PP^{(4)}$; x entspricht dabei der Zahl der Protonen.

$PP^{(x)}$	Energiedifferenzen in $\text{kJ}\cdot\text{mol}^{-1}$	
	$E_{PP^{(x)}} - E_{PP^{(x+1)}}$	$E_{PP^{(x)}} - E_{PP^{(4)}}$
$PP^{(1)}$	1243,22	3594,88
$PP^{(2)}$	1192,97	2351,66
$PP^{(3)}$	1158,69	1158,69
$PP^{(4)}$	-	0

Auf Basis dieser Beobachtungen kann die Hydrolyse für geringere Protonierungsgrade nur dann stärker energetisch begünstigt sein, als für höhere Protonierungsgrade – wie es in Tabelle 4.7 berechnet wurde –, wenn die Instabilität des Diphosphatproduktes mit fallendem Protonierungsgrad langsamer zunimmt, als die Instabilität des Substrates DMAPP mit fallendem Protonierungsgrad zunimmt. Genau dies ist Tabelle 4.8 und Tabelle 4.9 ebenfalls zu entnehmen.

Im Rahmen einer Diplomarbeit^[245] wurde die Energetik des H_2O -Angriffes bei verschiedenen Protonierungen des Geranyldiphosphats (GPP) untersucht. Dabei wurde die semiempirische Methode PM3 verwendet; die Berechnungen erfolgten in Gasphase. Die Verwendung der PM3-Methode im Zusammenhang mit Diphosphaten ist jedoch kritisch zu bewerten, da sie einen sehr stumpfen $\text{PO}^{\text{M}}\text{P}$ -Winkel von etwa 162° erzeugt und damit über 20° von den in Röntgenstrukturen beobachteten Werten abweicht (siehe Abschnitt 4.3.3). Dies hat zur Folge, dass die $\text{O}^\#$ -Sauerstoffatome des Diphosphats zu weit voneinander entfernt sind, um die für Gasphasenstrukturen typischen intramolekularen Wasserstoffbrücken ausbilden zu können. Die Analyse zeigte einen positiven Wert für die Hydrolyse des vollständig deprotonierten Diphosphats, jedoch negative Werte für die drei verbliebenen Protonierungsgrade. Die Werte des zweifach und dreifach protonierten GPP sind dabei fast identisch und haben die am stärksten negative Ausprägung. Als Resultat dieser Analysen ergibt sich demzufolge eine stärkere energetische Begünstigung der Hydrolyse mit steigendem Protonierungsgrad, was im starken Gegensatz zu den in Tabelle 4.7 gezeigten Ergebnissen steht. Es bleibt in diesem Zusammenhang zu beachten, dass die in der vorliegenden Arbeit berechneten Strukturen unter Einbeziehung eines impliziten Wassermodells ermittelt wurden und die erhaltenen $\text{PO}^{\text{M}}\text{P}$ -Winkel gut mit den in Röntgenstrukturen beobachteten übereinstimmen (siehe Abschnitt 4.3.3). Zusätzlich wurde die Methode B3LYP verwendet, die den semiempirischen Ansätzen methodisch überlegen ist. Damit ist davon auszugehen, dass die Genauigkeit der hier präsentierten Strukturen und der mit ihnen verbundenen Reaktionsenergien höher ist, als die der in ^[245] gezeigten Strukturen. Jedoch bestimmt nicht die Reaktionsenergie allein den Ablauf einer Reaktion. Auch die Aktivierungsenergie hat einen erheblichen Anteil. Sie wird im nachfolgenden Abschnitt näher betrachtet.

4.6.3 Angriff des H₂O an das C¹-Atom des Dimethylallyldiphosphats

Zunächst wurden – wie in Abschnitt 4.6.1 beschrieben – schrittweise Annäherungen des O^{H₂O} an das C¹ des DMAPP berechnet (Startabstand 3 Å, Schrittweite 0,2 Å), um den Angriff des H₂O an C¹ zu simulieren. Zusätzlich wurde jeweils in einem Bereich um das dabei entstehende Energiemaximum (für diese Konformation gelte: Distanz(O^{H₂O}, C¹) = x Å) eine weitere Reaktionskoordinate mit einer Schrittweite von 0,05 Å berechnet. Als Startstruktur diente hierbei jeweils der Vorgänger mit Distanz(O^{H₂O}, C¹) = $(x + 0,2)$ Å; die Rechnung wurde mit Erreichen des O^{H₂O}C¹-Abstandes des Nachfolgers mit Distanz(O^{H₂O}, C¹) = $(x - 0,2)$ Å beendet. Die Verwendung einer engeren Schrittweite dient dabei der Erhöhung der Genauigkeit des berechneten Pseudoübergangszustandes (PÜZ). Aber auch ein so berechneter PÜZ stellt nur eine Annäherung des wahren Übergangszustandes dar, da der PÜZ zum Einen auf der Fixierung des O^{H₂O}C¹-Abstandes beruht und zum Anderen – aufgrund der Wahl der großen Schrittweite von 0,05 Å – nicht gesichert ist, dass nicht ein noch instabilerer Zustand entlang der Reaktionskoordinate existiert. Die ermittelten PÜZ sind jedoch – trotz dieser Einschränkungen – zur groben Annäherung der wahren Übergangszustände geeignet, da sie für alle Rechnungen unter Verwendung der gleichen Parameter und Randbedingungen (z.B. Startdistanz für O^{H₂O}C¹, Schrittweite, Richtung des Angriffs) ermittelt wurden und somit die relativen Verhältnisse wahrer Übergangszustände zueinander widerspiegeln sollten. Wie bereits in den Abschnitten 3.3.5 und 4.6.1 angesprochen, gibt es verschiedene Möglichkeiten, Übergangszustände mittels quantenmechanischer Methoden exakt zu bestimmen. Im hier vorliegenden Fall wurde sowohl der QST2-, als auch der QST3-Ansatz verwendet.^[171] Beide Methoden scheiterten jedoch. Es kam entweder zum Abbruch der Rechnung oder es entstand eine Struktur mit Distanz(O^{H₂O}, C¹) $\gg 3$ Å, was mit den Ergebnissen der berechneten Reaktionskoordinate nicht zu vereinbaren ist. Die wahrscheinliche Ursache für das Scheitern beider Ansätze liegt in starken Unterschieden in den Konformationen der Startstrukturen (Vorgänger und Nachfolger des PÜZ) – auch für die Strukturen der Reaktionskoordinate mit Schrittweite 0,05 Å. Eine mögliche Lösung wäre das iterative Berechnen weiterer Reaktionskoordinaten mit engeren Schrittweiten in Bereichen um den angenäherten Übergangszustand PÜZ, um die Konformationsunterschiede einzuschränken.

Abbildung 4.43 zeigt das Beispiel einer Start-, sowie einer Pseudoübergangsstruktur am Beispiel des vollständig protonierten DMAPP⁽³⁾. Im PÜZ wird eine Drehung des organischen

Restes in Richtung H_2O deutlich, jedoch ist die Ausrichtung noch günstiger für die Ausbildung der Bindung zum Diphosphat. Zusätzlich liegt die $\text{C}^1\text{O}^{\text{R}}$ -Bindung mit $1,55 \text{ \AA}$ verlängert vor – die optimale Bindungslänge einer C-O-Bindung liegt um $1,43 \text{ \AA}$. Das kationische Intermediat der Reaktion wird – vermutlich aufgrund der großen Schrittweite von $0,05 \text{ \AA}$ – während der berechneten Reaktionskoordinate nicht sichtbar: Im Schritt, der dem PÜZ folgt, ist der Dimethylallylalkohol bereits ausgebildet.

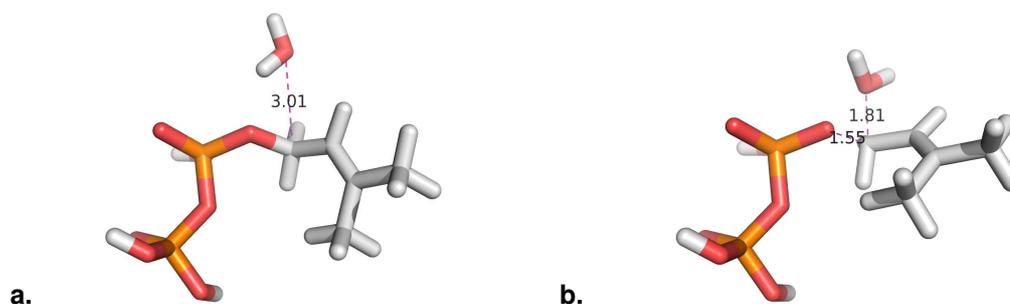


Abbildung 4.43: Startstruktur (a.) und Struktur des berechneten Pseudoübergangszustandes (b.) für den Angriff von H_2O an C^1 am Beispiel von $\text{DMAPP}^{(3)}$. Die Richtung des Angriffs ist in (a.) durch eine violette Linie angedeutet. Die dargestellten Distanzen sind in Å angegeben.

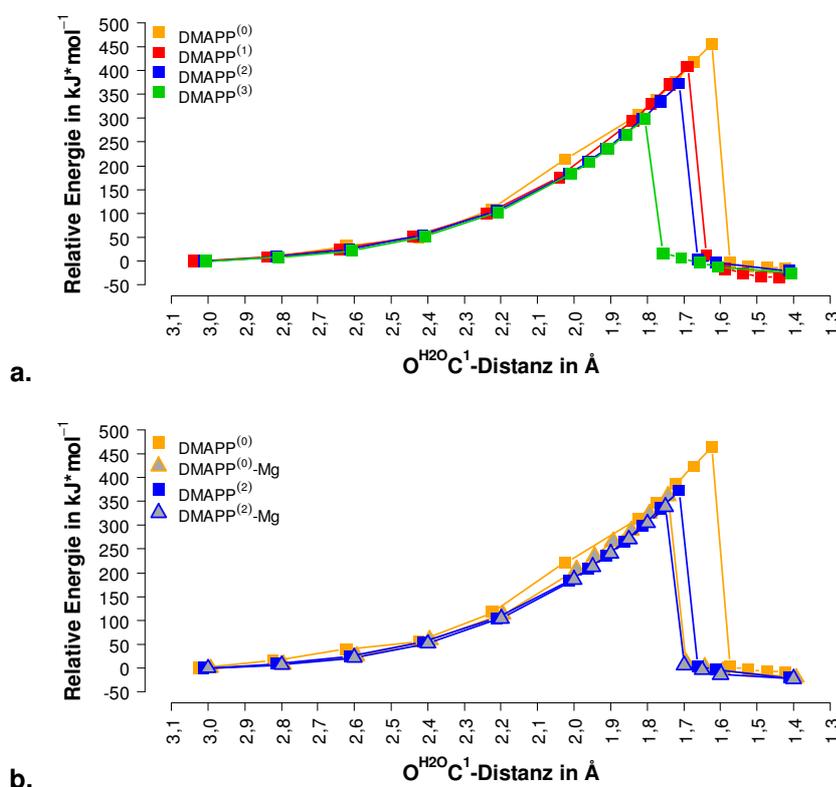


Abbildung 4.44: Auftragung der berechneten Reaktionskoordinate des Wasserangriffs an das C^1 -Atom verschiedener Protonierungsgrade (a.) und Metallionenkomplexierungen (b.) des DMAPP . Für die Metallionenkomplexe ist jeweils als Referenz noch einmal die Kurve für den Angriff an die metallfreie Struktur gleichen Protonierungsgrades angegeben. Die Energien der Ausgangskonformationen (Wasser- DMAPP -Komplex) wurde dabei jeweils auf 0 gesetzt, um eine bessere Vergleichbarkeit der Kurvenverläufe zu gewährleisten.

Tabelle 4.10: Aktivierungsenergien E_A verschiedener Protonierungsgrade und Metallionenkomplexierungen von DMAPP.

Protonierung bzw. Metallkomplexierung	E_A in $\text{kJ}\cdot\text{mol}^{-1}$
DMAPP ⁽⁰⁾	463,84
DMAPP ⁽¹⁾	409,24
DMAPP ⁽²⁾	385,42
DMAPP ⁽³⁾	306,06
DMAPP ⁽⁰⁾ -Mg-Komplex	373,57
DMAPP ⁽²⁾ -Mg-Komplex	341,23

Abbildung 4.44 zeigt die berechneten Energiediagramme („Reaktionskoordinaten“). Die zugehörigen totalen Energien können Tabelle 9.10 (Anhang) entnommen werden. Es wird deutlich, dass die Aktivierungsenergie mit steigendem Protonierungsgrad und durch Metallionenkomplexierungen des DMAPP abnimmt. Auch der $\text{O}^{\text{H}_2\text{O}}\text{C}^1$ -Abstand muss für geringere Protonierungsgrade bzw. für Strukturen ohne Metallkomplexierung geringer sein, um die energetisch maximale Struktur zu erreichen, d.h. die Annäherung des Wassers an das C^1 -Atom muss bereits weiter fortgeschritten sein. Damit zeigt sich auch anhand der Aktivierungsenergien der hier betrachteten Modellreaktion, dass durch die Komplexierung von Metallionen Zustände höherer Protonierung simuliert werden können (vgl. auch Abbildung 4.20 und Abbildung 4.40). So entspricht die Aktivierungsenergie für den DMAPP⁽⁰⁾-Mg-Komplex etwa der von DMAPP⁽²⁾. Der Energieverlauf aller Kurven ist bis zum Erreichen des $\text{O}^{\text{H}_2\text{O}}\text{C}^1$ -Abstandes von ca. 1,8 Å sehr ähnlich. Reaktionsenergien ΔE , die aus dem Diagramm abgelesen werden, stimmen nicht mit den in Tabelle 4.7 dargestellten Werten überein. Ursache hierfür sind die bereits in Abschnitt 4.6.2 angesprochenen intermolekularen Wechselwirkungen zwischen den Hydrolysesubstraten DMAPP und H_2O , sowie den Hydrolyseprodukten DMAOH und Diphosphat, die zur Verzerrung der wahren Reaktionsenergien führen. Weiterhin zeigt Tabelle 4.10 die Aktivierungsenergien E_A , die – wie beschrieben – approximiert sind. Die zugehörigen Energien, die zur Berechnung benötigt werden, können Tabelle 9.10 (Anhang) entnommen werden.

Experimentelle Ergebnisse (siehe Abschnitt 1.3.1) zeigten, dass in saurer Umgebung und ohne Ausbildung von Metallionenkomplexen keine Hydrolyse stattfindet. Jedoch führt bereits das Vorliegen von (1:1)-Verhältnissen aus dikationischem Metallion und Prenyldiphosphat dazu, dass unter sauren Bedingungen eine spontane Hydrolyse beobachtet werden kann. Daher würde zumindest für die in Tabelle 4.10 gezeigte Aktivierungsenergie für den DMAPP⁽²⁾-Mg-Komplex eine geringere Barriere erwartet werden als für DMAPP⁽³⁾. Möglicherweise kommt es jedoch in Zusammenhang mit Metallionen zur Ausbildung von Multimeren (Abbildung 4.45), die auch bei (1:1)-Verhältnissen von DMAPP und Metallion zu einer zweifachen Komplexierung führen und damit das beobachtete Verhalten erklären könnten.

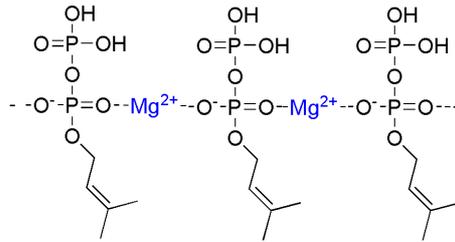


Abbildung 4.45: Schematische Darstellung der Ausbildung multimerer DMAPP-Mg-Komplexe, die eine Hydrolyse auch bei Vorliegen von (1:1)-Verhältnissen von DMAPP und Mg^{2+} erklären könnten. Das zentrale DMAPP ist durch zwei Metallionen komplexiert. Gestrichelte Linien zeigen die koordinativen Bindungen an. Koordinierungen der Metallionen über und unter der Bildebene, um den typischen Oktaeder zu erhalten, wurden aus Gründen der Übersichtlichkeit vernachlässigt.

Um die berechneten Aktivierungsbarrieren besser einschätzen zu können, wurde unter Verwendung der Eyring-Gleichung die Halbwertszeit für die Zersetzung des DMAPP-Substrates ermittelt. Die zugehörigen Formeln sind in Formel 4.1 angegeben. Auch für Enzyme wird die Eyring-Gleichung im Zusammenhang mit der Ermittlung der Aktivierungsenergie eingesetzt^[262]. Jedoch ist ihre Anwendung aufgrund der Abweichung der Wirkungsweise von Enzymen von der (der Eyring-Gleichung zugrundeliegenden) Kollisionstheorie stark eingeschränkt.

$$\text{a. } k = \frac{k_b * T}{h} * \exp(-E_A * R^{-1} * T^{-1}) \qquad \text{b. } t_{\frac{1}{2}} = \frac{\ln 2}{k}$$

Formel 4.1: Gleichungen zur Berechnung der Geschwindigkeitskonstanten (a., Eyringgleichung) und der Halbwertszeit (b.). k – Geschwindigkeitskonstante (s^{-1}); k_b – Boltzmannkonstante ($1,3806504 * 10^{-23} J * K^{-1}$); T – Temperatur (K); h – Plancksches Wirkungsquantum ($6,62606896 * 10^{-34} J * s$); E_A – Aktivierungsenergie ($J * mol^{-1}$); R – allgemeine Gaskonstante ($8,314472 J * K^{-1} * mol^{-1}$); $t_{\frac{1}{2}}$ – Halbwertszeit (s).

Damit ergibt sich für die Zersetzung des vollständig protonierten DMAPP⁽³⁾ bei Raumtemperatur (23 °C) unter Betrachtung der Aktivierungsenergie von $306,06 kJ * mol^{-1}$ für den Angriff von H_2O an C^1 (siehe Tabelle 4.10) eine Halbwertszeit des Substrates von:

$$t_{\frac{1}{2}} = 3,5 * 10^{33} \text{ Jahre.}$$

Neben H_2O -Angriffen spielt bei der Hydrolyse jedoch zusätzlich der Angriff von H_3O^+ eine Rolle, der diese Halbwertszeit wesentlich herabsetzen kann. Im Enzym können die hohen Aktivierungsbarrieren beispielsweise durch die Erzeugung bestimmter Substratkonformationen erniedrigt werden, wie bereits in Abschnitt 4.5 gezeigt wurde. Weiterhin ermöglichen Enzyme eine Stabilisierung kationischer Übergangszustände oder Intermediate, die während

des Aktivierungsprozesses entstehen. Brandt *et al.*^[51] konnten mit Hilfe von *ab-initio*-Berechnungen zur Wechselwirkungsenergie zwischen einem Dimethylallylkation und Phenol (als Modell für die Aminosäure Tyrosin) einen Betrag von $58 \text{ kJ} \cdot \text{mol}^{-1}$ ermitteln.

4.6.4 Angriff des H_2O an das Phosphoratom des P_α des Dimethylallyldiphosphats

Es wurden quantenchemische Berechnungen wie in Abschnitt 4.6.1 beschrieben zum nukleophilen Angriff des H_2O an das Phosphoratom des P_α des DMAPP durchgeführt. Abbildung 4.46 zeigt das Beispiel einer verwendeten Startstruktur. Bei der quantenchemischen Berechnung eines solchen Angriffs wie in Abschnitt 4.6.1 angegeben, zeigt sich jedoch, dass eine Annäherung des $\text{O}^{\text{H}_2\text{O}}$ an das Phosphoratom nur unter stetiger Energiezufuhr möglich ist. Die zugehörigen Energiewerte können Tabelle 9.12 (Anhang) entnommen werden. Für keinen der betrachteten Protonierungsgrade des DMAPP (siehe Abbildung 4.42) ist eine Abspaltung des Diphosphats zu beobachten. Zusätzlich zeigt sich, dass es bei Nachoptimierung des entstehenden Produktkomplexes ohne Fixierung des $\text{PO}^{\text{H}_2\text{O}}$ -Abstandes zur Rückbildung der Ausgangsstoffe DMAPP und H_2O kommt. Es ist daher zu erwarten, dass eine Diphosphatabspaltung von beliebigen Protonierungsgraden des DMAPP nicht durch einen Angriff von H_2O an das Phosphoratom des P_α erfolgt. Aufgrund dieser Beobachtungen und der Tatsache, dass die Komplexierung von Magnesiumionen durch ein DMAPP eines bestimmten Protonierungsgrades zu einer Verlängerung der $\text{C}^1\text{O}^{\text{R}}$ - und eher zu einer Verkürzung der $\text{O}^{\text{R}}\text{P}$ -Bindung führt (siehe Abschnitt 4.4.4), wurden entsprechende Angriffe für DMAPP-Metallkomplexe nicht berechnet.

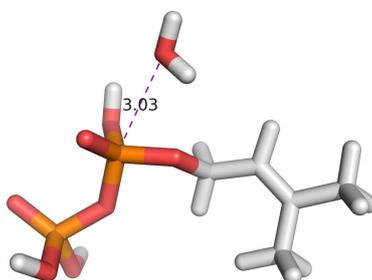


Abbildung 4.46: Startstruktur des Angriffes von H_2O an das Phosphoratom des P_α am Beispiel von $\text{DMAPP}^{(3)}$. Die Richtung des Angriffes ist durch eine violette Linie angedeutet. Distanzen sind in Å angegeben.

4.6.5 Zusammenfassung und Diskussion

Das Kapitel verwendet die Reaktion eines Wassermoleküls mit Dimethylallyldiphosphat (DMAPP) unter Bildung von Dimethylalkohol und anorganischem Diphosphat als Modell, um mittels quantenchemischer Methoden den Einfluss der Diphosphatprotonierung, sowie der Metallionenkomplexierung auf den Aktivierungsprozess zu studieren. Das Wassermolekül simuliert dabei das Nukleophil des während der Aktivierung freiwerdenden organischen Kations. Es wurden sowohl Reaktions- als auch Aktivierungsenergien betrachtet. Die Berechnungen zeigten wider Erwarten stärker negative Reaktionsenergien je geringer der Protonierungsgrad. Damit zeigt sich die Reaktion als umso stärker energetisch begünstigt, je geringer der Protonierungsgrad ist. Analysen zu Aktivierungsenergien anhand der Modellreaktion zeigten für den Angriff des Wassermoleküls am C¹-Atom des organischen Substrates mit nachfolgender Spaltung der C¹O^R-Bindung hingegen geringere Barrieren für stärker protonierte Substrate. Es konnte weiterhin gezeigt werden, dass Metallionen hier den gleichen Effekt auf die Aktivierungsenergie ausüben, wie höhere Protonierungsgrade. In keiner der Berechnungen zum Angriff des Wassermoleküls an das Phosphoratom des P_α verschiedener Protonierungsgrade des DMAPP war eine Diphosphatabspaltung zu beobachten. Demzufolge scheint eine Spaltung der PO^R-Bindung durch einen Wasserangriff nicht zu erfolgen. Die betrachtete Reaktion ist damit ein gutes Modell für die Enzymreaktion, da dort ebenfalls ausschließlich eine Spaltung der C¹O^R-Bindung stattfindet. Die berechneten Aktivierungsbarrieren sind sowohl für metallionenfrie Strukturen, als auch für die betrachteten Metallionenkompexe DMAPP⁽⁰⁾-Mg und DMAPP⁽²⁾-Mg zu hoch, um eine spontane Reaktion bei Raumtemperatur zu ermöglichen. Während dieses Ergebnis gut mit der metallfreien experimentellen Hydrolyse übereinstimmt, steht es in Kontrast zu den Beobachtungen bei Vorhandensein von Metallionen: Hier zeigt sich – auch für ein (1:1)-Verhältnis von Metallion und Prenyldiphosphat – eine Umsetzung in saurer Umgebung bei Raumtemperatur. Es wird daher angenommen, dass es *in vitro* zu einer Ausbildung von Multimeren kommt, die auch bei einem (1:1)-Verhältnis eine zweifache Komplexierung ermöglichen.

Es bleibt anzumerken, dass die berechneten Aktivierungsenergien weder unmittelbar auf die enzymfreie Hydrolyse noch auf die enzymatische Aktivierung übertragbar sind, da weder Nebenreaktionen, wie eine Aktivierung des O^R durch Säuren, noch Einflüsse der Proteinbindetasche auf die DMAPP-Konformation oder die Stabilisierung des entstehenden Intermediates durch die Modellreaktion berücksichtigt wurden. Die berechneten Aktivierungsenergien ermöglichen lediglich einen Vergleich der Effekte verschieden starker Protonierungen und Metallionenkomplexierungen des Diphosphats auf die Kinetik der Aktivierung.

5 ZUSAMMENFASSUNG UND AUSBLICK

Die vorliegende Arbeit versucht, Einblicke in die Aktivierung durch Diphosphate in Proteinen zu liefern. Die Bedeutung dieser Untersuchungen wird bei Betrachtung der Vielfalt der Metabolite klar, für die dieser Prozess Grundlage ihrer Biosynthese bildet. Allein die Gruppe der Terpenoide und Meroterpenoide fasst 55.000 Verbindungen. Ihre Rolle erstreckt sich dabei sowohl über Primär- als auch Sekundärstoffwechsel und hat sich ihren Weg bis in die pharmazeutische Anwendung gebahnt. Aber auch eine Bedeutung im Zusammenhang mit der chemischen Synthese ist nicht zu vernachlässigen – könnten die gewonnenen Erkenntnisse doch Voraussetzungen für die Herstellung von Katalysatoren liefern, die die Erzeugung aktivierter sp^3 -Kohlenstoffzentren unter milden Bedingungen ermöglichen.

Die Grundlage der vorliegenden Arbeit bildet die Protein Data Bank, deren Röntgenstrukturen auf diphosphathaltige Substrate, Substratanaloga oder Inhibitoren untersucht wurden. Eine aufwändige Nachbereitung der Strukturen lieferte einen nichtredundanten Datensatz von 15 aktivierenden und 414 nichtaktivierenden Bindestellen, auf dessen Basis Hypothesen aufgestellt und mittels quantenchemischer Methoden geprüft oder erweitert wurden. Ziel war es dabei, möglichst verschiedene Gesichtspunkte der Aktivierung zu berücksichtigen, um ein umfassendes Bild zu erhalten. Die Effekte der Diphosphatprotonierung und -metallionenkomplexierung wurden dabei aufgrund ihrer Bedeutung in der enzymfreien Hydrolyse allylischer Diphosphate verstärkt in die Betrachtungen einbezogen.

Zunächst wurde durch einen Clustering-Ansatz mittels des Algorithmus OPTICS^[168;169] versucht, typische Motive aktivierender Bindestellen zu detektieren, sowie evolutionäre Zusammenhänge aktivierender und nichtaktivierender Bindestellen aufzuklären. Obwohl es aufgrund von Ungenauigkeiten in den zugrundeliegenden Röntgenstrukturen und der stark variierenden Zahl von Wechselwirkungspartnern des Diphosphats nicht gelang, eine für die Detektion von Motiven geeignete Repräsentation der Bindestellen zu finden, wurde doch Eines deutlich: Die Beteiligung von Brønstedsäuren und lewissauren Metallionen an der Diphosphatbindung stellt auch in Proteinen ein notwendiges, wenn auch kein hinreichendes Kriterium für die Aktivierung dar.

Nachfolgend erzeugte Statistiken, die die Erkennung aktivierter und nichtaktivierter Diphosphate in der Proteinbindetasche gegenüberstellen, zeigten, dass in der Gesamtheit der berücksichtigten aktivierenden Bindestellen saure Gruppen (Brønstedsäuren und lewissaure Metallionen) und Wasserstoffbrückendonoren nahezu gleichhäufig vertreten sind, während in

nichtaktivierenden Bindestellen ein (1:5)-Verhältnis vorliegt. Eine Bevorzugung für bestimmte saure Gruppen durch aktivierte Diphosphate konnte dabei nicht festgestellt werden. Gleichzeitig wurde dargelegt, dass in aktivierenden Bindetaschen pro Bindetasche tendentiell mehr saure Gruppen an der Diphosphatbindung beteiligt sind, als in nichtaktivierenden Bindetaschen, wobei keine Bevorzugung für die Bindung von P_{α} oder P_{β} des Diphosphats vorliegt. Zusätzlich konnte mittels quantenchemischer Berechnungen am Beispiel von Dimethylallyldiphosphat (DMAPP) gezeigt werden, dass Aminosäuren mit ihrer geringen Säurestärke tatsächlich in der Lage sind, selbst im Metallkomplex befindliches Diphosphat zu protonieren. Es kann damit von einer Protonenübertragung von Aminosäuren auf das Diphosphat in der Proteinbindetasche ausgegangen werden. Interessanterweise wurde eine Häufung von Wechselwirkungen des zentralen Sauerstoffatoms des Diphosphats (O^M) in aktivierenden Bindestellen festgestellt. Die für die Zugänglichkeit des O^M -Atoms notwendige Diphosphatkonformation ist dabei nicht die energetisch günstigste, wie durch quantenchemische Berechnungen anhand von DMAPP ermittelt wurde; die geringfügige Destabilisierung kann jedoch wahrscheinlich durch die gehäuften Wechselwirkungen von O^M kompensiert werden. Um die Bedeutung saurer Gruppen im Rahmen der Aktivierung besser zu verstehen, wurden die lewissauren Metallionen und Brønstedtsäuren getrennt betrachtet und gleichzeitig nur Wechselwirkungen von aktivierenden und nichtaktivierenden Bindestellen gegenübergestellt, die ähnliche Liganden binden. Dabei wurde sowohl für die NDP-Zucker-Liganden als auch für Prenyldiphosphate deutlich, dass die Brønstedtsäuren auch in nichtaktivierenden Bindestellen häufig auftreten. Auffällig ist jedoch, dass ausschließlich in aktivierenden Bindestellen eine Beteiligung von Metallionen an der Diphosphaterkennung vorliegt. Die Komplexbildung durch Metallionen stellt somit also einen wesentlichen, die Aktivierung bestimmenden Faktor dar. Interessanterweise existieren jedoch auch aktivierende Bindestellen, die metallionenunabhängig sind (z.B. NDP-Zucker-umsetzende Glycosyltransferasen vom Typ GT-B). Es stellt sich somit die Frage nach dem Effekt, den die Komplexbildung von Metallionen auf die Aktivierung ausübt bzw. wie metallionenunabhängige Enzyme diesen möglicherweise imitieren können.

Es erfolgte daher eine Analyse der Konformation aktivierter Diphosphate in der Bindetasche getrennt für jede der drei möglichen Ligandengruppen (NDP-Zucker, Prenyldiphosphate, Pterindiphosphate). Ein direkter Vergleich mit den Konformationen nichtaktivierter Diphosphate gleichen Ligandentyps lieferte hierbei die Grundlage. Für NDP-Zucker-Liganden der metallionenunabhängigen GT-B-Glycosyltransferasen wurde dabei die Ausbildung einer U-Form beobachtet, die vor allem auf intramolekularen Wechselwirkungen zwischen Diphosphat und Zuckerrest beruht und ähnlich wie Metallionen eine Stabilisierung der am Diphosphat entstehenden negativen Ladung bewirkt. Ähnliche, in Glycosidasen beobachtete

Wechselwirkungen zeigten, dass diese wesentlich zur Herabsetzung der Aktivierungsbarriere beitragen können.^[257;258] Die Analyse der Diphosphatkonformationen der Prenyl-liganden wurde durch quantenchemische Analysen anhand von DMAPP gestützt. Dabei zeigte sich insbesondere die Bedeutung des Torsionswinkels CCCO, der bei einer Ausprägung von 90° die Ausbildung des mesomeriestabilisierten Zustandes des entstehenden Allylkations fördert. Es konnte außerdem gezeigt werden, dass dieser Torsionswinkel stark durch Protonierungen und Metallionenkomplexierung des Diphosphats beeinflusst wird: Beide bewirken in den Grundzuständen eine Verschiebung von CCCO in Richtung 90° und damit eine Vorjustierung dieses für die Aktivierung bedeutsamen Winkels. Diese Vorjustierung ist vermutlich auch die Ursache für die beobachtete Ausprägung des Torsionswinkels um $\pm 120^\circ$ in den Röntgenstrukturen prenyldiphosphataktivierender Bindestellen. Auch für aktivierte Pterindiphosphatstrukturen könnte dieser Winkel eine Rolle spielen, jedoch konnte dies anhand der verfügbaren Röntgenstruktur nicht nachgewiesen werden.

Neben der Erkennung des Diphosphats, seiner Konformation in der Bindetasche und dem jeweiligen Zusammenhang mit Protonierung und Metallionenkomplexierung des Diphosphats wurde auch die Energetik der Aktivierungsreaktion analysiert. Die Hydrolyse von DMAPP diente hierbei als Modellreaktion der quantenchemischen Berechnungen. Das Wassermolekül übernimmt dabei die Funktion des Nucleophils. Die Berechnung der Reaktionsenergie für verschieden starke Protonierungen und Metallionenkomplexierungen des Diphosphats zeigte eine stärkere Begünstigung der Reaktion für geringere Protonierungsgrade und metallionenfreie Systeme. Jedoch ist die berechnete Reaktionsenergie in allen Fällen negativ und die Reaktion damit energetisch begünstigt. Weiterhin wurden Angriffe des Wassermoleküls an das C¹-Atom des organischen Restes, sowie an das Phosphoratom des P_α des Diphosphats für verschieden starke Protonierungen und z.T. auch für Metallionenkomplexierungen des Diphosphats berechnet. Es zeigte sich, dass eine Spaltung der C¹O^R-Bindung durch C¹-Angriff generell möglich ist, während die Spaltung der O^RP-Bindung durch den P-Angriff nicht beobachtet werden konnte. Dies steht auch in Übereinstimmung mit der Literatur^[136], die einen S_N1-Mechanismus mit C¹O^R-Spaltung angibt. Die berechneten Aktivierungsenergien für den C¹-Angriff sind dabei stark abhängig von Protonierung und Metallionenkomplexierung des Diphosphats: Je höher der Protonierungsgrad, desto geringer ist die Aktivierungsenergie; Metallionen setzen die Barriere ebenfalls herab. Interessanterweise zeigten quantenchemische Optimierungen der DMAPP-Struktur, dass bereits erhöhte Protonierungsgrade oder Metallionenkomplexierungen zu einer Verlängerung der C¹O^R-Bindung und einer gegenläufigen Verkürzung der O^RP-Bindung führen, was ebenfalls auf eine Spaltung der

C¹O^R-Bindung bei den vorliegenden erhöhten Protonierungsgraden bzw. Metallionenkomplexierungen hindeutet.

Nun stellt sich die Frage nach dem Vorteil, der dem Enzym durch Nutzung von Metallionenkomplexen im Vergleich zur Protonierung des Diphosphats durch saure Aminosäuren entsteht. Die möglichen Ursachen sind vielfältig. Im Falle der Undecaprenyldiphosphatsynthase^[57], einem prenyldiphosphatkonvertierenden Enzym, sichert die Abhängigkeit vom Metallion beispielsweise die Präsenz beider Substrate, bevor die Aktivierung eingeleitet wird: Das nichtaktivierte Substrat bringt das Metallion in die Bindetasche und erst dort erfolgt eine Übertragung auf das aktivierte Substrat. Zusätzlich bilden Metallionen Chelatkomplexe, die eine bestimmte Anzahl und Anordnung von Liganden erfordern. Damit ist die Bindung von Metallionen präziser und stabiler, als die einzelner Protonen. Weiterhin sind Protonen mobiler als Metallionen. Die Ursache ist zum Einen ihre Größe, zum Anderen sind sie in der Lage Protonenrelais zu erzeugen, was sie schwer kontrollierbar machen kann.

Insgesamt ist die Diphosphataktivierung ein sehr komplexer Prozess, der stark abhängig ist von der exakten Gestalt des aktivierten Liganden und der zu katalysierenden Reaktion. Es konnte für mehrere Faktoren ein Einfluss auf den Aktivierungsprozess nachgewiesen werden. Eine besondere Rolle spielen dabei die Diphosphatprotonierung und -metallionenkomplexierung, die verschiedene Effekte, wie die Verlängerung oder Verkürzung von Bindungen oder die Vorjustierung von Torsionswinkeln, mit sich bringen. Jedes Protein scheint jeweils die Faktoren auszuwählen, die entsprechend der vorliegenden Bedingungen optimal sind; ein allgemein gültiger Steuerungsmechanismus konnte nicht identifiziert werden. Die Ergebnisse liefern Ansätze für das Verständnis von Reaktionsmechanismen diphosphataktivierender Enzyme oder zur Herstellung synthetischer Katalysatoren. Weitere theoretische Berechnungen – auch im Wechselspiel mit experimentellen Ansätzen – können helfen, auf Basis der präsentierten Daten aufgestellte Hypothesen zu spezifizieren und zu prüfen.

6 SUMMARY AND OUTLOOK

This work tries to provide insights into the activation by diphosphates in proteins. The importance of the analyses becomes clear when regarding the diversity of the metabolites for that this process represents the basis of their biosynthesis. Just the group of terpenoids and meroterpenoids contains 55.000 compounds. Their role ranges over the primary as well as the secondary metabolisms and even blazed its trail into pharmaceutical applications. But also the importance in connection with the chemical synthesis does not have to be neglected: The obtained findings could serve as a basis for the synthesis of catalysts that allow the generation of activated sp^3 -carbon centers under mild conditions.

As base of this work serves the Protein Data Bank, whose X-ray structures were searched for diphosphate containing substrates, substrate analogues or inhibitors. An extensive post-processing of the structures resulted in a non-redundant dataset of 15 activating and 414 non-activating protein binding sites which were used to create hypotheses that were themselves tested or extended with quantum mechanical approaches. The aim thereby was to consider as different points of view of the activation as possible to get a comprehensive picture. Because of their importance in the enzyme-independent hydrolysis of allylic diphosphates, the effects of the diphosphate protonation and metal ion complexation were explicitly involved into the considerations.

At first it was tried to detect typical motifs of activating binding sites and to elucidate evolutionary connections between activating and non-activating binding sites by a clustering approach with the algorithm OPTICS^[168;169]. Although it was because of uncertainties in the underlying X-ray structures and the strongly differing number of groups interacting with the diphosphates not possible to get a representation of the binding sites, that was useful for the detection of motifs or evolutionary connections, one thing became very clear: The involvement of Brønsted acids and Lewis-acidic metal ions in diphosphate binding represents a required but not a sufficient criterion for activation in proteins, too.

Afterwards generated statistics, that oppose the recognition of activated to non-activated diphosphates in the protein binding site, showed that in the entirety of the considered activating binding sites acidic groups (Brønsted acids and Lewis-acidic metal ions) and hydrogen bond donors are almost equally abundant represented, whereas there is a (1:5) relationship in non-activating binding sites. A preference of activated diphosphates for special acidic groups was not observed. Simultaneously it was demonstrated that in

activating binding sites by trend more acidic groups per binding site are involved in diphosphate binding then in non-activating binding sites, but no preference for the binding of P_{α} or P_{β} of the diphosphate exists. Additionally it was shown by quantum mechanical calculations on the example of dimethylallyl diphosphate (DMAPP) that amino acids with their low acidic strength are actually able to protonate even diphosphate that is in complex with a metal ion. Therefore it can be assumed that protons are actually transferred from amino acids to diphosphate within the protein pocket. Interestingly an accumulation of interactions of the central oxygen atom of the diphosphate (O^M) was detected in activating binding sites. The diphosphate conformation needed for this accessibility of the O^M atom is not the energetically most favoured, what was shown by quantum mechanical calculations on DMAPP, but the marginal destabilisation can probably be compensated by the cumulative interactions of O^M . To better understand the impact of acidic groups in the context of activation, the Lewis-acidic metal ions and the Brønsted acids were in a next step considered separately and just interactions of activating and non-activating binding sites with similar ligands were compared. In doing so it became clear for NDP sugar ligands as well as for prenyl diphosphates that Brønsted acids are also frequent in non-activating binding sites. However it is striking that an involvement of metal ions in the diphosphate recognition exclusively occurs in activating binding sites. The complexation of metal ions therefore represents an important, the activation affecting factor. Interestingly there are activating binding sites that are independent of metal ions, too (e.g. NDP sugar converting glycosyl transferases of type GT-B). Hence the question about the effect rises that metal ion complexation has on activation and how metal ion independent enzymes can possibly imitate that.

Therefore an analysis of the conformation of activated diphosphates in the binding site followed where the three possible ligand groups (NDP sugars, prenyl diphosphates, pterin diphosphates) were treated separately. A comparison with the conformations of non-activated diphosphates of the same ligand type served as a basis. In metal independent glycosyl transferases of type GT-B the formation of U-shaped NDP sugar ligands was observed, that is particularly based on intramolecular interactions between diphosphate and sugar moiety and that causes a stabilisation of the negative charge forming at the diphosphate (similar to metal ions). For similar interactions in glycosidases it has been shown that they can considerably contribute to the lowering of the activation barrier^[257;258]. The analysis of diphosphate conformations of prenyl ligands was supported by quantum mechanical calculations based on DMAPP. Thereby the importance of the torsion angle CCCO arised: A value of 90° assists the formation of an allyl cation that is stabilised by mesomerism. Additionally it was shown that this torsion angle is strongly influenced by protonation and metal ion complexation of the diphosphate: Both of them cause a shifting of CCCO in direction to 90° in the ground states

and therefore a pre-adjustment of this meaningful angle. This pre-adjustment probably causes the observed values of CCCO around 120° in the X-ray structures of prenyl diphosphate activating binding sites, too. For the activated pterin diphosphate structures this torsion angle possibly is relevant also, but this could not be proven by the only available X-ray structure.

Beyond the recognition of diphosphate, its conformation in the binding pocket and their respective relation to protonation and metal ion complexation the energetics of the activation reaction was analysed, too. The hydrolysis of DMAPP was used as model reaction for the quantum mechanical calculations, where the water molecule serves as a nucleophile. The determination of the reaction energy for different protonation states and metal ion complexations of the diphosphate showed that the reaction is stronger favoured for lower degrees of protonation or metal ion free systems. The calculated values are less than zero in all cases and therefore the reactions are altogether energetically favoured. Furthermore attacks of the water molecule on the C^1 atom of the organic moiety, as well as on the phosphorus atom of P_α of the diphosphate were calculated for different diphosphate protonation states and to some extent for different metal ion complexations. It became clear, that a C^1O^R -bond cleavage by a C^1 -attack is generally possible, whereas the cleavage of the $O^R P$ -bond by a P-attack was not observed. This also agrees with the literature^[136], where a S_N1 -mechanism with C^1O^R -bond cleavage is described. The determined activation energies for the C^1 -attack strongly depend on protonation state and metal ion complexation of the diphosphate: The higher the degree of protonation, the lower is the activation energy; metal ions lower the barrier, too. Interestingly it was also shown by quantum mechanical calculations that already higher degrees of protonation or metal ion complexations cause an elongation of the C^1O^R -bond and a contrary shortening of the $O^R P$ -bond, what additionally hints on a cleavage of C^1O^R at higher degrees of protonation or metal ion complexations.

At this point it could be asked for the advantage, that the enzyme arises when using metal ions instead of protonating the diphosphate by acidic amino acids. The possible causes are versatile. In the case of undecaprenyl diphosphate synthase^[57], a prenyl diphosphate converting enzyme, e.g., the dependency on the metal ion ensures the presence of both substrates, before the activation is introduced: The non-activated substrate takes the metal ion into the binding site and only there the transfer to the activated substrate occurs. Another fact is the formation of chelate complexes by metal ions, what requires a certain number and positioning of ligands. Therefore the binding of metal ions is more precise and more stable, than the binding of single protons. Protons are furthermore more mobile than metal ions. The

reason is their size on the one hand and their ability to form proton relays that can make them harder to control on the other hand.

Altogether diphosphate activation is a very complex process that strongly depends on the exact shape of the activated ligand and the reaction to catalyse. For several factors an influence on the activation process was demonstrated. A special role is played by protonation and metal ion complexation of the diphosphate, that both bring different effects, like the elongation or shortening of bonds or the pre-adjustment of torsion angles, with them. Each protein seems to choose those factors that are optimal with respect to the underlying conditions; a general mechanism could not be identified. The results provide insights into the reaction mechanisms of diphosphate activating enzymes and basic knowledge for the creation of synthetic catalysts. Other theoretical calculations – perhaps in connection with experimental approaches – can help to specify and to prove new hypotheses build on the basis of the here presented data.

7 SOFT- UND HARDWARE, SKRIPTE

Nachfolgend ist die wichtigste, verwendete Software aufgelistet.

Name der Software	Ref.
CCG MOE™ Version 2007.0902	[158]
Gaussian® 03	[171]
R Version 2.4.1	[170]
ACD/pK _a DB® Version 10.01	[233]

Die Durchführung quantenchemischer Berechnungen erfordert leistungsfähige Computer. Folgende Rechenleistungen standen hierbei zur Verfügung.

- 1 Cluster, 8x Dual-Core AMD Opteron™ Modell 875, 16GB RAM
- 1 Cluster, 16x Dual-Core AMD Opteron™ Modell 8218, 64GB RAM
- 5 Cluster, je 8x Quad-Core AMD Opteron™ Modell 2356, je 32GB RAM

Nachfolgend sind die wesentlichen, im Rahmen dieser Arbeit implementierten Skripte aufgelistet, sowie eine kurze Beschreibung ihrer Funktion angegeben. Die Quelltexte, als auch die damit zusammenhängenden Daten können im Leibniz-Institut für Pflanzenbiochemie, Weinberg 3, 06120 Halle(Saale), Abteilung Natur- und Wirkstoffchemie, AG Computerchemie eingesehen werden.

Name des Skriptes	Beschreibung
<i>Extraktion der Diphosphatbindestellen aus der PDB^[52]; Ermittlung verschiedener Eigenschaften der Bindestellen; Speicherung in MOE™^[158]-Datenbank:</i>	
- extractPPbindingSites_db.svl	Finden von organischen Diphosphaten und definierten Mimikri in einem Molekül
- GetBindingSite3D.svl	Extraktion der 3D-Koordinaten der Diphosphatbindestellen, sowie verschiedene Informationen über Ligand, PDB-Klassifizierung usw. aus einer .pdb-Datei
- smp_db_GetBindingSites3D.svl	parallelierte Anwendung der beiden vorhergenannten Skripte auf einen Ordner mit 3D-Moleküldateien (z.B. *.pdb) und Speichern der Resultate in einer MOE™-Datenbank
- db_SetPhosphorusType.svl	Bestimmung von P _α und P _β für alle Diphosphate in einer MOE™-Datenbank
- addInfoToBSMDB_batch.svl	Ermittlung weiterer Eigenschaften von Bindestellen in einer MOE™-Datenbank, wie z.B. Klassifizierung von Ligand und Protein

Bestimmung einer nichtredundanten Menge von Bindestellen:

- db_GetCdhitCluster.svl Bestimmung des CD-HIT^[163]-Clusters für jede Bindestelle einer MOETM-Datenbank
 - db_GetRepresentatives.svl Bestimmung einer repräsentativen Bindestelle für jedes CD-HIT-Cluster einer MOETM-Datenbank
-

Weiterführende Analysen der Bindestellen:

- db_CountGroupedInteractions.svl Berechnung eines Eigenschaftsvektors, der die Anzahl definierter Wechselwirkungen des Diphosphats mit der Umgebung angibt, für jede Bindestelle einer MOETM-Datenbank
 - db_GetBlengthsBangles.svl Berechnung von Interaktionen, Abständen und Winkeln mit Unterscheidung der Sauerstofftypen O[#], O^R und O^M für Bindestellen einer MOETM-Datenbank
 - db_GetTorsionAngles.svl Berechnung von definierten Diphosphattorsionswinkeln von Strukturen in einer MOETM-Datenbank
 - misc_AnalyseBindingSites.svl Funktionensammlung zur Analyse von Diphosphatbindestellen
-

Allgemeine Funktionensammlungen:

- getInteractions_db.svl Funktionen zur Berechnung von Wechselwirkungen (H-Brücken, elektrostatische Wechselwirkungen), sowie zu ihrer Verwaltung in einer MOETM-Datenbank
 - miscellaneous.svl allgemein genutzte Funktionen
-

von G. Kirsten (MOETM-Support) zur Verfügung gestellte Skripte:

- dockutil_thiol_gk.svl,
ph4type_thiol_gk.svl Erweiterung grundlegender Skripte zur Berechnung von Wasserstoffbrücken um die Einbeziehung von Schwefel (z.B. in Form von Thiolgruppen)
-

Clustering von Bindestellen:

- optics.R Implementierung des OPTICS^[168;169]-Clusteringalgorithmus
-

8 LITERATURVERZEICHNIS

- [1.] Lairson, L. L., Henrissat, B., Davies, G. J., Withers, S. G. (2008). *Annu. Rev. Biochem.* 77, 521-555.
- [2.] Rademacher, T. W., Parekh, R. B., Dwek, R. A. (1988). *Annu. Rev. Biochem.* 57, 785-838.
- [3.] Zou, Y. Z., Li, C., Brunzelle, J. S., Nair, S. K. (2007). *Biochemistry* 46, 4294-4304.
- [4.] Bourne, Y., Henrissat, B. (2001). *Curr. Opin. Struct. Biol.* 11, 593-600.
- [5.] Zechel, D. L., Withers, S. G. (2000). *Acc. Chem. Res.* 33, 11-18.
- [6.] Coutinho, P. M., Deleury, E., Davies, G. J., Henrissat, B. (2003). *J. Mol. Biol.* 328, 307-317.
- [7.] Ünlügil, U. M., Rini, J. M. (2000). *Curr. Opin. Struct. Biol.* 10, 510-517.
- [8.] Liu, J., Mushegian, A. (2003). *Protein Sci.* 12, 1418-1431.
- [9.] Ünlügil, U. M., Zhou, S. H., Yuwaraj, S., Sarkar, M., Schachter, H., Rini, J. M. (2000). *EMBO J.* 19, 5269-5280.
- [10.] Rossmann, M. G., Moras, D., Olsen, K. W. (1974). *Nature* 250, 194-199.
- [11.] Breton, C., Imberty, A. (1999). *Curr. Opin. Struct. Biol.* 9, 563-571.
- [12.] Brown, K., Pompeo, F., Dixon, S., Mengin-Lecreux, D., Cambillau, C., Bourne, Y. (1999). *EMBO J.* 18, 4096-4107.
- [13.] Campbell, R. E., Mosimann, S. C., Tanner, M. E., Strynadka, N. C. J. (2000). *Biochemistry* 39, 14993-15001.
- [14.] Breton, C., Bettler, E., Joziassé, D. H., Geremia, R. A., Imberty, A. (1998). *J. Biochem.* 123, 1000-1009.
- [15.] Wiggins, C. A. R., Munro, S. (1998). *Proc. Natl. Acad. Sci. U. S. A.* 95, 7945-7950.
- [16.] Larivière, L., Gueguen-Chaignon, V., Morera, S. (2003). *J. Mol. Biol.* 330, 1077-1086.
- [17.] Cantarel, B. L., Coutinho, P. M., Rancurel, C., Bernard, T., Lombard, V., Henrissat, B. (2009). *Nucleic Acids Res.* 37, D233-D238.
- [18.] Igura, M., Maita, N., Kamishikiryo, J., Yamada, M., Obita, T., Maenaka, K., Kohda, D. (2008). *EMBO J.* 27, 234-243.
- [19.] Lovering, A. L., de Castro, L. H., Lim, D., Strynadka, N. C. J. (2007). *Science* 315, 1402-1405.
- [20.] Yuan, Y. Q., Barrett, D., Zhang, Y., Kahne, D., Sliz, P., Walker, S. (2007). *Proc. Natl. Acad. Sci. U. S. A.* 104, 5348-5353.

- [21.] Kozmon, S., Tvaroska, I. (2006). *J. Am. Chem. Soc.* 128, 16921-16927.
- [22.] Frick, D. N., Townsend, B. D., Bessman, M. J. (1995). *J. Biol. Chem.* 270, 24086-24091.
- [23.] Xia, Z. Y., Azurmendi, H. F., Lairson, L. L., Withers, S. G., Gabelli, S. B., Bianchet, M. A., Amzel, L. M., Mildvan, A. S. (2005). *Biochemistry* 44, 8989-8997.
- [24.] Legler, P. M., Massiah, M. A., Bessman, M. J., & Mildvan, A. S. (2000). *Biochemistry* 39, 8603-8608.
- [25.] Bessman, M. J., Frick, D. N., OHandley, S. F. (1996). *J. Biol. Chem.* 271, 25059-25062.
- [26.] Koshland, D. E. (1953). *Biol. Rev. Camb. Philos. Soc.* 28, 416-436.
- [27.] Phillips, D. C. (1966). *Sci. Am.* 215, 78-90.
- [28.] Phillips, D. C. (1967). *Proc. Natl. Acad. Sci. U. S. A.* 57, 484-495.
- [29.] Wolfenden, R., Lu, X. D., Young, G. (1998). *J. Am. Chem. Soc.* 120, 6814-6815.
- [30.] Christianson, D. W. (2008). *Curr. Opin. Chem. Biol.* 12, 141-150.
- [31.] Liang, P. H., Ko, T. P., Wang, A. H. J. (2002). *Eur. J. Biochem.* 269, 3339-3354.
- [32.] Sadre, R., Gruber, J., Frentzen, M. (2006). *FEBS Lett.* 580, 5357-5362.
- [33.] Ashby, M. N., Kutsunai, S. Y., Ackerman, S., Tzagoloff, A., Edwards, P. A. (1992). *J. Biol. Chem.* 267, 4128-4136.
- [34.] Forsgren, M., Attersand, A., Lake, S., Grunler, J., Swiezewska, E., Dallner, G., Climent, C. L. (2004). *Biochem. J.* 382, 519-526.
- [35.] Melzer, M., Heide, L. (1994). *Biochim. Biophys. Acta* 1212, 93-102.
- [36.] Ohara, K., Yamamoto, K., Hamamoto, M., Sasaki, K., Yazaki, K. (2006). *Plant Cell Physiol.* 47, 581-590.
- [37.] Croteau, R. (1987). *Chem. Rev.* 87, 929-954.
- [38.] Dewick, P. M. (1999). *Nat. Prod. Rep.* 16, 97-130.
- [39.] Dewick, P. M. (2002). *Nat. Prod. Rep.* 19, 181-222.
- [40.] Bouvier, F., Rahier, A., Camara, B. (2005). *Prog. Lipid Res.* 44, 357-429.
- [41.] Klayman, D. L. (1985). *Science* 228, 1049-1055.
- [42.] Wani, M. C., Taylor, H. L., Wall, M. E., Coggon, P., Mcphail, A. T. (1971). *J. Am. Chem. Soc.* 93, 2325-2327.
- [43.] Julsing, M. K., Koulman, A., Woerdenbag, H. J., Quax, W. J., Kayser, O. (2006). *Biomol. Eng.* 23, 265-279.

- [44.] Nicolaou, K. C., Yang, Z., Liu, J. J., Ueno, H., Nantermet, P. G., Guy, R. K., Claiborne, C. F., Renaud, J., Couladouros, E. A., Paulvannan, K., Sorensen, E. J. (1994). *Nature* 367, 630-634.
- [45.] Kirby, J., Keasling, J. D. (2009). *Annu. Rev. Plant. Biol.* 60, 335-355.
- [46.] Rohdich, F., Bacher, A., Eisenreich, W. (2005). *Biochem. Soc. Trans.* 33, 785-791.
- [47.] Withers, S. T., Keasling, J. D. (2007). *Appl. Microbiol. Biotechnol.* 73, 980-990.
- [48.] Roberts, S. C. (2007). *Nat. Chem. Biol.* 3, 387-395.
- [49.] Tarshis, L. C., Yan, M. J., Poulter, C. D., Sacchettini, J. C. (1994). *Biochemistry* 33, 10871-10877.
- [50.] Starks, C. M., Back, K. W., Chappell, J., Noel, J. P. (1997). *Science* 277, 1815-1820.
- [51.] Brandt, W., Bräuer, L., Günnewich, N., Kufka, J., Rausch, F., Schulze, D., Schulze, E., Weber, R., Zakharova, S., Wessjohann, L. (2009). *Phytochemistry* 70, 1758-1775.
- [52.] Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, T. N., Weissig, H., Shindyalov, I. N., Bourne, P. E. (2000). *Nucleic Acids Res.* 28, 235-242.
- [53.] Chen, A. J., Kroon, P. A., Poulter, C. D. (1994). *Protein Sci.* 3, 600-607.
- [54.] Koyama, T., Obata, S., Osabe, M., Takeshita, A., Yokoyama, K., Uchida, M., Nishino, T., Ogura, K. (1993). *J. Biochem.* 113, 355-363.
- [55.] Shimizu, N., Koyama, T., Ogura, K. (1998). *J. Biol. Chem.* 273, 19476-19481.
- [56.] Fujihashi, M., Zhang, Y. W., Higuchi, Y., Li, X. Y., Koyama, T., Miki, K. (2001). *Proc. Natl. Acad. Sci. U. S. A.* 98, 4337-4342.
- [57.] Guo, R. T., Ko, T. P., Chen, A. P. C., Kuo, C. J., Wang, A. H. J., Liang, P. H. (2005). *J. Biol. Chem.* 280, 20762-20774.
- [58.] Kharel, Y., Zhang, Y. W., Fujihashi, M., Miki, K., Koyama, T. (2001). *J. Biol. Chem.* 276, 28459-28464.
- [59.] Pan, J. J., Yang, L. W., Liang, P. H. (2000). *Biochemistry* 39, 13856-13861.
- [60.] Gabelli, S. B., McLellan, J. S., Montalvetti, A., Oldfield, E., Docampo, R., Amzel, L. M. (2006). *Proteins* 62, 80-88.
- [61.] Chang, S. Y., Ko, T. P., Chen, A. P. C., Wang, A. H. J., Liang, P. H. (2004). *Protein Sci.* 13, 971-978.
- [62.] Poulter, C. D., Rilling, H. C. (1978). *Acc. Chem. Res.* 11, 307-313.
- [63.] Rahier, A., Taton, M., Benveniste, P. (1990). *Biochem. Soc. Trans.* 18, 48-52.
- [64.] Sagami, H., Korenaga, T., Ogura, K., Steiger, A., Pyun, H. J., Coates, R. M. (1992). *Arch. Biochem. Biophys.* 297, 314-320.

- [65.] Steiger, A., Pyun, H. J., Coates, R. M. (1992). *J. Org. Chem.* 57, 3444-3449.
- [66.] Sanchez, V. M., Crespo, A., Gutkind, J. S., Turjanski, A. G. (2006). *J. Phys. Chem. B* 110, 18052-18057.
- [67.] Rilling, H. C., Epstein, W. W. (1969). *J. Am. Chem. Soc.* 91, 1041-1042.
- [68.] Altman, L. F., Ash, L., Rilling, H. C., Muscio, F., Epstein, W. W., Gregonis, D. E., Larsen, B. R., Kowerski, R. C. (1972). *J. Am. Chem. Soc.* 94, 3257-3259.
- [69.] Benetka, W., Koranda, M., Eisenhaber, F. (2006). *Monatsh. Chem.* 137, 1241-1281.
- [70.] Gelb, M. H., Brunsveld, L., Hrycyna, C. A., Michaelis, S., Tamanoi, F., Van Voorhis, W. C., Waldmann, H. (2006). *Nat. Chem. Biol.* 2, 518-528.
- [71.] Reid, T. S., Terry, K. L., Casey, P. J., Beese, L. S. (2004). *J. Mol. Biol.* 343, 417-433.
- [72.] Novick, P., Zerial, M. (1997). *Curr. Opin. Cell Biol.* 9, 496-504.
- [73.] Pfeffer, S. R. (1994). *Curr. Opin. Cell Biol.* 6, 522-526.
- [74.] Huang, C. C., Hightower, K. E., Fierke, C. A. (2000). *Biochemistry* 39, 2593-2602.
- [75.] Reiss, Y., Brown, M. S., Goldstein, J. L. (1992). *J. Biol. Chem.* 267, 6403-6408.
- [76.] Hightower, K. E., Huang, C. C., Casey, P. J., Fierke, C. A. (1998). *Biochemistry* 37, 15555-15562.
- [77.] Zhang, F. L., Casey, P. J. (1996). *Biochem. J.* 320, 925-932.
- [78.] Wessjohann, L. A., Zakharova, S., Schulze, D., Kufka, J., Weber, R., Braeuer, L., Brandt, W. (2009). *Chimia* 63, 340-344.
- [79.] Botta, B., Delle Monache, G., Menendez, P., Boffi, A. (2005). *Trends Pharmacol. Sci.* 26, 606-608.
- [80.] Bräuer, L., Brandt, W., Wessjohann, L. A. (2004). *J. Mol. Model.* 317-327.
- [81.] Bräuer, L., Brandt, W., Schulze, D., Zakharova, S., Wessjohann, L. (2008). *Chembiochem* 9, 982-992.
- [82.] Ohara, K., Muroya, A., Fijkushima, N., Yazaki, K. (2009). *Biochem. J.* 421, 231-241.
- [83.] Tello, M., Kuzuyama, T., Heide, L., Noel, J. P., Richard, S. B. (2008). *Cell. Mol. Life Sci.* 65, 1459-1463.
- [84.] Kuzuyama, T., Noel, J. P., Richard, S. B. (2005). *Nature* 435, 983-987.
- [85.] Haagen, Y., Unsold, I., Westrich, L., Gust, B., Richard, S. B., Noel, J. P., Heide, L. (2007). *FEBS Lett.* 581, 2889-2893.
- [86.] Li, S.-M. (2009). *Phytochemistry* 70, 1746-1757.
- [87.] Steffan, N., Grundmann, A., Yin, W. B., Kremer, A., Li, S. M. (2009). *Curr. Med. Chem* 16, 218-231.

- [88.] Metzger, U., Schall, C., Zocher, G., Unsold, I., Stec, E., Li, S. M., Heide, L., Stehle, T. (2009). *Proc. Natl. Acad. Sci. U. S. A.* 106, 14309-14314.
- [89.] Luk, L. Y. P., Tanner, M. E. (2009). *J. Am. Chem. Soc.* 131, 13932-13933.
- [90.] Edwards, D. J., Gerwick, W. H. (2004). *J. Am. Chem. Soc.* 126, 11432-11433.
- [91.] Bohlmann, J., Steele, C. L., Croteau, R. (1997). *J. Biol. Chem.* 272, 21784-21792.
- [92.] Dudareva, N., Martin, D., Kish, C. M., Kolosova, N., Gorenstein, N., Faldt, J., Miller, B., Bohlmann, J. (2003). *Plant Cell* 15, 1227-1241.
- [93.] Christianson, D. W. (2006). *Chem. Rev.* 106, 3412-3442.
- [94.] Wendt, K. U., Schulz, G. E., Corey, E. J., Liu, D. R. (2000). *Angew. Chem., Int. Ed. Engl.* 39, 2812-2833.
- [95.] Cane, D. E., Kang, I. (2000). *Arch. Biochem. Biophys.* 376, 354-364.
- [96.] Zhou, K., Peters, R. J. (2009). *Phytochemistry* 70, 366-369.
- [97.] Hyatt, D. C., Croteau, R. (2005). *Arch. Biochem. Biophys.* 439, 222-233.
- [98.] Zheng, W., Sun, F., Bartlam, M., Li, X. M., Li, R., Rao, Z. H. (2007). *J. Mol. Biol.* 366, 1447-1458.
- [99.] Lange, B. M., Rujan, T., Martin, W., Croteau, R. (2000). *Proc. Natl. Acad. Sci. U. S. A.* 97, 13172-13177.
- [100.] Anderson, M. S., Muehlbacher, M., Street, I. P., Proffitt, J., Poulter, C. D. (1989). *J. Biol. Chem.* 264, 19169-19175.
- [101.] Kaneda, K., Kuzuyama, T., Takagi, M., Hayakawa, Y., Seto, H. (2001). *Proc. Natl. Acad. Sci. U. S. A.* 98, 932-937.
- [102.] Baca, A. M., Sirawaraporn, R., Turley, S., Sirawaraporn, W., Hol, W. G. J. (2000). *J. Mol. Biol.* 302, 1193-1212.
- [103.] Babaoglu, K., Qi, J. J., Lee, R. E., White, S. W. (2004). *Structure* 12, 1705-1717.
- [104.] Levy, C., Minnis, D., Derrick, J. P. (2008). *Biochem. J.* 412, 379-388.
- [105.] Then, R., Angehrn, P. (1973). *J. Gen. Microbiol.* 76, 255-263.
- [106.] Matherly, L. H. (2001). *Prog. Nucleic Acid Res. Mol. Biol.* 67, 131-162.
- [107.] Domagk, G. (1935). *Dtsch. Med. Wochenschr.* 61, 250-253.
- [108.] Bock, L., Miller, G. H., Schaper, K. J., Seydel, J. K. (1974). *J. Med. Chem.* 17, 23-28.
- [109.] Brown, G. M. (1962). *J. Biol. Chem.* 237, 536-540.
- [110.] Shiota, T., Mccann, M. P., Disraely, M. N. (1964). *J. Biol. Chem.* 239, 2259-2266.
- [111.] Woods, D. D. (1940). *Br. J. Exp. Pathol.* 21, 74-90.

- [112.] Hughes, W. T. (1988). *Antimicrob. Agents Chemother.* 32, 623-625.
- [113.] Sköld, O. (2000). *Drug Resist. Updat.* 3, 155-160.
- [114.] Sköld, O. (2001). *Vet. Res.* 32, 261-273.
- [115.] Graening, T. (2007). *Nachr. Chem.* 55, 836-840.
- [116.] Shilov, A. E., Shulpin, G. B. (1987). *Russ. Chem. Rev.* 56, 442-464.
- [117.] O'Malley, S. J., Tan, K. L., Watzke, A., Bergman, R. G., Ellman, J. A. (2005). *J. Am. Chem. Soc.* 127, 13496-13497.
- [118.] Wessjohann, L. A., Sontag, B. (1996). *Angew. Chem., Int. Ed. Engl.* 35, 1697-1699.
- [119.] Wessjohann, L. A., Sontag, B., Dessoy, M. A. (1999). *Enzymatic C-C coupling: the development of aromatic prenylation for organic synthesis*. In: *Bioorganic Chemistry - Highlights and New Aspects* (Hrsg.: Diedrichsen, U., Lindhorst, T. K., Westermann, B., Wessjohann, L. A.), S. 79-88, 1. Auflage, Wiley-VCH Verlag GmbH, Weinheim.
- [120.] Tamemoto, K., Takaishi, Y., Chen, B., Kawazoe, K., Shibata, H., Higuti, T., Honda, G., Ito, M., Takeda, Y., Kodzhimatov, O. K., Ashurmetov, O. (2001). *Phytochemistry* 58, 763-767.
- [121.] Zakharova, S., Fulhorst, M., Luczak, L., Wessjohann, L. (2004). *Arkivoc* 79-96.
- [122.] Bunton, C. A., Llewellyn, D. R., Oldham, K. G., Vernon, C. A. (1958). *J. Chem. Soc.* 3574-3587.
- [123.] Bunton, C. A., Llewellyn, D. R., Oldham, K. G., Vernon, C. A. (1958). *J. Chem. Soc.* 3588-3594.
- [124.] Kufka, J. (15.10.2009), Leibniz-Institut für Pflanzenbiochemie Halle (Saale). *persönl. Kommunikation*.
- [125.] Tidd, B. K. (1971). *J. Chem. Soc. (B)* 1168-1176.
- [126.] Valenzuela, P., Cori, O. (1967). *Tetrahedron Lett.* 3089-3094.
- [127.] Vial, M. V., Rojas, C., Portilla, G., Chayet, L., Perez, L. M., Cori, O., Bunton, C. A. (1981). *Tetrahedron* 37, 2351-2357.
- [128.] Cramer, F., Rittersdorf, W. (1967). *Tetrahedron* 23, 3015-3022.
- [129.] Eggerer, H. (1961). *Chem. Ber.* 94, 174-185.
- [130.] Goodman, D. S., Popjak, G. (1960). *J. Lipid Res.* 1, 286-300.
- [131.] Sofer, S. S., Rilling, H. C. (1969). *J. Lipid Res.* 10, 183-187.
- [132.] Brems, D. N., Rilling, H. C. (1977). *J. Am. Chem. Soc.* 99, 8351-8352.
- [133.] Cane, D. E. (1980). *Tetrahedron* 36, 1109-1159.

- [134.] George-Nascimento, C., Pont-Lezica, R., Cori, O. (1971). *Biochem. Biophys. Res. Commun.* 45, 119-124.
- [135.] Rittersdorf, W., Cramer, F. (1967). *Tetrahedron* 23, 3023-3028.
- [136.] Agranoff, B. W., Eggerer, H., Henning, U., Lynen, F. (1959). *J. Am. Chem. Soc.* 81, 1254-1255.
- [137.] Achari, A., Somers, D. O., Champness, J. N., Bryant, P. K., Rosemond, J., Stammers, D. K. (1997). *Nat. Struct. Biol.* 4, 490-497.
- [138.] Lesburg, C. A., Zhai, G. Z., Cane, D. E., Christianson, D. W. (1997). *Science* 277, 1820-1824.
- [139.] Long, S. B., Casey, P. J., Beese, L. S. (1998). *Biochemistry* 37, 9612-9618.
- [140.] Park, H. W., Boduluri, S. R., Moomaw, J. F., Casey, P. J., Beese, L. S. (1997). *Science* 275, 1800-1804.
- [141.] King, H. L., Rilling, H. C. (1977). *Biochemistry* 16, 3815-3819.
- [142.] Saderholm, M. J., Hightower, K. E., Fierke, C. A. (2000). *Biochemistry* 39, 12398-12405.
- [143.] Pojer, F., Wemakor, E., Kammerer, B., Chen, H. W., Walsh, C. T., Li, S. M., Heide, L. (2003). *Proc. Natl. Acad. Sci. U. S. A.* 100, 2316-2321.
- [144.] Unsold, I. A., Li, S. M. (2005). *Microbiology-Sgm* 151, 1499-1505.
- [145.] Unsold, I. A., Li, S. M. (2006). *ChemBiochem* 7, 158-164.
- [146.] Schellenberger, A. (1998). *Biochim. Biophys. Acta* 1385, 177-186.
- [147.] Eppendorfer, S., König, S., Golbik, R., Neef, H., Lehle, K., Jaenicke, R., Schellenberger, A., Hübner, G. (1993). *Bio. Chem. Hoppe-Seyler* 374, 1129-1134.
- [148.] Lohmann, K. (1929). *Naturwissenschaften* 17, 624-625.
- [149.] Lipmann, F. (1941). *Adv. Enzymol.* 1, 99-162.
- [150.] Voet, D., Voet, J. G. (1995). *Thermodynamics of Phosphate Compounds*. In: *Biochemistry*, S. 429-434, 2. Auflage, John Wiley & Sons, New York.
- [151.] Alberty, R. A., Smith, R. M., Bock, R. M. (1951). *J. Biol. Chem.* 193, 425-434.
- [152.] Gajewski, E., Steckler, D. K., Goldberg, R. N. (1986). *J. Biol. Chem.* 261, 2733-2737.
- [153.] Hansia, P., Guruprasad, N., Vishveshwara, S. (2006). *Biophys. Chem.* 119, 127-136.
- [154.] Alberty, R. A. (1968). *J. Biol. Chem.* 243, 1337-1343.
- [155.] George, P., Rutman, R. J., Phillips, R. C. (1963). *Biochemistry* 2, 508-512.

- [156.] Stryer, L. (1994). *Der Stoffwechsel: Konzepte und Grundmuster*. In: Biochemie, S. 327-343, 3. Auflage, Spektrum Akademischer Verlag GmbH, Heidelberg, Berlin, Oxford.
- [157.] Nemukhin, A. V., Grigorenko, B. L., Shadrina, M. S. (2008). *Russ. J. Gen. Chem.* 78, 696-703.
- [158.] MOE (The Molecular Operating Environment) Version 2007.0902 (2007). Chemical Computing Group Inc., 1010 Sherbrooke Street West, Suite 910, Montreal, Canada H3A 2R7 (<http://www.chemcomp.com>).
- [159.] Daylight Chemical Information Systems Inc., 120 Vantis, Suite 550, Aliso Viejo, CA 92656 (<http://www.daylight.com>).
- [160.] Labute, P. (2007). Protonate 3D: Assignment of Macromolecular Protonation State and Geometry (<http://www.chemcomp.com/journal/proton.htm>). Letzter Zugriff: 29.10.2010.
- [161.] Labute, P. (2000). Probabilistic Receptor Potentials (<http://www.chemcomp.com/journal/cstat.htm>). Letzter Zugriff: 29.10.2010.
- [162.] Boström, J., Hogner, A., Schmitt, S. (2006). *J. Med. Chem.* 49, 6716-6725.
- [163.] Li, W. Z., Godzik, A. (2006). *Bioinformatics* 22, 1658-1659.
- [164.] Jain, A. K., Murty, M. N., Flynn, P. J. (2007). *ACM Computing Surveys* 31, 264-323.
- [165.] Kettenring, J. R. (2006). *Journal of Classification* 23, 3-30.
- [166.] Handl, J., Knowles, J., Kell, D. B. (2005). *Bioinformatics* 21, 3201-3212.
- [167.] Yeung, K. Y., Haynor, D. R., Ruzzo, W. L. (2001). *Bioinformatics* 17, 309-318.
- [168.] Ankerst, Mihael, Breunig, Markus M., Kriegel, Hans-Peter, Sander, Jörg (1999). OPTICS: Ordering Points to Identify the Clustering Structure (<http://www.dbs.informatik.uni-muenchen.de/Publikationen/Papers/OPTICS.pdf>). Letzter Zugriff: 29.10.2010.
- [169.] Daszykowski, M., Walczak, B., Massart, D. L. (2002). *J. Chem. Inf. Comput. Sci.* 42, 500-507.
- [170.] R (A Language and Environment for Statistical Computing) Version 2.4.1 (2006). R Foundation for Statistical Computing, Wien (<http://www.R-project.org>).
- [171.] Frisch, M. J., Trucks, G. W., Schlegel, H. B., Scuseria, G. E., Robb, M. A., Cheeseman, J. R., Montgomery, Jr. J. A., Vreven, T., Kudin, K. N., Burant, J. C., Millam, J. M., Iyengar, S. S., Tomasi, J., Barone, V., Mennucci, B., Cossi, M., Scalmani, G., Rega, N., Petersson, G. A., Nakatsuji, H., Hada, M., Ehara, M., Toyota, K., Fukuda, R., Hasegawa, J., Ishida, M., Nakajima, T., Honda, Y., Kitao, O., Nakai, H., Klene, M., Li, X., Knox, J. E., Hratchian, H. P., Cross, J. B., Bakken, V., Adamo, C., Jaramillo, J., Gomperts, R., Salvador, P., Dannenberg, J. J., Zakrzewski, V. G., Dapprich, S., Daniels, A. D., Strain, M. C., Farkas, O., Malick, D. K., Rabuck, A. D., Raghavachari, K., Foresman, J. B., Ortiz, J. V., Cui, Q., Baboul, A. G., Clifford, S., Cioslowski, J., Stefanov, B. B., Liu, G., Liashenko, A., Piskorz, P., Komaromi, I., Martin, R. L., Fox, D. J., Keith, T., Al-Laham, M. A., Peng, C. Y., Nanayakkara, A., Challacombe, M., Gill, P. M. W., Johnson, B., Chen, W., Wong,

- M. W., Gonzalez, C., Pople, J. A., Gaussian 03, Revision C.02 (2004). Gaussian, Inc., Wallingford CT (<http://www.gaussian.com/>).
- [172.] Schrödinger, E. (1926). *Ann. Phys.* 79, 361-376.
- [173.] Born, M., Jordan, P. (1925). *Z. Phys.* 34, 858-888.
- [174.] Born, M., Heisenberg, E., Jordan, P. (1926). *Z. Phys.* 35, 557-615.
- [175.] Heisenberg, W. (1925). *Z. Phys.* 33, 879-93.
- [176.] Schrödinger, E. (1926). *Ann. Phys.* 79, 734-756.
- [177.] Born, M. (1926). *Z. Phys.* 38, 803-827.
- [178.] Born, M., Oppenheimer, R. (1927). *Ann. Phys.* 389, 457-484.
- [179.] Fock, V. (1930). *Z. Phys.* 61, 126-148.
- [180.] Hartree, D. R. (1928). *Proc. Cambridge Philos. Soc.* 24, 89-110.
- [181.] Møller, C., Plesset, M. S. (2009). *Phys. Rev.* 46, 618-622.
- [182.] Pople, J. A., Binkley, J. S., Seeger, R. (1976). *Int. J. Quantum Chem.* 1-19.
- [183.] Clark, T. (2003). *Quantum Mechanics*. In: Handbook of Chemoinformatics - From Data to Knowledge (Hrsg.: Gasteiger, J.), S. 947-975, 1. Auflage, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.
- [184.] Thomas, L. H. (1927). *Proc. Camb. Phil. Soc.* 23, 542-548.
- [185.] Fermi, E. (1927). *Rend. Acad. Lincei* 6, 602-607.
- [186.] Hohenberg, P., Kohn, W. (1964). *Phys. Rev. B* 136, B864-B871.
- [187.] Gill, P. M. W. (1998). *Density Functional Theory (DFT), Hartree-Fock (HF), and the Self-consistent Field*. In: Encyclopedia of Computational Chemistry (Hrsg.: von Ragué Schleyer, P., Allinger, N. L., Kollman, P. A., Clark, T., Schaefer, H. F., Gasteiger, J.), S. 678-689, 1. Auflage, John Wiley & Sons, Baffins Lane, Chichester.
- [188.] Kohn, W., Sham, L. J. (1965). *Phys. Rev.* 140, A1133-A1138.
- [189.] Kohn, W., Sham, L. J. (1965). *Phys. Rev.* 137, A1697-A1705.
- [190.] Gunnarsson, O., Lundqvist, B. I. (1976). *Phys. Rev. B* 13, 4274-4298.
- [191.] von Barth, U., Hedin, L. (1972). *J. Phys. C* 5, 1629-1642.
- [192.] Becke, A. D. (1993). *J. Chem. Phys.* 98, 5648-5652.
- [193.] Stewart, J. J. P. (2007). *J. Mol. Model.* 13, 1173-1213.
- [194.] Kayi, H., Clark, T. (2007). *J. Mol. Model.* 13, 965-979.
- [195.] Kayi, H., Clark, T. (2009). *J. Mol. Model.* 15, 1253-1269.

- [196.] Kayi, H., Clark, T. (2009). *J. Mol. Model.* 15, 295-308.
- [197.] Kayi, H., Clark, T. (2010). *J. Mol. Model.* 16, 29-47.
- [198.] Winget, P., Horn, A. H. C., Selcuki, C., Martin, B., Clark, T. (2003). *J. Mol. Model.* 9, 408-414.
- [199.] Winget, P., Clark, T. (2005). *J. Mol. Model.* 11, 439-456.
- [200.] Young, D. (2001). *Semiempirical Methods*. In: Computational Chemistry - A Practical Guide for Applying Techniques to Real World Problems, S. 32-41, 1. Auflage, John Wiley & Sons, New York.
- [201.] Slater, J. C. (1930). *Phys. Rev.* 36, 57-64.
- [202.] Hehre, W. J., Stewart, R. F., Pople, J. A. (1969). *J. Chem. Phys.* 51, 2657-2664.
- [203.] Hehre, W. J., Ditchfield, R., Pople, J. A. (1972). *J. Chem. Phys.* 56, 2257-2261.
- [204.] Hariharan, P. C., Pople, J. A. (1972). *Chem. Phys. Lett.* 16, 217-219.
- [205.] Hariharan, P. C., Pople, J. A. (1973). *Theor. Chim. Acta* 28, 213-222.
- [206.] Clark, T., Chandrasekhar, J., Spitznagel, G. W., Schleyer, P. V. (1983). *J. Comput. Chem.* 4, 294-301.
- [207.] Miertus, S., Scrocco, E., Tomasi, J. (1981). *Chem. Phys.* 55, 117-129.
- [208.] Miertus, S., Tomasi, J. (1982). *Chem. Phys.* 65, 239-245.
- [209.] Young, D. (2001). *Solvation*. In: Computational Chemistry - A Practical Guide for Applying Techniques to Real World Problems, S. 206-215, 1. Auflage, John Wiley & Sons, New York.
- [210.] Peng, C. Y., Schlegel, H. B. (1993). *Isr. J. Chem.* 33, 449-454.
- [211.] Peng, C. Y., Ayala, P. Y., Schlegel, H. B., Frisch, M. J. (1996). *J. Comput. Chem.* 17, 49-56.
- [212.] Lanig, H. (2003). *Molecular Mechanics*. In: Handbook of Chemoinformatics - From Data to Knowledge (Hrsg.: Gasteiger, J.), S. 920-946, 1. Auflage, Wiley-VCH Verlag GmbH & Co. KGaA, Weinheim.
- [213.] Halgren, T. A., Nachbar, R. B. (1996). *J. Comput. Chem.* 17, 587-615.
- [214.] Halgren, T. A. (1996). *J. Comput. Chem.* 17, 616-641.
- [215.] Halgren, T. A. (1996). *J. Comput. Chem.* 17, 490-519.
- [216.] Halgren, T. A. (1996). *J. Comput. Chem.* 17, 520-552.
- [217.] Halgren, T. A. (1996). *J. Comput. Chem.* 17, 553-586.
- [218.] Cheatham, T. E., Cieplak, P., Kollman, P. A. (1999). *J. Biomol. Struct. Dyn.* 16, 845-862.

- [219.] Cornell, W. D., Cieplak, P., Bayly, C. I., Gould, I. R., Merz, K. M., Ferguson, D. M., Spellmeyer, D. C., Fox, T., Caldwell, J. W., Kollman, P. A. (1995). *J. Am. Chem. Soc.* 117, 5179-5197.
- [220.] Wang, J. M., Cieplak, P., Kollman, P. A. (2000). *J. Comput. Chem.* 21, 1049-1074.
- [221.] Copley, R. R., Barton, G. J. (1994). *J. Mol. Biol.* 242, 321-329.
- [222.] Unligil, U. M., Zhou, S. H., Yuwaraj, S., Sarkar, M., Schachter, H., Rini, J. M. (2000). *EMBO J.* 19, 5269-5280.
- [223.] Gibbons, B. J., Roach, P. J., Hurley, T. D. (2002). *J. Mol. Biol.* 319, 463-477.
- [224.] Zhang, Y., Xiang, Y., Van Etten, J. L., Rossmann, M. G. (2007). *Structure* 15, 1031-1039.
- [225.] Alfaro, J. A., Zheng, R. B., Persson, M., Letts, J. A., Polakowski, R., Bai, Y., Borisova, S. N., Seto, N. O. L., Lowary, T. L., Palcic, M. M., Evans, S. V. (2008). *J. Biol. Chem.* 283, 10097-10108.
- [226.] Gibson, R. P., Tarling, C. A., Roberts, S., Withers, S. G., Davies, G. J. (2004). *J. Biol. Chem.* 279, 1950-1955.
- [227.] Offen, W., Martinez-Fleites, C., Yang, M., Kiat-Lim, E., Davis, B. G., Tarling, C. A., Ford, C. M., Bowles, D. J., Davies, G. J. (2006). *EMBO J.* 25, 1396-1405.
- [228.] Martinez-Fleites, C., Proctor, M., Roberts, S., Bolam, D. N., Gilbert, H. J., Davies, G. J. (2006). *Chem. Biol.* 13, 1143-1152.
- [229.] Brazier-Hicks, M., Offen, W. A., Gershater, M. C., Revett, T. J., Lim, E. K., Bowles, D. J., Davies, G. J., Edwards, R. (2007). *Proc. Natl. Acad. Sci. U. S. A.* 104, 20238-20243.
- [230.] Long, S. B., Hancock, P. J., Kral, A. M., Hellinga, H. W., Beese, L. S. (2001). *Proc. Natl. Acad. Sci. U. S. A.* 98, 12948-12953.
- [231.] Wang, W., Dong, C., McNeil, M., Kaur, D., Mahapatra, S., Crick, D. C., Naismith, J. H. (2008). *J. Mol. Biol.* 381, 129-140.
- [232.] Kemp, L. E., Bond, C. S., Hunter, W. N. (2002). *Proc. Natl. Acad. Sci. U. S. A.* 99, 6591-6596.
- [233.] ACD/pK_a DB Version 10.01 (2006). Advanced Chemistry Development, Inc. Toronto, On, Kanada (<http://www.acdlabs.com>).
- [234.] Hosfield, D. J., Zhang, Y. M., Dougan, D. R., Broun, A., Tari, L. W., Swanson, R. V., Finn, J. (2004). *J. Biol. Chem.* 279, 8526-8529.
- [235.] Allen, F. H. (2002). *Acta Crystallogr., B* 58, 380-388.
- [236.] Casas, J. S., Castellano, E. E., Couce, M. D., Garcia-Tasende, M. S., Sanchez, A., Sordo, J., Taboada, C., Vazquez-Lopez, E. M. (2001). *Inorg. Chem.* 40, 946-951.
- [237.] Casas, J. S., Castellano, E. E., Couce, M. D., Ellena, J., Sanchez, A., Sanchez, J. L., Sordo, J., Taboada, C. (2004). *Inorg. Chem.* 43, 1957-1963.

- [238.] Cini, R., Pifferi, C. (1999). *J. Chem. Soc., Dalton Trans.* 699-710.
- [239.] Malandrinos, G., Louludi, M., Mitsopoulou, C. A., Butler, I. S., Bau, R., Hadjiliadis, N. (1998). *J. Biol. Inorg. Chem.* 3, 437-448.
- [240.] Pletcher, J., Sax, M. (1972). *J. Am. Chem. Soc.* 94, 3998-4005.
- [241.] Pletcher, J., Wood, M., Blank, G., Shin, W., Sax, M. (1977). *Acta Crystallogr., B* 33, 3349-3359.
- [242.] Pletcher, J., Blank, G., Wood, M., Sax, M. (1979). *Acta Crystallogr., B* 35, 1633-1637.
- [243.] Swaminathan, P., Sundaralingam, M. (1980). *Acta Crystallogr., B* 36, 2590-2597.
- [244.] Viswamitra, M. A., Post, M. L., Kennard, O. (1979). *Acta Crystallogr., B* 35, 1089-1094.
- [245.] Rausch, F. (2009), Leibniz-Institut für Pflanzenbiochemie Halle (Saale). Detaillierte Untersuchungen zum Katalysemechanismus verschiedener Monoterpensynthesen mittels kombinierter quanten- und molekülmechanischer Methoden. *Diplomarbeit*.
- [246.] Hirsch, A. K. H., Fischer, F. R., Diederich, F. (2007). *Angew. Chem. Int. Ed. Engl.* 46, 338-352.
- [247.] Bossemeyer, D. (1994). *Trends Biochem. Sci.* 19, 201-205.
- [248.] Kinoshita, K., Sadanami, K., Kidera, A., Go, N. (1999). *Protein Eng.* 12, 11-14.
- [249.] Schulz, G. E. (1992). *Curr. Opin. Struct. Biol.* 2, 61-67.
- [250.] Walker, J. E., Saraste, M., Runswick, M. J., Gay, N. J. (1982). *EMBO J.* 1, 945-951.
- [251.] Velloso, L. M., Bhaskaran, S. S., Schuch, R., Fischetti, V. A., Stebbins, C. E. (2008). *EMBO Rep.* 9, 199-205.
- [252.] Bertrand, J. A., Auger, G., Martin, L., Fanchon, E., Blanot, D., La Beller, D., van Heijenoort, J., Dideberg, O. (1999). *J. Mol. Biol.* 289, 579-590.
- [253.] *CRC Handbook of Chemistry and Physics* (Hrsg.: Lide, D. R., Haynes, W. M.), 90. Auflage (2009), Taylor & Francis Group, Boca Raton.
- [254.] Jolly, W. L. (1991). *Modern Inorganic Chemistry*, 2. Auflage, McGraw-Hill Education, New York.
- [255.] Franck, R. W. (1983). *Tetrahedron* 39, 3251-3252.
- [256.] Morera, S., Imberty, A., Aschke-Sonnenborn, U., Ruger, W., Freemont, P. S. (1999). *J. Mol. Biol.* 292, 717-730.
- [257.] Namchuk, M. N., Withers, S. G. (1995). *Biochemistry* 34, 16194-16202.
- [258.] Mccarter, J. D., Adam, M. J., Withers, S. G. (1992). *Biochem. J.* 286, 721-727.
- [259.] Biarrotte-Sorin, S., Maillard, A. P., Delettre, J., Sougakoff, W., Arthur, M., Mayer, C. (2004). *Structure* 12, 257-267.

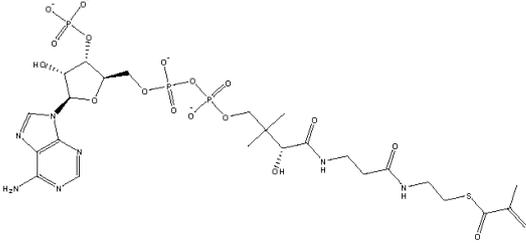
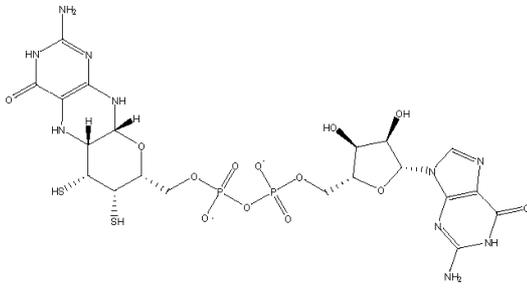
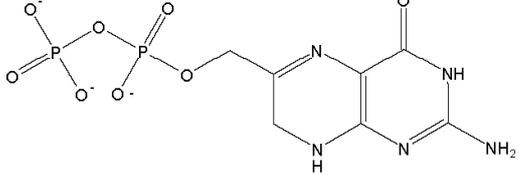
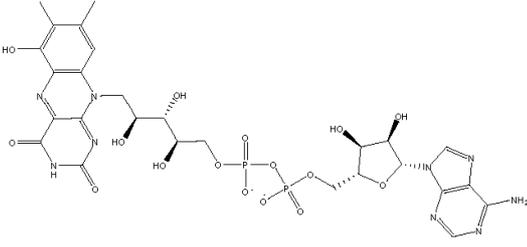
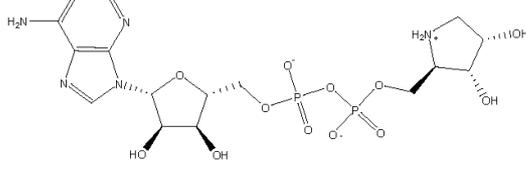
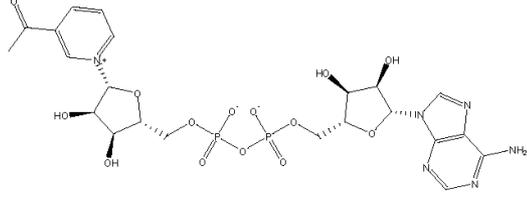
- [260.] Snook, C. F., Tipton, P. A., Beamer, L. J. (2003). *Biochemistry* 42, 4658-4668.
- [261.] Skarzynski, T., Mistry, A., Wonacott, A., Hutchinson, S. E., Kelly, V. A., Duncan, K. (1996). *Structure* 4, 1465-1474.
- [262.] Bisswanger, H. (1979). *pH- und Temperatur-Abhängigkeit von Enzymen*. In: *Theorie und Methoden der Enzymkinetik*, S. 125-131, 1. Auflage, Verlag Chemie, Weinheim.
- [263.] Feng, Z. K., Chen, L., Maddula, H., Akcan, O., Oughtred, R., Berman, H. M., Westbrook, J. (2004). *Bioinformatics* 20, 2153-2155.

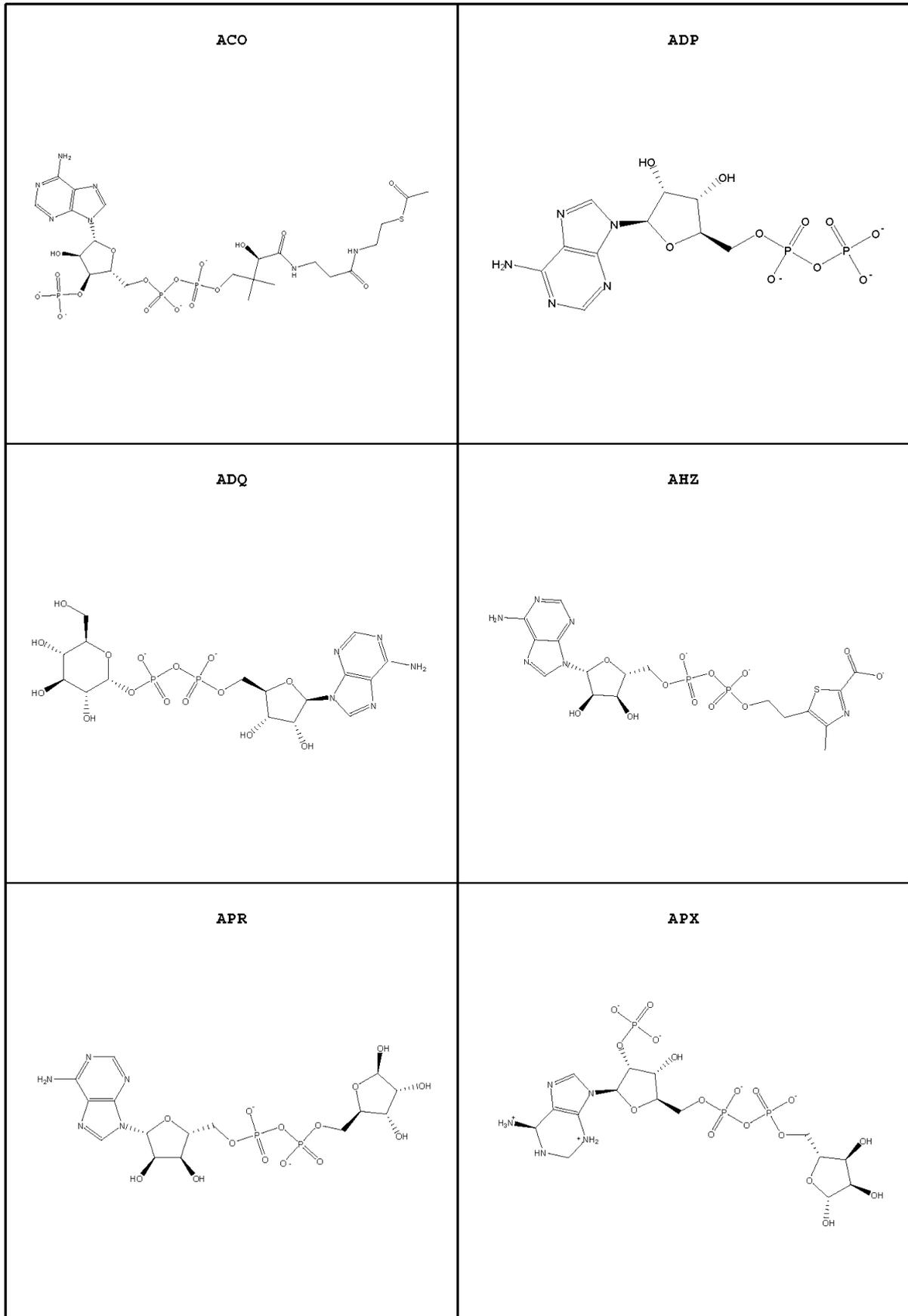
9 ANHANG

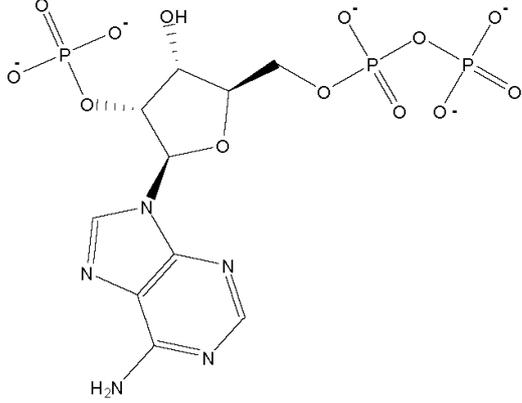
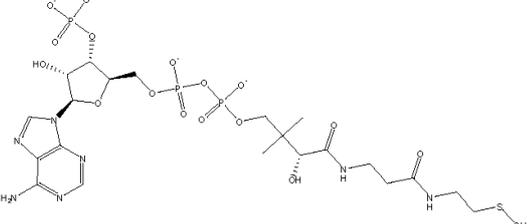
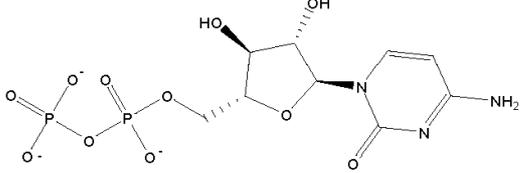
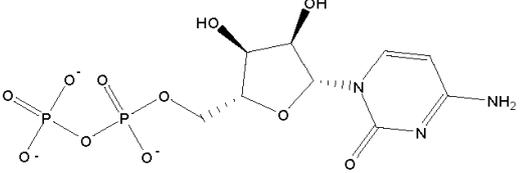
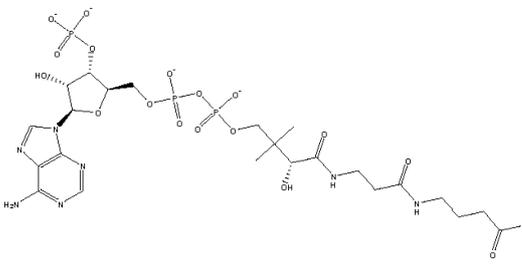
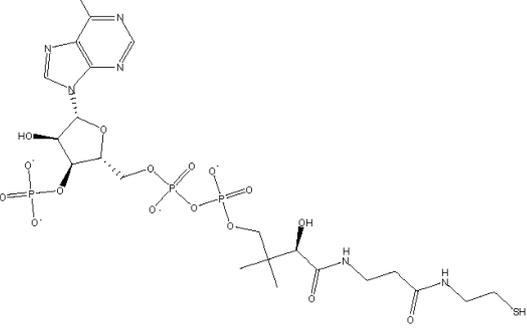
Tabelle 9.1: PDB-Kodierungen der in der vorliegenden Arbeit berücksichtigten Röntgenstrukturen. Aus einigen der PDB-Einträge konnte mehr als eine Bindestelle extrahiert werden. Dies führt zu einer Reduktion auf 400 Röntgenstrukturen, die die 429 betrachteten Diphosphatbindestellen enthalten.

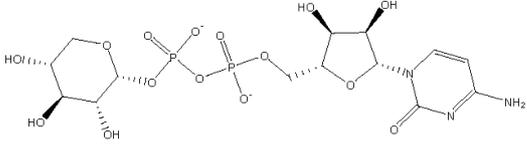
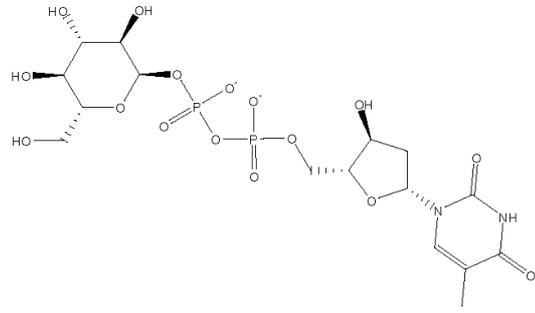
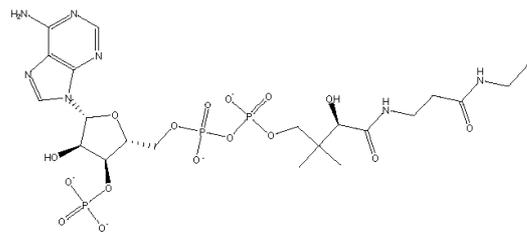
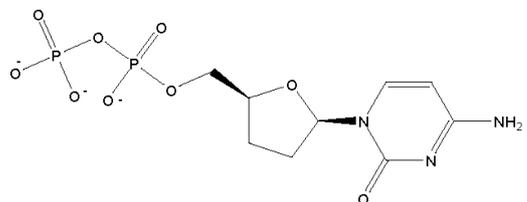
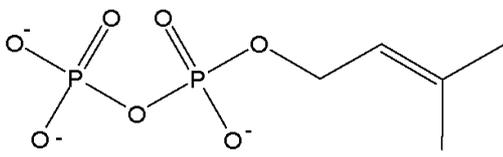
1AH4	1AJ2	1AJ8	1AOE	1B37	1BDB	1BXK	1C0P	1C14	1C1X
1CF3	1CJC	1COY	1CQX	1CYD	1D4A	1D7O	1DG7	1DGF	1DJL
1DOB	1DQS	1DSS	1DXY	1E1M	1E60	1E6U	1E6W	1E7W	1EBL
1EJ2	1EK6	1EMD	1EQ2	1EU1	1F0X	1F0Y	1F20	1FEC	1FK8
1FL2	1FMC	1FNB	1FOA	1FRB	1FY7	1G8K	1GCO	1GD1	1GEG
1GET	1GHE	1GPE	1GR0	1GSN	1GTE	1GVE	1GX1	1GY8	1H0H
1H16	1H2B	1H5Q	1HDO	1HYE	1I24	1IY8	1J39	1J5P	1J96
1JA9	1JAY	1JB9	1JCR	1JNR	1JR8	1JRA	1JU2	1JUV	1K4M
1K6X	1K87	1KDG	1KET	1KEW	1KOL	1KQF	1L7E	1LJ8	1LL2
1LQA	1LQT	1LQU	1LUA	1M6I	1MG5	1MI3	1MO9	1MUU	1MV8
1MX3	1N1E	1N4W	1N62	1N71	1N7H	1N9G	1NDC	1NP7	1NUE
1NUQ	1NVM	1NYT	1NYW	1O02	1O2A	1O2D	1O94	1O97	1OBB
1OC2	1OG1	1OJ7	1OQC	1ORR	1OWL	1OZH	1OZP	1P0F	1P1H
1P4N	1P77	1PL8	1PN0	1PO7	1PR9	1PT5	1PZG	1Q0Q	1Q16
1Q1R	1Q5M	1QGD	1QMG	1QR0	1QSR	1QYV	1R66	1RA8	1REQ
1RKX	1RM4	1RM6	1RP0	1RQJ	1RRM	1RSG	1RX0	1RX2	1RYI
1S2A	1S3Z	1S9F	1SBY	1SEP	1SNY	1SOW	1T2A	1T2D	1T3Q
1TCS	1TI6	1TI7	1TIQ	1TRB	1TRK	1U1I	1U3W	1U8F	1U8V
1UAE	1UDB	1UDC	1UMD	1UMK	1UQU	1UXG	1V3V	1V5E	1V5F
1V93	1V97	1VI0	1VJP	1VLB	1VLJ	1W4X	1W85	1WLW	1WMA
1WVG	1WWK	1WWZ	1X06	1X14	1X1T	1X7D	1XA4	1XG5	1XI2
1XJE	1XNG	1XQ6	1Y0P	1Y7T	1YB5	1YHL	1YNQ	1YOA	1YON
1YQZ	1Z4R	1Z82	1ZB6	1ZEM	1ZGD	1ZK4	1ZP3	1ZPD	1ZSX
1ZX9	2A9K	2AE2	2AG5	2AHV	2AO0	2APG	2B58	2B5V	2B69
2B9W	2BC0	2BD0	2BEI	2BEV	2BGM	2BGS	2BI7	2BJK	2BL9
2BRY	2BZS	2C0C	2C1Z	2C29	2C2S	2C31	2C3M	2C43	2C59
2C6Q	2C6Q	2CB8	2CD2	2CDU	2CFC	2CMJ	2CUL	2CVZ	2CWH
2CY0	2CY2	2CZC	2D1C	2D1Y	2D3M	2D4V	2DBQ	2DC1	2DFD
2DKH	2DKN	2DTE	2DVM	2EHQ	2EIX	2EP9	2EWD	2F1K	2FAV
2FR0	2FT0	2FZE	2FZJ	2G5C	2G76	2G82	2GAG	2GDZ	2GGS
2GIV	2GJC	2GN4	2GQT	2GQU	2GQW	2GR0	2GR2	2GSD	2H12
2H29	2H3U	2H4F	2H7C	2HEJ	2HQ9	2HRB	2HXV	2I3G	2I7N
2I8T	2ICK	2IGK	2IHT	2IID	2IVF	2IW1	2IXC	2IXI	2IXL
2IZZ	2J07	2J4D	2J6L	2JAE	2JAH	2JD1	2JDC	2JFG	2JHF
2JIF	2MBR	2NAP	2NSD	2NSY	2NVC	2NXW	2O23	2O28	2O2Q
2OB0	2OD7	2OZG	2P0W	2P5Y	2P72	2PA7	2PD6	2PD7	2PG0
2PV7	2PZM	2Q0K	2Q0L	2Q1S	2Q2V	2Q3E	2Q4W	2QA1	2QDX
2QTC	2R0C	2R6J	2RAF	2RC3	2RJ7	2UDP	2UUU	2UZ1	2V5Z
2VBF	2VCE	2VG0	2VG2	2VIG	2VJM	2YVJ	2YVW	2YWL	2YYY
2Z1M	2ZBA	3B2S	3B4W	3BEO	3C96	3DFR	4UAG	8DFR	9LDT

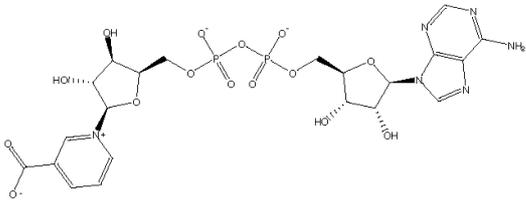
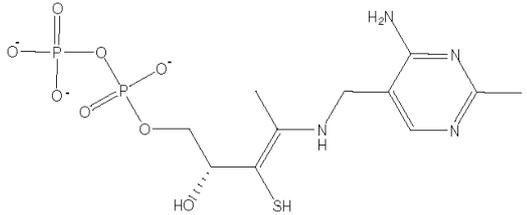
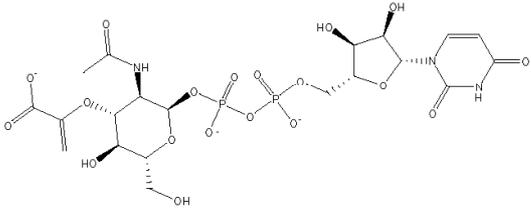
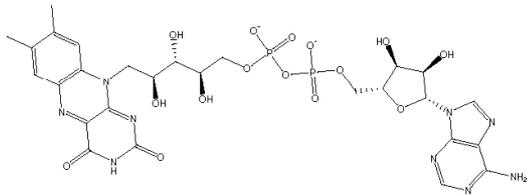
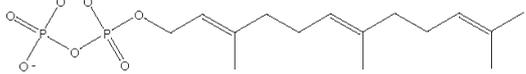
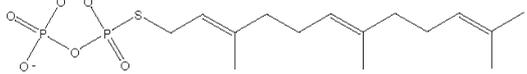
Tabelle 9.2: Strukturen und Dreibuchstabenkodierungen der in der vorliegenden Arbeit einbezogenen Diphosphatliganden. Die Strukturen wurden dem Ligand Depot^[263] entnommen und mit MOE^{TM[158]} dargestellt.

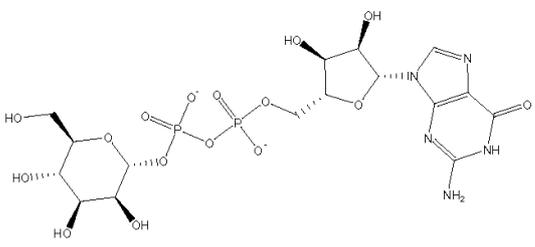
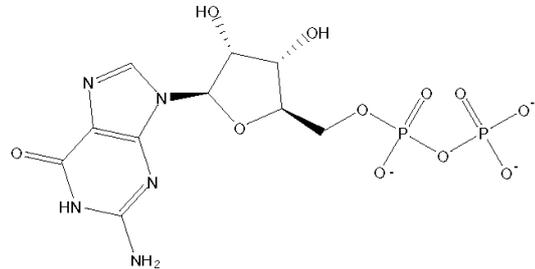
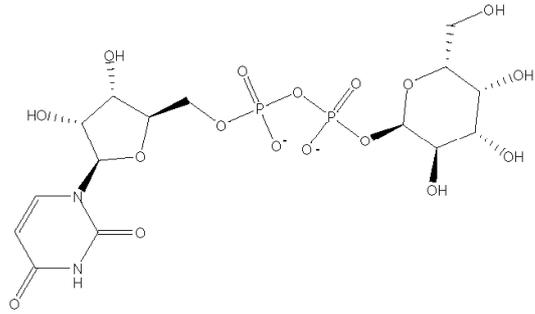
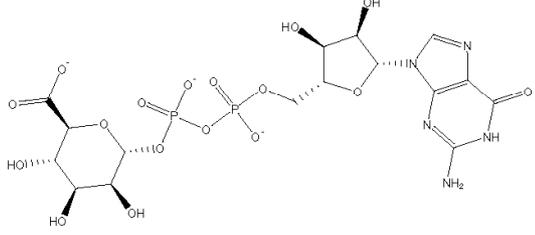
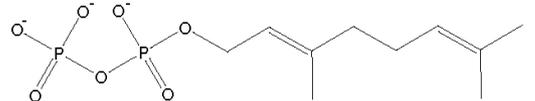
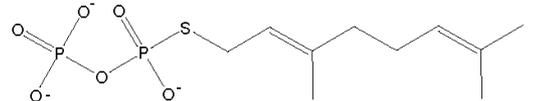
<p style="text-align: center;">2MC</p> 	<p style="text-align: center;">2MD</p> 
<p style="text-align: center;">2PH</p> 	<p style="text-align: center;">6FA</p> 
<p style="text-align: center;">A1R</p> 	<p style="text-align: center;">A3D</p> 

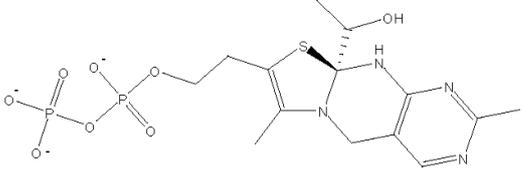
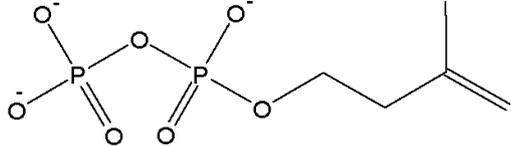
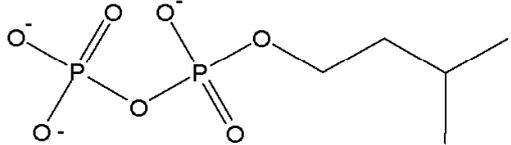
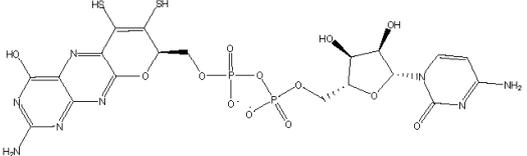
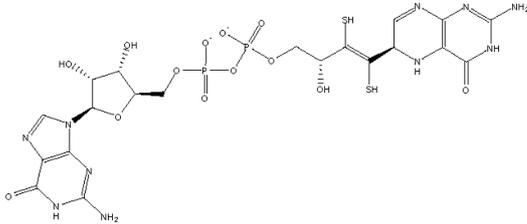
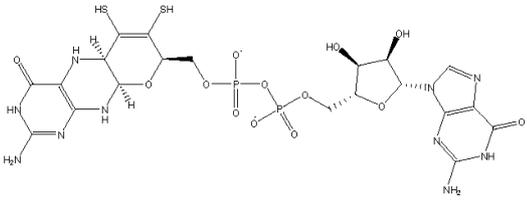


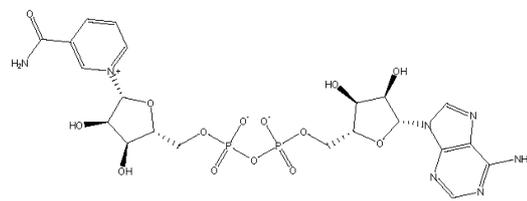
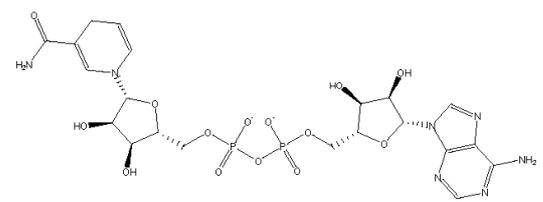
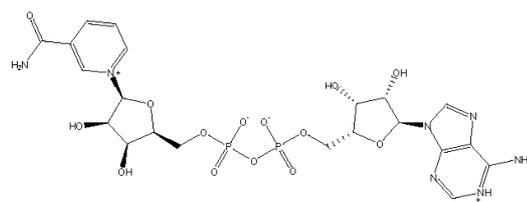
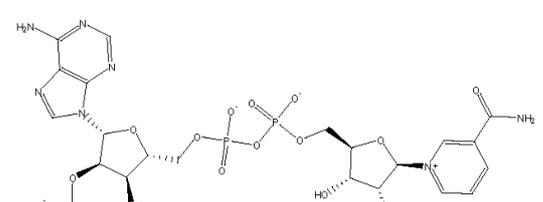
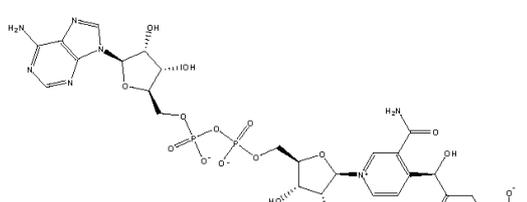
<p style="text-align: center;">ATR</p> 	<p style="text-align: center;">CAO</p> 
<p style="text-align: center;">CDF</p> 	<p style="text-align: center;">CDP</p> 
<p style="text-align: center;">CMX</p> 	<p style="text-align: center;">COA</p> 

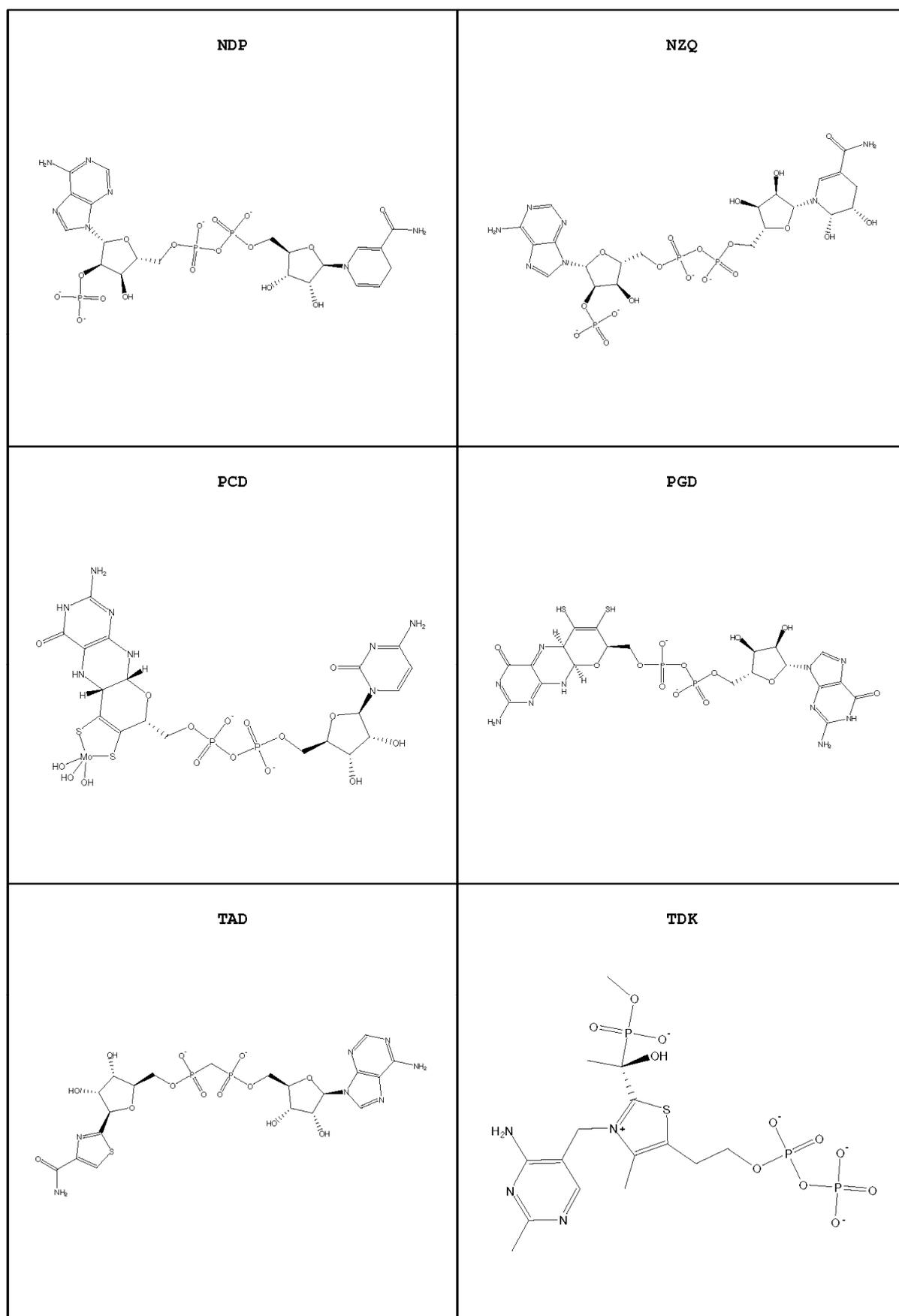
<p style="text-align: center;">CXY</p> 	<p style="text-align: center;">DAU</p> 
<p style="text-align: center;">DCA</p> 	<p style="text-align: center;">DCC</p> <p style="text-align: center;">Struktur siehe Seite 155</p>
<p style="text-align: center;">DDY</p> 	<p style="text-align: center;">DMA</p> 

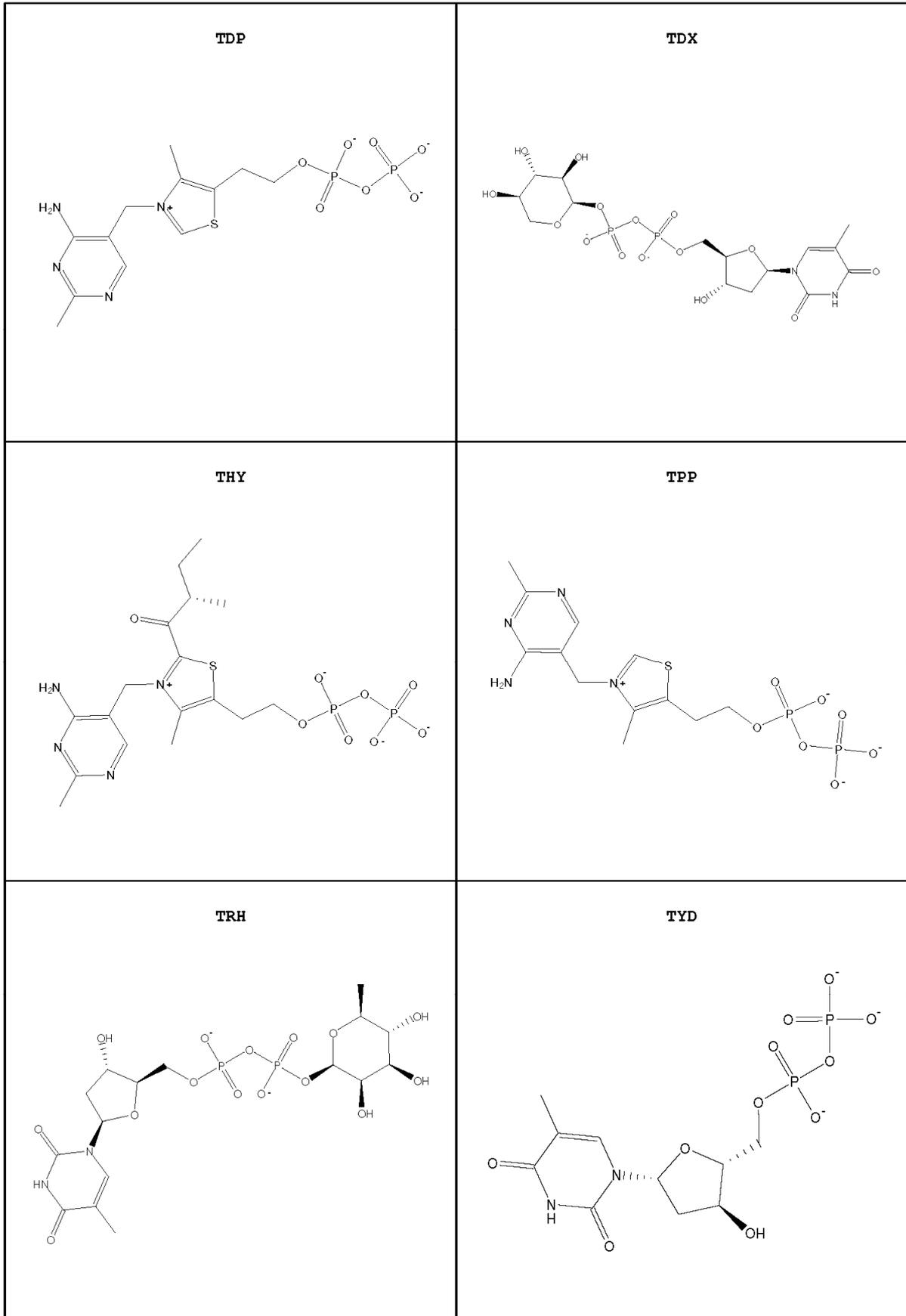
<p style="text-align: center;">DND</p> 	<p style="text-align: center;">DPX</p> 
<p style="text-align: center;">EPU</p> 	<p style="text-align: center;">FAD</p> 
<p style="text-align: center;">FPP</p> 	<p style="text-align: center;">FPS</p> 

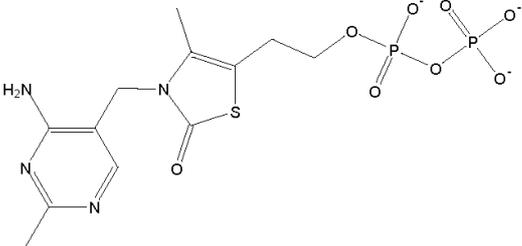
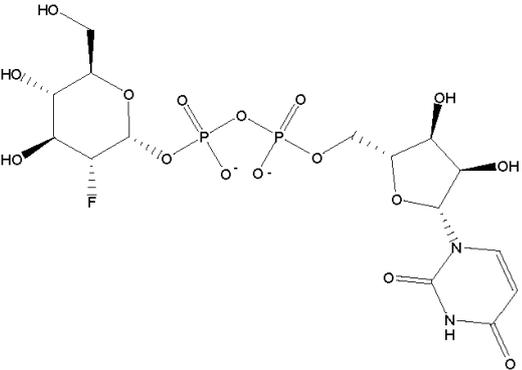
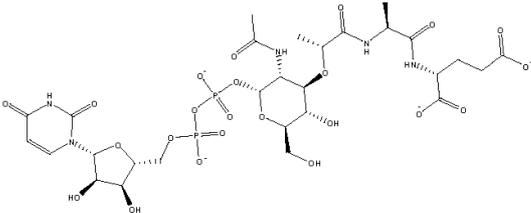
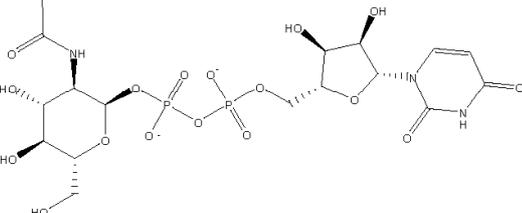
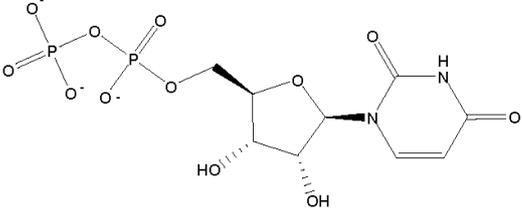
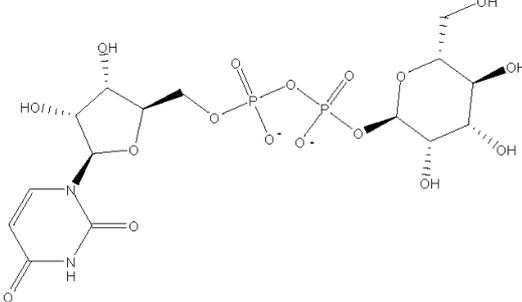
<p style="text-align: center;">GDD</p> 	<p style="text-align: center;">GDP</p> 
<p style="text-align: center;">GDU</p> 	<p style="text-align: center;">GDX</p> 
<p style="text-align: center;">GPP</p> 	<p style="text-align: center;">GST</p> 

<p style="text-align: center;">HE3</p> 	<p style="text-align: center;">IPE</p> 
<p style="text-align: center;">IPR</p> 	<p style="text-align: center;">MCN</p> 
<p style="text-align: center;">MD1</p> 	<p style="text-align: center;">MGD</p> 

<p style="text-align: center;">MYA</p> <p style="text-align: center;">Struktur siehe Seite 155</p>	<p style="text-align: center;">NAD</p> 
<p style="text-align: center;">NAI</p> 	<p style="text-align: center;">NAJ</p> 
<p style="text-align: center;">NAP</p> 	<p style="text-align: center;">NDE</p> 





<p style="text-align: center;">TZD</p> 	<p style="text-align: center;">U2F</p> 
<p style="text-align: center;">UAG</p> 	<p style="text-align: center;">UD1</p> 
<p style="text-align: center;">UDP</p> 	<p style="text-align: center;">UFM</p> 

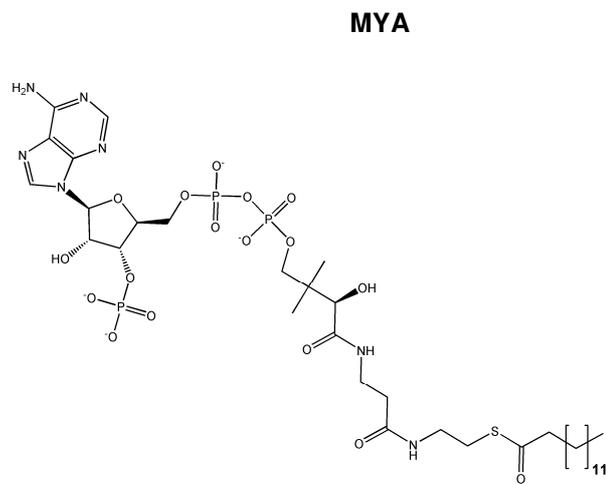
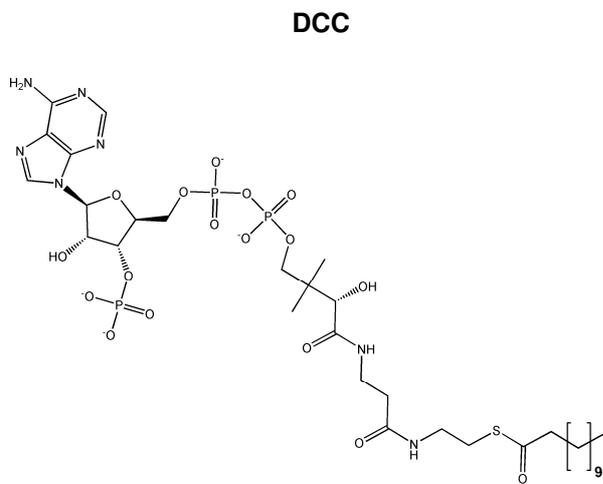
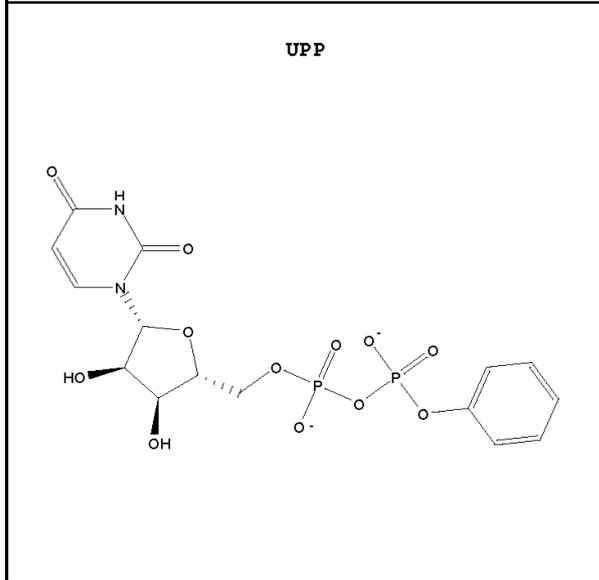
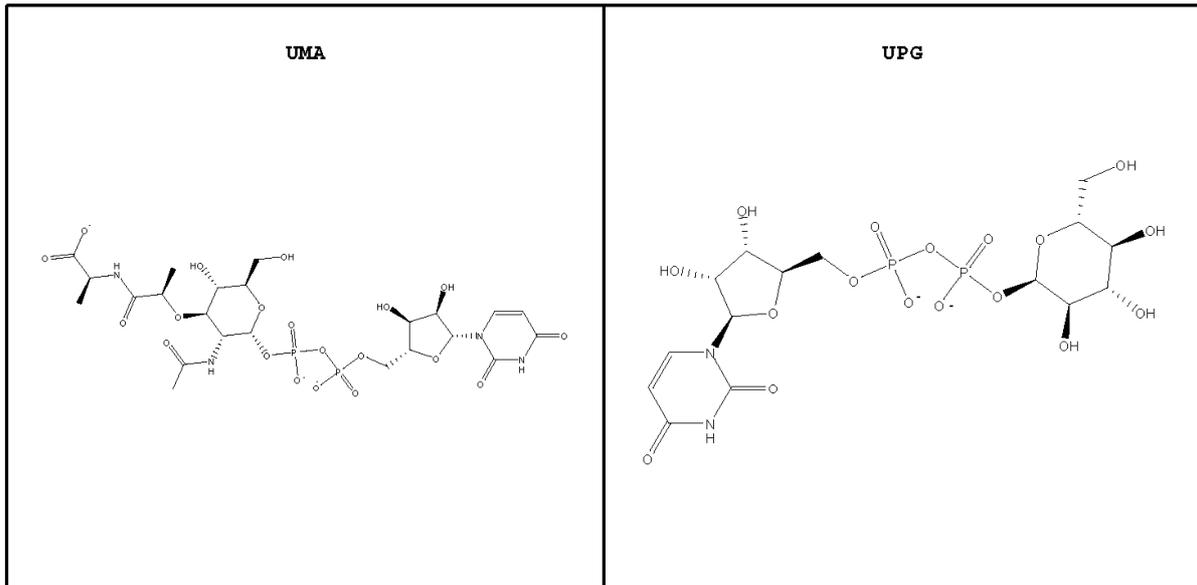


Tabelle 9.3: Totale Energien verschiedener Protonierungsgrade von DMAPP unter Verwendung der Methoden B3LYP, MP2 und HF, sowie dem Basissatz 6-31+G** und dem impliziten Solvationsmodell PCM.

Struktur	Energie in H		
	B3LYP	MP2	HF
DMAPP ⁽⁰⁾	-1405,92726459	-1403,03841990	-1400,71549853
DMAPP ⁽¹⁾	-1406,40990386	-1403,52058360	-1401,20395002
DMAPP ⁽²⁾	-1406,86518085	-1403,96708460	-1401,66348514
DMAPP ⁽³⁾	-1407,30713155	-1404,41453400	-1402,11874762

Tabelle 9.4: Totale Energien verschiedener Metallkomplexierungen von DMAPP unter Verwendung von B3LYP, 6-31+G** und PCM.

Struktur	Energie in H
DMAPP ⁽⁰⁾ -Mg-Komplex	-2063,96769370
DMAPP ⁽²⁾ -Mg-Komplex	-2141,32791485
DMAPP ⁽⁰⁾ -Mg ₃ -Komplex	-3303,48581866

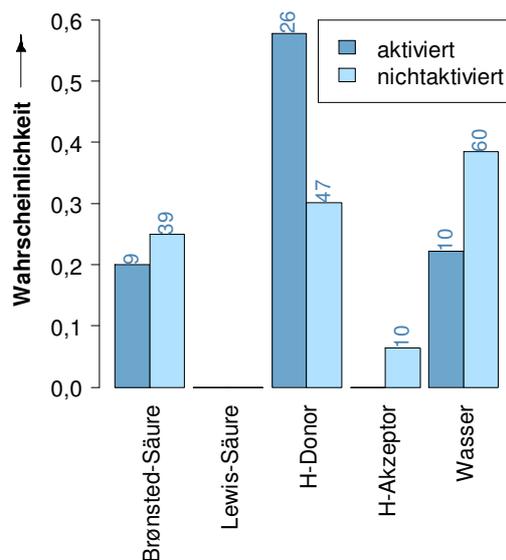


Abbildung 9.1: Verteilung der Häufigkeiten potentieller Protonendonoren (Brønsted-Säuren), Metallionen (Lewis-Säuren), Wasserstoffbrückendonoren und -akzeptoren, sowie Wassermolekülen in Bindestellen aktivierter (nur Enzyme vom GT-B-Typ) und nichtaktivierter NDP-Zucker. Die Zahlen oberhalb jedes Balkens geben die absolute Häufigkeit an.

Tabelle 9.5: Totale Energien verschiedener Protonierungsstufen und Metallkomplexierungen von DMAPP, sowie verschiedener Säuren und ihrer korrespondierenden Basen unter Verwendung von B3LYP, 6-31+G** und ohne Einbeziehung eines Solvatationsmodells.

Struktur (DMAPP)	Energie in H	
	DMAPP	DMAPP_H
DMAPP ⁽²⁾	-1406,75752882	-1407,25035834
DMAPP ⁽¹⁾	-1406,10490984	-1406,75752882
DMAPP ⁽⁰⁾	-1405,26490343	-1406,10490984
DMAPP ⁽²⁾ -Mg-Komplex	-2141,23950842	-2141,59453097
DMAPP ⁽⁰⁾ -Mg ₃ -Komplex	-3303,38379316	-3303,81441678
DMAPP ⁽¹⁾ -Mg-Komplex	-2064,30477490	-2064,78924094
DMAPP ⁽⁰⁾ -Mg-Komplex	-2063,64414199	-2064,30477490

Struktur	Energie in H	
	Säure (AH)	Base (A)
Ethylammonium	-135,553730006	-135,192282930
4-Ethylimidazol (protoniert)	-305,256005210	-304,875031520
1-Ethylguanidinium	-284,423033900	-284,023768300
Propansäure	-268,422808480	-267,858908010
<i>p</i> -Ethylphenol	-386,130766580	-385,564304510

Tabelle 9.6: Totale Energien von *n*-Propyldiphosphat mit verschiedenen Ausprägungen des COPO-Torsionswinkels unter Verwendung von B3LYP, 6-31+G** und PCM.

COPO in °	Energie in H
-60	-1328,47931024
180	-1328,47643935

Tabelle 9.7: Totale Energien der Strukturen der Reaktionskoordinate zur Verlängerung der C¹O^R-Bindung von DMAPP⁽³⁾ unter Verwendung von B3LYP, 6-31+G** und PCM.

Länge von C ¹ O ^R in Å	Energie in H
1,49	-1407,30709762
1,69	-1407,29964022
1,89	-1407,28905731
2,09	-1407,28197771
2,29	-1407,27803466
2,49	-1407,27552863
2,69	-1407,27421442
2,89	-1407,27306543
3,09	-1407,27206002
3,29	-1407,27163524
3,49	-1407,27134750
3,69	-1407,27105950

Tabelle 9.8: Totale Energien der aktivierten Struktur des DMAPP⁽³⁾ (C¹O^R-Distanz von 2.69 Å) bei verschiedenen Winkelausprägungen für die Torsionswinkel COPO und CCCO unter Verwendung von B3LYP, 6-31+G** und PCM.

Winkelausprägung in °	Energie in H	
	COPO	CCCO
0	-1407,27390880	1407,26987220
10	-1407,27388455	1407,26799800
20	-1407,27400782	1407,26796651
30	-1407,27379151	1407,26880519
40	-1407,27362832	1407,27076604
50	-1407,27377256	1407,27248458
60	-1407,27386203	1407,27335186
70	-1407,27399990	1407,27417147
80	-1407,27432248	1407,27484190
90	-1407,27439367	1407,27509958
100	-1407,27437259	1407,27505561
110	-1407,27429330	1407,27464407
120	-1407,27419350	1407,27369440
130	-1407,27401387	1407,27245457
140	-1407,27372737	1407,27087598
150	-1407,27369603	1407,26932956
160	-1407,27347862	1407,26927904
170	-1407,27312312	1407,26889817
180	-1407,27307118	1407,26889016
-170	-1407,27385772	-
-160	-1407,27385952	-
-150	-1407,27391558	-
-140	-1407,27390355	-
-130	-1407,27342257	-
-120	-1407,27370612	-
-110	-1407,27390311	-
-100	-1407,27399096	-
-90	-1407,27377674	-
-80	-1407,27397036	-
-70	-1407,27388752	-
-60	-1407,27354808	-
-50	-1407,27369805	-
-40	-1407,27402397	-
-30	-1407,27370337	-
-20	-1407,27309861	-
-10	-1407,27316509	-

Tabelle 9.9: Totale Energien verschiedener Protonierungsgrade und Metallkomplexierungen von DMAPP^(x) und des durch Hydrolyse daraus entstehenden anorganischen Diphosphats PP^(x+1) unter Verwendung von B3LYP, 6-31+G** und PCM. Die Energien für H₂O und DMAOH unter denselben Bedingungen sind -76,44843654 H bzw. -271,79091506 H.

Struktur	Energie in H	
	DMAPP ^(x)	PP ^(x+1)
DMAPP ⁽⁰⁾	-1405,92726459	-1210,59883377
DMAPP ⁽¹⁾	-1406,40990386	-1211,07235204
DMAPP ⁽²⁾	-1406,86518085	-1211,52672886
DMAPP ⁽³⁾	-1407,30713155	-1211,96805059
DMAPP ⁽⁰⁾ -Mg-Komplex	-2063,96769370	-1868,63393698
DMAPP ⁽²⁾ -Mg-Komplex	-2141,32791485	-1945,98689434

Tabelle 9.10: Totale Energien der Strukturen entlang der Reaktionskoordinate des Angriffs eines Wassermoleküls an C¹ verschiedener Protonierungsgrade und Metallkomplexierungen von DMAPP unter Verwendung von B3LYP, 6-31+G** und PCM. BL – Länge der O^{H₂O}-C¹-Bindung, die während der Berechnung verkürzt wird

DMAPP ⁽⁰⁾		DMAPP ⁽¹⁾		DMAPP ⁽²⁾	
BL in Å	Energie in H	BL in Å	Energie in H	BL in Å	Energie in H
3,02	-1482,37497509	3,04	-1482,85846068	3,01	-1483,30960058
2,82	-1482,37196699	2,84	-1482,85519978	2,81	-1483,30617090
2,62	-1482,36283747	2,64	-1482,84890059	2,61	-1483,29991966
2,42	-1482,35641994	2,44	-1482,83876471	2,41	-1483,28871406
2,22	-1482,33316741	2,24	-1482,82055565	2,21	-1483,26945418
2,02	-1482,29334342	2,04	-1482,79121215	2,01	-1483,23991450
1,82	-1482,25816528	1,84	-1482,74610789	1,96	-1483,23019626
1,77	-1482,24593805	1,79	-1482,73241189	1,91	-1483,22033300
1,72	-1482,23101688	1,74	-1482,71763867	1,86	-1483,20841140
1,67	-1482,21632428	1,69	-1482,70303723	1,81	-1483,19598378
1,62	-1482,20116066	1,64	-1482,85341198	1,76	-1483,18220183
1,57	-1482,37630213	1,59	-1482,86447753	1,71	-1483,16770868
1,52	-1482,37921419	1,54	-1482,86810738	1,66	-1483,30823254
1,47	-1482,38025485	1,49	-1482,87087606	1,61	-1483,31089586
1,42	-1482,38068297	1,44	-1482,87152863	1,41	-1483,31732768

DMAPP ⁽³⁾		DMAPP ⁽⁰⁾ -Mg-Komplex		DMAPP ⁽²⁾ -Mg-Komplex	
BL in Å	Energie in H	BL in Å	Energie in H	BL in Å	Energie in H
3,01	-1483,75063107	2,99	-2140,42145504	2,999728	-2217,77003267
2,81	-1483,74778070	2,79	-2140,41825677	2,799728	-2217,76728407
2,61	-1483,74223528	2,59	-2140,41169665	2,599728	-2217,76164333
2,41	-1483,73095684	2,39	-2140,39877103	2,399728	-2217,75039246
2,21	-1483,71162877	2,19	-2140,37823257	2,199728	-2217,73031178
2,01	-1483,68095959	1,99	-2140,34272387	1,999728	-2217,69917952
1,96	-1483,67171189	1,94	-2140,33187225	1,949728	-2217,68923240
1,91	-1483,66111118	1,89	-2140,32015093	1,899728	-2217,67832930
1,86	-1483,64926909	1,84	-2140,31158778	1,849728	-2217,66724449
1,81	-1483,63683846	1,79	-2140,29802941	1,799728	-2217,65436676
1,76	-1483,74433528	1,74	-2140,28379791	1,749728	-2217,64124289
1,71	-1483,74811961	1,69	-2140,41676696	1,699728	-2217,76743339
1,66	-1483,75198878	1,64	-2140,42054237	1,649728	-2217,77137184
1,61	-1483,75575652	1,59	-2140,42350012	1,599728	-2217,77508036
1,41	-1483,75991421	1,39	-2140,42889598	1,399728	-2217,77846151

Tabelle 9.11: Totale Energien der Struktur des ohne Abstandsfixierung nachoptimierten DMAPP-H₂O-Startkomplexes und des Pseudoübergangszustandes entlang der Reaktionskoordinate des Angriffs eines Wassermoleküls an C¹ für verschiedene Protonierungsgrade und Metallkomplexierungen von DMAPP unter Verwendung von B3LYP, 6-31+G und PCM.**

Struktur	Energie in H	
	DMAPP-H ₂ O-Startkomplex	Pseudoübergangszustand
DMAPP ⁽⁰⁾	-1482,37782873	-1482,20115749
DMAPP ⁽¹⁾	-1482,85890834	-1482,69247585
DMAPP ⁽²⁾	-1483,31450807	-1483,19583677
DMAPP ⁽³⁾	-1483,75341067	-1483,63682510
DMAPP ⁽⁰⁾ -Mg-Komplex	-2140,42608248	-2140,29481698
DMAPP ⁽²⁾ -Mg-Komplex	-2217,77121097	-2217,64124289

Tabelle 9.12: Totale Energien der Strukturen entlang der Reaktionskoordinate des Angriffs eines Wassermoleküls an das Phosphoratom des P_α für verschiedene Protonierungsgrade von DMAPP unter Verwendung von B3LYP, 6-31+G und PCM. BL – Länge der O^{H₂O}-C¹-Bindung, die während der Berechnung verkürzt wird**

DMAPP ⁽⁰⁾		DMAPP ⁽¹⁾		DMAPP ⁽²⁾	
BL in Å	Energie in H	BL in Å	Energie in H	BL in Å	Energie in H
3,004586	-1482,36520587	3,001665	-1482,85128482	2,99483	-1483,30489786
2,804586	-1482,35778261	2,801665	-1482,84741463	2,79483	-1483,30056159
2,604586	-1482,34616998	2,601665	-1482,83960950	2,59483	-1483,29424566
2,404586	-1482,33650222	2,401665	-1482,83191228	2,39483	-1483,28925283
2,204586	-1482,32246829	2,201665	-1482,82097033	2,19483	-1483,28037337
2,004586	-1482,30410869	2,001665	-1482,80788262	1,99483	-1483,26996296
1,804586	-1482,28082296	1,801665	-1482,79003551	1,79483	-1483,26894592
1,604586	-1482,23785948	1,601665	-1482,77769816	1,59483	-1483,24913750
1,404586	-1482,13877924	1,401665	-1482,70459845	1,39483	-1483,16897480

DMAPP ⁽³⁾	
BL in Å	Energie in H
3,033993	-1483,75078795
2,833993	-1483,74835180
2,633993	-1483,74372312
2,433993	-1483,73577460
2,233993	-1483,73121082
2,033993	-1483,72228869
1,833993	-1483,70959725
1,633993	-1483,68246230
1,433993	-1483,60902087

CURRICULUM VITAE

ANGABEN ZUR PERSON

NAME	Diana Schulze
GEBURTSDATUM	19. 05.1982
GEBURTSORT	Nordhausen

BERUFLICHER WERDEGANG

SEIT 01.2010	Klinikum der Friedrich-Schiller-Universität Jena <i>angestellt als Bioinformatikerin im Bereich der medizinischen Dokumentation</i>
05.2006 - 12.2009	Leibniz-Institut für Pflanzenbiochemie Halle (Saale) <i>Promotion im Bereich „Computerchemie/Chemoinformatik“</i> <i>Promotionsthema: Die Diphosphataktivierung in Proteinen – Betrachtungen aus verschiedenen Blickwinkeln</i> Betreuer: PD Dr. W. Brandt, Prof. Dr. L. A. Wessjohann
10.2000 - 05.2006	Martin-Luther-Universität Halle/Wittenberg <i>Studiengang: Bioinformatik (Diplom)</i> <i>Vertiefungsrichtung: Biochemie/Biotechnologie</i> <i>Diplomthema: Charakterisierung von Diphosphatbindestellen in Proteinen und Homologiemodellierung eines Proteins als Beitrag zur Untersuchung der Evolution prenylierender Enzyme</i> Betreuer: PD Dr. W. Brandt, Prof. Dr. L. A. Wessjohann Abschlussnote: sehr gut, mit Auszeichnung
03.2005	Praktikum Enzymkinetik & Enzymologie (ganztägig)
10.2003 - 03.2004	Praktikum „Molecular Modelling“ (4h/Woche)
09.2003	Praktikum Molekularbiologie (ganztägig)
04.2003 - 06.2003	Praktikum Datenbankprogrammierung (Oracle® SQL) (3h/Woche)
1992 - 2000	Gymnasium „Friedrich Schiller“ Bleicherode Abschluss: Abitur
1988 - 1992	Grundschule „Adolf Diesterweg“ Niedergebra

VERÖFFENTLICHUNGEN

PUBLIKATIONEN:

- D. Schulze, J. Kufka, W. Brandt, L. A. Wessjohann; Activation of diphosphates: Several answers to one question; in Vorbereitung
- W. Brandt, J. Kufka, R. Muntendam, O. Kayser, D. Schulze, E. Schulze, L. A. Wessjohann; Stabilization of cationic intermediates in prenyl transfer: A new functional role of methionine besides aromatic amino acids; in Vorbereitung
- L. Bräuer, W. Brandt, D. Schulze, S. Zakharova, L. A. Wessjohann: A structural model of the membrane-bound aromatic prenyltransferase UbiA from *E. coli*; ChemBioChem 9, 982-92 (2008)
- L. A. Wessjohann, S. Zakharova, D. Schulze, J. Kufka, R. Weber, L. Bräuer, W. Brandt: Enzymatic C-C-Coupling Prenylation: Bioinformatics – Modelling – Mechanism – Protein-Redesign – Biocatalytic Application; Chimia 63, 340-44 (2009)
- W. Brandt, L. Bräuer, N. Günnewich, J. Kufka, F. Rausch, D. Schulze, E. Schulze, R. Weber, S. Zakharova, L. A. Wessjohann; Molecular and structural reasons of metabolic diversity mediated by prenyl converting enzymes; Phytochemistry 70, 1758-75 (2009)

VORTRÄGE:

- Molecular Modelling Workshop, Erlangen (2008)
- Doktorandenworkshop "Naturstoffe: Chemie, Biologie und Ökologie", Bayreuth (2009)

POSTERBEITRÄGE:

- 3. German Conference on Cheminformatics, Goslar (2007)
- Molecular Modelling Workshop, Erlangen (2007)
- CCG European User Group Meeting, Berlin (2007)
- 4. German Conference on Cheminformatics, Goslar (2008)
- BioIT World Conference & Expo, Boston (2009)

EIDESSTATTLICHE ERKLÄRUNG

Hiermit erkläre ich, dass ich die vorliegende Arbeit selbstständig und ohne fremde Hilfe angefertigt habe und mich dabei keiner anderen als den von mir ausdrücklich gekennzeichneten Quellen und Hilfsmittel bedient habe.

Die Dissertation wurde nur der Naturwissenschaftlichen Fakultät II (Chemie, Physik und Mathematik) der Martin-Luther-Universität Halle-Wittenberg vorgelegt; eine Bewerbung um den Doktorgrad zu einem früheren Zeitpunkt ist nicht erfolgt.

Gera, im Dezember 2010