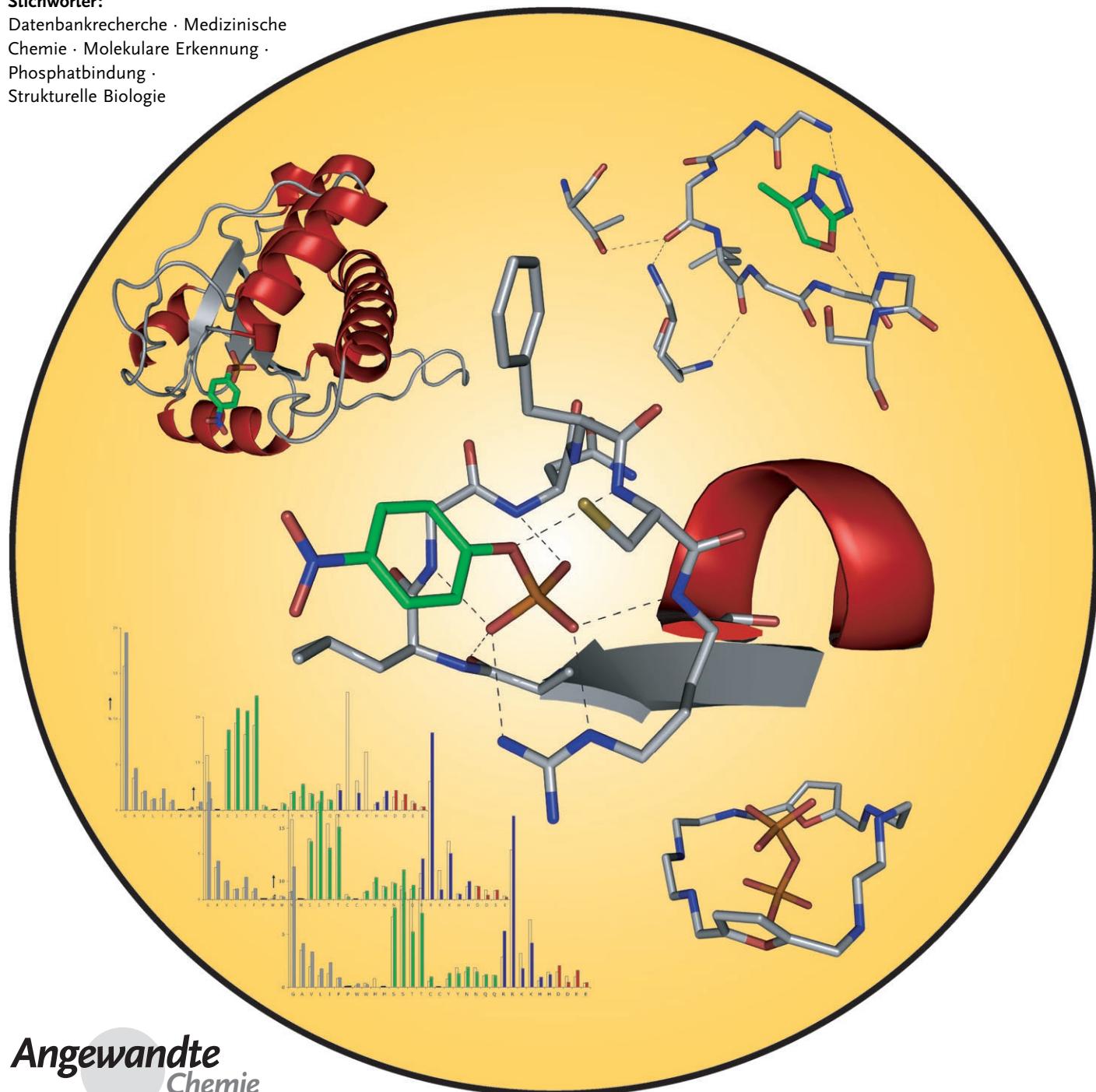


# Molekulare Erkennung von Phosphaten in der Strukturbioologie

Anna K. H. Hirsch, Felix R. Fischer und François Diederich\*

**Stichwörter:**

Datenbankrecherche · Medizinische Chemie · Molekulare Erkennung · Phosphatbindung · Strukturelle Biologie



Die medizinisch-chemische Forschung hat sich in den letzten Jahren Proteinen zugewandt, die wie Proteinkinasen und Proteinphosphatasen eine Phosphatbindungsstelle aufweisen. Dieser Aufsatz versucht, mithilfe von Datenbankrecherchen einen Überblick über die wichtigsten Prinzipien der molekularen Erkennung von Phosphatgruppen durch Enzyme zu geben. 3003 Kristallstrukturen von Proteinen mit gebundenem organischem Phosphat aus der RCSB-Proteindatenbank wurden im Hinblick auf H-Brücken zwischen Protein und Ligand untersucht. Die bekannten Bindungsmotive für Phosphate werden behandelt und Ähnlichkeiten bei der Komplexierung von Phosphaten durch synthetische Rezeptoren hervorgehoben. Eine Analyse der Häufigkeit, mit der spezifische Aminosäuren in unterschiedlichen phosphatbindenden Enzymklassen H-Brücken eingehen, führt zu einer für die Enzymklasse charakteristischen Verteilung. Der Aufsatz zeigt, dass bei strukturbasierter Ligandenentwicklung und -optimierung die Umgebung der Phosphatbindungsstelle sorgfältig mitberücksichtigt werden sollte und weist gleichzeitig auf Alternativen hin, wie diese Bindungstaschen zu füllen sind.

## 1. Einleitung

In den vergangenen Jahren haben wir einen Ansatz zur Entschlüsselung der molekularen Erkennungsprinzipien in chemischen und biologischen Systemen verfolgt, der Strukturuntersuchungen an Proteinen und synthetischen Rezeptoren, biologische und Wirt-Gast-Bindungsstudien, Datenbankrecherchen in der Cambridge Structural Database (CSD) und der PDB, Rechnungen sowie Gasphasenuntersuchungen berücksichtigt. Die Bedeutung dieses Ansatzes, der das strukturbasierte Ligandendesign und dessen Optimierung deutlich erleichtert, wurde schon früher mit der Analyse von aromatischen Wechselwirkungen,<sup>[1]</sup> orthogonalen multipolaren Wechselwirkungen<sup>[2]</sup> und Kation-π-Wechselwirkungen in aromatischen Kästen in den aktiven Taschen von Enzymen dokumentiert.<sup>[3]</sup> Bei diesen Untersuchungen haben wir eine beachtliche Kompetenz in der PDB-Datenbankrecherche erworben, von der die vorliegende Studie profitiert.

Kürzlich wurde unser Interesse an neuen Zielenzymen<sup>[4]</sup> in der Malariaforschung geweckt, und wir begannen mit der Entwicklung von Inhibitoren gegen IspF (2C-Methyl-D-erythritol-2,4-cyclodiphosphat-Synthase, *ybgB*), eines der sieben Enzyme des mevalonatfreien Biosyntheseweges.<sup>[5]</sup> Dieser Pfad wird von Malaria-Plasmoden und anderen Parasiten, nicht jedoch vom Menschen, zur Biosynthese von Isoprenoiden verwendet. Auf dem Weg von Pyruvat und D-Glyceraldehyd-3-phosphat zu den Isoprenoidvorstufen Isopentylidiphosphat und Dimethylallyldiphosphat setzen diese Enzyme kleine Mono- und Diphosphate um. Die Phosphatgruppen machen einen Großteil der molekularen Masse dieser eher hydrophilen Substrate aus. Bisher wurden nur auf Phosphaten oder Phosphonaten basierende Liganden mit bescheidenen bis mäßigen Affinitäten für einige der Enzyme (DXS, IspC, IspF) des mevalonatfreien Biosyntheseweges veröffentlicht.<sup>[6]</sup>

## Aus dem Inhalt

<b>1. Einleitung</b>	343
<b>2. Bekannte Phosphatbindungsmotive</b>	344
<b>3. Statistische Auswertung</b>	347
<b>4. Phosphatbindung von Proteinkinasen und Phosphatasen</b>	351
<b>5. Zusammenfassung und Schlussfolgerungen</b>	355

Diese Tatsache veranlasste uns, die molekulare Erkennung von Phosphaten in biologischen Systemen genauer zu studieren. Ein besseres Verständnis von Phosphatbindungsstellen sollte die Entwicklung neuartiger Phosphatanaloga für die medizinisch-chemische Forschung ermöglichen.

Tatsächlich sind Phosphate in der Biologie allgegenwärtig, und nahezu die Hälfte aller bekannten Proteine wechselwirkt mit Partnern, die einen derartigen Rest enthalten. Sie sind ein integraler Bestandteil von Erkennungsprozessen zwischen Proteinen, Nucleinsäuren, Cofaktoren und Antikörpern. Das Binden einer Phosphorylgruppe ist für eine Vielzahl von biologischen Prozessen wie Metabolismus und Biosynthese, Genregulation, Signaltransduktion, Muskelkontraktion oder Antikörperresistenz von essenzieller Bedeutung. Phosphatbindung verleiht Enzymen zusätzliche Stabilität wie im Fall der Asp-Aminotransferase<sup>[7]</sup> und leitet die Ablagerung von Eisen in Apoferritin ein.<sup>[8]</sup>

Protein-Phosphorylierung ist unerlässlich für die Regulation der transmembranalen und intrazellulären Signaltransduktion.<sup>[9,10]</sup> Proteinkinasen (PK)<sup>[11,12]</sup> katalysieren den Transfer einer endständigen Phosphatgruppe von Adenosintriphosphat (ATP) auf den Rest einer spezifischen Aminosäure wie Serin, Threonin oder Tyrosin; umgekehrt verläuft dieser Prozess bei den Proteinphosphatasen (PP),<sup>[13]</sup> die die Phosphatgruppe von einer Aminosäure abspalten. Vertreter

[\*] A. K. H. Hirsch,<sup>[4]</sup> F. R. Fischer,<sup>[4]</sup> Prof. Dr. F. Diederich  
Laboratorium für Organische Chemie  
ETH Zürich  
Hönggerberg, HCI, 8093 Zürich (Schweiz)  
Fax: (+41) 44-632-1109  
E-Mail: diederich@org.chem.ethz.ch

[†] A.K.H.H. und F.R.F. haben zu gleichen Teilen zu diesem Aufsatz beigetragen.

 Hintergrundinformationen zu diesem Beitrag sind im WWW unter <http://www.angewandte.de> zu finden oder können beim Autor angefordert werden.

beider Enzymklassen zählen zu den wichtigsten Zielproteinen im Kampf gegen Krebs und Fettleibigkeit. Während bei der Entwicklung niedermolekularer PK-Inhibitoren in vielen Fällen die Besetzung der stark polaren Bindungsstelle für das Triphosphat von ATP vermieden wird,<sup>[14,15]</sup> besetzen viele PP-Inhibitoren die Phosphatbindungstelle mit einer phosphatimitierenden Gruppe.<sup>[16,17]</sup> Die Einführung von Phosphatgruppen würde dem Liganden jedoch unvorteilhafte pharmakokinetische Eigenschaften wie Instabilität gegen Hydrolyse und geringe Membrangängigkeit verleihen. Zwar wurde bereits eine Vielzahl von phosphatimitierenden Strukturen, wie Vanadate, konjugierte Anionen von Säuren wie Carbonsäuren, Quadratsäure, Oxamsäure, Difluormethylensulfonsäuren, Difluormethylenphosphonsäuren und andere saure funktionelle Gruppen beschrieben, allerdings ist die Suche nach wirkungsvollen Phosphatanaloga mit den erwünschten Bindungsmotiven und physikochemischen Eigenschaften noch immer nicht abgeschlossen. Wir hoffen, dass dieser Aufsatz dazu beiträgt, diese Suche zu erleichtern.

Unseres Wissens gibt es bisher keine umfassende Auswertung der Phosphatbindung in Proteinen, die die Gesamtzahl der in der PDB hinterlegten Röntgenkristallstrukturen berücksichtigt.<sup>[18]</sup> Im vorliegenden Aufsatz analysieren wir die molekulare Erkennung von Phosphatgruppen unter Verwendung aller in der Datenbank vorhandenen Protein-Ligand-Komplexe. Da der Schwerpunkt auf Wechselwirkungen zwischen Proteinen und wirkstoffähnlichen Molekülen liegt, wurden DNA/RNA enthaltende Strukturen nicht berücksichtigt.

Die Übersicht beginnt mit einer Diskussion der bereits bekannten Phosphatbindungsmodi wie der P-Schleife (P-loop).<sup>[19–21]</sup> In einigen Fällen werden wir auf eine Analogie zwischen den biologischen Bindungsmotiven und denjenigen bei der Komplexierung von Phosphaten durch synthetische Rezeptoren hinweisen. Das Gebiet der molekularen Erkennung von Anionen durch synthetische Rezeptoren ist bereits in zahlreichen Übersichtsartikeln zusammengefasst worden.<sup>[22]</sup> In den folgenden Abschnitten präsentieren wir eine statistische Auswertung der Röntgenkristallstrukturen von Proteinen mit gebundenen phosphathaltigen Liganden. Dabei werden einige Punkte hervorgehoben, z.B. 1) die Art der Aminosäuren, die an der Erkennung beteiligt sind, 2) die Rolle von Wechselwirkungen mit den Seitenketten von Aminosäuren gegenüber Wechselwirkungen mit dem Prote-

inrückgrat, 3) der Einfluss von Metallionen, 4) die Rolle basischer Aminosäuren, 5) die Rolle spezifischer Bindungsmotive wie der P-Schleife sowie 6) charakteristische Bindungsmotive für ausgewählte Enzymklassen. Der letzte Teil des Aufsatzes ist der Bindung von Phosphaten durch PK und PP im Hinblick auf deren Bedeutung als Zielenzyme in der medizinischen Chemie gewidmet.

## 2. Bekannte Phosphatbindungsmotive

1974 identifizierten Rossmann et al. ein gemeinsames Proteinfaltungsmuster in Dinucleotid-bindenden Proteinen, das als Rossman-Faltung bekannt wurde und später auch in Mononucleotid-bindenden Proteinen gefunden wurde.<sup>[23]</sup> Charakteristisch dafür ist eine parallele  $\beta$ -Faltblattstruktur, deren Strände rechtsgängig durch  $\alpha$ -Helices verbunden sind. Die ständig wachsende Zahl an Protein-Ligand-Röntgenkristallstrukturen führte zur Identifizierung von charakteristischen Sequenz-Fingerabdrücken, die sich als nützlich für die Identifizierung der Funktion neuer Proteine erwiesen.

### 2.1. Gly-reiche Sequenzen

Ausgehend von der Entdeckung der Rossmann-Faltung konnten zwei konservierte Sequenzen identifiziert werden, die als Fingerabdrücke für Mono- und Dinucleotid-Bindung betrachtet werden können.<sup>[21,24]</sup> Sie werden als Gly-reiche Sequenzen bezeichnet, bei denen X für eine beliebige Aminosäure steht und alternative Seitenketten an einer spezifischen Position (wie S, T) in Klammern angegeben sind:<sup>[21,25]</sup>

- GXGXXG für die Bindung von Dinucleotiden
- GXXGXGK(S,T) oder GXXX für die Bindung von Monucleotiden.

### 2.2. Dinucleotid-bindende Proteine

Diese Proteine binden Nicotinamidadenindinucleotid (NAD), das entsprechende Phosphat (NADP) oder Flavinadenindinucleotid (FAD). Das Gly-reiche Fragment befindet sich in einem engen Turn zwischen einem  $\beta$ -Faltblatt und einer  $\alpha$ -Helix einer Rossmann-Faltung. Ausnahmslos werden



Anna K. H. Hirsch, geboren 1982 in Trier, studierte Chemie an der University of Cambridge und führte ihre Masterarbeit in der Gruppe von Prof. Steven V. Ley durch. Im Rahmen eines Auslandsaufenthaltes am Massachusetts Institute of Technology arbeitete sie in der Gruppe von Prof. Timothy F. Jamison. Sie ist seit 2004 Doktorandin in der Arbeitsgruppe von Prof. François Diederich an der ETH Zürich und beschäftigt sich mit dem Design und der Synthese von Inhibitoren von IspE.



Felix R. Fischer wurde 1980 in Deutschland geboren. Er studierte Chemie an der Ruprecht-Karls-Universität Heidelberg und erhielt sein Diplom 2004. Seine Diplomarbeit fertigte er unter der Leitung von Prof. Rolf Gleiter an. Seit 2005 ist er Doktorand in der Arbeitsgruppe von Prof. François Diederich an der ETH Zürich und widmet sich dem Design und der Synthese von Modellsystemen zur Messung biologisch relevanter multipolarer Wechselwirkungen.

dabei die Phosphatgruppen durch das positiv polarisierte N-terminale Ende eines Helixdipols stabilisiert.<sup>[26,27]</sup> Dabei sind die konservierten Gly-Reste gleich in mehrfacher Hinsicht von Bedeutung: Sie stellen ausreichend Platz für die Komplexierung des sperrigen Diphosphats zur Verfügung und ermöglichen gleichzeitig einen engen Turn. Eine der Ausnahmen für dieses sehr verbreitete Bindungsmuster sind Aladolreduktasen, die ein alternatives NADP-Bindungsmotiv verwenden.<sup>[28]</sup>

### 2.3. Mononucleotid-bindende Proteine

Ein Vergleich von 491 Mononucleotidbindungsstellen, die von Kinoshita et al. 1999 in der PDB gefunden wurden, führte zur Identifizierung einer in allen Strukturen konservierten Sequenz aus vier Aminosäureresten, GXXX, die als „structural P-loop“ (strukturelle P-Schleife) bezeichnet wurde.<sup>[29]</sup> Diese Sequenz schließt eine Reihe schon vorher beschriebener Motive wie die P-Schleife und die konservierte Sequenz GXGXXG in Proteinkinasen ein. 13 Proteinsuperfamilien lassen sich diesem Motiv zuordnen, während andere Sequenzen nur von gerade zwei Superfamilien geteilt werden und einige andere keine einzige konservierte Sequenz aufweisen.

### 2.4. P-Schleife

Ursprünglich von Walker et al. als Motiv A bezeichnet, wird die P-Schleife (P-loop) häufig in ATP- und GTP-bindenden Proteinen gefunden.<sup>[30]</sup> Zusätzlich kann eine weitere, weniger konservierte Sequenz, bekannt als Motiv B, vorhanden sein.

Die konservierte Sequenz für die P-Schleife ist GXXXXGK(S,T). Innerhalb einer Proteinfamilie lässt sich diese Sequenz noch weiter verfeinern, da bestimmte Merkmale in einer Familie oft geteilt werden. Wittinghofer et al. verglichen sieben ausgesuchte ATP- und GTP-bindende Proteinfamilien aus der Swissprot-Datenbank. Ein Beispiel für eine verfeinerte konservierte Sequenz ist GXPGXGKGT für Adenylatkinasen, bei denen zwischen dem konservierten Lys- und Thr-Rest ein Gly-Rest eingeführt ist.<sup>[20]</sup> Der konservierte Lys-Rest ist in allen Fällen anzutreffen, und es wird

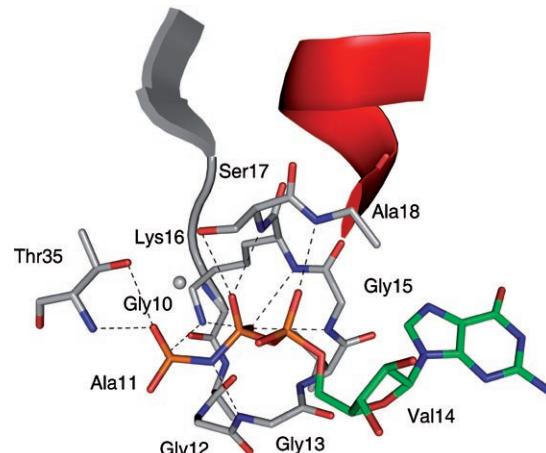


François Diederich, gebürtiger Luxemburger, studierte Chemie an der Universität Heidelberg (1971–1977) und promovierte 1979 bei Prof. Heinz A. Staab. Er forschte als Postdoktorand bei Prof. Orville L. Chapman an der UCLA (1979–1981) und habilitierte anschließend in Heidelberg am Max-Planck-Institut für medizinische Forschung (1981–1995). Anschließend wechselte er an die UCLA, wo 1989 Full Professor of Organic and Bioorganic Chemistry wurde. 1992 nahm er den Ruf an die ETH Zürich an. Er erforscht unter anderem dendritische Modellsysteme für globuläre Proteine, synthetische und biologische Rezeptoren und nichtpeptidische Enzymhemmer.

postuliert, dass er sowohl für die Konformation der P-Schleife als auch für die Stabilisierung des  $\beta$ - und  $\gamma$ -Phosphats von Bedeutung ist. Zusätzlich wird eine Beteiligung am Transfer des endständigen Phosphats vermutet. Die Überlagerung einer Reihe von P-Schleifen-haltigen Proteinen zeigte, dass die Positionen des  $\alpha$ - und  $\beta$ -Phosphats nahezu identisch sind.<sup>[28]</sup> Die beiden konservierten Gly-Reste nehmen Konformationen ein, die für andere Aminosäuren mit Seitenketten energetisch ungünstig wären.

Die im Unterschied zu Dinucleotid-Bindungsmotiven verhältnismäßig lange P-Schleife verbindet ein  $\beta$ -Faltblatt mit einer  $\alpha$ -Helix. Ebenso wie im Fall der Dinucleotid-bindenden Proteine findet man die Schleife am N-terminalen Ende einer  $\alpha$ -Helix. In manchen Fällen wird die P-Schleife auch als „gigantisches Anionenloch“ bezeichnet.<sup>[19]</sup> In erweitertem Sinn wurde sie kürzlich auch als „Nest“ beschrieben. Dieses wird durch eine Sequenz aus drei bis sechs Aminosäuren definiert, deren aufeinander folgende Amidfunktionen im Peptidrückgrat Anionen wie Phosphate oder Eisen-Schwefel-Cluster binden.<sup>[31]</sup>

Ein Beispiel für ein Protein, das eine P-Schleife enthält, ist p21, das Produkt des H-ras-Onkogens.<sup>[32]</sup> Pai et al. lösten die Röntgenkristallstruktur von p21 im Komplex mit dem langsam hydrolysierenden GTP-Analogon 5'-Guanyl- $\beta,\gamma$ -amido-triphosphat (GppNp) und einem Mg<sup>2+</sup>-Kation (Abbildung 1,



**Abbildung 1.** Ausschnitt aus der Röntgenkristallstruktur von p21, das das GTP-Analogon GppNp mittels einer P-Schleife bindet (PDB-Code: 5P21, 1.35 Å Auflösung).<sup>[33]</sup> Gestrichelte Linien entsprechen H-Brücken von weniger als 3.2 Å Länge (Abstand zwischen den Schweratomen). Ligandgerüst: grün, C: grau, O: rot, N: blau, P: orange. Die Auswahl der H-Brückenlänge und die Farbcodierung gelten für den gesamten Aufsatz, soweit nicht anders vermerkt.

PDB-Code: 5P21).<sup>[33]</sup> Die konservierte P-Schleife in der Phosphatbindungstasche erstreckt sich von Aminosäure 10 bis 18 mit der Sequenz GXXXXGKS. Die Seitenkette von Thr35 geht eine zusätzliche Wechselwirkung mit dem  $\gamma$ -Phosphat ein. Die konservierten Aminosäuren Gly10 und Gly15 liegen im Ramachandran-Plot in einer Konformation vor, die nur von Gly eingenommen werden kann. Die Phosphatgruppen sind von einem positiven elektrostatischen Feld umgeben, das von den Amidfunktionen im Peptidrückgrat

der Reste 13–18 erzeugt wird, die alle auf die Phosphatgruppen ausgerichtet sind und ionische Wasserstoffbrücken eingehen.

## 2.5. Die neue P-Schleife

Ein neues Nucleotid-bindendes Motiv wurde für die GHMP-Superfamilie identifiziert (GHMP=Galactosekinase, Homoserinkinase, Mevalonatkinese, Phosphomevalonatkinese). Es wurde „novel P-loop“ genannt und unterscheidet sich von der klassischen P-Schleife dadurch, dass es ADP/ATP in der unüblichen *syn*-Konformation bindet.<sup>[34]</sup> Nucleotide werden in der Regel in der *anti*-Konformation gebunden. Ein hochgradig konserviertes Motiv, ursprünglich Motiv 2 genannt, konnte als PXXXGLGSSAA in einer Schleife zwischen einem  $\beta$ -Faltblatt und einer  $\alpha$ -Helix identifiziert werden. Diese Schleife bildet eine riesige Anionenbindungsstelle, die sich wiederum am N-terminalen Ende einer  $\alpha$ -Helix befindet.

Die neue und die klassische P-Schleife haben einige Gemeinsamkeiten: Beide befinden sich zwischen einem  $\beta$ -Faltblatt und einer  $\alpha$ -Helix, und beide nutzen den stabilisierenden Effekt des Dipols der Helix und ionische H-Brücken zu Amidfunktionen im Peptidrückgrat zur Phosphatbindung. Die Struktur und die Sequenz unterscheiden sich jedoch: Die neue P-Schleife ist zwei Aminosäuren länger, und ihr fehlt der konservierte Lys/Arg-Rest. Es kann postuliert werden, dass die längere Schleife mehr ionische H-Brückenbindungen des Typs  $P-O^- \cdots H-N$  eingeht, die das Fehlen der positiv geladenen Lys/Arg-Seitenkette kompensieren.

## 2.6. Proteinkinasen

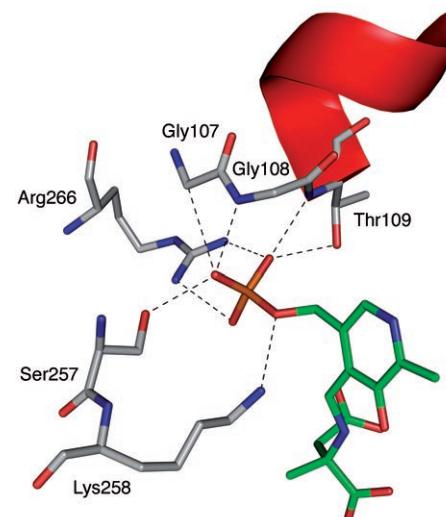
Eine Überlagerung der Aminosäuresequenz des katalytisch aktiven Bereichs von 65 Proteinkinasen führte zur Identifizierung einer konservierten Sequenz GXGXXG,<sup>[28,35]</sup> die identisch mit der von Dinucleotid-bindenden Proteinen ist. Strukturell gesehen haben Proteinkinasen jedoch Phosphatbindungsstellen, die mehr denen von Mononucleotid-bindenden Proteinen ähneln.<sup>[21]</sup> Es scheint klar, dass sowohl die Gly-reiche Anionenbindungsstelle als auch der vicinale Lys-Rest für den Phosphoryltransfer von essenzieller Bedeutung sind, da sie sich unabhängig voneinander für zwei verschiedene Proteinfaltungen – die klassische P-Schleife und die Proteinkinasenfaltung – entwickelt haben.

## 2.7. Das C<sup>a</sup>NN-Motiv

Denesyuk et al. fanden ein weiteres Anionenbindungs-motiv, ausgehend von ihrer ursprünglich in Pyridoxal-5'-phosphat(PLP)-bindenden Proteinen entdeckten, als „binding cup“ bezeichneten Bindungsstelle für Phosphatgruppen.<sup>[36]</sup> Eine Strukturanalyse aller repräsentativen Proteinfaltungen in der Families-of-Structurally-Similar-Proteins-(FSSP)-Datenbank<sup>[37]</sup> deckte ein Motiv auf, das bei 62 verschiedenen Faltungen gefunden wird. Es erkennt sowohl freie

Phosphat- und Sulfatanionen als auch Phosphatgruppen in Nucleotiden und Cofaktoren. Das Motiv enthält ein C<sup>a</sup>-Atom und zwei N-Atome aus dem Peptidrückgrat und findet sich üblicherweise in für die Funktion wichtigen Proteinregionen.

Der Komplex der Aspartyltransferase aus dem Cytosol von Schweinen mit PLP ist ein Beispiel für ein solches Bin-dungselement (Abbildung 2, PDB-Code: 1AJS).<sup>[38]</sup> Die



**Abbildung 2.** Ausschnitt aus der Röntgenkristallstruktur der cytosolischen Asp-Aminotransferase vom Schwein im Komplex mit Pyridoxal-5'-phosphat über das C<sup>a</sup>NN-Strukturmotiv (EC 2.6.1.1, PDB-Code: 1AJS, 1.60 Å Auflösung).<sup>[38]</sup>

Phosphateinheit von PLP wird durch Wechselwirkungen mit dem C<sup>a</sup>NN-Element an Ort und Stelle gehalten: Dabei handelt es sich um eine (sehr schwache) C–H $\cdots$ O–P-Brücke zu Gly107 (Abstand zwischen den Schweratomen 3.39 Å) und zwei starke (ionische) H-Brücken zu den N-Atomen von Gly108 und Thr109 aus dem Peptidrückgrat. Zusätzlich wird die Phosphatgruppe sowohl durch H-Brücken mit den Seitenketten von Thr109 und Ser257 als auch durch Ionenpaarung (begleitet von ionischen H-Brücken) mit Lys258 und Arg266 stabilisiert.

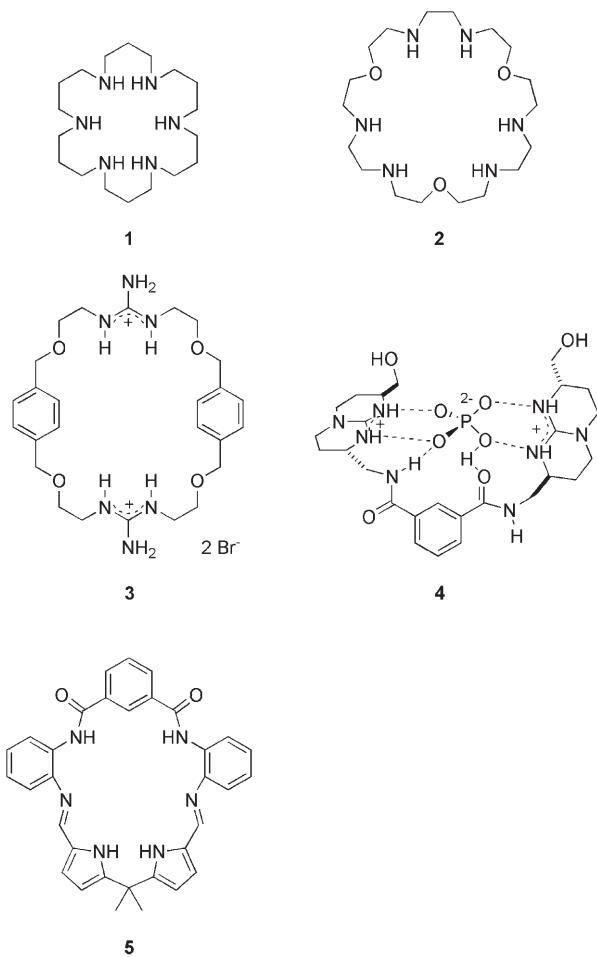
In etwa 90 % der Fälle konnte gezeigt werden, dass das Motiv phosphatspezifisch ist. Auch wenn sich dieses Motiv deutlich von anderen Anionenbindungsstellen unterscheidet, nutzt es doch auch Aminosäuren ohne oder nur mit kleinen Seitenketten, insbesondere Gly, und bildet H-Brücken mit dem Peptidrückgrat zur Phosphatbindung.<sup>[39]</sup>

Fasst man die Befunde aus den Abschnitten 2.1–2.7 zusammen, so ergibt sich, dass Nucleotid-bindende Proteine oft spezifische Faltungen und verschiedene charakteristische Struktureigenschaften aufweisen: Gly als Teil einer Schleife, benachbarte Lys-Seitenketten, die am Phosphoryltransfer beteiligt sind, und die Nähe eines positiv polarisierten N-Terminus einer  $\alpha$ -Helix. Die hier vorgestellten Proteinfaltungen und Fingerabdruck-Sequenzen sind allerdings nicht die einzigen Phosphatbindungs-motive in Proteinen. Actin, HSP70 (heat shock protein 70) und Zuckerkinasen – wie

Hexokinase – binden Phosphate mit Resten aus zwei  $\beta$ -Haarnadelstrukturen.

### 2.8. Einige Vergleiche mit synthetischen Phosphatrezeptoren<sup>[22]</sup>

Um Phosphate durch synthetische Rezeptoren in wässrigen Lösungen zu binden, bedarf es mehrerer Ladungswechselwirkungen, die von H-Brücken begleitet werden. Stabile Komplexe bilden sich mit einer Reihe vollständig protonierter makrocyclischer Polyamine wie **1**·6H<sup>+</sup> oder **2**·6H<sup>+</sup> (Abbildung 3), die durch Lehn et al. eingeführt wurden.<sup>[22a]</sup> Die



**Abbildung 3.** Synthetische Rezeptoren für Phosphatanionen.

Bindungsstärke steigt mit der Zahl der Ladungswechselwirkungen im Fall von **1**·6H<sup>+</sup> von Adenosinmonophosphat (AMP; lg  $K_{\text{ass}} = 3.4$  in 0.1M wässr. Me<sub>4</sub>NCl) über Adenosindiphosphat (ADP; 6.5) zu ATP (8.9). Die Selektivität hingegen ist relativ niedrig, und andere Anionen wie Oxalat (3.8), Sulfat (4.0) und Citrat (4.7) werden ebenso gebunden.<sup>[40a]</sup> Dabei falten sich die Makrocyclen um das sperrige Phosphatanion und bilden nestartige Komplexe mit optimierten H-Brücken.<sup>[40b,c]</sup>

Rezeptoren, die auf Guanidiniumionen basieren, sind intensiv untersucht worden. So komplexiert der Makrocyclus

**3** PO<sub>4</sub><sup>3-</sup> in H<sub>2</sub>O mit lg  $K_{\text{ass}} = 1.7$  und in MeOH/H<sub>2</sub>O 9:1 mit lg  $K_{\text{ass}} = 3.1$ .<sup>[41]</sup> Eine Reihe von klammerartigen Mono- und Bis-Guanidinium-Rezeptoren komplexiert sehr wirksam Oxoanionen in polaren Lösungsmitteln; bei Komplex **4**, der sich in Me<sub>2</sub>SO bildet, wird angenommen, dass zwei Guanidiniumgruppen HPO<sub>4</sub><sup>2-</sup> in einer tetraedrischen Geometrie koordinieren.<sup>[22c,42]</sup> Ein allgemeiner Vergleich legt nahe, dass die Fähigkeit eines einzelnen oder zweier primärer Ammoniumionen (RNH<sub>3</sub><sup>+</sup>), mit einem Phosphatanion in Wechselwirkung zu treten, deutlich geringer ist als die eines einzelnen Guanidiniumrestes.<sup>[22]</sup> Die Phosphaterkennung durch Guanidiniumionen profitiert vom Chelateffekt und der Bildung konvergierender ionischer H-Brücken mit begünstigenden sekundären elektrostatischen Wechselwirkungsmustern. Wir glauben, dass in Analogie dazu der Beitrag einer wechselwirkenden Arg-Seitenkette zur freien Bindungsenergie in Protein-Phosphat-Komplexen deutlich größer ist als der einer Lys-Seitenkette.

Auch in nur geringfügig oder nicht konkurrierenden Lösungsmitteln (in Bezug auf die Fähigkeit zur Bildung von H-Brücken) kann eine Komplexierung durch Amide und heterocyclische NH-Reste erreicht werden, die auf das gebundene Phosphation ausgerichtet sind. Dies konnte von Sessler et al. anhand von Makrocyclus **5** demonstriert werden, von dem angenommen wird, dass er sich in MeCN nestförmig um ein H<sub>2</sub>PO<sub>4</sub><sup>2-</sup>-Anion anordnet ( $K_{\text{ass}} = 340\,000 \text{ L mol}^{-1}$ ), das dabei von den Wechselwirkungen mit den beiden Amiden und den Pyrrol-NH-Gruppen profitiert.<sup>[43]</sup>

Sowohl die Anordnung neutraler NH-Reste um das Phosphation in **5** als auch das Umschließen des Oxoanions durch die protonierten -NH<sup>+</sup>-Reste der makrocyclischen Polyamine **1**·6H<sup>+</sup> und **2**·6H<sup>+</sup> führen zu Bindungsgeometrien, die starke Ähnlichkeit mit den in Abbildung 1 gezeigten P-Schleifen haben. Tatsächlich zeigt die dreidimensionale Darstellung der Phosphatbindung durch P-Schleifen bei einer Reihe von Proteinkomplexen eine erstaunliche, nahezu makrocyclische Anordnung der NH-Reste um das gebundene Phosphat. Die Schleife legt sich derart um den Phosphatrest, dass sich optimale H-Brücken bilden können. Mit anderen Worten, das Anion optimiert die Geometrie der Bindungsstelle (siehe dazu den Komplex eines Triphosphatanalogons mit dem Enzym IspE (PDB-Code: 1OJ4) aus dem mevalonatfreien Biosyntheseweg von Isoprenoiden, Abbildung 6, Abschnitt 3.3).<sup>[44]</sup>

Die Natur ist in der Lage, die räumliche Anordnung einer Bindungsstelle derart zu optimieren, dass sogar eine Unterscheidung aufgrund geringster Differenzen in Größe und Ladung – wie zwischen Sulfat und Phosphat – möglich ist.<sup>[45]</sup>

### 3. Statistische Auswertung

Anionenbindungsstellen wurden schon zuvor anhand von Informationen aus Röntgenkristallstrukturen von Proteinkomplexen mit Phosphaten und Sulfaten analysiert, was jedoch auf eine relativ kleine Zahl an Strukturen beschränkt war (< 70 Strukturen).<sup>[28,44]</sup> Außerdem wurde eine Reihe von Analysen einzelner Enzymklassen oder spezifischer Nucleotide wie ATP durchgeführt.<sup>[46]</sup> Unseres Wissens gibt es aller-

dings keine umfassende Auswertung der Phosphatbindung in Proteinen, die auf der Gesamtheit der Röntgenkristallstrukturen in der PDB basiert.

Für ein besseres Verständnis der Art und Weise, wie Phosphatgruppen in Proteinen gebunden werden, haben wir eine umfassende PDB-Suche durchgeführt.<sup>[18]</sup> Unsere Analyse basiert auf den in der Datenbank verfügbaren Atomkoordinaten. Wenn mehr als eine Röntgenkristallstruktur verfügbar war, wurde diejenige mit der besseren Auflösung berücksichtigt. Im Fall von oligomeren Proteinen wurde die Analyse auf eine Untereinheit beschränkt, da homologe Untereinheiten im Allgemeinen den gleichen Bindungsmodus aufweisen.

Die Wechselwirkungen von Phosphatgruppen in Proteinen über kurze Entfernung wurden mit dem Programm Relibase ausgewertet.<sup>[47,48]</sup> Dabei wurden folgende Suchparameter definiert (Abbildung 4): Die Suche wurde auf an ein C-Atom gebundene  $\alpha$ -Phosphate beschränkt, um Röntgenkristallstrukturen mit freien Phosphaten auszuschließen. Wegen der verwendeten Kristallisationsbedingungen sind freie Phosphate häufig an Stellen cokristallisiert, die nicht der aktiven Tasche entsprechen. Weil die freien Valenzen am O-Atom des  $\alpha$ -Phosphats undefiniert sind, wurden auch Strukturen mit Di-, Tri- und Pentaphosphaten berücksichtigt. Weiterhin beschränkten wir den Abstand zwischen dem O-Atom des Phosphats und einem H-Atom des Proteins auf 1.75–3.00 Å. Wir wählten diese relativ unkonventionelle Beschreibung einer H-Brücke, um alle Arten von H-Brücken-Donoren (HO, HN, HS) innerhalb eines Suchvorgangs zu berücksichtigen. Zuletzt untersuchten wir für die Auswertung alle möglichen H-Brücken bis zu einem maximalen Abstand zwischen den Schweratomen von 3.20 Å.

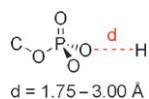


Abbildung 4. Parameter für die Relibase-Suche.

### 3.1. Die Strukturen des gesamten Datensatzes („Alle“)

Am 14. Februar 2006 waren 35 144 Strukturen in der PDB hinterlegt. Die Aminosäurenverteilung für den gesamten Datensatz ist in Abbildung 1SI in den Hintergrundinformationen gezeigt.<sup>[49]</sup> Eine Gesamtzahl von 14 590 Einträgen wies die gesuchten Strukturmerkmale auf. Von diesen entsprachen 3003 der Abstandseinschränkung. Sie bilden den gesamten Datensatz „Alle“ und zeigen eine Gesamtzahl von 19 713 H-Brücken zu 5520 Phosphatgruppen, was einem Durchschnitt von 3.6 H-Brücken pro Phosphatgruppe entspricht. Alle 3003 Strukturen wurden einzeln visualisiert und untersucht. Das Säulendiagramm in Abbildung 2SI zeigt die Prozentzahl der jeweiligen Aminosäuren, die an der Phosphaterkennung beteiligt sind. Man beachte, dass Aminosäuren, die Seitenketten mit H-Brücken-Donoren oder -Acceptoren haben, doppelt aufgeführt sind, wobei der erste Prozentsatz die Beteiligung der Rückgrat-NH-Gruppe und der zweite die Beteiligung der

Seitenkette beschreibt. Der Vergleich der Aminosäurenverteilung in der gesamten PDB mit der Verteilung derjenigen Aminosäuren, die an der Phosphatbindung teilnehmen, zeigt wie erwartet große Unterschiede auf.

### 3.2. Ein erster Datensatz: Weglassen der identifizierten Metallionen („Alle-M“)

Da anionische Phosphatgruppen Ionenbindungen mit Metallionen eingehen können und diese Coulomb-Wechselwirkungen H-Brückeneffekte sicherlich überlagern würden, beschlossen wir, Strukturen auszuschließen, in denen Metallionen an das Phosphat binden. Diese Einschränkung ließ immer noch 2456 Einträge übrig (von 3003): Metallionen kommen also weitaus seltener in der Phosphaterkennung vor, als man gemeinhin annimmt. Dieser erste Datensatz wird „Alle-M“ genannt. Der Vergleich der an den H-Brücken zu Phosphaten beteiligten Aminosäuren aller Strukturen („Alle“) mit dem Datensatz „Alle-M“ (Abbildung 2SI) deckt ein nahezu identisches Aminosäremuster auf. Bemerkenswert ist der große Anteil der Gly-Reste, der polaren Reste Ser und Thr und erwartungsgemäß auch der basischen Reste Lys und Arg. Das häufige Vorkommen von Gly-Resten stimmt mit der Beobachtung überein, dass Gly-reiche Schleifen wichtig für Phosphatbindungs motive sind (siehe Abschnitt 2): Gly-Reste ermöglichen dank ihrer Konformation das Falten und Wickeln der Schleife um das gebundene Phosphat. Mehr als die Hälfte der Einträge des Datensatzes haben eine an der Phosphatbindung beteiligte Lys- oder Arg-Seitenkette. Der Gesamtanteil an Aminosäuren mit unpolaren (Summe von 25%) und basischen (Summe von 28%) Seitenketten ist recht ähnlich. Überraschend ist die relativ niedrige Zahl an Tyr-Resten (verglichen mit Ser- und Thr-Resten), die an H-Brücken beteiligt ist, da die OH-Gruppe von Tyr sauer ist und deshalb der bessere H-Brücken-Donor sein sollte.<sup>[50]</sup> Wahrscheinlich führen sterische Faktoren zu einem Vorzug für Ser und Thr gegenüber Tyr.

Angesichts der recht kleinen Zahl an Metallionen-komplexierenden Phosphatgruppen und der äußerst ähnlichen Aminosäureverteilung in der An- und Abwesenheit solcher Ionen beschlossen wir, den Datensatz „Alle-M“ für alle künftigen Vergleiche zu verwenden.

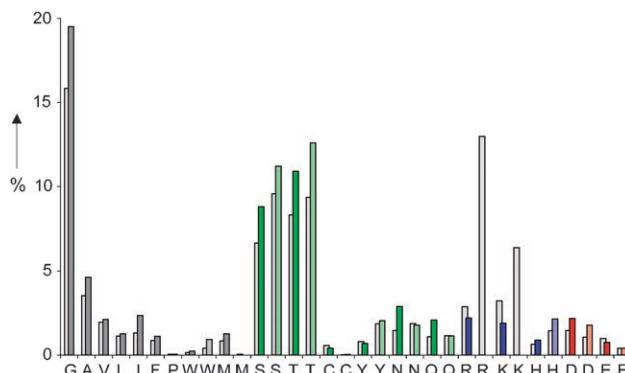
### 3.3. Ein zweiter Datensatz: Phosphatbindungsstellen ohne Ionenbindung („Alle-M-Lys/Arg“)

Aus medizinisch-chemischer Perspektive war es von besonderem Interesse, zu erforschen, in welchem Ausmaß Phosphate auch an neutralen Erkennungsstellen gebunden werden, ohne Mitwirkung von Ionenbindungen mit Metallionen und/oder protonierten basischen Aminosäureseitenketten. In solchen Phosphatbindungsstellen sollten H-Brücken die wichtigste Wechselwirkung sein und könnten von Leitstrukturresten mit ausgeprägten, mehrfachen H-Brücken-Acceptoren belegt werden.

Der Ausschluss der Röntgenkristallstrukturen aus dem ersten Datensatz, die an der Phosphatbindung beteiligte Arg/

Lys-Seitenketten aufweisen, führte zu der überraschend hohen Zahl von 1070 Strukturen, die insgesamt 5303 H-Brücken zu 1668 Phosphaten aufweisen, was einem Durchschnitt von 3.2 H-Brücken pro Phosphatgruppe entspricht. Fast ein Drittel aller Phosphatbindungsstellen verwendet keine Ionenpaarbindungen!

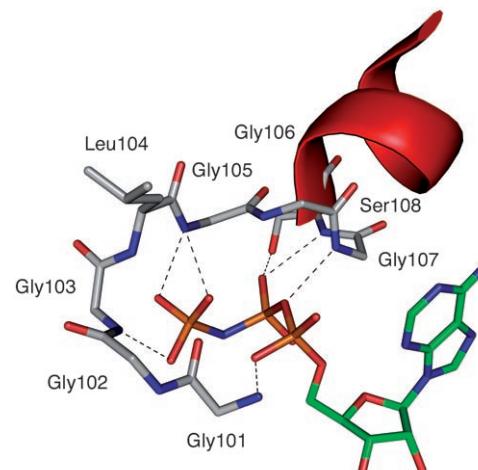
Ein Vergleich der beiden Datensätze „Alle-M“ und „Alle-M-Lys/Arg“ (Abbildung 5) zeigt, dass der Mangel an basischen Resten durch eine Zunahme polarer Reste wie Ser,



**Abbildung 5.** Säulendiagramm des Vergleichs der an den H-Brücken zu Phosphatgruppen beteiligten Aminosäurereste (Einbuchstabencode) des ersten Datensatzes „Alle-M“ (hintere hellgraue Säulen) mit dem zweiten Datensatz „Alle-M-Lys/Arg“ (farbige vordere Säulen). Aminosäuren mit Seitenketten, die H-Brücken bilden können, sind doppelt aufgeführt, wobei sich die erste Säule auf Rückgrat-NH- und die zweite, schraffierte Säule auf Seitenketten-Wechselwirkungen bezieht. Die Aminosäuren sind in Klassen unterteilt und haben folgende Farbcodierung: grau: unpolär, grün: polar, blau: basisch, rot: sauer. Diese Farbcodierung gilt für den gesamten Aufsatz, soweit nicht anders vermerkt.

Thr, His (das protoniert sein kann) und Asn, aber auch unpolärer Reste wie Gly und auch Ala und Ile kompensiert wird. Dieser Trend scheint in Einklang mit der Beschreibung der neuen P-Schleife zu sein (Abschnitt 2.5). In diesem Zusammenhang wird postuliert, dass das Fehlen eines konservierten Lys-Restes durch eine längere Schleife mit mehr konservierten Resten kompensiert wird. Außer den Aminosäureseitenketten tragen Rückgrat-NH-Gruppen über H-Brücken und über das durch sie induzierte positiv geladene elektrostatische Feld zur Phosphatbindung bei.

Ein repräsentatives Beispiel eines Proteins des Datensatzes „Alle-M-Lys/Arg“ ist der ternäre Komplex der 4-Diphosphocytidyl-2C-methyl-D-erythritol-Kinase (IspE), ihres Substrates und des nicht hydrolysierbaren ATP-Analogs 5'-Adenyl-β,γ-amidotriphosphat (AppNp) in Abbildung 6 (PDB-Code: 1OJ4).<sup>[51]</sup> IspE ist eines der sieben Enzyme des mevalonatfreien Biosynthesewegs für die Synthese der Isoprenoid-Vorstufen Isopentenyldiphosphat und Dimethylallyldiphosphat, der in den meisten Bakterien und einigen Parasiten – nicht jedoch im Menschen – vorkommt. IspE zeigt eine für Kinasen charakteristische Zwei-Domänen-Faltung einer Substrat- und einer ATP-Bindungsdomäne. In diesem Komplex ist die ATP-Bindungsstelle durch eine lange, Gly-reiche Schleife gekennzeichnet, die nur einen an



**Abbildung 6.** Ausschnitt aus der Röntgenkristallstruktur des Enzyms IspE, das das hydrolyseresistente ATP-Analogon AppNp ohne Verwendung eines Metallions oder einer positiv geladenen Seitenkette bindet (EC 2.7.1.148, PDB-Code: 1OJ4, 2.01 Å Auflösung).<sup>[51]</sup>

der Phosphatbindung beteiligten polaren Ser-Rest, aber eine große Zahl an Gly-NH-Gruppen im Rückgrat enthält (Reste 101–107), die auf das gebundene Anion zeigen. Ein weiteres Beispiel für eine Struktur des zweiten Datensatzes findet sich in den Hintergrundinformationen (Abbildung 3SI).<sup>[52]</sup>

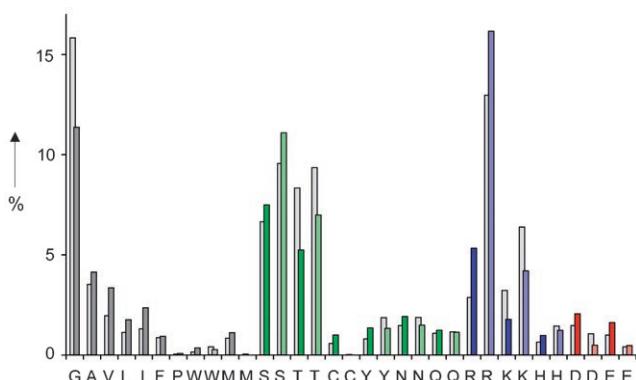
### 3.4. Datensatz „Alle-M“ in verschiedenen Enzymklassen

Der Datensatz „Alle-M“ wurde in verschiedene Enzymklassen aufgeteilt, unter Verwendung der in der PDB verwendeten Beschreibung (Tabelle 1). Nur die vier umfangreichsten Klassen sowie statistisch signifikante Änderungen werden besprochen.

**Tabelle 1:** Unterteilung des ersten Datensatzes „Alle-M“ in Enzymklassen.

Enzymklasse	Zahl	Verteilung [%]
Oxidoreduktasen	787	32
andere	520	21
Transferasen	430	18
Lyasen	245	10
Hydrolasen	149	6
Isomerasen	119	5
Signaltransduktion	77	3
Elektronentransport	73	3

Aus den Säulendiagrammen, die die zwei umfangreichsten Gruppen (Oxidoreduktasen und Transferasen) im Datensatz „Alle-M“ vergleichen, wird klar ersichtlich, dass jede Enzymklasse eine charakteristische Aminosäurenverteilung hat (Abbildung 7 und Abbildung 4SI). Konzentriert man sich auf die bedeutendste Gruppe, die Oxidoreduktasen, so fällt ein recht geringer Anteil an Gly-Resten auf. Andererseits nimmt der relative Anteil an unpolaren Resten wie Ala, Val, Leu, Ile und Met zu. Dieser Zuwachs an unpolaren Resten

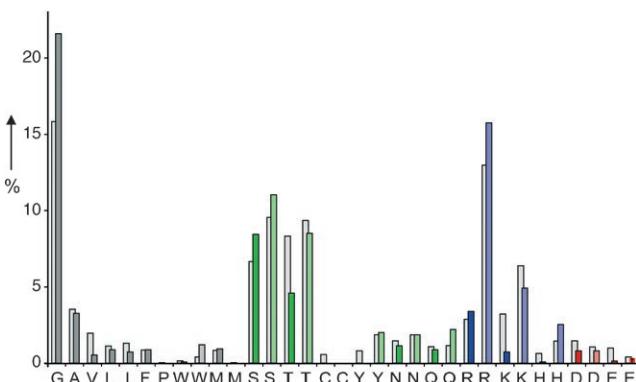


**Abbildung 7.** Säulendiagramm des Vergleichs der an den H-Brücken zu Phosphatgruppen beteiligten Aminosäurereste (Einbuchstabencode) des ersten Datensatzes „Alle-M“ (hintere hellgraue Säulen) mit dessen Oxidoreduktasen (farbige vordere Säulen). Aminosäuren mit Seitenketten, die H-Brücken bilden können, sind doppelt aufgeführt, wobei sich die erste Säule auf Rückgrat-NH- und die zweite schraffierte Säule auf Seitenketten-Wechselwirkungen bezieht.

könnte essenziell zur Feineinstellung der Umgebungspolarität für effiziente Elektronentransferprozesse sein. Eine Abnahme an Thr-Resten wird durch einen Zuwachs an Ser-Resten kompensiert. Gleichermaßen kompensiert ein Zuwachs an Arg-Resten eine geringere Zahl an Lys-Resten.

Bei der zweithäufigsten Klasse, den Transferasen, findet sich beim Vergleich mit dem gesamten Datensatz keine signifikante Veränderung der unpolaren Reste. Ein umgekehrtes Verhalten zeigt sich hingegen im Fall der polaren Aminosäuren: Transferasen enthalten eine erhöhte Zahl an Thr- und eine verminderte Zahl an Ser-Resten. Transferasen und Oxidoreduktasen haben fast die gleiche Verteilung basischer Reste.

Die zwei nächsthäufigsten Klassen, die Lyasen und die Isomerasen, werden mit dem gesamten Datensatz „Alle-M“ in den Abbildungen 8 und 5SI verglichen. Die Lyasen zeigen einen Zuwachs in der Zahl an Gly-Resten, der mit einer Abnahme des Anteils der restlichen unpolaren Aminosäuren



**Abbildung 8.** Säulendiagramm des Vergleichs der an den H-Brücken zu Phosphatgruppen beteiligten Aminosäurereste (Einbuchstabencode) des ersten Datensatzes „Alle-M“ (hintere hellgraue Säulen) mit seinen Lyasen (farbige vordere Säulen). Aminosäuren mit Seitenketten, die H-Brücken bilden können, sind doppelt aufgeführt, wobei sich die erste Säule auf Rückgrat-NH- und die zweite schraffierte Säule auf Seitenketten-Wechselwirkungen bezieht.

einherzugehen scheint. Wiederum zeigen die polaren Reste Ser und Thr entgegengesetzte Trends, und Gleches trifft auf die Arg- und Lys-Reste zu.

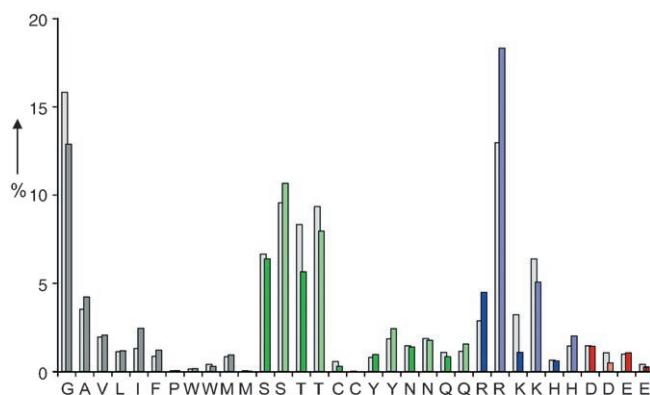
In der Klasse der Isomerasen zeigt die unpolare Untergruppe ein etwas anderes Verhalten. Der Zuwachs an Leu- und Ile-Resten gleicht eine reduzierte Zahl an Ala- und Val-Resten aus. Die Zahl an Ser- und Thr-Resten – normalerweise sehr wichtige Phosphatbindungsreste – ist verringert, die Menge der Tyr- und Asn- sowie Gln-Reste ungewöhnlicherweise erhöht. Wiederum könnte dies auf eine ersetzende Rolle schließen lassen. Die basischen Reste zeigen wieder ein typisches Verhalten, mit einem Anstieg der Zahl an Lys-Resten bei gleichzeitiger Abnahme der Arg-Reste.

Aus dieser Analyse scheint ein allgemeiner Trend vorzugehen: Der Anteil der verschiedenen, an der Phosphatbindung beteiligten Unterklassen von Aminosäureresten (unpolär, polar, basisch) ist mehr oder weniger konstant. Offenbar hat allerdings jede Enzymklasse eine klare Präferenz für die Natur der Aminosäuren in jeder Untergruppe, die sie zur Phosphatbindung verwendet. Zum Beispiel ist es häufig der Fall, dass sich eine verminderte Zahl an Thr-Resten in einer erhöhten Zahl an Ser-Resten widerspiegelt. Ähnliche Trends finden sich für die unpolaren und basischen Reste. Auf die Untergruppe der sauren Aminosäurereste scheint dies nicht zuzutreffen, allerdings ist die Zahl der an der Phosphatbindung beteiligten sauren Aminosäuren recht klein, was die Identifizierung statistisch signifikanter Veränderungen erschwert. Als Resultat dieses Verhaltens ergeben sich sehr charakteristische Verteilungen der für die Phosphatbindung verwendeten Aminosäuren, die als Fingerabdrücke für die Identifizierung einer bestimmten Enzymklasse genutzt werden können.

### 3.5. Ein dritter Datensatz: Abwesenheit von Metallionen und Schleife („Alle-M-loop“)

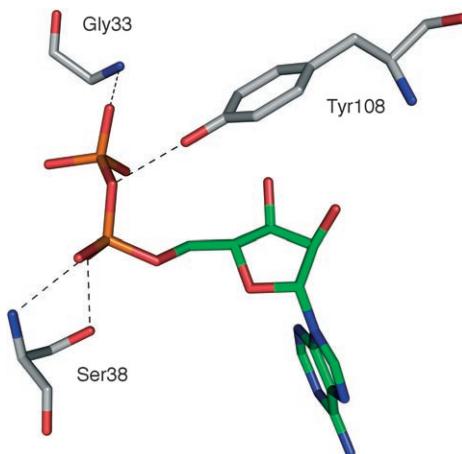
Da eine Schleife ein häufiges Strukturelement für Phosphatbindung ist, war der nächste logische Schritt die Untersuchung, wie Phosphate in Abwesenheit sowohl eines Metallions als auch einer Schleife gebunden werden. Zu diesem Zweck haben wir eine Schleife als Folge von mindestens drei aufeinander folgenden Aminosäuren definiert, die an H-Brücken zu Phosphaten beteiligt sind. Fälle, bei denen der erste und dritte, aber nicht der zweite Rest eine H-Brücke bilden, sind inbegriﬀen. Der Abzug der betroffenen Einträge führte zu einem dritten Datensatz: „Alle-M-loop“. Er enthält eine Gesamtzahl von 1675 Einträgen mit 8428 H-Brücken zu 2740 Phosphatgruppen, was einem Durchschnitt von 3.1 H-Brücken pro Phosphatgruppe entspricht.

Abbildung 9 zeigt einen Vergleich des ersten Datensatzes („Alle-M“) mit dem dritten Datensatz („Alle-M-loop“). Eine Abnahme der Zahl an Gly-Resten ist offensichtlich, während sich gleichzeitig die Zahl an Ala- und Ile-Resten erhöht. Der Zuwachs an Ser- und Tyr-Resten könnte eine geringere Zahl an Thr-Resten kompensieren. Schließlich könnte in der basischen Untergruppe ein Anstieg in der Beteiligung von His- und Arg-Resten der reduzierten Zahl von Lys-Resten entgegenwirken.



**Abbildung 9.** Säulendiagramm des Vergleichs der an den H-Brücken zu Phosphatgruppen beteiligten Aminosäurereste (Einbuchstabencode) des ersten Datensatzes „Alle-M“ (hintere hellgraue Säulen) mit dem dritten Datensatz „Alle-M-loop“ (farbige vordere Säulen). Aminosäuren mit Seitenketten, die H-Brücken bilden können, sind doppelt aufgeführt, wobei sich die erste Säule auf Rückgrat-NH- und die zweite schraffierte Säule auf Seitenketten-Wechselwirkungen bezieht.

Ein veranschaulichendes Beispiel dieses dritten Datensatzes ist der Komplex der Riboflavin-Kinase von *S. pombe* mit einem der Produkte, ADP (Abbildung 10, PDB-Code: 1N07).<sup>[53]</sup> Die Phosphatgruppen von ADP werden von drei isolierten Resten gebunden, einer Tyr-Seitenkette, einem Gly-Rückgrat und einem Ser-Rest, der sowohl über eine Amidgruppe des Rückgrates als auch über seine Seitenkette wechselwirkt. Selbst wenn der Gly- und der Ser-Rest Teil



**Abbildung 10.** Ausschnitt aus der Röntgenkristallstruktur der Riboflavin-Kinase, die ADP ohne die Hilfe eines Metallions oder einer Schleife bindet (EC 2.7.1.26, PDB-Code: 1N07, 2.45 Å Auflösung).<sup>[53]</sup>

einer strukturellen Schleife sind, entspricht sie doch nicht unserer Definition einer Phosphatbindungsstelle, da zu viele Reste dazwischen liegen, die nicht zur Phosphatbindung beitragen.

Die Resultate dieser statistischen Analyse sind in Tabelle 2 zusammengefasst. Es gibt kaum Unterschiede zwischen dem Datensatz „Alle“ und dem ersten Datensatz („Alle-M“), was nochmals unsere Entscheidung bestätigt, ausschließlich letzteren Datensatz zu verwenden. Der zweite Datensatz zeigt eine klare Abnahme der basischen Reste, da nur noch His-Reste und Rückgrat-Amide von Lys und Arg vorhanden sind. Dies findet seine Entsprechung in einem starken Zuwachs an polaren und zu einem gewissen Ausmaß auch an unpolaren Resten. Wie erwartet, bleiben die sauren Reste mehr oder weniger unverändert.

Der dritte Datensatz zeigt eine geringfügige Zunahme an basischen Resten, was insofern plausibel ist, als die basischen Reste die Abwesenheit einer Schleife kompensieren könnten. Im dritten Datensatz findet sich folglich offenbar ein ähnliches Verhalten wie bei den verschiedenen Enzymklassen, d.h., es gibt eine gegenseitige Kompensation von Aminosäureresten einer Gruppe (unpolar, polar, basisch).

#### 4. Phosphatbindung von Proteinkinasen und Phosphatasen

Schon in den frühen fünfziger Jahren wurde die Entdeckung gemacht, dass Enzymaktivität durch Phosphorylierung und Dephosphorylierung reguliert werden kann.<sup>[10,54]</sup> Die Proteinkinasen und Proteinphosphatasen haben nicht nur entgegengesetzte Funktionen, sondern sie wirken auch zusammen, um Zellwachstum und -differenzierung zu regulieren.<sup>[55]</sup> Eine Störung dieses Gleichgewichtes führt fast unweigerlich zu einer Krankheit. Folglich sind sowohl die Proteinkinasen als auch die Proteinphosphatasen wichtige Zielverbindungen in der medizinischen Chemie. Im letzten Teil dieses Aufsatzes konzentrieren wir uns daher auf die Mechanismen der Phosphatbindung durch Proteinkinasen und Proteinphosphatasen, mit dem Ziel, das Verständnis ihrer Erkennungsstellen zu verbessern, was der Entwicklung von Leitstrukturen zugute kommen sollte.

##### 4.1. Proteinkinasen

Die Proteinkinasen, eine Untergruppe der Kinasen,<sup>[56]</sup> phosphorylieren OH-Gruppen eines Proteinsubstrats. Prote-

**Tabelle 2:** Zusammenfassung der statistischen Auswertung der verschiedenen Datensätze.

Datensatz	H-Brücken pro Phosphat	H-Brücken [%] <sup>[a]</sup>				Schleife	Einträge [%] <sup>[b]</sup> Arg-, Lys-Seitenkette
		unpolar	polar	basisch	sauer		
„Alle“	3.6	5023 (25)	8418 (43)	5533 (28)	717 (4)	1198 (40)	1733 (58)
„Alle-M“	3.6	4107 (26)	6734 (43)	4344 (27)	614 (4)	943 (38)	1434 (58)
„Alle-M-Arg/Lys“	3.2	1751 (33)	2858 (54)	372 (7)	267 (5)	344 (32)	–
„Alle-M-loop“	3.1	2141 (25)	3356 (40)	2659 (32)	272 (3)	–	908 (54)

[a] Zahl an H-Brücken, die von jeder Klasse von Aminosäureresten gebildet werden; in Klammern steht der prozentuale Anteil, den diese Untergruppe darstellt. [b] Zahl an Einträgen; in Klammern steht der prozentuale Anteil, den dies darstellt.

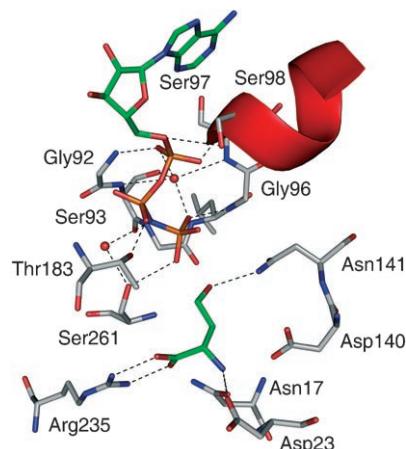
inserinthreoninkinasen (PSTKs) sind die häufigsten, gefolgt von den Proteintyrosinkinasen (PTKs) und zuletzt den Doppelspezifitätinkinasen (DSKs), die alle drei Aminosäuren phosphorylieren können. Andere Klassifizierungen nutzen die Information über die spezifischen Protein-Ligand-Wechselwirkungen in der ATP-Bindungstasche oder Sequenzalignments, um die Proteinkinasen in Unterfamilien einzuteilen.<sup>[11,57]</sup>

Dank einer Fülle sowohl struktureller als auch biochemischer Studien gehören die Kinasen wohl zu den am besten untersuchten Enzymen. Ihre Diversität und Substratspezifität sind bemerkenswert angesichts Tatsache, dass sie im Prinzip alle die gleiche Reaktion katalysieren, nämlich den Transfer der  $\gamma$ -Phosphorylgruppe von ATP auf ein Substrat. Einige gemeinsame Strukturmerkmale der Kinasen wurden identifiziert und ausführlich besprochen.<sup>[58]</sup> Dazu zählt das Vorhandensein eines oder manchmal auch zweier Metallkationen ( $Mg^{2+}$  oder  $Mn^{2+}$ ), ein Nucleotidbindungsmotiv, z.B. eine P-Schleife oder eine Gly-reiche Sequenz, ein positiv geladener Rest (normalerweise Lys) und der positive N-Terminus einer  $\alpha$ -Helix. Der Mechanismus des Kinase-katalysierten Phosphoryltransfers wurde intensiv erforscht.<sup>[59]</sup> Etwa 1.7 % des menschlichen Genoms dienen der Expression der Proteinkinasen.<sup>[12]</sup>

Die Suche nach effizienten und selektiven Inhibitoren für Proteinkinasen als den vielleicht wichtigsten Medikamentenzielen des 21. Jahrhunderts<sup>[60]</sup> ist eine der größten gegenwärtigen Aufgaben der Pharmaindustrie.<sup>[10,13]</sup> Ein Beispiel eines erfolgreich entwickelten niedermolekularen Wirkstoffes ist Gleevec, das die Abl(Abelson)-PTK des Fusionsproteins Bcr-Abl inhibiert und gegen Knochenmarksleukämie<sup>[61]</sup> und Magen-Darm-Tumoren eingesetzt wird.<sup>[62]</sup> Außerdem wurde Sorafenib für die Behandlung des metastasierenden Nierenzellenkarzinoms entwickelt.<sup>[63]</sup> Ein letztes Beispiel für eine erfolgreiche Entwicklung ist der monoklonale Antikörper Herceptin für die Behandlung von metastasierendem Brustkrebs, der an die Her2/neu-Rezeptor-PTK bindet.<sup>[64]</sup>

#### 4.2. Beispiele von Phosphatbindungsstellen in Proteinkinasen

Die Proteinkinasen verwenden eine große Vielfalt an Phosphatbindungsmotiven, wie die folgenden Beispiele verdeutlichen. Die Homoserinkinase (HSK) ist ein Mitglied der GHMP-Kinase-Superfamilie und katalysiert den ersten selektiven Schritt in der Biosynthese von Thr, die Bildung von *O*-Phospho-L-homoserin aus L-Homoserin und ATP. Die Röntgenkristallstrukturen einer Reihe ternärer Komplexe von HSK mit Homoserin oder Thr (einem Feedback-Inhibitor) und dem ATP-Analogon AppNp wurden mit Auflösungen zwischen 1.8 und 2.0 Å gelöst (Abbildung 11, PDB-Code: 1H72).<sup>[34,65]</sup> Sowohl Homoserin als auch AppNp werden in einer 13 Å tiefen Tasche gebunden, die von einer  $\alpha$ -Helix und drei relativ flexiblen Schleifenstrukturen gebildet wird. Das Homoserin befindet sich am unteren Ende der Kavität und bildet unter anderem eine starke Salzbrücke zu Arg235 (2.71 und 2.65 Å). Im oberen Teil der Tasche bindet AppNp am N-Terminus der  $\alpha$ -Helix. Das Triphosphat nimmt eine *sc,ac,sc,sp*-Konformation ein (beginnend auf der Ribose-

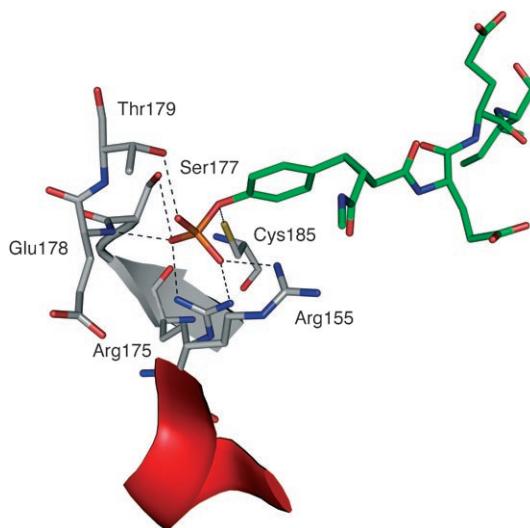


**Abbildung 11.** Ausschnitt aus der Röntgenkristallstruktur der Homoserinkinase im Komplex mit dem ATP-Analogon AppNp und Homoserin (EC 2.7.1.39, PDB-Code: 1H72, 1.80 Å Auflösung).<sup>[65]</sup> Kristall-H<sub>2</sub>O: rote Kugeln.

Seite; *sc* = synclinal, *ac* = anticlinal, *sp* = synperiplanar) und wird durch fünf direkte und drei H<sub>2</sub>O-vermittelte H-Brücken gebunden. Das  $\alpha$ -Phosphat bildet chelatisierende H-Brücken zur Rückgrat-NH-Gruppe von Ser98 (2.47 Å) und zur OH-Gruppe von dessen Seitenkette (3.02 Å). Gly 92 geht eine H-Brücke mit einem der  $\alpha$ -Phosphat-O-Atome ein (2.57 Å), während ein stark koordiniertes H<sub>2</sub>O-Moleköl sowohl ein  $\alpha$ - (2.91 Å) als auch das  $\gamma$ -Phosphat-O-Atom (2.43 Å) bindet. Das  $\gamma$ -Phosphat bildet eine H-Brücke mit Gly 96 (2.55 Å), während die O-Atome des  $\beta$ -Phosphats lediglich eine starke H-Brücke mit der OH-Gruppe der Seitenkette von Thr183 (2.02 Å) eingehen. Eine bemerkenswerte Eigenschaft dieser Bindungstasche ist die Abwesenheit kationischer Aminosäurereste in der Triphosphatbindungstasche. Dieses Beispiel zeigt auch, wie die visuelle Auswertung aller Strukturen dabei half, den wichtigen Beitrag zur Phosphatbindung durch (kristallographisch lokalisierte) H<sub>2</sub>O-Moleküle zu identifizieren. Ferner scheint es recht häufig vorzukommen, dass das  $\beta$ -Phosphat die schwächsten Wechselwirkungen in einem Triphosphatkomplex eingeht.

Ein anderes Beispiel, das einige dieser Merkmale verdeutlicht, z.B. die Abwesenheit basischer Aminosäuren in direkter Nähe zum gebundenen Phosphat, liefert die Röntgenkristallstruktur des Komplexes der Pyruvatdehydrogenase-Kinase PDK3 mit L2 (innere Lipoyldomäne), Mg<sup>2+</sup> und ATP (Abbildung 6SI).<sup>[66]</sup>

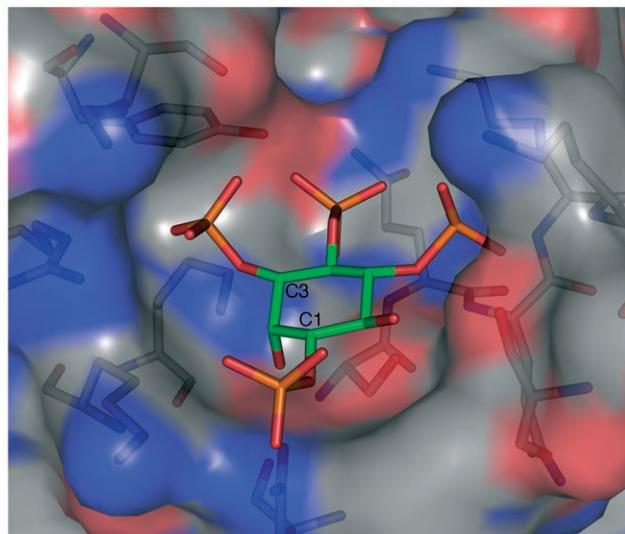
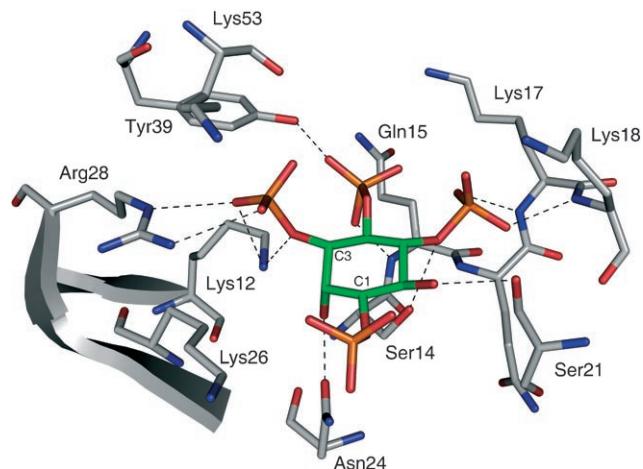
Die Src-Familie der PTKs sowie eine Reihe anderer Proteine, die an der intrazellulären Signaltransduktion beteiligt sind, enthalten alle eine hochkonservierte SH2-Domäne. Sie ist für die spezifische Erkennung Phosphotyrosin-haltiger Motive an aktivierten Zelloberflächenrezeptoren verantwortlich; folglich ist die SH2-Domäne entscheidend für die Signaltransduktion. Abbildung 12 zeigt die Röntgenkristallstruktur eines Komplexes der Src-Homologiedomäne der Src-Kinase mit einem Isoster des *O*-phosphorylierten YEEI-Tetrapeptids (PDB-Code: 1ISO).<sup>[67]</sup> Um die Topographie der Bindungstasche zu testen, wurde zur Versteifung der Peptidkette ein Cyclopropanrest eingeführt. Der Ligand bindet



**Abbildung 12.** Ausschnitt aus der Röntgenkristallstruktur der Src-Kinase im Komplex mit einem O-phosphorylierten YEEI-Tetrapeptid (EC 2.7.1.112, PDB-Code: 1IS0, 1.90 Å Auflösung).<sup>[67]</sup>

an der Oberfläche des Proteins, wobei nur das *O*-Phosphotyrosinfragment in eine flache Kavität reicht, die vom N-terminalen Ende einer  $\alpha$ -Helix und einer Schleife, die zwei  $\beta$ -Stränge miteinander verbindet, gebildet wird. Das Phosphat ist durch starke Salzbrücken mit den Seitenketten von Arg175 (2.69 und 2.78 Å) und Arg155 (2.90 Å) gebunden. Weitere H-Brücken werden mit der Rückgrat-NH-Gruppe von Glu178 (2.56 Å) und der OH-Gruppe der Seitenketten von Thr179 (2.69 Å) gebildet. Eine sechste, recht schwache Wechselwirkung kann man zwischen der SH-Gruppe von Cys185 (3.21 Å) und dem phosphorylierten Tyr-O-Atom ausmachen. Die Bevorzugung kationischer oder polarer Aminosäurereste sowie die flache Tasche an der Oberfläche stehen im Gegensatz zum vorher besprochenen Beispiel der HSK.

Brutons Tyrosinkinase ist ein wichtiges Protein für die Reifung von B-Zellen. Im Menschen führt eine einfache Punktmutation im Enzym zu X-verknüpfter Agammaglobulinämie, einer schweren Immunschwäche-krankheit. Das Protein enthält eine Domäne, die spezifisch Phosphatidyl-inositol-3,4,5-triphosphat bindet. In der Röntgenkristallstruktur des mit einer Auflösung von 2.4 Å gelösten dimeren Komplexes (Abbildung 13, PDB-Code: 1B55)<sup>[68]</sup> ist der Phosphatidylinositol-3,4,5-triphosphat-Ligand in einer flachen Tasche an der Enzymoberfläche gebunden. Man kann Wechselwirkungen mit den Aminosäuren einer von Arg28 bis Tyr39 reichenden Schleife beobachten, die zwei  $\beta$ -Stränge miteinander verbindet. Die freien OH-Gruppen (an C2 und C6) des Liganden bilden H-Brücken mit den Seitenketten von Ser21 (2.81 Å) und Asn24 (2.32 Å). Die C1-Phosphorylgruppe geht keine direkten H-Brückenkontakte mit dem Protein ein, profitiert aber vom positiv polarisierten Umfeld, das durch die Nähe der Seitenkette von Lys26 induziert wird. Das Phosphat an C3 hingegen ist über Salzbrücken fest an Arg28 (3.08 und 3.18 Å) und Lys12 (2.67 und 2.96 Å) gebunden. Eine seltene H-Brücke findet sich zwischen der OH-



**Abbildung 13.** Oben: Ausschnitt aus der Röntgenkristallstruktur von Brutons Tyrosinkinase im Komplex mit Phosphatidylinositol-3,4,5-triphosphat (EC 2.7.1.112, PDB-Code: 1B55, 2.40 Å Auflösung).<sup>[68]</sup> Unten: Elektrostatisches Potential, das die elektropositive Umgebung von C1–C5 des Liganden am Eingang zur schalenförmigen Bindungstasche zeigt.

Gruppe von Tyr39 und der Phosphatgruppe an C4 (2.38 Å), des Weiteren gibt es eine Wechselwirkung mit der Rückgrat-NH-Gruppe von Gln15 (2.39 Å). Die Rückgrat-NH-Gruppen von Lys127 (2.73 Å) und Lys18 (2.81 Å) bilden H-Brücken zum C5-Phosphat sowie eine H-Brücke mit der Seitenketten-OH-Gruppe von Ser14 (3.04 Å). Insgesamt kann man beobachten, dass die Proteinoberfläche am Boden der schalenförmigen Bindungsdomäne ein recht negatives elektrostatisches Potential hat (Ser14, Ser21, Asn24), während der Rand der Schale von positiv geladenen Aminosäureresten gebildet wird (Lys12, Lys17, Lys18, Lys26, Arg28, Lys53). Der Rand ist jedoch nicht vollständig positiv geladen und weist eine Lücke an der unphosphorylierten C6-Position auf. Folglich passt das Oberflächenpotential des Proteins komplementär zur Ladungsverteilung des Phosphatidylinositol-3,4,5-triphosphat-Liganden, was zur Selektivität gegenüber unterschiedlich phosphorylierten Derivaten führt.

## 4.3. Andere Phosphatbindungsstellen in Kinasen

In den Hintergrundinformationen werden einige weitere Beispiele der Phosphaterkennung durch Kinasen besprochen, um die strukturelle Vielfalt noch besser zu illustrieren:

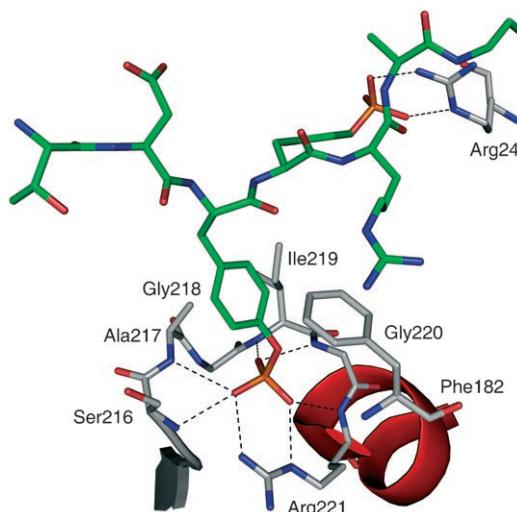
- Die Röntgenkristallstruktur eines ternären Komplexes der Adenylylkinase von Hefe, Bis(adenyl)-5'-pentaphosphat ( $\text{Ap}_5\text{A}$ ) und einem  $\text{Mg}^{2+}$ -Ion (Abbildung 7SI, PDB-Code: 2AKY)<sup>[69]</sup> zeigt, dass das Pentaphosphat in einem gigantischen Anionenloch mit der Konsensussequenz GXXGXK gebunden ist.<sup>[21]</sup> Adenylylkinasen sind allgegenwärtige Enzyme, die den Transfer einer Phosphorylgruppe von ATP auf AMP katalysieren, wobei zwei Moleküle ADP entstehen. Dieser Prozess ist  $\text{Mg}^{2+}$ -abhängig.
- Im Komplex der Glycerokinase mit ADP, Glycerol-3-phosphat und einem  $\text{Mn}^{2+}$ -Ion (Abbildung 8SI, PDB-Code: 1GLD)<sup>[70]</sup> sind beide Phosphate in einer 15 Å tiefen Tasche an der Grenzfläche zwischen den N- und C-terminalen Domänen an das  $\text{Mn}^{2+}$ -Ion gebunden.
- Im Komplex der Riboflavin kinase mit ADP, Flavinmononucleotid (FMN) und einem  $\text{Mg}^{2+}$ -Ion (Abbildung 9SI, PDB-Code: 1P4M)<sup>[71]</sup> reichen die beiden Phosphate in eine tiefe Kavität, um an das Metallion zu koordinieren. Die Riboflavin kinase ist ein Enzym, das die Phosphorylierung von Riboflavin (Vitamin  $\text{B}_2$ ) zu FMN katalysiert.
- Ein schönes Beispiel für die Bindung von ADP in einer Tasche, die aus einer Schleife besteht, die einen  $\beta$ -Strang und eine  $\alpha$ -Helix miteinander verknüpft, findet sich in der Röntgenkristallstruktur des Komplexes der menschlichen Desoxycytidinkinase mit ADP, einem  $\text{Mg}^{2+}$ -Ion und der Medikamentenvorstufe („Prodrug“) Gemcitabine (Abbildung 10SI, PDB-Code: 1P62).<sup>[72]</sup> Die Desoxycytidinkinase katalysiert die Phosphorylierung natürlicher Desoxyribonucleoside, wie Desoxycytidin, Desoxyguanosin und Desoxyadenosin, sowie zahlreicher synthetischer Nucleosidanaloga, die als Medikamentenvorstufen in der antiviralen und Krebschemotherapie eingesetzt werden.

## 4.4. Proteinphosphatasen

Die Proteinphosphatasen werden nach ihrer Substrat- und Strukturspezifität in Proteinserinthreoninphosphatasen, Proteintyrosinphosphatasen und Doppelspezifitätphosphatasen eingeteilt. Die letzten zwei Klassen sind miteinander verwandt und zeigen hohe Sequenzhomologie. Der Mechanismus der Dephosphorylierung, der katalytische Asp-, Arg- und His-Reste einbezieht, ist weitgehend aufgeklärt.<sup>[73,74]</sup> Die Proteinphosphatasen sind erst seit kurzem interessante Ziele in der Medikamentenentwicklung.<sup>[15,17,64,75]</sup> Zwar werden auch andere Phosphatasen erforscht,<sup>[76]</sup> den Proteinphosphatasen gilt aber das größte Interesse als potenzielle Wirkstoff-Targets.

## 4.5. Beispiele für Phosphatbindungsstellen in Proteinphosphatasen

Die Proteintyrosinphosphatase PTP1B ist für die Dephosphorylierung der Phosphotyrosinreste des Insulinrezeptors IRK verantwortlich, wodurch sie den Insulin-Signaltransduktionsweg negativ reguliert. Die Struktur von PTP1B im Komplex mit einem substrathähnlichen, diphosphorylierten Modellpeptid wurde mit einer Auflösung von 2.4 Å bestimmt (Abbildung 14, PDB-Code: 1G1H).<sup>[77]</sup> Einer der *O*-phos-



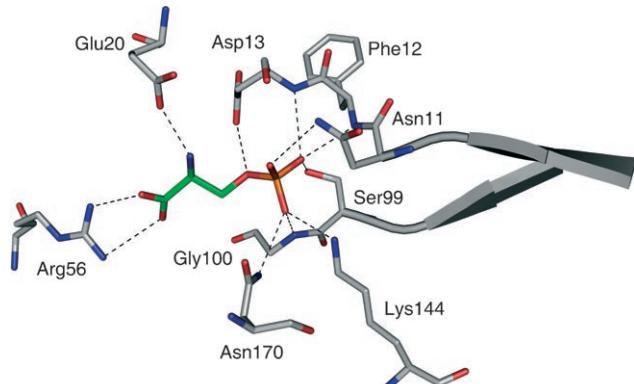
**Abbildung 14.** Ausschnitt aus der Röntgenkristallstruktur der Proteintyrosinphosphatase PTP1B im Komplex mit einem diphosphorylierten Modellpeptid (EC 3.1.3.48, PDB-Code: 1G1H, 2.40 Å Auflösung).<sup>[77]</sup>

phorylierten Tyr-Reste ist in einer 7 Å tiefen Tasche gebunden, in der er durch das Dipolmoment des N-terminalen Endes einer  $\alpha$ -Helix stabilisiert wird. Die Schleife, die diese Helix mit einem  $\beta$ -Faltblatt verbindet, bildet eine starke, einem Polyazamakrocyclus ähnelnde Bindungsstelle für die Phosphatgruppe, die sechs H-Brücken zu den Rückgrat-NH-Gruppen von Ser216 (3.00 Å), Ala217 (3.26 Å), Ile219 (2.96 Å), Gly220 (2.89 Å) und Arg221 (2.99 Å) sowie eine Salzbrücke zur Guanidiniumgruppe der Arg221-Seitenkette (2.92 und 3.04 Å) eingeht. Die Tasche ist durch den aromatischen Ring von Phe182, der eine  $\pi$ -Stapelwechselwirkung mit der aromatischen Seitenkette des phosphorylierten Tyr-Restes eingeht, von Wechselwirkungen mit dem Lösungsmittel abgeschirmt. Das zweite *O*-phosphorylierte Tyr in dem kurzen Peptid ist nur an der Proteinoberfläche durch die Seitenkette von Arg24 (3.20 und 3.23 Å) gebunden.

Die Röntgenkristallstruktur einer Proteintyrosinphosphatase ist als Komplex mit *p*-Nitrophenylphosphat im Bild des Inhaltsverzeichnisses gezeigt (EC 3.1.3.48, PDB-Code: 1D1Q, 1.70 Å Auflösung).<sup>[78]</sup>

Die Phosphoserinphosphatase (PSP) gehört zu einer großen Enzymklasse, die die Phosphoesterhydrolyse über ein Phosphoaspartat-Zwischenprodukt katalysiert. PSP ist wohl an der Regulation der Steady-State-Konzentration von L-Serin im Gehirn beteiligt. Die Röntgenkristallstruktur des binären Komplexes der PSP mit *O*-phosphoryliertem L-Serin wurde mit einer Auflösung von 1.9 Å gelöst (Abbildung 15, PDB-Code: 1L7P).<sup>[79]</sup> Nach der Bindung faltet sich das Protein vollständig um das Substrat herum, sodass dieses wirksam vom Lösungsmittel abgeschirmt ist. Die Phosphatgruppe bildet H-Brücken mit drei Rückgrat-NH-Gruppen von Phe12

(2.81 Å), Asp13 (3.09 Å) und Gly100 (2.84 Å). Außerdem



**Abbildung 15.** Ausschnitt aus der Röntgenkristallstruktur der Phosphoserinphosphatasen im Komplex mit O-phosphoryliertem L-Serin (EC 3.1.3.3, PDB-Code: 1L7P, 1.90 Å Auflösung).<sup>[79]</sup>

gehen einige Seitenketten von Asp13 (2.82 Å), Asn11 (3.25 Å), Ser99 (2.59 Å), Lys144 (2.68 Å) und Asn170 (3.09 Å) H-Brücken mit den Phosphat-O-Atomen ein. Die Carboxylatgruppe von L-Serin ist durch eine Salzbrücke mit Arg56 (2.92 und 3.15 Å) verbunden, während die Ammoniumgruppe durch die Seitenkette von Glu20 (2.50 Å) stabilisiert wird.

## 5. Zusammenfassung und Schlussfolgerungen

Die molekulare Erkennung von Phosphaten in den aktiven Zentren von Enzymen wurde bisher nicht umfassend beschrieben, obwohl großes Interesse an der Entwicklung von Wirkstoffen besteht, die an Enzyme mit phosphathaltigen Substraten, vor allem an Proteinkinasen und Proteinphosphatasen, binden. Wir geben hier einen entsprechenden Überblick, wobei wir uns auf die wichtigste Wechselwirkung zwischen Phosphat und Rezeptor, nämlich die H-Brücke, konzentrieren und dabei von der reichhaltigen Strukturinformation der PDB profitieren. Zuerst fassen wir die bekannten Phosphatbindungsmotive zusammen, z.B. die Gly-reiche Schleife und die P-Schleife, und streichen die Analogie mit der Komplexierung durch Oligoazamakrocyclen hervor: Begünstigt durch die konformative Flexibilität der Gly-Reste in diesen Schleifen organisieren die Phosphat-Gastmoleküle den Rezeptor, indem sich die Schleife um das Anion wickelt und H-Brücken mit den konvergierenden Rückgrat-NH-Resten gebildet werden. Dies erinnert an das Prinzip der Komplexierung von Ionen durch flexible Rezeptoren: Auch in diesem Fall organisiert der Gast seinen Wirt.

Die darauf folgende statistische Auswertung liefert eine Reihe unerwarteter Resultate: Unter den 3003 berücksichtigten Strukturen, die alle einzeln analysiert wurden, zeigt die bemerkenswert hohe Zahl von 2456 Einträgen eine Phosphationenbindung ohne Mitwirkung eines Metallions. Noch bemerkenswerter ist die Tatsache, dass ein Drittel des Datensatzes „Alle“ – nämlich 1070 Strukturen – eine Phosphationenkomplexierung ohne Beteiligung eines Metallions oder Gegenwart basischer (protonierter) Seitenketten von Arg-

oder Lys-Resten innerhalb des entscheidenden Abstandes für H-Brücken (< 3.2 Å) zeigt. Die Analyse der Aminosäurenverteilung in verschiedenen Klassen phosphatbindender Enzyme (Oxidoreduktasen, Transferasen, Lyasen und Isomerasen) liefert sehr charakteristische Verteilungen der an der Phosphatbindung beteiligten Aminosäuren, die man als Fingerabdrücke der verschiedenen Klassen betrachten kann.

Der Aufsatz endet mit Beispielen von Phosphatbindung durch Proteinkinasen und Proteinphosphatasen, die zu den interessantesten Targets der gegenwärtigen Wirkstoffentwicklung gehören. Während viele Proteinkinase-Inhibitoren, die an die ATP-Bindungsstelle binden, die Triphosphat-Bindungsstelle vermeiden, binden die meisten Proteinphosphatase-Inhibitoren an die Monophosphatbindungsstelle über einen sauren Rest, der nach Deprotonierung das anionische Phosphat imitiert. Da mehr als ein Drittel aller Phosphat-bindungsstellen keine Metallionen oder basische (protonierte) Arg- oder Lys-Reste innerhalb des typischen Abstandes für H-Brücken enthält, scheint es, dass auch niedermolekulare Wirkstoffe mit neutralen Substituenten in Betracht gezogen werden sollten, um diese „neutralen“ Bindungtaschen zumindest teilweise zu füllen. Diese befinden sich häufig tief im Innern des Proteins. Kleine heteroalicyclische oder heteroaromatische wirkstoffähnliche Reste, die ausgedehnte H-Brücken-Acceptoreinheiten in der Peripherie aufweisen, sollten besonders gut geeignet sein, um mit den konvergierenden H-Brückendonorien an der Phosphat-Erkennungsstelle zu wechselwirken. Man kann davon ausgehen, dass die Gly-reichen Schleifen, die häufig die Bindungtaschen dominieren, ausreichend flexibel sind, um sich um solche Reste zu winden. Wir testen gegenwärtig in einem strukturbasierten Ansatz diese Theorie der Bindung an die Gly-reiche Schleife der ATP-Bindungsstelle von IspE mit kleinen H-Brücken-akzeptierenden Heterocyclen (Abbildung 6). Beispiele solcher Heterocyclen, die in die Phosphatbindungsstelle von IspE modelliert sind, werden in den Hintergrundinformationen geliefert (Abbildung 11SI und Abbildung 12SI). Selbst wenn eine Lys-Seitenkette an der Phosphatbindung beteiligt ist, könnten neutrale Liganden geeignet sein, da sich Lys-Seitenketten häufig in eine andere Position drehen können. Nur wenn Arg-Seitenketten in die Phosphatbindungsstelle zeigen und ihre Ladung nicht durch einen nahe liegenden Asp- oder Glu-Rest kompensiert wird, kann es schwierig sein, die Phosphatbindungsstelle mit neutralen Resten zu füllen. Man kann generell sagen, dass die Natur der Phosphatbindungsstelle stark von ihrer Lage abhängt: An der Oberfläche dominieren eher kationische Reste, während tief im Innern neutrale Aminosäuren überwiegen.

Dieser Aufsatz verdeutlicht den Wert einer detaillierten, alle intermolekularen Wechselwirkungen zwischen individuellen Atomen berücksichtigenden Untersuchung der Phosphatbindungsstelle für die strukturbasierte Entwicklung und Optimierung von Leitstrukturen; nur eine derartige, umfassende Analyse ermöglicht es, alle Möglichkeiten für einen innovativen Phosphatersatz zu nutzen.

In den Hintergrundinformationen finden sich Säulendiagramme mit der Verteilung der an der molekularen Erkennung von Phosphaten beteiligten Aminosäuren, Vergleiche der Gesamtzahl der Strukturen mit ausgewählten Unter-

gruppen und einzelnen Enzymklassen sowie ausgewählte Beispiele für Phosphatbindungsstellen aus der RCSB-Proteinbank.

Diese Arbeit wurde von der ETH-Forschungskommission, Hoffmann-La Roche AG, Basel, und Chugai Pharmaceuticals unterstützt. Wir danken Jörg Klein, Christian Kramer und Fabian Weibel für ihren wertvollen Beitrag zur PDB-Suche. Wichtige Anregungen für diesen Aufsatz kamen von Diskussionen bei Roche und Chugai, die dankend anerkannt werden. Wir danken Dr. W. Bernd Schweizer (ETH Zürich) sowie der Relibase-Gruppe in Cambridge für Hilfe mit der Relibase-Suche.

Eingegangen am 21. August 2006

- [1] E. A. Meyer, R. K. Castellano, F. Diederich, *Angew. Chem. 2003*, **115**, 1244–1287; *Angew. Chem. Int. Ed.* **2003**, **42**, 1210–1250.
- [2] R. Paulini, K. Müller, F. Diederich, *Angew. Chem. 2005*, **117**, 1820–1839; *Angew. Chem. Int. Ed.* **2005**, **44**, 1788–1805.
- [3] a) K. Schärer, M. Morgenthaler, R. Paulini, U. Obst-Sander, D. W. Banner, D. Schlatter, J. Benz, M. Stihle, F. Diederich, *Angew. Chem. 2005*, **117**, 4474–4479; *Angew. Chem. Int. Ed.* **2005**, **44**, 4400–4404; b) J. C. Ma, D. A. Dougherty, *Chem. Rev. 1997*, **97**, 1303–1324.
- [4] C. M. Crane, J. Kaiser, N. L. Ramsden, S. Lauw, F. Rohdich, W. Eisenreich, W. N. Hunter, A. Bacher, F. Diederich, *Angew. Chem. 2006*, **118**, 1082–1087; *Angew. Chem. Int. Ed.* **2006**, **45**, 1069–1074.
- [5] a) F. Rohdich, S. Hecht, A. Bacher, W. Eisenreich, *Pure Appl. Chem. 2003*, **75**, 393–405; b) M. Rohmer, M. Knani, P. Simonin, B. Sutter, H. Sahm, *Biochem. J. 1993*, **295**, 517–524; c) M. K. Schwarz, Dissertation, ETH Zürich, Nr. 10951, **1994**; d) S. T. J. Broers, Dissertation, ETH Zürich, Nr. 10978, **1994**.
- [6] J. Wiesner, R. Ortmann, H. Jomaa, M. Schlitzer, *Angew. Chem. 2003*, **115**, 5432–5451; *Angew. Chem. Int. Ed.* **2003**, **42**, 5274–5293, zit. Lit.
- [7] J. H. Martinez-Liarte, A. Iriarte, M. Martinez-Carrion, *Biochemistry 1992*, **31**, 2712–2719.
- [8] Y. G. Cheng, N. D. Chasteen, *Biochemistry 1991*, **30**, 2947–2953.
- [9] T. Hunter, *Cell 1995*, **80**, 225–236.
- [10] T. Hunter, *Cell 2000*, **100**, 113–127.
- [11] S. K. Hanks, T. Hunter, *FASEB J. 1995*, **9**, 576–596.
- [12] G. Manning, D. B. Whyte, R. Martinez, T. Hunter, S. Sudarsanam, *Science 2002*, **298**, 1912–1934.
- [13] Z.-Y. Zhang, *Annu. Rev. Pharmacol. Toxicol. 2002*, **42**, 209–234.
- [14] a) A. J. Bridges, *Chem. Rev. 2001*, **101**, 2541–2571; b) M. E. M. Noble, J. A. Endicott, L. N. Johnson, *Science 2004*, **303**, 1800–1805.
- [15] K. Grosios, P. Traxler, *Drugs Future 2003*, **28**, 679–697.
- [16] a) R. H. van Huijsduijnen, A. Bombrun, D. Swinnen, *Drug Discovery Today 2002*, **7**, 1013–1019; b) R. E. Honkanen, T. Golden, *Curr. Med. Chem. 2002*, **9**, 2055–2075; c) G. Liu, *Curr. Med. Chem. 2003*, **10**, 1407–1421.
- [17] L. Bialy, H. Waldmann, *Angew. Chem. 2005*, **117**, 3880–3906; *Angew. Chem. Int. Ed.* **2005**, **44**, 3814–3839.
- [18] H. M. Berman, J. Westbrook, Z. Feng, G. Gilliland, T. N. Bhat, H. Weissig, I. N. Shindyalov, P. E. Bourne, *Nucleic Acids Res. 2000*, **28**, 235–242.
- [19] D. Dreusicke, G. E. Schulz, *FEBS Lett. 1986*, **208**, 301–304.
- [20] M. Saraste, P. R. Sibbald, A. Wittinghofer, *Trends Biochem. Sci. 1990*, **15**, 430–434.
- [21] G. E. Schulz, *Curr. Opin. Struct. Biol. 1992*, **2**, 61–67.
- [22] Übersichtsartikel zur Anionen- und besonders Phosphationenkomplexierung durch synthetische Rezeptoren: a) J.-M. Lehn, *Angew. Chem.* **1988**, **100**, 91–116; *Angew. Chem. Int. Ed. Engl. 1988*, **27**, 89–112; b) M. P. Mertes, K. Bowman Mertes, *Acc. Chem. Res. 1990*, **23**, 413–418; c) F. P. Schmidtchen, M. Berger, *Chem. Rev. 1997*, **97**, 1609–1646; d) P. D. Beer, P. A. Gale, *Angew. Chem. 2001*, **113**, 502–532; *Angew. Chem. Int. Ed.* **2001**, **40**, 486–516; e) C. A. Iloudis, J. W. Steed, *J. Supramol. Chem. 2001*, **1**, 165–187; f) J. M. Llinares, D. Powell, K. Bowman-James, *Coord. Chem. Rev. 2003*, **240**, 57–75; g) K. Bowman-James, *Acc. Chem. Res. 2005*, **38**, 671–678; h) J. L. Sessler, P. A. Gale, W.-S. Cho, *Anion Receptor Chemistry*, Royal Society of Chemistry, Cambridge, **2006**.
- [23] M. G. Rossmann, D. Moras, K. W. Olsen, *Nature 1974*, **250**, 194–199.
- [24] W. Möller, R. Amens, *FEBS Lett. 1985*, **186**, 1–7.
- [25] T. W. Traut, *Eur. J. Biochem. 1994*, **222**, 9–19.
- [26] a) W. G. J. Hol, P. T. Vanduijnen, H. J. C. Berendsen, *Nature 1978*, **273**, 443–446; b) R. K. Wierenga, M. C. H. De Maeyer, W. G. J. Hol, *Biochemistry 1985*, **24**, 1346–1357; c) B. E. Bernstein, P. A. M. Michels, W. G. J. Hol, *Nature 1997*, **385**, 275–278.
- [27] R. R. Copley, G. J. Barton, *J. Mol. Biol. 1994*, **242**, 321–329.
- [28] D. Bossemeyer, *Trends Biochem. Sci. 1994*, **19**, 201–205.
- [29] K. Kinoshita, K. Sadanami, A. Kidera, N. Go, *Protein Eng. 1999*, **12**, 11–14.
- [30] J. E. Walker, M. Saraste, M. J. Runswick, N. J. Gay, *EMBO J. 1982*, **1**, 945–951.
- [31] a) E. J. Milner-White, M. J. Russell, *Origins Life Evol. Biosphere 2005*, **35**, 19–27; b) J. D. Watson, E. J. Milner-White, *J. Mol. Biol. 2002*, **315**, 171–182.
- [32] J. Feuerstein, R. S. Goody, M. R. Webb, *J. Biol. Chem. 1989*, **264**, 6188–6190.
- [33] E. F. Pai, U. Krengel, G. A. Petsko, R. S. Goody, W. Kabsch, A. Wittinghofer, *EMBO J. 1990*, **9**, 2351–2359.
- [34] T. Zhou, M. Daugherty, N. V. Grishin, A. L. Ostermann, H. Zhang, *Structure 2000*, **8**, 1247–1257.
- [35] S. K. Hanks, A. M. Quinn, T. Hunter, *Science 1988*, **241**, 42–52.
- [36] A. I. Denesyuk, K. A. Denessiouk, T. Korpela, M. S. Johnson, *J. Mol. Biol. 2002*, **316**, 155–172.
- [37] L. Holm, C. Sander, *Science 1996*, **273**, 595–602.
- [38] S. Rhee, M. M. Silva, C. C. Hyde, P. H. Rogers, C. M. Metzler, D. E. Metzler, A. Arnone, *J. Biol. Chem. 1997*, **272**, 17293–17302.
- [39] K. A. Denessiouk, M. S. Johnson, A. I. Denesyuk, *J. Mol. Biol. 2005*, **345**, 611–629.
- [40] a) M. W. Hosseini, J.-M. Lehn, *Helv. Chim. Acta 1987*, **70**, 1312–1319; b) Q. Lu, R. J. Motekaitis, J. J. Reibenspies, A. E. Martell, *Inorg. Chem. 1995*, **34**, 4958–4964; c) A. C. Warden, M. Warren, M. T. W. Hearn, L. Spiccia, *Inorg. Chem. 2004*, **43**, 6936–6943.
- [41] B. Dietrich, T. M. Fyles, J.-M. Lehn, L. G. Pease, D. L. Fyles, *J. Chem. Soc. Chem. Commun. 1978*, 934–936.
- [42] F. P. Schmidtchen, *Tetrahedron Lett. 1989*, **30**, 4493–4496.
- [43] J. L. Sessler, E. Katayev, G. D. Pantos, Y. A. Ustynyuk, *Chem. Commun. 2004*, 1276–1277.
- [44] P. Chakrabarti, *J. Mol. Biol. 1993*, **234**, 463–482.
- [45] a) J. W. Pflugrath, F. A. Quiocio, *Nature 1985*, **314**, 257–260; b) H. Luecke, F. A. Quiocio, *Nature 1990*, **347**, 402–406; c) J. J. He, F. A. Quiocio, *Science 1991*, **251**, 1479–1481.
- [46] a) K. L. Longenecker, P. J. Roach, T. D. Hurley, *J. Mol. Biol. 1996*, **257**, 618–631; b) N. Kobayashi, N. Go, *Eur. Biophys. J. 1997*, **26**, 135–144; c) K. A. Denessiouk, J. V. Lehtonen, M. S. Johnson, *Protein Sci. 1998*, **7**, 1768–1771.
- [47] PDB-Recherchen wurden mithilfe von Relibase V. 1.3.2. (August 2005)<sup>[48]</sup> und der am 14. Februar 2006 aktualisierten PDB<sup>[18]</sup> durchgeführt; Urheberrechte: M. Hendlich 1994–1999 und Cambridge Crystallographic Data Centre 1999–2005, Union Road, Cambridge CB2 1EZ, Großbritannien.

- [48] a) M. Hendlich, A. Bergner, J. Günther, G. Klebe, *J. Mol. Biol.* **2003**, *326*, 607–620; b) J. Günther, A. Bergner, M. Hendlich, G. Klebe, *J. Mol. Biol.* **2003**, *326*, 621–636.
- [49] Wir danken S. Robinson von Relibase herzlich für ihre Unterstützung bei der Bestimmung der Aminosäureverteilung über die gesamte PDB hinweg.
- [50] P. R. Rablen, J. W. Lockman, W. L. Jorgensen, *J. Phys. Chem. A* **1998**, *102*, 3782–3797.
- [51] L. Miallau, M. S. Alphey, L. E. Kemp, G. A. Leonard, S. M. McSweeney, S. Hecht, A. Bacher, W. Eisenreich, F. Rohdich, W. N. Hunter, *Proc. Natl. Acad. Sci. USA* **2003**, *100*, 9173–9178.
- [52] W. Shi, N. R. Munagala, C. C. Wang, C. M. Li, P. C. Tyler, R. H. Furneaux, C. Grubmeyer, V. L. Schramm, S. C. Almo, *Biochemistry* **2000**, *39*, 6781–6790.
- [53] S. Bauer, K. Kemter, A. Bacher, R. Huber, M. Fischer, S. Steinbacher, *J. Mol. Biol.* **2003**, *326*, 1463–1473.
- [54] E. G. Krebs, J. A. Beavo, *Annu. Rev. Biochem.* **1979**, *48*, 923–959.
- [55] W. Vogel, R. Lammers, J. Hunag, A. Ullrich, *Science* **1993**, *259*, 1611–1614.
- [56] a) S. Cheek, H. Zhang, K. Ginalski, N. V. Grishin, *BMC Struct. Biol.* **2005**, *5*, 6; b) S. Cheek, H. Zhang, N. V. Grishin, *J. Mol. Biol.* **2002**, *320*, 855–881.
- [57] T. Naumann, H. Matter, *J. Med. Chem.* **2002**, *45*, 2366–2378.
- [58] a) J. R. Knowles, *Annu. Rev. Biochem.* **1980**, *49*, 877–919; b) W. W. Cleland, A. C. Hengge, *FASEB J.* **1995**, *9*, 1585–1594.
- [59] a) A. S. Mildvan, *Proteins Struct. Funct. Genet.* **1997**, *29*, 401–416; b) Y.-W. Xu, S. Morera, J. Janin, J. Cherfils, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 3579–3583; c) I. Schlichting, J. Reinstein, *Biochemistry* **1997**, *36*, 9290–9296; d) S. D. Lahiri, G. Zhang, D. Dunaway-Mariano, K. N. Allen, *Science* **2003**, *299*, 2067–2071; e) J. Knowles, *Science* **2003**, *299*, 2002–2003.
- [60] P. Cohen, *Nat. Rev. Drug Discovery* **2002**, *1*, 309–315.
- [61] a) B. J. Druker, S. Tamura, E. Buchdunger, S. Ohno, G. M. Segal, S. Fanning, J. Zimmermann, N. B. Lydon, *Nat. Med.* **1996**, *2*, 561–566; b) B. J. Druker, N. B. Lydon, *J. Clin. Invest.* **2000**, *105*, 3–7; c) T. Schindler, W. Bornmann, P. Pellicena, W. T. Miller, B. Clarkson, J. Kuriyan, *Science* **2000**, *289*, 1938–1942; d) B. Okram, A. Nagle, F. J. Adrián, C. Lee, P. Ren, X. Wang, T. Sim, Y. Xie, X. Wang, G. Xia, G. Spraggan, M. Warmuth, Y. Liu, N. S. Gray, *Chem. Biol.* **2006**, *13*, 779–786; e) S. W. Cowman-Jacob, V. Guez, G. Fendrich, J. D. Griffin, D. Fabbro, P. Furet, J. Liebetanz, J. Mestan, P. W. Manley, *Mini-Rev. Med. Chem.* **2004**, *4*, 285–299.
- [62] A. T. van Oosterom, I. Judson, J. Verweij, S. Stroobants, E. D. di Paola, S. Dimitrijevic, M. Martens, A. Webb, R. Sciot, M. Van Glabbeke, S. Silberman, O. S. Nielsen, *Lancet* **2001**, *358*, 1421–1423.
- [63] a) P. T. C. Wan, M. J. Garnett, S. M. Roe, S. Lee, D. Niculescu-Duvaz, V. M. Good, C. M. Jones, C. J. Marshall, C. J. Springer, B. Barford, R. Marais, *Cell* **2004**, *116*, 855–867; b) T. Ahmad, T. Eisen, *Clin. Cancer Res.* **2004**, *10*, 6388s–6392s.
- [64] a) H.-S. Cho, K. Mason, K. X. Ramyar, A. M. Stanley, S. B. Gabelli, D. W. Denney, Jr., D. J. Leahy, *Nature* **2003**, *421*, 756–760; b) L. K. Shawver, D. Slamon, A. Ullrich, *Cancer Cell* **2002**, *1*, 117–123; c) D. J. Slamon, B. Leyland-Jones, S. Shak, H. Fuchs, V. Paton, A. Bajamonde, T. Fleming, W. Eiermann, J. Wolter, M. Pegram, J. Baselga, L. Norton, *N. Engl. J. Med.* **2001**, *344*, 783–792; d) D. J. Slamon, G. M. Clark, S. G. Wong, W. J. Levin, A. Ullrich, W. L. McGuire, *Science* **1987**, *235*, 177–182.
- [65] S. S. Krishna, T. Zhou, M. Daugherty, A. Osterman, H. Zhang, *Biochemistry* **2001**, *40*, 10810–10818.
- [66] a) M. Kato, J. L. Chuang, S. C. Tso, R. M. Wynn, D. T. Chuang, *EMBO J.* **2005**, *24*, 1763–1774; b) R. Dutta, M. Inouye, *Trends Biochem. Sci.* **2000**, *25*, 24–28.
- [67] J. P. Davidson, O. Lubman, T. Rose, G. Waksman, S. F. Martin, *J. Am. Chem. Soc.* **2002**, *124*, 205–215.
- [68] E. Baraldi, K. Djinovic, Carugo, M. Hyvönen, P. Lo Surdo, A. M. Riley, B. V. L. Potter, R. O'Brien, J. E. Ladbury, M. Saaraste, *Structure* **1999**, *7*, 449–460.
- [69] U. Abele, G. E. Schulz, *Protein Sci.* **1995**, *4*, 1262–1271.
- [70] J. H. Hurley, H. R. Faber, D. Worthylake, N. D. Meadow, S. Roseman, D. W. Pettigrew, S. J. Remington, *Science* **1993**, *259*, 673–677.
- [71] S. Karthikeyan, Q. Zhou, F. Mseeh, N. V. Grishin, A. L. Osterman, H. Zhang, *Structure* **2003**, *11*, 265–273.
- [72] E. Sabini, S. Ort, C. Monnerjahn, M. Konrad, A. Lavie, *Nat. Struct. Biol.* **2003**, *10*, 513–519.
- [73] Übersichtsartikel: a) K. Hinterding, D. Alonso-Diaz, H. Waldmann, *Angew. Chem.* **1998**, *110*, 716–780; *Angew. Chem. Int. Ed.* **1998**, *37*, 688–749; b) D. Barford, *Trends Biochem. Sci.* **1996**, *21*, 407–412; c) E. B. Fauman, M. A. Saper, *Trends Biochem. Sci.* **1996**, *21*, 413–417.
- [74] a) H. L. Schubert, E. B. Fauman, J. A. Stuckey, J. E. Dixon, M. A. Saper, *Protein Sci.* **1995**, *4*, 1904–1913; b) K. L. Guan, J. E. Dixon, *J. Biol. Chem.* **1991**, *266*, 17026–17030; c) J. M. Denu, J. E. Dixon, *Proc. Natl. Acad. Sci. USA* **1995**, *92*, 5910–5914.
- [75] M. Elchebly, P. Payette, E. Michaliszyn, W. Cromlish, S. Collins, A. L. Loy, D. Normandin, A. Cheng, J. Himms-Hagen, C.-C. Chan, C. Ramachandran, M. J. Gresser, M. L. Tremblay, B. P. Kennedy, *Science* **1999**, *283*, 1544–1548.
- [76] Eine aktuelle Untersuchung: M. J. Begley, G. S. Taylor, M. A. Brock, P. Ghosh, V. L. Woods, J. E. Dixon, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 927–932.
- [77] A. Salmeen, J. N. Andersen, M. P. Myers, N. K. Tonks, D. Barford, *Mol. Cell* **2000**, *6*, 1401–1412.
- [78] S. Wang, L. Tabernero, M. Zhang, E. Harms, R. L. Van Etten, C. V. Stauffacher, *Biochemistry* **2000**, *39*, 1903–1914.
- [79] W. Wang, H. S. Cho, R. Kim, J. Jancarik, H. Yokota, H. H. Nguyen, I. V. Grigoriev, D. E. Wemmer, S. H. Kim, *J. Mol. Biol.* **2002**, *319*, 421–431.