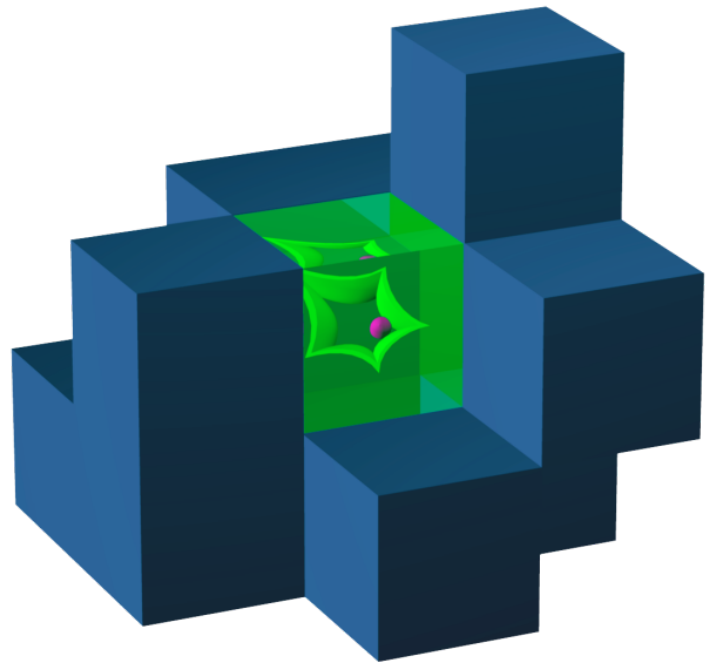




TECHNISCHE
UNIVERSITÄT
DRESDEN

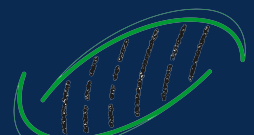
clsd
computational
life science lab
Dresden



Abschlussbericht für Phase II der BMBF Innovationsinitiative ForMaT

7.11.2008

Prof. Dr. Michael Schroeder
Biotechnologisches Zentrum
TU Dresden, 01062 Dresden
Tel: 0351 463 40060



Zusammenfassung.....	3
Kurze Darstellung zu Aufgabenstellung.....	4
Aufgabenstellung.....	4
Voraussetzungen, unter denen das Vorhaben durchgeführt wurde.....	4
Planung und Ablauf des Vorhabens.....	5
Wissenschaftlicher und technischer Stand, an den angeknüpft wurde, insbesondere.....	5
Zusammenarbeit mit anderen Stellen.....	6
Verwendung der Zuwendung und Ergebnisse.....	7
Vorgegebene Ziele.....	7
Übersicht über Arbeitspakete und Balkendiagramm.....	7
Arbeitspaketbeschreibungen.....	8
TA 1: ProWiki.....	17
Übersicht zur wissenschaftlich-technischen Arbeit für TA1.....	17
TA 2: Powergraph.....	18
Übersicht zur wissenschaftlich-technischen Arbeit für TA2.....	18
TA 3: DURIN.....	19
Übersicht zur wissenschaftlich-technischen Arbeit für TA3.....	19
Wichtigste Positionen des zahlenmäßigen Nachweises.....	21
Notwendigkeit und Angemessenheit der geleisteten Arbeit.....	21
Voraussichtlicher Nutzen und Verwertbarkeit des Ergebnisses.....	21
Fortschritts auf dem Gebiet des Vorhabens bei anderen Stellen.....	22
Erfolgte oder geplante Veröffentlichungen der Ergebnisse.....	22
Anlage.....	29

Zusammenfassung

Das computational life science lab CLSD der TU Dresden hat drei Technologieansätze für die biomedizinische Wirkstoff-Forschung entwickelt: In TA1 konnte eine neuartige Methode zur Vorhersage von Wirkstoff-Zielprotein Beziehungen entwickelt werden, in TA2 wurde ein Verfahren zur Vorhersage von Biomarkern entwickelt und in TA3 wurde ein Verfahren zur Analyse von Rasterkraftmikroskopiedaten entwickelt. Für jedes der drei TA wurde je ein Patent eingereicht. Insgesamt wurden 22 wissenschaftliche Artikel zu den Themen veröffentlicht, davon 13 in Top-Zeitschriften mit Impaktfaktor 4 und besser. CLSD hat insgesamt 12 Preise gewonnen. Darunter renommierte Auszeichnungen wie den Deutschen Innovationpreis und den Red Dot Design Award. Die Resultate des Projektes wurden auf mehr als 50 Treffen, Konferenzen und Messen präsentiert, darunter die Cebit, Bioanalytics und Biotechnica. Bzgl. der Verwertung der Resultate arbeitet die TU bzgl. TA1 eng mit der Transinsight GmbH, bzgl. TA2 mit der Qualitytype AG zusammen. Für TA3 wurde ein Geschäftsplan entwickelt, der u.a. auf dem renommierten nationalen Science4life-Businessplan Wettbewerb ausgezeichnet wurde.

Kurze Darstellung zu Aufgabenstellung

Aufgabenstellung

Ziel des CLSD war es, Softwarelösungen zur Wirkstoffforschung zu entwickeln, die den Workflow akademischer und industrieller Forscher unterstützt und Analysen erleichtert. Die Ergebnisse basieren auf drei Technologien, die Protein-Wirkstoff-Krankheits-Zusammenhänge aus Literatur und Patenten extrahieren, mit Proteinnetzen Nebeneffekte modellieren und die Struktur von Membranproteinen als Wirkstoffziel untersuchen. Die Integration dieser Technologien und ihre enge Verzahnung mit Ergebnissen aus Experimenten stellen einen neuartigen Ansatz dar, um die ersten Phasen der Wertschöpfungskette in der Wirkstoffentwicklung zu unterstützen. Neben der Weiterentwicklung der Technologien wurde ein Innovationslabor zur schnelleren kommerziellen Verwertung der Entwicklungen aufgebaut. Um die Nachhaltigkeit des Innovationslabors und die Verwertung voranzutreiben, wurde ein Verwertungskonzept begonnen umzusetzen und ein umfangreiches Transfernetzwerk aufgebaut.

Mit dem oben beschriebenen Forschungsvorhaben wurden drei wesentliche Aufgaben verfolgt.

- a) Erfassung von Verwertungsmöglichkeiten und Ausschöpfung des Innovationspotenzials Proteinsuchmaschine (TA1: PROWIKI)
- b) Erfassung von Verwertungsmöglichkeiten und Ausschöpfung des Innovationspotenzials einer Software zur Analyse krankheitsbezogener Hochdurchsatzdaten (TA2: POWERGRAPH)
- c) Erfassung von Verwertungsmöglichkeiten und Ausschöpfung des Innovationspotenzials einer Software zur Hochdurchsatzanalyse von Proteinstrukturen und Biomolekülen aus SMFS Messungen (TA3: DURIN)

Voraussetzungen, unter denen das Vorhaben durchgeführt wurde

Das Vorhaben wurde im Zeitraum von April 2009 bis März 2011 durchgeführt. Für die jeweiligen Teilprojekte wurden mehrere wissenschaftliche Mitarbeiter der Forschungsgruppe Bioinformatik in das Projektteam integriert. Die Mitarbeiter wurden nach ihren Qualifikationen und Forschungsschwerpunkten auf die jeweiligen Teilprojekte eingesetzt (meist keine vollen Stellen).

Team:

Forschungsgruppe TA1: Dr. Liliana Barrio-Alvers, Dr. Heiko Dietze, Dr. Thomas Wächter, Dr. Dimitra Alexopoulou, Conrad Plake, Ruslan David

Forschungsgruppe TA2: Dr. Dr. Christof Winter, Dr. Bingding Huang, Dr. Loic Alain Royer, Matthias Reimann, Janine Roy, Rainer Winnenburg

Forschungsgruppe TA3: Dr. William Andreopoulos, Dr. Frank Dressel, Dr. Annalisa Marsico, Anne Tuukkanen

Betriebswirt: Joscha Köllner

Ort: Biotechnologisches Zentrum der TU Dresden, BioCampus

Erfahrungen/Publicationen: Für die jeweiligen Forscherteams wurden die Mitarbeiter der Forschungsgruppe Bioinformatik je nach Erfahrungen und Kompetenzen ausgewählt, die die technische Weiterentwicklung der Teilprojekte sowie die Ausschöpfung der Innovationspotenziale dieser Teilprojekte vorangetrieben haben und weiterhin vorantreiben.

Planung und Ablauf des Vorhabens

Um das InnoLab CLSD zu realisieren, enthielt der Arbeitsplan sieben Arbeitspakete (siehe Abb. 11). In TA1-3 wurden die Ideen und Algorithmen für die drei Technologienansätze entwickelt. Im Arbeitspaket DI (Design und Implementierung) wurden die Resultate aus TA1-3 umgesetzt. Dazu wurde die Architektur der Systeme designed und implementiert, sowie softwaretechnisch getestet und dokumentiert. Die Arbeitspakete TA1-3 erhielten Anforderungen aus dem Arbeitspaket A, in dem zwei unabhängige Anwendungen aus der Krebs- und Stammzellforschung bearbeitet wurden. Diese beiden Anwendungen dienten am Ende des Projektes wiederum der inhaltlichen Evaluierung der in DI realisierten Software. Neben diesen fünf technischen/wissenschaftlichen Arbeitspaketen gab es das Arbeitspaket InnoLab, in dem die Transferstruktur realisiert wurde und das Managementarbeitspaket M, zur Verwaltung des Projektes. Diese Abhängigkeiten sind in Abb. XX illustriert.

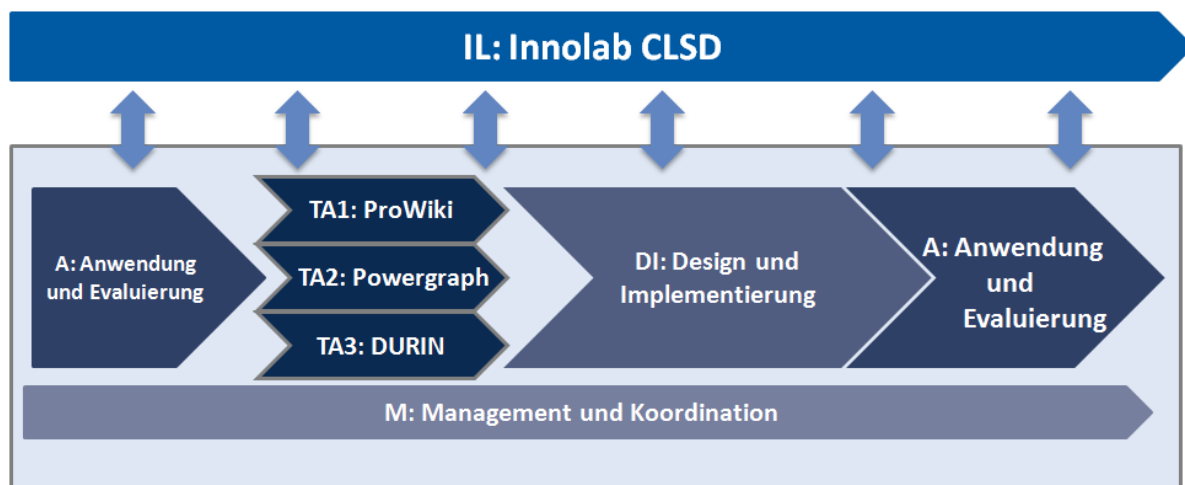


Abb. 1. Abhängigkeiten der Arbeitspakete. Die Anwendung (A) bildet den Ausgangspunkt für die Technologieansätze TA1-3. Diese wurden im Arbeitspaket DI implementiert und dann in Arbeitspaket A inhaltlich evaluiert. Alle Aktivitäten wurden durch das Paket InnoLab und Management betreut.

Wissenschaftlicher und technischer Stand, an den angeknüpft wurde, insbesondere

Angabe bekannter Konstruktionen, Verfahren und Schutzrechten die für die Durchführung des Vorhabens benutzt wurden

Für jeden Technologieansatz wurde ein Patent wie unten aufgeführt eingereicht. Diese Patentanmeldungen enthalten umfangreiche Informationen zu verwandten Schutzrechten, Verfahren, etc.

TU Dresden. **Datenverarbeitungssystem**. Eingereichtes Patent (German Patent Application Number DE 10 2009 043 708.8), 2009

TU Dresden. **Verfahren und Mittel zur Vorhersage der Überlebensdauer beim Pankreaskarzinom durch Analyse von Biomarkern**. Eingereichtes Patent (German Patent Application Number DE 10 2009 054 793.2), 2009

nanometis and TU Dresden. **Automatic analysis of bio-molecules**. Eingereichtes Patent (European Patent Application Number EPO 9180269.4), 2009

Angabe der verwendeten Fachliteratur sowie der benutzten Informations- und Dokumentationsdienste

Als Hauptdokumentationsdienst wird die biomedizinische Literaturdatenbank PubMed sowie Freepatentsonline für Patente genutzt. Konkret liegen allen Arbeiten eine Vielzahl von Artikeln zugrunde. Die Liste der publizierten 22 Papers findet sich unten. Sie enthalten jeweils Referenzen der genutzten Literatur.

Zusammenarbeit mit anderen Stellen

Das Kontaktnetzwerk zu Industrie, Wissenschaft und Transferorganisationen nimmt im CLSD einen zentralen Bestandteil ein. Um eine markt- und anwendungsorientierte Entwicklung von Technologieansätzen und Ideen zu gewährleisten, wurden Wirtschaftsvertreter aus dem Kontaktnetzwerk ausgewählt. Die Vertreter stammen vorrangig aus der Pharma- und Biotechnologie-Branche. Sie bewerteten Technologieansätze und Ideen und nahmen so am Transferprozess teil.

Zusammenfassend wurden mit dem CLSD nachhaltige Strukturen zum Wissens- und Technologietransfer geschaffen. Die Basis bilden Wissenschaftler am BIOTEC und Studenten aus den Wirtschaftswissenschaften, die zukünftig das Potential von Ideen evaluieren und Verwertungskonzepte erstellen können. Dabei können Sie auf ein Netz von Transferpartnern wie GWT, PIZ, Dresden exists, sowie Industrie- und Finanzkontakte zugreifen.

Verwendung der Zuwendung und Ergebnisse

Vorgegebene Ziele

Übersicht über Arbeitspakete und Balkendiagramm

Für jedes Arbeitspaket wurde zu Beginn des Projektes Ziele, Teilaufgaben und Meilensteine definiert. Folgende Tabelle fasst diese Aufgaben mit Anfangs- und Endmonat zusammen. Jedem Arbeitspaket liegt eine detaillierte Planung zugrunde, die drei Gruppen (TA1-3) und der BWLer (InnoLab) umfasst. Der zeitlich geplante Ablauf des Projektes ist aus einem Balkendiagramm (Gantt-Chart) in Abb. XX ersichtlich.

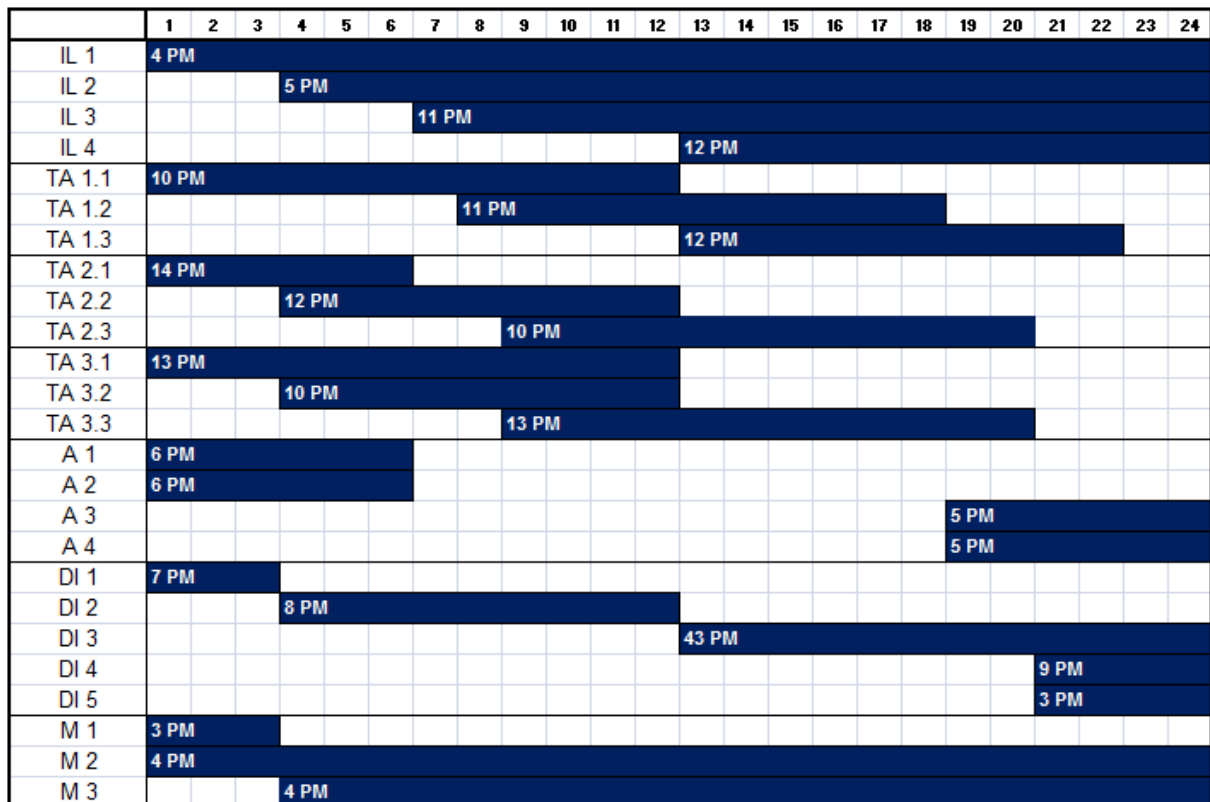


Abb. XX. Balkendiagramm mit zeitlichen Abhängigkeiten der Arbeitspakete und Aufgaben

Arbeitspaket		Start	Ende	PM
IL 1	Aufbau InnoLab Struktur und Teaching	1	24	4
IL 2	Unternehmensbeirat und Kunden-Netzwerk	4	24	5
IL 3	Scouting, Technologiemonitoring und Ideenmining	7	24	11
IL 4	Bewertung und Verwertungskonzept	13	24	12
TA 1.1	Entitäten-Erkennung	1	12	10

TA 1.2	Relationsextraktion	8	18	11
TA 1.3	Termgenerierung und Ontologieeditor	13	22	12
TA 2.1	Vorhersage von Proteininteraktionen	1	6	14
TA 2.2	Rekonstruktion von Interaktionskaskaden	4	12	12
TA 2.3	Modulerkennung	9	20	10
TA 3.1	Effiziente Algorithmen für Rasterkraftmikroskopie	1	12	13
TA 3.2	Strukturelle Interaktionen	4	12	10
TA 3.3	Motive in Membranproteinen	9	20	13
A 1	Anwendung Pankreaskrebs	1	6	6
A 2	Anwendung Stammzellen	1	6	6
A 3	Evaluierung Pankreaskrebs	19	24	5
A 4	Evaluierung Stammzellen	19	24	5
DI 1	Datenbank Design und Datensammlung	1	3	7
DI 2	Softwarearchitektur und -Design	4	12	8
DI 3	Implementierung/Programmierung	13	24	43
DI 4	Testen und Evaluierung	21	24	9
DI 5	Erstellung der Handbücher/Manuals	21	24	3
M 1	Webseite	1	3	3
M 2	Berichte und Dokumentation	1	24	4
M 3	Öffentlichkeitsarbeit	4	24	4

Arbeitspaketbeschreibungen

Arbeitspaket	InnoLab			Start	1
Leitung	IL			Ende	24
Gruppe	TA 1	TA 2	TA 3	IL	
Personenmonate	4	4	4	20	

Ziel: Aufbau nachhaltiger Strukturen zum Technologietransfer durch Technologie-Scouts und Transfermoderatoren

InnoLab 1: Aufbau InnoLab Struktur und Teaching: Aufbau der Transferstrukturen und detaillierte Konzeption des Scouting und der Transfermoderation. Kontakt zu Partnern wie GWT, Patentinformationszentrum, Dresden exists werden weiter ausgebaut. Konkreter Zeitplan für Technologiemoderatoren wird erstellt und mit der Analytikum GmbH koordiniert. Dokumentation der Strukturen und des Vorgehens. Evaluierung des Verwertungskonzeptes aus Phase 1. Zur Qualifizierung der Transfermoderatoren wird ein Teaching-Programm von der Analytikum GmbH genutzt bestehend aus aka-

demischen Weiterbildungskursen (Innovationsmanagement, Schutzrechtsmanagement, Controlling, Innovationsfinanzierung, Markt & Marketing, Geschäftsplanung, Projektmanagement, Soft Skills), Praktikum (Betreuung/Zusammenarbeit bei Markt-/Wettbewerbsanalyse, Geschäftsmodellentwicklung, Businessplanseminare, Businessplanwettbewerbe, Präsentationen, Mentorentätigkeiten, Gründungsvorbereitung, Gründertreffs, etc.) und Unternehmer-Workshops.

Für das Paket sind vier Personenmonate vorgesehen, wovon jeweils ein Personenmonat auf alle drei Teams und den BWLer (InnoLab) fällt. Die Laufzeit dieses Paketes zieht sich über die gesamten zwei Jahre.

InnoLab 2: Unternehmensbeirat und Kunden-Netzwerk: Der Beirat wird gepflegt und es wird Kontakt gehalten sowie das Netz potentieller Kunden weiter ausgebaut. Dazu werden am Ende des 1. Jahres eine größere Konferenz und/oder kleinere Veranstaltungen in mehreren Städten (vgl. Phase 1, Dresden, Berlin, Cambridge) organisiert. Desweiteren werden Reisen zu den jeweiligen wissenschaftlichen und industriellen Kunden und Partnern unternommen, um eventuelle Kollaborationen anzubahnen.

Für das Paket werden fünf Personenmonate veranschlagt. Den Projektteams wird jeweils ein Personenmonat dafür eingeräumt, dem BWLer (InnoLab) werden dazu in Summe zwei Personenmonate zugesprochen. Die Laufzeit dieses Pakets ist von Monat 4 bis Monat 24 angesetzt.

InnoLab 3: Scouting, Technologiemonitoring und Ideenmining: Kern des InnoLab bilden die Aktivitäten des Scouting und der Transfermoderation. Der Scout bringt Ideen, potentielle Kunden und Investoren zusammen. Die Transfermoderatoren arbeiten in kleinen Wissenschaftler-Teams und entwickeln gemeinsam mit Wirtschaftsstudenten erste Einschätzungen für deren Ideen und Konzepte zur Verwertung. Hierbei leistet der SAP Lehrstuhl/Dresden exists Unterstützung bei der Zusammenführung der Teams und der Organisation der Zusammenarbeit. Des Weiteren wird die Patentierung der Verfahren in TA1-3 formal betreut und vorangetrieben.

Für das Paket sind elf Personenmonate veranschlagt. Team 1, Team 2 und Team 3 haben jeweils zwei Personenmonate dafür zur Verfügung. Der BWLer (InnoLab) hat fünf Personenmonate dafür zur Verfügung. Dieses Paket beginnt im siebten Monat und wird bis zum Ende der Projektlaufzeit durchgeführt.

InnoLab 4: Bewertung und Verwertungskonzept: Insbesondere für die Technologieansätze 1-3 wird ein detailliertes Verwertungskonzept erstellt und bewertet. Dieses bildet dann die Basis für das weitere Vorgehen in Form von Spin-Off und/oder Kooperationen. Dieses Arbeitspaket wird dem BWLer (InnoLab) zugeordnet. Er hat dafür zwölf Personenmonate zur Verfügung und beginnt im 13. Monat mit der Bewertung und Erstellung der Verwertungskonzepte. Dieses Arbeitspaket endet im 24. Monat des Gesamtvorhabens. Die Arbeit mündet in die tatsächliche Verwertung der Technologieansätze nach dem ForMaT-Projekt.

Meilensteine:

Monat 12: Beirat und Kontaktnetz erweitert und gefestigt; Patentierungsprozess begonnen; Teaching begonnen

Monat 24: Verwertungskonzept liegt vor; Transferstruktur arbeitet nachhaltig; Teaching abgeschlossen

Arbeitspaket TA 1: ProWiki

Arbeitspaket	TA 1: ProWiki			Start	1
Leitung	TA 1			Ende	22
Partner	TA 1	TA 2	TA 3	IL	
Personenmonate	33	0	0	0	

Ziel: Tagesaktuelle Extraktion aller Fakten über Protein-Wirkstoff-Krankheitszusammenhänge aus der wissenschaftlichen Literatur und aus Patenten.

TA1.1: Entitäten-Erkennung: Es werden Algorithmen und Technologien zur Erkennung von biologischen Entitäten im Text entwickelt. Es wird eine allgemeine Bibliothek für Proteinnamen, Wirkstoffe und Krankheiten erstellt. Mit Hilfe von Text-Mining-Verfahren werden Textstellen markiert, welche mit hoher Wahrscheinlichkeit eine biologische Entität beschreiben. Die markierten Textstellen müssen mit Einträgen der erstellten Bibliothek in Zusammenhang gebracht werden. Ziel ist es, bei steigender Genauigkeit und Vollständigkeit vollautomatisch Textstellen mit spezifischen Proteinnamen zu assoziieren. Zur Sicherung der Qualität wird ein umfassender Korpus von manuell annotierten Dokumenten zusammengetragen, der es ermöglicht, die Qualität der Methode ständig zu überprüfen. Der zweite Forschungsschwerpunkt liegt auf der automatischen Assoziation der Proteine mit Wirkstoffen und Krankheiten.

Das Arbeitspaket 1 des TA1 wird ausschließlich vom ersten Projektteam realisiert. Dazu werden zehn Personenmonate eingeräumt. Die Laufzeit wird auf 12 Monate ab Beginn der gesamten Projektlaufzeit gesetzt.

TA1.2: Relationsextraktion: Es soll erforscht werden, wie in der Literatur beschriebene Protein-Wirkstoff, Wirkstoff-Krankheits- und Protein-Krankheitszusammenhänge automatisch extrahiert werden können. Dazu soll eine Ontologie für Beziehungen aufgebaut werden. Die Ontologie wird neben den Interaktionen auch Eigenschaften der Proteine, Wirkstoffe, Krankheiten enthalten. Es wird untersucht werden, wie die Fakten in wissenschaftlichen Publikationen und Patenten beschrieben werden und welche beschreibenden Informationen darüber hinaus vorliegen. Es wird ein Sequenzalignment-Ansatz von Hakenberg et al. (2008) weiterentwickelt.

Das Paket wird in Überlappung mit dem ersten Paket aus TA1 vom Projektteam 1 realisiert. Das Arbeitspaket umfasst einen Arbeitsaufwand von elf Personenmonaten und beginnt in Monat 8. Das Arbeitspaket wird in Monat 18 abgeschlossen.

TA1.3: Termgenerierung und Ontologieeditor: Unterstützung bei der kollaborativen Erstellung der Ontologie durch einen webbasierten Ontologieeditor mit integrierter Vorschlagsplattform für neue Terminologie. Innerhalb der geplanten Wissensplattform soll ein Editor verfügbar sein, der die gemeinsame Entwicklung einer Ontologie für Krankheiten und Proteinfamilien und Genen ermöglicht. Durch die enge Integration soll es möglich werden, zum einen automatische Unterstützung beim

Finden dieser neuen Terme und Definitionen zu geben und zum anderen die Indexierung der Dokumente mit Ontologietermen ständig überprüfen zu können.

Nach Abschluss des ersten Arbeitspaketes für TA1 wird TA1.3 beginnen (Monat 13) und in Monat 22 enden. Es werden zur Erfüllung des Paketes 12 Personenmonate veranschlagt, die auf Projektteam 1 entfallen.

Meilensteine:

Monat 12: Erkennung von Proteinen, Wirkstoffen und Krankheitsbezeichnern; Patent

Monat 24: Ontologieeditor incl. Erkennung von Proteinennamen, Wirkstoffen und Krankheitsbezeichnern, sowie Relationen zwischen ihnen.

Arbeitspaket TA 2: Powergraph

Arbeitspaket	TA 2: Powergraph			Start	1
Leitung	TA 2			Ende	20
Partner	TA 1	TA 2	TA 3	IL	
Personenmonate	0	36	0	0	

Ziel: Weiterentwicklung strukturbasierter Vorhersage von Interaktionen und Erkennung funktionaler Module, so dass sie eng verzahnt mit krankheitsbezogenen Genexpressionsdaten eingesetzt werden können.

TA2.1: Vorhersage von Proteininteraktionen: Verschiedene Methoden zur strukturbasierten Vorhersage von Interaktionen werden entwickelt. Der Algorithmus wird mit verschiedenen Methoden zur Erkennung der Strukturvorlagen getestet, wie BLAST und Threading. Alternativ wird noch eine Methode entwickelt, um Sequenzmotive der Interaktionsschnittstellen zu generieren, die durch Hidden Markov Modelle repräsentiert werden. Zur Evaluierung wird ein Benchmark-Datensatz aufgesetzt mit bekannten, verlässlichen Interaktionen. Die Erfolgsquote der Vorhersagen wird auf dem Benchmark geprüft.

Das Arbeitspaket wird von Projektteam 2 von Monat 1 bis Monat 6 erfüllt. Dazu werden 14 Personenmonate als Arbeitszeit angesetzt.

TA2.2: Rekonstruktion von Interaktionskaskaden: Ziel ist es, mittels der Vorhersagen aus 2.1 und bekannter Interaktionen aus Datenbanken wie BioGrid, MIPS, HPRD, etc. ganze Pfade interagierender Proteine zu rekonstruieren. Dabei werden neben physikalischen Interaktionen auch regulatorische Transkriptionsnetze berücksichtigt.

Bereits während der Arbeit an TA2.1 wird die Rekonstruktion von Interaktionskaskaden beginnen (Monat 4) und in Monat 12 enden. Die Erfüllung des Pakets wird von Projektteam 2 sichergestellt und mit zwölf Personenmonaten angesetzt.

TA2.3: Modulerkennung: Es werden verschiedene Clustering-Verfahren geprüft, um Module zu finden. Hauptziel ist, einen existierenden Algorithmus weiterzuentwickeln und dessen Erfolgsquote und

Schnelligkeit zu verbessern. Zur Evaluierung werden Benchmarks definiert sowie statistische Verfahren zur Charakterisierung der Module an Hand von Genontologie-Annotationen.

Die Modulerkennung wird 10 Personenmonate in Anspruch nehmen. Beginn ist in Monat 9. Das Paket wird von Projektteam 2 realisiert und endet in Monat 20.

Meilensteine:

Monat 12: Evaluierte Methode zur Vorhersage von Interaktionen und Interaktionskaskaden; Patent

Monat 24: Evaluierte Methode zur Modulerkennung

Arbeitspaket TA 3: DURIN

Arbeitspaket	TA 3: DURIN			Start	1
Leitung	TA 3			Ende	20
Partner	TA 1	TA 2	TA 3	IL	
Personenmonate	0	0	36	0	

Ziel: Entwicklung eines effizienten und qualitativ hochwertigen Verfahrens zur Analyse von Hochdurchsatz-Rasterkraftmikroskopiedaten sowie weiterführender strukturbasierter Analysen

TA3.1: Effiziente Algorithmen für Rasterkraftmikroskopie: Basierend auf einem Algorithmus der zwar akkurat, aber langsam ist, wird ein neues Verfahren realisiert, das effizient tausende Kraftkurven in kurzer Zeit bearbeitet. Dazu werden statistische Verfahren zur Maximafindung mit effizienten Clustering-Verfahren kombiniert. Zur Bewertung wird ein Benchmark-Datensatz definiert. Die Algorithmen werden mithilfe von den Ergebnissen aus der Anwendung bei der Stammzellenforschung und Pankreaskrebs evaluiert und abgeglichen.

Für das Paket 1 beim TA3 werden 13 Personenmonate benötigt. Das Projektteam 3 wird dieses Paket in den ersten zwölf Monaten realisieren.

TA3.2: Strukturelle Interaktionen: Um die in TA3.1 untersuchten Strukturen weiter zu analysieren, werden eine Datenbank und ein Verfahren entwickelt, um intramolekulare und intermolekulare Interaktionen zu bestimmen und mit den Daten aus 3.1 abzugleichen. Dazu werden systematisch alle bekannten Strukturen in der ProteinDataBank, PDB, indiziert und Schnittstellen zu Wirkstoffen und Hot-Spot-Aminosäuren identifiziert. Diese Daten werden in einer Datenbank abgelegt. Die Algorithmen werden mithilfe von den Ergebnissen aus der Anwendung bei der Stammzellenforschung und Pankreaskrebs evaluiert und abgeglichen.

Für das Paket 2 werden zehn Personenmonate von Projektteam 3 eingebracht. Strukturelle Interaktionen werden bereits während der Erarbeitung von effizienten Algorithmen für die Rasterkraftmikroskopie in Algorithmen umgesetzt. Dazu werden zehn Personenmonate veranschlagt. Die Laufzeit ist von Monat 4 bis Monat 12 angesetzt.

TA3.3: Motive in Membranproteinen: Es werden Algorithmen entwickelt, die aus den Proteinstrukturen lineare Motive, die wenige Aminosäuren lang sind, und eine bestimmte Funktion haben, ableiten. Die Motive werden in einer Datenbank abgelegt und in allen verfügbaren Sequenzdaten werden Motive annotiert. Die Datenbank wird an die Analyse in 3.1 gekoppelt. Die Algorithmen werden mit Hilfe von den Ergebnissen aus der Anwendung bei der Stammzellenforschung und Pankreaskrebs evaluiert und abgeglichen.

Das Arbeitspaket wird von Projektteam 3 realisiert und umfasst 13 Personenmonate zur Leistungserstellung. Die Laufzeit ist Monat 9 bis Monat 20.

Meilensteine:

Monat 12: Effizienter und akkurater Algorithmus zur Analyse der Rasterkraftmikroskopiedaten liegt vor

Monat 24: Datenanalyse ist gekoppelt an Analyse mit Strukturinteraktionen und -motiven

Arbeitspaket A: Anwendung und Evaluierung

Arbeitspaket	A Anwendung und Evaluierung			Start	1
Leitung	TA 1/2/3			Ende	24
Partner	TA 1	TA 2	TA 3	IL	
Personenmonate	8	8	6	0	

Ziel: Definition von Anforderungen an die Technologieansätze und Evaluierung an Hand von praktischen und exemplarischen Fragestellungen aus der Krebs- und Stammzellforschung

A 1: Anwendung Pankreaskrebs: Die Pilarsky-Gruppe am Klinikum der TU Dresden führt Microarray-Studien im Bauchspeicheldrüsenkrebs durch. Ziel ist es, neue Wirkstoffe zu entwickeln und Biomarker zu finden, die mit Überlebensdauer korrelieren und es so ermöglichen, eine optimale Therapie zu wählen. Die Gruppe hat umfangreiche Experimente mit über 50 Gewebeproben durchgeführt. Die Daten liegen dem CLSD bereits vor. Es werden Anforderungen an die Software definiert, um die beiden Probleme Wirkstoffe/Biomarker optimal zu unterstützen. Die Schroeder und Pilarsky-Gruppe haben zusammen zwei Artikel publiziert, so dass bereits ein großes Vorwissen auf beiden Seiten vorliegt.

Basierend auf diesen Screens wird in Aufgabe A1 ein Anwendungsszenario erarbeitet, dass die drei Phasen Target-Identifizierung, Wirkstoff-Screening und Lead-Optimierung abdeckt. Es werden relevante Daten gesammelt und aufbereitet. Insbesondere werden relevante Membranproteine identifiziert, die in TA3 experimentell und rechnergestützt untersucht werden. Dazu wird die entsprechende Technologieplattform des BIOTEC genutzt.

Für Arbeitspaket A 1 werden sechs Personenmonate veranschlagt. Davon entfallen 2 Personenmonate jeweils auf Team 1, Team 2 und Team 3. Die Laufzeit wird von Monat 1 bis Monat 6 angesetzt.

A 2: Anwendung Stammzellen: Die Storch-Gruppe am Klinikum der TU Dresden arbeitet an neurodegenerativen Krankheiten wie Parkinson und führt dazu u.a. Microarray-Studien durch. In einer Zusammenarbeit der Schroeder- und Storch-Gruppe konnte bereits der Hypoxia-Inducing Factor HIF1A als Meisterregulator identifiziert werden, sowie eine Reihe von Proteinen, die zur Signalverarbeitung wichtig sind inkl. einiger Membranproteine. Diese Vorarbeiten haben zu zwei Manuskripten geführt, die gegenwärtig eingereicht sind.

Basierend auf diesen Vorarbeiten wird A 2 ein Anwendungsszenario definieren und insbesondere alle relevanten Daten sammeln und aufbereiten. Dazu werden existierende Microarray Experimente verarbeitet und neue Daten mittels Rasterkraftmikroskopie für relevante Proteine generiert.

Das Arbeitspaket wird von allen drei Teams zu je zwei Personenmonaten erfüllt. Die Laufzeit ist Monat 1 bis Monat 6.

A 3: Evaluierung Pankreaskrebs: Die Technologieansätze werden mittels der Anforderungen bewertet. Insbesondere ergibt sich dadurch bei positivem Verlauf eine allgemeine Fallstudie, die Kunden vorgelegt werden kann.

Die Evaluation Pankreaskrebs wird fünf Personenmonate beanspruchen. Team 1 und Team 2 bringen je 2 Personenmonate ein, Team 3 bringt einen Personenmonat ein. Die Evaluation beginnt in Monat 19, d.h. wenn sämtliche vorhergehende Arbeitspakete weitestgehend abgeschlossen sind. Das Paket wird in Monat 24 abgeschlossen.

A 4: Evaluierung Stammzellen: Die Technologieansätze werden mittels der Anforderungen bewertet. Insbesondere ergibt sich dadurch bei positivem Verlauf eine allgemeine Fallstudie, die Kunden vorgelegt werden kann.

Die Evaluation Stammzellen wird fünf Personenmonate beanspruchen. Team 1 und Team 2 bringen je 2 Personenmonate ein, Team 3 bringt einen Personenmonat ein. Die Evaluation beginnt in Monat 19, d.h. wenn sämtliche vorhergehende Arbeitspakete weitestgehend abgeschlossen sind. Das Paket wird in Monat 24 abgeschlossen.

Meilensteine:

Monat 6: Definition der Anforderungen exemplarischer Problemstellungen aus der Krebs- und Stammzellforschung

Monat 24: Evaluierung und Definition von Fallstudien der beiden Anwendungen

Arbeitspaket DI: Design und Implementierung

Arbeitspaket	DI: Design und Implementierung	Start	1
---------------------	---------------------------------------	--------------	----------

Leitung	TA 1/2/3			Ende	24
Partner	TA 1	TA 2	TA 3	IL	
Personenmonate	25	22	23	0	

Ziel: Ziel ist die professionelle Umsetzung der Verfahren und Algorithmen aus TA1-3 in ein modulares Softwaresystem sowie formales Testen und Evaluierung der Software.

DI 1: Datenbank Design und Datensammlung: Alle drei TA nutzen umfangreiche existierende Datenbestände aus Datenbanken wie UniProt, ProteinDataBank, Medline, Patent Datenbanken, SCOP und weitere. Dazu wird ein Datenbankserver konfiguriert und die Daten eingespielt und gepflegt. Es werden Skripte geschrieben, um die Daten tagesaktuell vorzuhalten. Zusätzlich werden spezifische Datenbanken für die jeweiligen TA aufgesetzt.

Projektteam 1 wird dafür fünf Personenmonate aufbringen. Team 2 und 3 werden jeweils einen Personenmonat leisten. Das Datenbank Design und die Datensammlung wird in den ersten drei Projektmonaten realisiert.

DI 2: Softwarearchitektur und –design: Ziel ist es, die Software aus den TA professionell zu entwickeln und dabei Industriestandards einzusetzen. Insbesondere steht die modulare Architektur der jeweiligen TA im Vordergrund mit klar definierten Schnittstellen (API), um die Software später einfach in bestehende Systeme einzubinden. Es werden Coding-Standards festgelegt, an die sich alle Programmierer halten müssen und ein Trackingsystem installiert, um Fehler und Features effizient zu verwalten. Die Software wird nach softwaretechnischen Standards designed.

Für das Paket werden von Team 1 und 3 jeweils drei Personenmonate aufgebracht. Team 2 leistet zwei Personenmonate. Die Laufzeit für DI 2 ist Monat 4 bis Monat 12.

DI 3: Implementierung/Programmierung: Diese Aufgabe nimmt einen Großteil des Arbeitspaketes ein. Hier werden die Algorithmen aus TA1-3 mit der in DI 2 definierten Architektur implementiert.

Dazu werden alle Projektteams insgesamt 43 Personenmonate in der Laufzeit Monat 13 bis Monat 24 einbringen. Auf Team 1 entfallen 13 Personenmonate, auf die anderen beiden Teams jeweils 15 Personenmonate.

DI 4: Testen und Evaluierung: Um hochwertige Resultate zu erzielen, ist die Testphase sehr wichtig. Es werden funktionale und nicht-funktionale Eigenschaften geprüft und eine Bibliothek von JUnit Tests aufgesetzt, um systematisch bei der Entwicklung definierte Tests zu prüfen.

Für diese Phase werden neun Personenmonate eingeplant. Jedes Projektteam bringt drei Personenmonate auf. Das Arbeitspaket wird zwischen Monat 21 und 24 realisiert.

DI 5: Erstellung der Handbücher/Manuals: Um die Software auch nach Ende der Phase 2 weiter zu nutzen, ist Dokumentation wichtig. Darüber hinaus werden Manuals geschrieben, die die Funktionalität und das GUI der Software beschreiben und für Interessenten und Kunden geeignet ist.

Dafür werden alle drei Teams in den letzten vier Monaten jeweils einen Personenmonat aufbringen.

Meilensteine:

Monat 6: Alle Daten werden in Datenbanken vorgehalten

Monat 12: Design und Architektur der Software ist abgeschlossen

Monat 24: Lauffähige, getestete, dokumentierte Software liegt vor

Arbeitspaket M: Management

Arbeitspaket	M: Management			Start	1
Leitung	IL			Ende	24
Partner	TA 1	TA 2	TA 3	IL	
Personenmonate	2	2	3	4	

Ziel: Das Arbeitspaket hat zwei Ziele: Neben der Verwaltung des Projektes und der Berichterstellung dient es der Öffentlichkeitsarbeit durch eine Website und beispielsweise Präsenz auf Konferenzen und Tagungen.

M 1: Webseite: Entwicklung einer Webpräsenz. Das Arbeitspaket umfasst drei Personenmonate, die von Projektteam 3 (1 PM) und dem BWLer (2 PM) geleistet werden. Die Website wird in den ersten drei Monaten entwickelt.

M 2: Berichte und Dokumentationen: Erstellung aller notwendigen Berichte und Dokumentationen. Die Dokumentation der Tätigkeiten wird von allen Projektteams und dem BWLer (InnoLab) mit je einem Personenmonat über die gesamte Projektlaufzeit hinweg geleistet.

M 3: Öffentlichkeitsarbeit: Präsenz des Projektes auf geeigneten Tagungen, Treffen und Konferenzen. Die Öffentlichkeitsarbeit wird ab dem vierten Projektmonat bis zum Ende der Projektlaufzeit von allen vier Teams zu gleichen Teilen (je 1 PM) geleistet.

ERREICHTE ZIELE – wissenschaftlich-technischen Ergebnisse

TA 1: ProWiki

Zusammenfassung: In den letzten 24 Monaten wurden im Rahmen der wissenschaftlichen Arbeit 10 Publikationen veröffentlicht und ein Patent angemeldet. Die geplanten Meilensteine wurden erreicht und es mussten bisher keine Modifizierungen der Entwicklungs- und Verwertungsziele vorgenommen werden. Die Technologiefortschritte wurden im Betrachtungszeitraum auf 18 unterschiedlichen Konferenzen, Workshops, Tagungen und anderweitigen Veranstaltungen vorgestellt.

Übersicht zur wissenschaftlich-technischen Arbeit für TA1

A. Meilensteinplanung laut Projektantrag 01/2009

Arbeitspaket	Bezeichnung	Inhalte	Laufzeit	Erfüllung
[TA1.1]	Entitäten-Erkennung	- [TA1.1.1] Entitäten-Bibliothek mit Proteinnamen, Wirkstoffen und Krankheiten - [TA1.1.2] Entitäten-Extraktion - [TA1.1.3] Dokumenten-Korpus	M1-M3 M5-M9 M11-M12	100% 100% 100%
[TA1.2]	Relationsextraktion	- [TA1.2.1] Relationsextraktion wissenschaftlicher Publikationen erfolgt - [TA1.2.2] Relationsextraktion Patentbeschreibungen - [TA1.2.3] Beziehungsontologie	M8-M11 M12-M13 M14-M18	100% 100% 100%
[TA1.3]	Termgenerierung und Ontologieeditor	- [TA1.3.1] Webbasierter Ontologieeditor - [TA1.3.2] Integrierte Vorschlagsplattform - [TA1.3.3] Indexierung	M13-M18 M18-M22 M19-M20	100% 100% 100%
[A1]	Anwendung Pankreaskrebs	- Automatische Extraktion von Literaturdaten	M1-M6	100%
[A2]	Anwendung Stammzellen	- Definition von Anforderungen, Entwurf von Szenarien	M1-M6	100%
[A3]	Evaluierung Pankreaskrebs	- Entwicklung Fallstudie Pankreaskrebs	M19-M24	100%
[A4]	Evaluierung Stammzellen	- Entwicklung Fallstudie Neurodegeneration	M19-M24	100%
[DI]	Design und Implementierung	- [DI1.1] Konfiguration des Datenbankservers - [DI1.2] projektspezifische Datenbanken - [DI2] Aktualisierung - [DI3] Implementierung/Programmierung - [DI4] Testen und Evaluierung - [DI5] Erstellung der Handbücher	M1-M2 M3-M3 M4-M12 M13-M24 M21-M24 M21-M24	100% 100% 100% 100% 100%

[M]	Management	- [M2] Berichte und Dokumenta- tion - [M3] Öffentlichkeitsarbeit	M1-M24 M4-M24	100% 100%
[IL]	Aufbau Innolab Struktur und Teaching	- [IL1.3] Teaching-Programm (Retreat) - [IL2.2] Kollaborationsanbah- nungen, Firmen und Konferenzen - [IL3.2] Ideenevaluierung und Verwertungskonzepte - [IL3.3] Patentierung	M1-M24 M4-M24 M7-M24 M7-M24	100% 100% 100% 100%

Meilensteine Monat 12: Erkennung von Proteinen, Wirkstoffen und Krankheitsbezeichnern; Patent

TA 2: Powergraph

Zusammenfassung: In den letzten 24 Monaten wurden im Rahmen der wissenschaftlichen Arbeit 9 Publikationen veröffentlicht und ein Patent angemeldet. Die geplanten Meilensteine wurden erreicht und es mussten bisher keine Modifizierungen der Entwicklungs- und Verwertungsziele vorgenommen werden. Die Technologiefortschritte wurden im Betrachtungszeitraum auf 14 unterschiedlichen Konferenzen, Workshops, Tagungen und anderweitigen Veranstaltungen vorgestellt.

Übersicht zur wissenschaftlich-technischen Arbeit für TA2

A. Meilensteinplanung laut Projektantrag 01/2009

Ar- beits- paket	Bezeichnung	Inhalte	Lauf- zeit	Erfül- lung
[TA2.1]	Vorhersage von proteininter- aktionen	- [TA2.1.1] Methodenentwick- lung zur strukturbasierten Vorhersage	M1-M6	100%
		- [TA2.1.2] Methodenentwick- lung zur sequenzbasierten Vor- hersage	M3-M6	100%
		- [TA2.1.3] Generierung eines Interaktions-Benchmark- Datensatzes	M1-M2 M5-M6	100% 100%
		- [TA2.1.4] Evaluation der vorhergesagten Interaktionen auf Benchmark-Datensatz		
[TA2.2]	Rekonstruktion von Interakti- onskaskaden	- [TA2.2.1] Rekonstruktion von Proteininteraktionsnetzwerken	M4-M8	100%
		- [TA2.2.2] Rekonstruktion von regulatorischen Netzwerken	M4-M8	100%
		- [TA2.2.3] Verbindung von verschiedenen Netzwerken	M9-M12	100%
[TA2.3]	Modulerkennung	- [TA2.3.1] Modulerkennung durch Clustering	M9-M16	100%
		- [TA2.3.2] Evaluierung der Modulerkennung	M17-M20	100%
[A1]	Anwendung Pan- kreaskrebs	- Regulatorische Netze mit Krebsbezug	M1-M6	100%
[A2]	Anwendung Stammzellen	- Definition der Anforderungen an Powergraph	M1-M6	100%
[A3]	Evaluierung Pankreaskrebs	- Entwicklung Fallstudie Pankre- askrebs	M19-M24	100%
[A4]	Evaluierung	- Entwicklung Fallstudie Neuro-	M19-M24	100%

[A4]	Evaluierung Stammzellen	- Entwicklung Fallstudie Neurodegeneration	M19-M24	100%
[DI]	Design und Implementierung	- [DI1.2] Projektspezifische Datenbanken - [DI2] Aktualisierung - [DI3] Implementierung/Programmierung - [DI4] Testen und Evaluierung - [DI5] Erstellung der Handbücher	M3-M3 M4-M12 M13-M24 M21-M24 M21-M24	100% 100% 100% 100% 100%
[M]	Management	- [M2] Berichte und Dokumentation - [M3] Öffentlichkeitsarbeit	M1-M24 M4-M24	100% 100%
[IL]	Aufbau Innolab Struktur und Teaching	- [IL1.3] Teaching-Programm (Retreat) - [IL2.2] Kollaborationsanbahnungen, Firmen und Konferenzen - [IL3.2] Ideenevaluierung und Verwertungskonzepte - [IL3.3] Patentierung	M1-M24 M4-M24 M7-M24 M7-M24	100% 100% 100% 100%

Meilensteine Monat 12: Evaluierte Methode zur Vorhersage von Interaktionen und Interaktionskaskaden; Patent

TA 3: DURIN

Zusammenfassung: In den letzten 24 Monaten wurden im Rahmen der wissenschaftlichen Arbeit 3 Publikationen veröffentlicht und ein Patent angemeldet. Die geplanten Meilensteine wurden erreicht und es mussten bisher keine Modifizierungen der Entwicklungs- und Verwertungsziele vorgenommen werden. Die Technologiefortschritte wurden im Betrachtungszeitraum auf 27 unterschiedlichen Konferenzen, Workshops, Tagungen und anderweitigen Veranstaltungen vorgestellt.

Übersicht zur wissenschaftlich-technischen Arbeit für TA3

A. Meilensteinplanung laut Projektantrag 01/2009

Arbeitspaket	Bezeichnung	Inhalte	Laufzeit	Erfüllung
[TA3.1]	Effiziente Algorithmen für Rasterkraftmikroskopie	- [TA3.1.1] Mustererkennung - [TA3.1.2] Vergleich der Muster - [TA3.1.3] Benchmarkdatensatz - [TA3.1.4] Evaluierung	M1-M3 M5-M9 M11-M12	100% 100% 100%
[TA3.2]	Strukturelle Interaktionen	- [TA3.2.1] Indizierung - [TA3.2.2] Datenbank - [TA3.2.3] Evaluierung	M8-M11 M12-M13 M14-M18	100% 100% 100%
[TA3.3]	Motive in Membranproteinen	- [TA3.3.1] Motive - [TA3.3.2] Suche - [TA3.3.3] Datenbank - [TA3.3.4] Evaluierung	M13-M18 M18-M22 M19-M20	100% 100% 100%
[A1]	Anwendung Pankreaskrebs	- Experimentelle und rechnergestützte Untersuchung von Membranproteinen	M1-M6	100%
[A2]	Anwendung	- Formulierung von Anforderungen	M1-M6	100%

[A2]	Anwendung Stammzellen	- Formulierung von Anforderungen an Durin	M1-M6	100%
[A3]	Evaluierung Pankreaskrebs	- Entwicklung Fallstudie Pankreaskrebs	M19-M24	0%
[A4]	Evaluierung Stammzellen	- Entwicklung Fallstudie Neurodegeneration	M19-M24	0%
[DI]	Design und Implementierung	- [DI1.2] Projektspezifische Datenbanken	M3-M3	100%
		- [DI2] Festlegung und Konfiguration von Entwicklungsumgebungen	M4-M12	100%
		- [DI3] Effiziente Algorithmen für AFM, Strukturelle Interaktionen, Motive in Membranproteinen	M13-M24	100%
		- [DI4] Testen und Evaluierung	M21-M24	100%
		- [DI5] Erstellung der Handbücher	M21-M24	100%
[M]	Management	- [M1] Website	M1-M3	100%
		- [M2] Berichte und Dokumentation	M1-M24	100%
		- [M3] Öffentlichkeitsarbeit	M4-M24	100%
[IL]	Aufbau Innolab Struktur und Teaching	- [IL1.3] Teaching-Programm (Retreat)	M1-M24	100%
		- [IL2.2] Kollaborationsanbahnungen, Firmen und Konferenzen	M4-M24	100%
		- [IL3.2] Ideenevaluierung und Verwertungskonzepte	M7-M24	100%
		- [IL3.3] Patentierung	M7-M24	100%

Meilensteine Monat 12: Effizienter und akkurater Algorithmus zur Analyse der Rasterkraftmikroskopiedaten

Wichtigste Positionen des zahlenmäßigen Nachweises

Die wesentlichen Ausgabenposten des Projektes sind Personal- und Rechnerkosten. Die Anwendungen der TA1-3 sind rechenaufwendig, so dass eine Cluster-Infrastruktur mit 20 mächtigen Servern (Dell Quad-Core mit 16GB Hauptspeicher) sowie 2 Servern mit bis zu 48GB Hauptspeicher zum Projektanfang angeschafft wurde.

Notwendigkeit und Angemessenheit der geleisteten Arbeit

Die mit dem CLSD verbundenen Ziele der Forschungsgruppe um Prof. Schroeder konnten im Rahmen des ForMaT-Projektes unter sehr guten Bedingungen realisiert werden. Die Schaffung nachhaltiger Strukturen und die Verwertung der Technologieansätze konnten mit der Forschungsgruppe zusammen vorangetrieben werden. Dessen Ergebnisse lieferten einen Mehrwert für die Gruppe, den Standort und die Region. Dieses ist insbesondere durch 3 Patenteinreichungen und 22 Publikationen in Top-Zeitschriften, sowie Kollaborationen mit Firmen und dem Geschäftsplan für eine Ausgründung dokumentiert.

Voraussichtlicher Nutzen und Verwertbarkeit des Ergebnisses

Für TA1 gibt es eine enge Zusammenarbeit mit der Transinsight GmbH. Die Transinsight GmbH entwickelt Produkte und Lösungen für semantische Suchtechnologien in den Lebenswissenschaften. Kunden von Transinsight sind u.a. Roche (D), Unilever (UK), das Bundesinstitut für Risikobewertung (D), AbCam (UK), European Center for Disease Control (S), BASF (D), Wintershall (D), StatoilHydro (NO). Transinsight ist ein Start-up, das 2006 als Leutturmprojekt des HTGF gewürdigt wurde.. Dr. Michael R. Alvers ist Transinsights Geschäftsführer. Er hat über 18 Jahre Managementenerfahrung in Technologiefirmen wie Definiens AG und interActive Systems GmbH in Positionen als Direktor für Bioinformatik, Direktor für Produktmarketing und Direktor für Geschäfts- und Produktentwicklung.

Für TA2 gibt es eine enge Zusammenarbeit mit Qualitype. Die Dresdner Qualitype AG verfügt über umfangreiche Erfahrung in der Entwicklung von Software für die Bereiche human- und veterinärmedizinische Diagnostik, pharmazeutische Produktentwicklung, Qualitätssicherung in der Nahrungsmittelindustrie, Forensik und Bioinformatik. Die Produktpalette reicht hier von Einzelplatzsystemen bis zu Client-Server-Anwendungen bzw. internetbasierten Systemen oder kompletten Laborinformationsmanagementsystemen (LIMS). Dr. Frank Götz ist promovierter Biologe und verfügt über eine langjährige Erfahrung im Bereich der Bio- und Chemo-Informatik. Er ist Vorstand und Inhaber der Qualitype AG, Leiter der Arbeitsgruppe von Biosaxony und arbeitete u.a. für Firmen wie Merck KGaA, Boehringer Ingelheim und Mermaid Pharmaceuticals GmbH.

Für TA3 wurde ein Geschäftsplan für eine Ausgründung der nanometis entwickelt. Dieser Geschäftsplan gewann:

- den 1. Platz beim FutureSax Businessplan Wettbewerb. Leipzig. 6.2009
- den 2. Platz beim Science4Life Venture Cup 2009, Berlin, 6.2009
- den 3. Platz beim Businessplan Wettbewerb Medizinwirtschaft 2009, Essen, 12.2009

- den 3. Platz beim Thüringer Elevator Pitch. Bad Köstritz. 2.2010

Es wurden Gespräche mit Investoren wie dem HTGF, TGFS und dem LSI Bonn geführt.

Fortschritt auf dem Gebiet des Vorhabens bei anderen Stellen

Die drei Patente und 22 Publikationen zeigen, dass CLSD stetig den Fortschritt des Gebietes verfolgt hat. Die Patente und Artikel werden begutachtet, so dass sichergestellt ist, dass sie jeweils den State-of-the-art kennen und erweitern.

Erfolgte oder geplante Veröffentlichungen der Ergebnisse

TA1: Technologievorstellung auf Konferenzen, Workshops, etc.

Datum	Veranstaltung	Aktivität	Ort	Personen
März 2010	GeneOntology meeting	Talk	San Francisco, US	Michael Schroeder
März 2010	Analytica	Messestand	München, DE	Dr. Frank Dressel, Dr. Bill Andreopoulos, Dr. Liliana Barrio-Alvers
März 2010	Cebit	Messestand	Hannover, DE	Dr. Liliana Barrio-Alvers
Januar 2010	Computational systems biology approaches for cancer research and biomarker discovery	Talks	Hangzhou, CN	Prof. Dr. Michael Schroeder, Dr. Bingding Huang, Dr. Christof Winter, Loic Royer
November 2009	Ontology workshop	Talks	Leipzig, DE	Prof. Dr. Michael Schroeder, Thomas Wächter, Heiko Dietze, Conrad Plake, Rainer Winnenburg
November 2009	SWAT4LS	Keynote Talk	Amsterdam, NL	Prof. Dr. Michael Schroeder
November 2009	Retreat	Workshop	Bad Schandau, DE	Alle
Oktober 2009	PHPPBC	Invited Talk	Dresden, DE	Prof. Dr. Michael Schroeder, Dr. Christof Winter, Dr. Annalisa Marsico, Janine Roy, Loic Royer, Anne Tuukkanen, Conrad Plake, Rainer Winnenburg, Thomas Wächter
Oktober 2009	Mittweida Informatik Tag	Talks	Mittweida, DE	Dr. Frank Dressel, Matthias Reimann, Prof. Dr. Dirk Labudde
Oktober 2009	Minds at work, iSoft conference	Keynote Talk	London, UK	Prof. Dr. Michael Schroeder
September 2009	German Conference on Bioinformatics	Talk	Halle, DE	Prof. Dr. Michael Schroeder, Janine Roy, Loic Royer, Conrad Plake

September 2009	GO meeting	Talk	Cambridge, UK	Prof. Dr. Michael Schroeder, Thomas Wächter
August 2009	World Congress on Alternatives and Animal Use in the Life Sciences	Talk	Rom, IT	Prof. Dr. Michael Schroeder, Thomas Wächter
Juli 2009	Optimised Lead Discovery	Invited Talk	Berlin, DE	Prof. Dr. Michael Schroeder
Juni 2009	ISMB/ECCB2009	Talk	Stockholm, SE	Prof. Dr. Michael Schroeder, Prof. Dr. Dirk Labudde, Anne Tuukkanen
Juni 2009	Visual Document Analytics	Talk	Leipzig, DE	Prof. Dr. Michael Schroeder, Dimitra Alexopoulou, Rainer Winnenburg, Heiko Dietze, Thomas Wächter, Conrad Plake
Mai 2009	Saechsischer Biotechnologietag 2009	Poster	Leipzig, DE	Dr. Christof Winter, Anne Tuukkanen, Conrad Plake, Matthias Reimann, Loic Royer, Rainer Winnenburg
April 2009	3rd International Biocuration Conference	Talks	Berlin, DE	Thomas Wächter, Rainer Winnenburg, Conrad Plake

TA2: Technologievorstellung auf Konferenzen, Workshops, etc.

Datum	Veranstaltung	Aktivität	Ort	Personen
Januar 2010	Computational systems biology approaches for cancer research and biomarker discovery	Talks	Hangzhou, CN	Prof. Dr. Michael Schroeder, Dr. Bingding Huang, Dr. Christof Winter, Loic Royer
November 2009	JCB		Jena, DE	Janine Roy
November 2009	Visual analytics SPP meeting	Talk	Braunschweig, DE	Matthias Reimann
November 2009	Retreat	Workshop	Bad Schandau, DE	Alle
Oktober 2009	PHPPBC	Invited Talk	Dresden, DE	Prof. Dr. Michael Schroeder, Dr. Christof Winter, Dr. Annalisa Marsico, Janine Roy, Anne Tuukkanen, Loic Royer, Conrad Plake, Rainer Winnenburg
Oktober 2009	Mittweida Informatik Tag	Talk	Mittweida, DE	Prof. Dr. Dirk Labudde, Dr. Frank Dressel, Matthias Reimann

Oktober 2009	Bioinformatics meeting of SPP Visual Analytics	Talks	Leipzig, DE	Prof. Dr. Michael Schroeder, Matthias Reimann
September 2009	German Conference on Bioinformatics	Talk	Halle, DE	Prof. Dr. Michael Schroeder, Janine Roy, Loic Royer, Conrad Plake
September 2009	Workshop Bioinformatik & Sysbio	Workshop	Berlin, DE	Prof. Dr. Dirk Labudde, Joscha Köllner
Juli 2009	Optimised Lead Discovery	Invited Talk	Berlin, DE	Prof. Dr. Michael Schroeder
Juni 2009	ISMB/ECCB2009	Talk	Stockholm, SE	Prof. Dr. Michael Schroeder, Prof. Dr. Dirk Labudde, Anne Tuukkanen
Juni 2009	Visual Document Analytics	Talk	Leipzig, DE	Prof. Dr. Michael Schroeder, Dimitra Alexopoulou, Rainer Winnenburger, Heiko Dietze, Thomas Wächter, Conrad Plake
Mai 2009	Sächsischer Biotechnologietag 2009	Poster	Leipzig, DE	Dr. Christof Winter, Anne Tuukkanen, Conrad Plake, Matthias Reimann, Rainer Winnenburger, Loic Royer
April 2009	3rd International Biocuration Conference	Talks	Berlin, DE	Thomas Wächter, Rainer Winnenburger, Conrad Plake

TA3: Technologievorstellung auf Konferenzen, Workshops, etc.

Datum	Veranstaltung	Aktivität	Ort	Personen
März 2010	Kooperationsgespräch GWT-TUD GmbH	Präsentation	Dresden, DE	Joscha Köllner, Dr. Frank Dressel
März 2010	Analytica	Messestand	München, DE	Dr. Frank Dressel, Dr. Liliana Barrio-Alvers, Dr. Bill Andreopoulos, Joscha Köllner
März 2010	Kooperationsgespräch	Präsentation	Düsseldorf, DE	Prof. Dr. Dirk Labudde, Dr. Bill Andreopoulos, Joscha Köllner
März 2010	Science4Life	Präsentation	Berlin, DE	Joscha Köllner
Februar 2010	6. Thüringer Elevator Pitch	Präsentation	Bad Köstritz, DE	Dr. Frank Dressel, Dr. Bill Andreopoulos
Februar 2010	Kooperationen zwischen Wissenschaft und Wirtschaft	Workshop	Berlin, DE	Joscha Köllner
Februar 2010	12th Annual Linz Winter Workshop	Workshop	Linz, AT	Prof. Dr. Dirk Labudde, Dr. Frank Dressel

Januar 2010	LifeScience Inkubator	Präsentation	Bonn, DE	Dr. Frank Dressel, Joscha Köllner
Januar 2010	Workshop/ Kooperationsstreffen	Präsentation	Marburg, DE	Dr. Bill Andreopoulos
Januar 2010	Messevorbereitung Biotechnica	Workshop	Berlin, DE	Dr. Frank Dressel, Joscha Köllner
Januar 2010	Computational systems biology approaches for cancer research and biomarker discovery	Talks	Hangzhou, CN	Prof. Dr. Michael Schroeder, Dr. Bingding Huang, Dr. Christof Winter, Loic Royer
Dezember 2009	Startbahn Ruhr Businessplanwettbewerb Medizinwirtschaft	Wettbewerb	Essen, DE	Joscha Köllner, Dr. Bill Andreopoulos
November 2009	Retreat	Workshop	Bad Schandau, DE	Alle
Oktober 2009	PHPPBC	Invited Talk	Dresden, DE	Prof. Dr. Michael Schroeder, Dr. Christof Winter, Dr. Annalisa Marsico, Janine Roy, Loic Royer, Rainer Winnenburg, Thomas Wächter, Anne Tuukkanen
Oktober 2009	Mittweida Informatik Tag	Talks	Mittweida, DE	Prof. Dr. Dirk Labudde, Dr. Frank Dressel, Matthias Reimann
Oktober 2009	Biotechnica	Messestand	Hannover, DE	Dr. Frank Dressel, Dr. Bill Andreopoulos, Joscha Köllner
Oktober 2009	Science4Life Messe	Messestand	Frankfurt, DE	Dr. Frank Dressel, Dr. Bill Andreopoulos, Joscha Köllner
Oktober 2009	What is a macromolecular Complex? Shades of Meaning Across Cellular, Systems, and Structural Biology	Talks	Amsterdam, NL	Dr. Annalisa Marsico, Anne Tuukkanen
Oktober 2009	Kooperationsgespräche Merck KGaA	Präsentation	Darmstadt, DE	Prof. Dr. Dirk Labudde, Joscha Köllner
September 2009	German Conference on Bioinformatics	Talk	Halle, DE	Prof. Dr. Michael Schroeder, Janine Roy, Loic Royer, Conrad Plake
September 2009	Seed4money	Präsentation	Frankfurt, DE	Joscha Köllner
September 2009	Kooperationsgespräche Sanofi Aventis	Präsentation	Frankfurt, DE	Prof. Dr. Dirk Labudde, Joscha Köllner
Juli 2009	Optimised Lead Discovery	Invited Talk	Berlin, DE	Prof. Dr. Michael Schroeder

Juni 2009	Science4Life Venture Cup	Wettbewerb	Frankfurt, DE	Joscha Köllner, Dr. Bill Andreopoulos
Juni 2009	FutureSAX Businessplan-Wettbewerb	Wettbewerb	Dresden, DE	Joscha Köllner, Prof. Dr. Dirk Labudde, Dr. Frank Dressel
Juni 2009	ISMB/ECCB2009	Talk	Stockholm, SE	Prof. Dr. Michael Schroeder, Prof. Dr. Dirk Labudde, Anne Tuukkanen
Mai 2009	Saechsischer Biotechnologietag 2009	Poster	Leipzig, DE	Dr. Christof Winter, Anne Tuukkanen, Conrad Plake, Matthias Reimann, Rainer Winnenburg, Loic Royer

Patent und Publikationen zu TA1 ProWiki

Patente:

TU Dresden. **Datenverarbeitungssystem**. Filed Patent (German Patent Application Number DE 10 2009 043 708.8), 2009

TU Dresden. **Verfahren und Mittel zur Vorhersage der Überlebensdauer beim Pankreaskarzinom durch Analyse von Biomarkern**. Filed Patent (German Patent Application Number DE 10 2009 054 793.2), 2009

nanometis and TU Dresden. **Automatic analysis of bio-molecules**. Filed Patent (European Patent Application Number EPO 9180269.4), 2009

Publikationen:

Old Friends in new Guise: Repositioning of known Drugs with Structural Bioinformatics. Joachim Haupt and Michael Schroeder. Briefings in Bioinformatics, 2011

Molecular Fragmentation Query Language: a Novel Informatics Concept for High-Throughput Shotgun Lipidomics. Ronny Herzog, Dominik Schwudke, Kai Schuhman, Stefan Bornstein, Michael Schroeder, and Andrej Shevchenko. Accepted for Genome Biology, 2011

Whole blood genome-wide expression profiling and network analysis suggest MELAS master regulators. Susanne Mende, Loic Royer, Alexander Herr, Janet Schmiedel, Marcus Deschauer, Thomas Klopstock, Vladimir S. Kostic, Michael Schroeder, Heinz Reichmann, and Alexander Storch. Accepted for Neurological Research, 2011, Impact factor 2009: 1.8

Predicting drug-target interactions with ontological profile similarity. Conrad [Plake](#) and Michael [Schroeder](#). Submitted, 2010

Semi-automated ontology generation within OBO-Edit. Thomas [Wächter](#) and Michael [Schroeder](#). In Bioinformatics, 2010, Impact factor 2008: 4.3

SuperCYP: a comprehensive database on Cytochrome P450 enzymes including a tool for analysis of CYP-drug interactions. Saskia Preissner, Katharina Kroll, Mathias Dunkel, Christian Senger, Gady Goldsobel, Daniel Kuzman, Stefan Guenther, Rainer [Winnenburg](#), Michael [Schroeder](#), and Robert Preissner. Nucleic Acids Research, 38(Database issue):D237--D243, 2010, Impact factor 2008: 6.9, [[paper](#)]

GoGene: gene annotation on the fast lane. Conrad [Plake](#), Loic [Royer](#), Rainer [Winnenburg](#), Jörg [Hakenberg](#), and Michael [Schroeder](#). Nucleic Acid Research, 37(2):W300--4, 2009, Impact factor 2008: 6.9, [[paper](#)]

GoWeb: A semantic search engine for the life science web. Heiko [Dietze](#) and Michael [Schroeder](#). BMC Bioinformatics, 10(Suppl 10), 2009, Accepted. Impact factor 2008: 3.8, [[paper](#)]

A User-Centered Evaluation Framework for the Sealife Semantic Web Browsers. Helen Oliver, Gayo Diallo, Ed de Quincey, Patty Kostkova, Gawesh Jawaheer, Dimitra [Alexopoulou](#), Bianca Habermann, Robert Stevens, Simon Jupp, Khaled Khelif, Michael [Schroeder](#), and Gemma Madle. BMC Bioinformatics, 10(Suppl 10), 2009, Impact factor 2008: 3.8, [[paper](#)]

Go3R - Semantic Internet Search Engine for Alternative Methods to Animal Testing. Ursula G. Sauer, Thomas [Wächter](#), Barbara Grune, Andreas [Doms](#), Michael R. Alvers, Horst Spielmann, and Michael [Schroeder](#). ALTEX, 26(1):17--31, 2009, [[paper](#)]

GoPubMed: Exploring PubMed with Ontological Background Knowledge. Heiko [Dietze](#), Dimitra [Alexopoulou](#), Michael R. Alvers, Liliana Barrio-Alvers, Bill [Andreopoulos](#), Andreas [Doms](#), Jörg [Hakenberg](#), Jan Mönnich, Conrad [Plake](#), Andreas Reichsuck, Loic [Royer](#), Thomas [Wächter](#), Matthias Zschunke, and Michael [Schroeder](#). In Stephen A. Krawetz, editor, Bioinformatics for Systems Biology. Humana Press, 2009, Cited by 8 TU Dresden.

Network compression as a quality measure for protein interaction networks. Loic [Royer](#), A Francis Stewart, and Michael [Schroeder](#). Submitted, 2010

A Genome-wide expression profiling and functional network analysis reveal master regulators of neuroectodermal conversion of mesenchymal stem cells. Martina Maisel, Hans-Jörg Habisch, Loic [Royer](#), Alexander Herr, Javorina Milosevic, Stefan Liebau, Rolf Brenner, Johannes Schwarz, Michael [Schroeder](#), and Alexander Storch. Experimental Cell Research, 2010

Earlier differentiation of mesenchymal stem cells cultivated on smooth tantalum compared with titanium surfaces. Claudia Stiehler, Cody Bünger, Rupert Overall, Loic [Royer](#), Michael [Schroeder](#), Morten Foss, Flemming Besenbacher, Mogens Kruhoffer, Moustapha Kassem, Klaus-Peter Günther, and Maik Stiehler. Submitted, 2010

Examination of Apoptosis Signaling in Pancreatic Cancer by Computational Signal Transduction Analysis. Felix Rückert, Gihan Dawelbait, Christof Winter, Arndt Hartmann, Axel Denz, Ole Ammerpohl, Michael Schroeder, Hans Konrad Schackert, Bence Sipos, Günter Klöppel, Holger Kalthoff, Hans-Detlev Saeger, Christian Pilarsky, and Robert Grützmann. PLoS One, 5(8):e12243, 2010, Impact factor 2009: 4.4

Xenopus meiotic microtubule-associated protein interaction network. Vincent Gache, Patrice Waridel, Christof [Winter](#), Aurelie Juhem, Michael [Schroeder](#), Andrej Shevchenko, and Andrei V. Popov. PLoS One, 5(2):e9248, 2010, [[paper](#)]

Triangle network motifs predict complexes by complementing high-error interactomes with structural information. Bill [Andreopoulos](#), Christof [Winter](#), Dirk [Labudde](#), and Michael [Schroeder](#). BMC Bioinformatics, 10(196), 2009, Impact factor 2008: 3.8, [[paper](#)]

A Rapid changes of mRNA binding protein levels following glucose and IBMX stimulation of insulinoma INS-1 cells. Christin Süß, Cornelia Czupalla, Christof [Winter](#), Theresia Pursche, Klaus Knoch, Michael [Schroeder](#), Bernard Hoflack, and Michele Solimena. Molecular and Cellular Proteomics, 8(3):393--408, 2009, Impact factor 2008: 8.8, [[paper](#)]

A Identifying hot cancer and novel novel cell-cycle genes from literature. Loic [Royer](#), Conrad [Plake](#), and Michael [Schroeder](#). In Peter Stadler and Stefan Posch, editors, Proceedings of the German Conference on Bioinformatics. Gesellschaft für Informatik, 2009

Structural fragment clustering reveals novel structural and functional motifs in alpha-helical transmembrane proteins. Annalisa [Marsico](#), Andreas [Henschel](#), Christof [Winter](#), Anne [Tuukkanen](#), Boris Vassilev, Kerstin Scheubert, and Michael [Schroeder](#). To appear in BMC Bioinformatics, 2010, Impact factor 2008: 3.8

MeMotif: Structural fragment clustering reveals novel structural and functional motifs in alpha-helical transmembrane proteins. Annalisa [Marsico](#), Kerstin Scheubert, Andreas [Henschel](#), Christof [Winter](#), Rainer [Winnenburg](#), Anne [Tuukkanen](#),

Boris Vassilev, and Michael Schroeder. Nucleic Acids Research, 38(Database issue):D181--D189, 2010, Impact factor 2008: 6.9, [[paper](#)]

Improved mutation tagging with gene identifiers applied to membrane protein stability prediction. Rainer Winnenburg, Conrad Plake, and Michael Schroeder. BMC Bioinformatics, 10(Suppl 8), 2009, Impact factor 2008: 3.8, [[paper](#)]

A. Preise:

Simone Daminelli, Joachim Haupt. 1st Poster prize at Biotec Forum. 5.2011

Janine Roy and Matthias Reimann. 2nd Poster prize at Saxon Biotechnology Day SBTT. 5.2011

GoPubMed gewinnt Deutschen Innovationspreis (43/2000). 3.2011

Nomination of Dr. Loic Royer for PhD thesis award of German Society of Computing. 2.2011

Dr. Christof Winter wins Klaus-Tschira Prize for understandable science. 14.10.2010

GoPubMed gewinnt Deutschen Industriepreis 2010 in der Kategorie Forschung und Entwicklung, Hannover. 19.4.2010

Dr. Andreas Doms gewinnt TU Dresden Georg Helm Dissertationspreis. Dresden. 10.2009

GoPubMed gewinnt Red Dot Design Award. Wuppertal, Germany, 8.2009

nanometis gewinnt 1.Platz beim FutureSax Businessplan Wettbewerb. Leipzig. 6.2009

nanometis gewinnt 2.Platz beim Science4Life Venture Cup 2009, Berlin, 6.2009

nanometis gewinnt 3.Platz beim Businessplan Wettbewerb Medizinwirtschaft 2009, Essen, 12.2009

nanometis gewinnt 3.Platz beim Thüringer Elevator Pitch. Bad Köstritz. 2.2010