

## Kapitel 4

# Entwicklung von Datenbankanwendungen

### 4.1 InvariomTool

#### 4.1.1 Motivation

Das von Birger Dittrich et al. eingeführte Invariommodell [27, 28] ermöglicht es, das Modell der unabhängigen Atome (IAM) durch ein asphärisches Modell zu ersetzen, welches Bindungseffekte beschreiben kann. Da beim Invariommodell im Vergleich zum IAM keine weiteren Parameter verfeinert werden müssen, benötigt man hierzu nicht die für die klassische Multipolverfeinerung notwendige hohe Auflösung von  $>1.1 \text{ \AA}^{-1}$  in  $\frac{\sin\theta}{\lambda}$ . Es genügt eine Auflösung von  $0.6 \text{ \AA}^{-1}$  in  $\frac{\sin\theta}{\lambda}$ . Beim Invariommodell werden die Multipolpopulationen aus theoretischen Optimierungen von kleinen Modellverbindungen, welche dieselbe chemische Umgebung besitzen wie das jeweilige Atom, erhalten. Diese Multipolpopulationen sind Teil einer Datenbank. Um diese Datenbank aufzubauen und auch auszulesen, wurde in Zusammenarbeit mit Birger Dittrich eine Anwendung entworfen, die diese Aufgaben automatisiert. InvariomTool analysiert dabei die Struktur anhand von Bindungsabständen und ermittelt so den systematischen Namen des für jedes Atom notwendigen Invarioms. Diese Invariome werden dann in der Datenbankdatei gesucht und die Multipolpopulationen zusammen mit den Informationen zur Wahl des lokalen Koordinatensystems in für die weitere Verfeinerung mit XD [29] nötige Dateien geschrieben.

### 4.1.2 Ermittlung von Konnektivität und Bindungsordnung

Um die chemische Umgebung, das heißt die nächsten beziehungsweise übernächsten Nachbarn, eines Atoms zu ermitteln, ist es nötig, die Geometrie der untersuchten Struktur zu ermitteln. Dazu werden zunächst die Koordinaten eingelesen, wie sie nach der konventionellen Verfeinerung vorliegen. Die Abstände der Atome zueinander werden mit Kovalenzradien von Einfachbindungen aus der Literatur verglichen. Da eine Elektronegativitätsdifferenz  $EN$  zwischen den Atomen die Bindungslängen erheblich verkürzen kann, wird dieser Verkürzung in folgender Formel Rechnung getragen [30].

$$0.85 \cdot d \leq r_{Kovalenz}(atom1) + r_{Kovalenz}(atom2) - 0.08 \cdot |\Delta(EN)| \quad (4.1)$$

Ist die um eine skalierte Elektronegativitätsdifferenz verringerte Summe der Kovalenzradien der beteiligten Atome kleiner gleich deren um 15% verringerten tatsächlichen Abstands zueinander, so werden diese als aneinander gebunden betrachtet. Im folgenden Schritt werden die formalen Bindungsordnungen aller gefundenen Bindungen zugeordnet. Dazu wird Ungleichung 4.1 wie folgt modifiziert:

$$\chi = [r_{Kovalenz}(atom1) + r_{Kovalenz}(atom2) - 0.08 \cdot |\Delta(EN)|] - d. \quad (4.2)$$

Der so erhaltene Bindungsordnungsfaktor  $\chi$ , der desto größer wird, je kürzer der Atomabstand ist, wird dann mit anpassbaren Bindungsordnungsgrenzen verglichen. Die typischen Bindungsordnungsgrenzen sind 0.0847 Å zwischen einfachen und 1.5-fachen Bindungen, 0.184 Å zwischen 1.5-fachen und zweifachen Bindungen und 0.27 Å zwischen zweifachen und dreifachen Bindungen. Bindungen zu Wasserstoff werden immer als einfach betrachtet. So erhält jedes Atom eine Liste mit der Anzahl seiner direkt gebundenen Nachbarn, der Atomsorte derselben und des jeweiligen Bindungstyps. Diese Liste wird zunächst nach abfallendem  $\chi$  und dann nach absteigender Ordnungszahl der Atomsorten sortiert.

### 4.1.3 Zuordnung von systematischen Namen

Eine Besonderheit beim Multipoltransfer durch Invariome ist die eindeutige und nachvollziehbare Nomenklatur der einzelnen Invariome. Diese folgt klaren Regeln: Das Elementsymbol des betreffenden Atoms steht am Anfang des Invariomnamens, beginnend mit Großbuchstaben. Dann folgen die Atomsymbole der Nachbaratome in Kleinbuchstaben, angeführt von deren Bindungsord-

nung. Die Reihenfolge ist dieselbe wie die im vorherigen Abschnitt beschriebene. Bei Wasserstoffatomen und bei Bindungsordnungen von 1.5 werden auch die übernächsten Nachbarn in den Namen mit einbezogen. Diese folgen dann in eckigen Klammern [] dem Atom, an welches sie gebunden sind. Sie sind ebenfalls in Kleinbuchstaben mit davorstehenden Bindungsordnungen zu schreiben. Sind die Invariome chiral, was der Fall ist, wenn sie vier unterschiedliche nächste Nachbarn besitzen, dann ist das Chiralitätssymbol nach Cahn, Ingold und Prelog [31] in Großbuchstaben, gefolgt von einem Bindestrich, dem Invariomnamen voranzustellen. Sind die dem Invariom entsprechenden Modellverbindungen geladen und dieses nicht aus der Konnektivität ersichtlich, so wird das Vorzeichen, durch einen Unterstrich getrennt, dem Invariomnamen angefügt. Das hier besprochene Programm InvariomTool ist in der Lage, den Namen nach diesen Regeln automatisch zu erstellen. Nur die Ladungen können derzeit nicht automatisch ermittelt werden. Sind die Atomabstände der konventionellen Struktur ungenügend oder liegen partielle Fehlorderungen vor, so scheidet die automatische Nomenklatur und der Nutzer muss die Namen manuell vorgeben. Beispiele für die Invariomnomenklatur sind dem Experimentellen Teil im Anschluss zu entnehmen.

#### 4.1.4 Suche der Invariome in der Datenbank

Sind die Invariomnamen aller Atome in der untersuchten Verbindung bekannt, wird die Datenbankdatei, eine einfache Textdatei, geöffnet und nach den passenden Einträgen durchsucht. Ein Datenbankeintrag ist wie folgt aufgebaut:

```
Invariomname
Pv P00 P11 P1-1 P10 P20 P21 P2-1 P22 P2-2
P30 P31 P3-1 P32 P3-2 P33 P3-3 P40 P41 P4-1
P42 P4-2 P43 P4-3 P44 P4-1
Symmetrie: Punktgruppen-Symmetrie-Symbol nach Hermann-Maugin
KS: 1.Achse:Atomname1 2.Achse:Atomname2 AX1:Elementsymbol1 chi1 AX2:Elementsymbol2 chi2
Kappa=kappa,kappa0,kappa1,kappa2,kappa3,kappa4
```

PNN sind die einzelnen Multipolparameter. Die Zeile, die mit KS: beginnt, verschlüsselt die für die lokalen Koordinatensysteme notwendigen Informationen.

Atomnamen können auch Dummyatome sein, welche entsprechend berechnet werden müssen. Der Datenbankeintrag einer kohlenstoffgebundenen O-H-Gruppe O1c1h sieht beispielsweise so aus.

```
O1c1h
  6.0097  0.0000 -0.0606 -0.1069  0.0000  0.1179  0.0000  0.0000 -0.0295  0.0499
  0.0000 -0.0160 -0.0254  0.0000  0.0000  0.0658 -0.0222  0.0117  0.0000  0.0000
  0.0056  0.0021  0.0000  0.0000  0.0141  0.0152
Symmetrie: m
KS: X:C(3) Y:H(2) AX1:C-0.02 AX2:H0.02
Kappa=1.004698;=1.000000;=1.000000;=1.000000;=1.000000;=1.000000
```

Die in der Datenbank gefundenen Informationen werden dann in die entsprechenden XD-Dateien `xd.mas` und `xd.inp` übertragen. Dabei werden auch für das lokale Koordinatensystem eventuell benötigte Dummyatome korrekt über Kreuzprodukte oder Vektorsummen berechnet.

#### 4.1.5 Durchforstung von Verzeichnisbäumen nach passenden Dateien zum Datenbankaufbau

Sind die gewünschten Invariome noch nicht in der Datei zu finden, wird ein Verzeichnisbaum, in dem sich die Ergebnisse theoretischer Rechnungen an zahlreichen Modellverbindungen befinden, durchforstet. Befinden sich dort bereits passende Ergebnisse, werden diese automatisch der Datenbank hinzugefügt. Um dies zu ermöglichen, muss der Verzeichnisbaum wie folgt geordnet sein: Jede Optimierung einer Modellverbindung muss sich in einem Unterverzeichnis befinden, welches den richtigen Invariomnamen trägt. Die XD-Verfeinerungen der theoretischen Strukturaktoren sollen sich in Unterverzeichnissen mit dem Namen `xd_k` für Verfeinerungen ohne  $\kappa'$ -Parameter oder `xd_kk` für Verfeinerungen mit  $\kappa'$ -Parametern befinden.

#### 4.1.6 Anpassung der Monopolladungen

Führt man den Multipoltransfer durch, stellt man fest, dass die Summe der Monopolladungen bis auf eine kleine Abweichung der Anzahl der Valenzelektronen in der asymmetrischen Einheit entspricht. Um einen geladenen Kristall zu vermeiden, muss die Summe der Monopolladungen aber exakt der Summe der Valenzelektronen entsprechen. InvariomTool kennt mehrere Verfahren, um die Monopolladungen dementsprechend zu korrigieren.

1. Das einfachste Verfahren addiert jedem Atom den Durchschnittswert der Abweichung zwischen der Summe der Monopolladungen  $P_v$  und der Summe der Valenzelektronen  $V$  hinzu. Im Folgenden ist  $N_A$  die Anzahl der Atome in der asymmetrischen Einheit.

$$P'_v = \left( \frac{\sum_{i=1}^{N_A} V_i - \sum_{i=1}^{N_A} P_{vi}}{N_A} \right) + P_v \quad (4.3)$$

2. Da meist die Wasserstoffatome durch ihre hohe Polarisierbarkeit den Großteil der Ladungen tragen, wird bei einem weiteren Verfahren der oben genannte Unterschied nur auf die  $n_h$  Wasserstoffatome verteilt.

$$P'_v(H_{n_h}) = \left( \frac{\sum_{i=1}^{N_A} V_i - \sum_{i=1}^{N_A} P_{vi}}{n_h} \right) + P_v(H_{n_h}) \quad (4.4)$$

3. Bei diesem Verfahren wird die Allred-Rochow-Elektro negativität ( $EN$ ) [32] zur Gewichtung der Korrekturen benutzt. Dazu wird eine durchschnittliche Elektro negativität  $av.EN$  aus allen Atomen der asymmetrischen Einheit berechnet. Es werden zwei Fälle unterschieden: Wenn die Summe der Monopolladungen größer als die Summe der Valenzelektronen der neutralen Atome ist und wenn dies umgekehrt der Fall ist. Im ersten Fall wird der abzuziehende Unterschied so gewichtet, dass bei relativ hoher  $EN$  die Monopolladung stärker verringert wird als bei relativ geringer. Im zweiten Fall wird der dazuzuaddierende Unterschied so gewichtet, dass bei relativ hoher  $EN$  die Monopolladung schwächer vergrößert wird als bei relativ geringer.

$$\sum_{i=1}^{N_A} V_i - \sum_{i=1}^{N_A} P_{vi} < 0 : \quad (4.5)$$

$$P'_v(H_{n_h}) = \left[ \left( \frac{\sum_{i=1}^{N_A} V_i - \sum_{i=1}^{N_A} P_{vi}}{n_h} \right) \frac{EN_i}{av.EN} \right] + P_v(H_{n_h}) \quad (4.6)$$

$$\sum_{i=1}^{N_A} V_i - \sum_{i=1}^{N_A} P_{vi} > 0 : \quad (4.7)$$

$$P'_v(H_{n_h}) = \left[ \left( \frac{\sum_{i=1}^{N_A} V_i - \sum_{i=1}^{N_A} P_{vi}}{n_h} \right) \frac{av.EN}{EN_i} \right] + P_v(H_{n_h}) \quad (4.8)$$

### 4.1.7 Nutzung von InvariomTool zum Invariomtransfer

InvariomTool lässt sich ganz einfach zum Invariomtransfer verwenden. Dazu benötigt man lediglich einen atomar aufgelösten Einkristall-Röntgendatensatz und eine abgeschlossene konventionelle Strukturverfeinerung. Da InvariomTool derzeit noch ausschließlich auf das XD-Programmpaket [29] zugeschnitten ist, wird eine lizenzierte Version davon benötigt. Um die Einfachheit des Verfahrens zu beschreiben, soll hier in fünf Schritten der Gang einer Invariomverfeinerung beschrieben werden:

1. Konvertierung der Strukturlösungsdateien in das XD-Format. Hierzu wird das XD-Programm XDINI in derselben Weise aufgerufen, wie es zu Beginn einer klassischen Multipolverfeinerung üblich ist. XDINI schreibt die Dateien `xd.mas`, `xd.inp` und `xd.hkl`.
2. Nun wird das Programm InvariomTool das erste Mal aufgerufen. Befindet sich im Arbeitsverzeichnis noch keine Datei mit dem Namen `invtool.ini`, so bricht InvariomTool mit einer Fehlermeldung ab. Zuvor schreibt es jedoch eine Datei `example4invtool.ini`, welche die Standardwerte für die Bindungsordnungsgrenzen, den benötigten Pfad zur Datenbankdatei, und die Behandlungsart von Monopolladungen enthält. Nachdem der Nutzer diese Werte entsprechend angepasst hat, muss er sie nur noch in `invtool.ini` umbenennen.
3. Es folgt der zweite Programmaufruf von InvariomTool. Bei diesem werden alle zur weiteren Verfeinerung nötigen Informationen inklusive der zu transferierenden Multipolparameter in die Dateien `xd.mas.invtool` und `xd.res.invtool` geschrieben. Mit dem Programm Molecool kann man sich von der Korrektheit der lokalen Koordinatensysteme überzeugen. Zur weiteren Verfeinerung sollte man diese Dateien in `xd.mas` und `xd.inp` umbenennen.
4. Verfeinerung: Mit dem XD-Programm XDLSM führt man dann die Verfeinerung der Ortskoordinaten und Temperaturparameter durch. Dabei beginnt man von Anfang an mit dem Multipollevel  $l = 4$ . Nach der Verfeinerung ist man eigentlich schon fertig.
5. Theoretische X—H-Abstände: InvariomTool bietet die Möglichkeit, die X—H-Abstände aus den theoretischen Optimierungen von Modellverbindungen zu übertragen. Diese sind bis auf 1–4% dieselben wie die aus gemittelten Neutronenexperimenten erhaltenen. InvariomTool schreibt die

dazu notwendigen „RESET BOND“-Befehle in eine Datei mit dem Namen `resetbond.txt`. Der Nutzer muss diese nur noch an die passende Stelle in `xd.mas` kopieren und die Verfeinerung der Ortskoordinaten der Wasserstoffatome dort ausschalten. Nach einer weiteren Verfeinerung mit XDLSM ist der Invariomtransfer komplett vollendet.

#### 4.1.8 Nutzung von InvariomTool zur Automatisierung klassischer Multipolverfeinerung

Die Durchführung einer klassischen Multipolverfeinerung ist ein zeitaufwendiges Unterfangen. Lokale Koordinatensysteme müssen passend zu der lokalen Symmetrie, die man verwenden will, aufgestellt werden. Dabei sind manchmal auch Dummyatome zu berechnen. Dies ist meist sehr aufwendig und es schleichen sich sehr schnell Fehler ein, die schwer nachzuvollziehen sind. Für die X—H Bindungen sind in den International Tables for Crystallography passende gemittelte Neutronenabstände zu finden und in die Verfeinerung einzufügen. Da bei der Wahl der lokalen Koordinatensysteme oft mehrere Möglichkeiten bestehen, ist das Einführen von chemischen *constraints* meist sehr kritisch, da dann die Koordinatensysteme identisch gewählt werden müssen.

Die Funktionen von InvariomTool lassen sich auch dazu nutzen, die oben genannten Probleme elegant zu vermeiden. Das Prozedere ist in [33] beschrieben und dem im vorherigen Abschnitt sehr ähnlich. Statt der Datei `xd.mas.invariomtool` wird die Datei `xd.mas.HasSameSym` verwendet. Diese Datei ermöglicht die Verfeinerung der Multipolparameter entsprechend der Symmetrie und chemischer *constraints*.

#### 4.1.9 Programmdetails

InvariomTool ist in C++ geschrieben und in über 30 Unterprogramme und Funktionen gegliedert. InvariomTool besteht aus 3219 durch Semikola getrennten Anweisungen. Die ausführbaren Dateien sind bei den beiden Autoren Christian B. Hübschle und Birger Dittrich zu beziehen.

