Dissertation ETH No. 11422

Zeolite Structure Determination from Powder Data: Computer-based Incorporation of Crystal Chemical Information

Dissertation

submitted to the
Swiss Federal Institute of Technology
Zurich
for the degree of Doctor of Natural Sciences

Presented by
Ralf Wilhelm Grosse-Kunstleve
dipl. Mineralogist (Ruhr-Universität Bochum)
born 3 April 1964 in Germany

Accepted on the recommendation of
Prof. Dr. M. Dobler
Dr. L.B. McCusker
Prof. Dr. W. Steurer

1996

Abstract

Many technologically and industrially important materials, including zeolites, are synthesized and used in polycrystalline form. Since the crystal structures of such phases often determine their useful properties, it is essential that methods to study their structures be developed. Rietveld refinement techniques can be used to extract structural details from a powder diffraction pattern, provided an approximate structure is known. However, if a structural model is not available, its determination from powder diffraction data is a non-trivial matter. The structural information contained in the diffracted intensities is obscured by systematic or accidental overlap of reflections in the powder pattern. As a consequence, the application of structure determination techniques which are very successful for single crystal data (primarily direct methods) is, in general, limited to simpler structures. Since zeolite structures tend to be too complex for this approach, the determination of their structures is still very much dominated by model building. An idea frequently suggested has been to include the active use of the same information on which model building is founded in an automated procedure. The realization of this idea was the predominant goal of this research project.

The FOCUS method, which incorporates the use of some of the chemical information used in model building into the structure determination process, has been developed. FOCUS combines automatic Fourier recycling (using integrated intensities extracted from a powder pattern and random starting phases) with a specialized framework search specific to zeolite structures, which can be described as 3-dimensional 4-connected topologies. The capabilities of FOCUS have been demonstrated with six test examples of medium to high complexity (zeolite topologies **DOH**, **LEV**, **RSN**, **AFR**, **LTA**, **EMT**). The proportion of overlapping reflections in these examples ranged from 15% to 83%, but the correct topology could be recovered in all cases. Furthermore, the examples show that the most frequently occurring topology produced by FOCUS is, in general, the correct solution, and that the procedure for extracting the intensities from the powder pattern plays a vital role in the outcome of a solution attempt.

The method was then applied to three novel zeolite structures – the two zincosilicates VPI-9 and VPI-10, and the beryllosilicate B2 – and a promising model was obtained in all cases. Preliminary Rietveld refinements of the VPI-9 and VPI-10 structures indicate that the proposed models are correct. The structure of VPI-9 has since been confirmed with a full Rietveld refinement, and the Structure Commission of the International Zeolite Association has assigned the code **VNI** to that topology. Refinements for VPI-10 and B2 are in progress.

Experience gathered during the course of this project shows that the approach of using chemical and geometrical knowledge can compensate for some of the information that is lost as a result of the overlap problem. At the same time, there is an intrinsic disadvantage: any method based on assumptions of certain structural properties is also limited to materials which conform to these assumptions. However, from the outset it has been foreseen that the basic idea of using crystal chemical information should also be applicable to other classes of materials. Two examples which show the consequences of relaxing the structural assumptions are presented. It was found that the computing time requirements of FOCUS grow very rapidly with the number of different possible connectivity types. Suggestions for further developments to overcome this problem are outlined.

Zusammenfassung

Viele technologisch und industriell wichtige Materialien, eingeschlossen Zeolithe, werden in polykristalliner Form synthetisiert und eingesetzt. Die nützlichen Eigenschaften dieser Phasen werden oft durch ihre Kristallstruktur bestimmt, woraus sich ein hoher Stellenwert für die Entwicklung von Verfahren zur Untersuchung dieser Strukturen ergibt. Ist ein Strukturmodell näherungsweise bekannt, lassen sich mit der Rietveld Technik strukturelle Details aus einem Pulverdiagramm ableiten. Die Bestimmung einer unbekannten Kristallstruktur aus Pulverdiffraktionsdaten ist jedoch keine leicht zu lösende Aufgabe, da die Strukturinformation, die in den abgebeugten Intensitäten enthalten ist, durch systematische oder zufällige Überlappungen im Pulverdiagramm verrauscht wird. Daraus ergibt sich, dass die Anwendbarkeit von Methoden zur Strukturbestimmung, die für Einkristalldaten sehr erfolgreich sind (vor allem Direkte Methoden), in der Regel auf einfachere Strukturen beschränkt ist. Auch Zeolithe tendieren dazu, für solche Lösungsverfahren zu komplex zu sein. Deshalb wird die Strukturbestimmung in diesem Bereich nach wie vor sehr von Modellbau beherrscht. Eine oft vorgeschlagene Idee ist die Einbeziehung derselben Information, auf die Modellbau begründet ist, in ein automatisiertes Verfahren. Die Umsetzung dieser Idee war das vorherrschende Ziel dieser Forschungsarbeit.

Die FOCUS Methode wurde entwickelt, bei der einige der Informationen in den Strukturbestimmungsprozess einfliessen, die auch im Modellbau benutzt werden. FOCUS kombiniert automatisiertes Fourier Recycling mit einer spezialisierten Gerüststruktursuche, die auf Zeolithe abzielt, welche als dreidimensional vierfachverknüpfte Topologien beschreibbar sind. Die Möglichkeiten von FOCUS werden anhand von sechs Testbeispielen mit mittlerer bis hoher Komplexität (Zeolith Topologien **DOH**, **LEV**, **RSN**, **AFR**, **LTA**, **EMT**) veranschaulicht. Der Anteil überlappender Reflexe in diesen Beispielen reicht von 15% bis 83%, doch in allen Fällen konnte die gesuchte Topologie zurückgewonnen werden. Darüber hinaus zeigen die Beispiele, dass die von FOCUS am häufigsten erzeugte Topologie in der Regel die gesuchte ist und dass die Art und Weise, wie die integrierten Intensitäten aus dem Pulverdiagramm extrahiert werden, eine überaus wichtige Rolle spielt.

Bei der Anwendung der Methode auf drei neuartige Zeolithstrukturen – die zwei Zinksilikate VPI-9 und VPI-10, sowie das Berylliumsilikat B2 – konnte in allen Fällen ein vielversprechendes Modell gefunden werden. Vorläufige Rietveld Verfeinerungen der Strukturen von VPI-9 und VPI-10 deuten darauf hin, dass die vorgeschlagenen Modelle richtig sind. Inzwischen konnte die Struktur von VPI-9 durch eine vollständige Rietveld Verfeinerung

bestätigt werden, und die Struktur-Kommission der Internationalen Zeolith-Vereinigung hat dieser Topologie offiziell den Kode **VNI** zugewiesen. Verfeinerungen für VPI-10 und B2 sind in Bearbeitung.

Die im laufe dieses Projektes gesammelte Erfahrung zeigt, dass die Benutzung chemischen und geometrischen Vorwissens teilweise den Informationsverlust ausgleichen kann, der durch das Überlappungsproblem entsteht. Diesem Ansatz ist jedoch auch ein wichtiger Nachteil zu eigen: eine Methode, die auf bestimmten strukturellen Annahmen basiert, ist gleichzeitig auch auf Materialien beschränkt, die diesen Annahmen entsprechen. Es wurde jedoch von Anfang an angenommen, dass die grundlegende Idee – die Einbeziehung kristallchemischer Information – auch auf andere Stoffklassen anwendbar sein sollte. An zwei Beispielen wird gezeigt, wie sich die Lockerung der strukturellen Annahmen auswirkt. Dabei stellte sich heraus, dass die Rechenzeitanforderungen von FOCUS sehr stark mit der Zahl unterschiedlicher Verknüpfungsmöglichkeiten ansteigen. Skizzenhaft werden Weiterentwicklungsmöglichkeiten vorgestellt, die diesem Problem begegnen.

Table of Contents

1	**Introduction**	**1**
1.1	Zeolites	1
1.2	Polycrystalline diffraction techniques	1
1.3	Structure determination from single crystal diffraction data	3
1.4	Structure determination from powder diffraction data	3
1.5	Extending the limits imposed by powder diffraction data	4
2	**Approaches investigated**	**5**
2.1	Anomalous scattering	5
2.2	Genetic algorithm & automatic Fourier recycling	6
2.3	Automatic Fourier recycling & topology search	7
2.4	Coordination sequences	7
3	**The FOCUS method**	**8**
3.1	The FOCUS environment	8
3.2	Automatic Fourier recycling	12
3.2.1	Overview	12
3.2.2	The automatic Fourier recycling loop	13
3.3	Topology search	18
3.3.1	Creation of the bondlists	18
3.3.2	The backtracking procedure	20
3.3.3	Selecting truly 3-dimensional frameworks	20
3.3.4	Modified topology search: “two color” frameworks	21
3.4	Sorting of topologies	22
3.4.1	Determination of a CS: a node counting algorithm	22
3.4.2	Determination of a LC: modification of the node counting algorithm	23
3.4.3	Combined evaluation of multiplicities, LC’s and CS’s	24
4	**Applications**	**26**
4.1	The test structure Dodecasil-1H (DOH)	28
4.1.1	Preparation	28
4.1.2	The FOCUS input file	30
4.1.3	Tests with varying reflection usage and recycling strategy	36
4.2	The test structure NU-3 (LEV)	39
4.3	The test structure RUB-17 (RSN)	46
4.4	The test structure SAPO-40 (AFR)	51
4.5	The test structure Zeolite-A (LTA)	57
4.6	The test structure EMC-2 (EMT)	61
4.7	The structure of the zincosilicate zeolite VPI-9	67
4.7.1	Preparation and application of FOCUS	67
4.7.2	Verification of the structure: preliminary Rietveld refinement	71
4.8	The structure of the zincosilicate zeolite VPI-10	79
4.8.1	Preparation and application of FOCUS	79
4.8.2	Verification of the structure: preliminary Rietveld refinement	86
4.9	The proposed structure of the beryllosilicate B2	90
4.9.1	Preparation and structure determination	90
4.9.2	Review of the three models suggested for B2	98

5 Further aspects 103
5.1 Review VPI-9 anomalous scattering 103
5.2 Searching for non-tetrahedral node connectivities 105
5.2.1 Searching for interrupted frameworks 105
5.2.2 Searching for 3, 4, and 6-fold connectivities 107
5.3 Backtracking on a grid 110
5.4 (Mis)using FOCUS for the generation of hypothetical topologies 112
5.5 Possible developments of FOCUS algorithm 116

6 Conclusions 118

7 Appendix 119
7.1 F-weighted phase changes 119
7.2 Symmetrically equivalent bonds 119
7.3 Geometrical evaluation of tetrahedral node connectivities 120
7.4 Analytic integration of electron density peaks 121
7.5 Internally stored X-ray wavelengths [Å] and their keywords 122
7.6 DLS prescriptions used 122
7.7 FOCUS source code 122

8 References 123

[83] Acknowledgments 128

Curriculum Vitae 129

1 Introduction

1.1 Zeolites

Zeolites are a class of microporous materials with 3-dimensional framework structures, that have many applications as ion exchangers, as molecular sieves, as absorbents and as shape selective catalysts in the petrochemical industry [1, 2, 3]. The useful properties of these inorganic "host-guest" structures are closely related to the topologies of the host frameworks, because they define the size and shape of the pore openings (absorption and molecular sieving), the kinds of cation sites available (cation selectivity), the dimensionality of the channel system (reactant and product diffusion in catalysis) and the dimensions of the cages and channels (space available for reactant intermediates in catalysis). The framework generally has the composition TO_2, where T is a tetrahedrally coordinated atom (T-atom), and each T-atom is linked to four neighboring T-atoms via oxygen bridges. Classically, the T-atom is Si or Al, but this definition can be relaxed to include a number of additional elements (e.g. P, Ga, B, Be, Zn, etc.). The channels and/or cages of the frameworks are usually filled with inorganic or organic guest species, which can be removed by heating, a process which is called calcination, or exchanged with other cations or molecules.

All unique and confirmed framework topologies (connectivity of T-atoms without reference to chemical composition) are assigned structure type codes by the Structure Commission of the International Zeolite Association (authorized by IUPAC). These are published on a regular basis in the "Atlas of Zeolite Structure Types" [4]. The 1992 edition lists 332 zeolite and zeolite-like materials with 85 structure type codes for the unique topologies. A structure type code consists of three capital letters, for example **EMT** for the "type material" EMC-2, or **LTA** for the type material Linde-A. To illustrate a typical zeolite framework structure, a wire-frame plot of the topology **LTA** is shown in Fig. 1-1. This framework has a 3-dimensional, 8-ring channel system and two types of cages. The smaller cage is known as the sodalite or β-cage and the larger as an α-cage. Both are also found in other zeolite topologies.

1.2 Polycrystalline diffraction techniques

Since most synthetic zeolites are only available as polycrystalline powders, and methods for the synthesis of high quality single crystals are unknown, powder diffraction techniques must be applied. Knowledge of the structural properties of zeolites is fundamental to the understanding of their chemical or catalytic behavior, so the development of powder diffraction techniques beyond simple phase identification and unit cell determination has

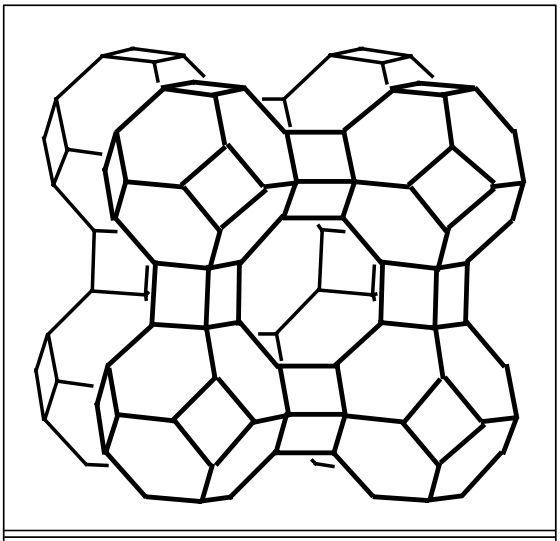


Figure 1-1: The topology of **LTA**
(O atoms have been omitted for clarity)

become an essential branch of zeolite research.

At present, the most important technique for the derivation of structural parameters from powder diffraction patterns is the whole-profile intensity fitting procedure introduced by H.M. Rietveld for neutron data in 1969 [5]. The Rietveld structure refinement technique, can be, and often is, combined with difference Fourier analyses to complete the structural model. In the past two decades, the complexity limit for structures suitable for Rietveld refinement has been raised by both methodological and experimental improvements. The most important methodological advance with respect to zeolite research is probably the inclusion of geometric "observations" in the refinement procedure [6], but the introduction of more elaborate mathematical descriptions for the peak shapes found in X-ray powder profiles [7] has also had significant impact. On the experimental side, the main advance is the increasing availability of synchrotron radiation, which allows very high resolution powder diffraction data to be measured. The higher resolution increases the information content of the powder pattern [8].

The application of a Rietveld-like technique to a powder pattern for which a unit cell and space group, but no structural model is available, was a natural extension of the whole-profile approach. In 1981, G.S. Pawley [9] developed a technique for the extraction of symmetry-allowed reflection intensities through least squares refinement, and recently Sivia & David [10] have enhanced this technique with a Bayesian approach. However, the major breakthrough was probably the introduction of a pragmatic and robust procedure for the iterative adjustment of the intensities by A. Le Bail. [11] in 1988. Using one of these techniques, integrated intensities can be extracted from a powder pattern in order to generate a pseudo single crystal data set.

1.3 Structure determination from single crystal diffraction data

From the intensities of diffracted X-rays, it is straightforward to compute the amplitudes of the Fourier transform of the diffracting electron density. What cannot be measured, or computed from measurable data, are the phases of the Fourier coefficients. However, both amplitudes and phases are needed to reconstruct the electron density. This is the origin of the so called “phase problem” of crystal structure determination.

Direct methods of structure solution, which were conceived in the 1950s and have continued to develop with the availability of ever increasing computing power [12], have made the solution of the phase problem largely a routine task – given that a macroscopic single crystal can be used in the diffraction experiment. Nowadays, several thousand single crystal structure solutions are published every year.

1.4 Structure determination from powder diffraction data

As mentioned before, the powder diffraction method is frequently the only option for the investigation of a specific zeolite structure. In contrast to single crystal methods, almost any crystalline material can be investigated with powder methods. On the other hand, the robustness of the method is encumbered with a substantial disadvantage: with single crystal data, the diffracted intensities are distributed in three dimensions, whereas with powder data they are projected onto a single dimension. The projection causes information, which is separated in three dimensions and which plays a vital role in the structure solution process, to be obscured. This inherent projection in powder diffractometry causes systematic or accidental overlap of the diffraction peaks that are to be measured, so that in addition to the phase problem, an overlap or “partitioning” problem is introduced.

Nevertheless, at present the predominant technique used for structure determination from powder data is the extraction of integrated intensities, followed by the application of direct methods [13]. However, in comparison to real single crystal data, the data set obtained by extraction is generally of substantially lower quality. Both structure complexity and unit cell size are important limiting factors: a more complex structure involves the determination of more variables, and a larger unit cell exacerbates the overlap problem. The necessarily inexact partitioning of the overlapping intensities invalidates the statistical assumptions on which direct methods are based, so the likelihood of a successful application decreases dramatically with increasing overlap. The importance of this phenomenon is reflected by the fact that

10000s of single crystal structure solutions per year are accompanied by only about 30 solutions based on powder data [13]. Furthermore, the structures solved from powder data tend to be less complex.

1.5 Extending the limits imposed by powder diffraction data

In the past few years, several attempts have been made to improve the quality of the pseudo single crystal data by skillful manipulations [14, 15, 16, 17] (see page 11), and to improve direct methods themselves by modifying them specifically for application to powder data. For example, Cascarano et al. [18] permute the moduli of reflections which are judged to be important for the direct methods procedure. However, the large discrepancy between the complexity of structures that can be refined and that that can be determined remains. Since zeolite structures tend to lie near or over this complexity limit, model building still plays a very important role in the determination of their structures (e.g. [19, 20]). An idea frequently suggested has been to learn from model building. In other words, it should be possible to use the same information on which model building is founded actively in an automated procedure. Geometrical and chemical knowledge gleaned from known structures related to the one to be solved and from chemical analysis should lend themselves to such an approach. In particular, the types and numbers of atoms expected, their expected coordination numbers, typical bond distances and angles, and minimum distances between non-bonded atoms could be used. The implementation of this idea is the predominant feature of this study.

2 Approaches investigated

2.1 Anomalous scattering

While most zeolite-like materials contain only light elements, the Rb,K-zincosilicate VPI-9 [21, 22, 23] has two potential anomalous scatterers (Rb and Zn). In the hope of exploiting this fact, data sets were collected at three different wavelengths (near the Rb and Zn absorption edges and in between the two) at the NSLS synchrotron facility in Brookhaven, N.Y. [24]. Intensity differences were apparent (Fig. 2-1), and an attempt was made to apply the method described by Prandl in his paper entitled “Phase Determination and Patterson Maps from Multiwavelength Powder Data” [25] to these measurements.

Two main problems arose: the relative scaling of the three synchrotron powder patterns collected for VPI-9 proved to be a nontrivial matter, and, more importantly, Prandls idealization of fully resolved peaks was violated to a large extent (for a further details, see section 5.1). As a result, the errors made in extracting the integrated intensities from the powder patterns outweighed the effects due to anomalous dispersion, so this theme was not pursued further.

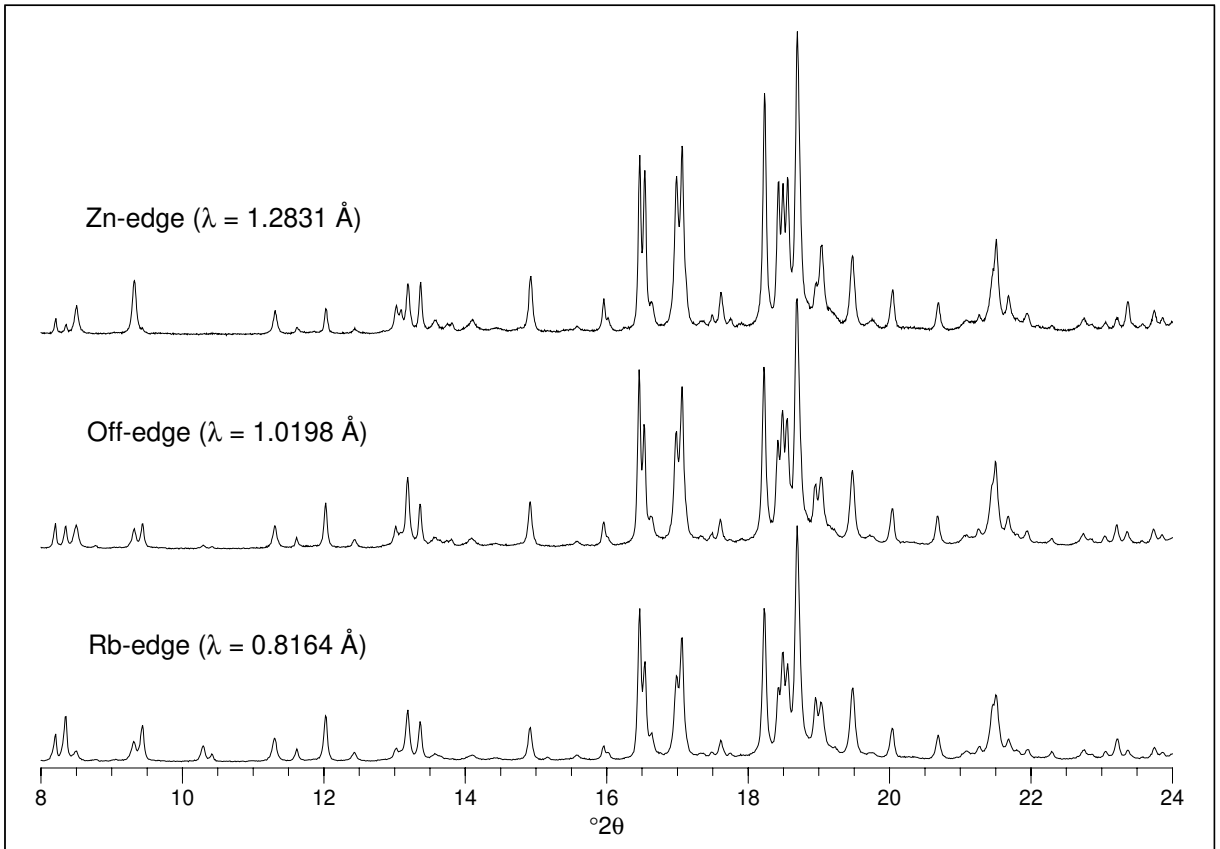


Figure 2-1: Section of the data collected on VPI-9 at three different wavelengths.
For comparison purposes, the 2θ scale has been normalized to 1.0 Å for all data sets.

2.2 Genetic algorithm & automatic Fourier recycling

Next, the "phase-problem" was tackled using a very different technique, involving the combination of a "genetic algorithm" [26] with automatic Fourier recycling. The first tests were carried out with 1-dimensional, centrosymmetric "atomic" arrays. At first, the "data-sets" (i.e. the structure amplitudes) were simulated. The phase (sign) determination process was initialized by generating a certain number of random start-phase sets, the "gene pool". Each phase set was then subjected to an automatic Fourier recycling procedure which assigned the atoms by correlating atomic weight and electron density peak height, and also considering a minimum interatomic distance of 1.0 Å. After all phase sets had been recycled, these were treated as "genomes" and subjected to the recombination functions of the genetic algorithm. The "fitness" of a gene was determined by the R_F-value. The resulting new "generation" of phase sets was again used for Fourier recycling and recombination, until convergence of the "gene pool" was obtained.

Since the 1-dimensional tests looked promising, the procedures were expanded from one dimension to three dimensions. Then it became clear that the simple R_F-value is essentially useless as fitness measure for more complex structures. A genetic algorithm requires a good fitness measure to work, and therefore the applicability of a modified version of the integral

$$\int_V \rho^3 dV$$

(where ρ is the electron density in electrons per $Å^3$, and V the volume of the unit cell) [27], which is well known from direct methods, was investigated. Although this provided a significant improvement over the R_F criterion, it was still not sufficient to indicate a "good" phase set unambiguously. Furthermore, the number of phases which have to be considered in the determination of a zeolite structure of typical complexity proved to be well above 100. For the simplest case of a centrosymmetric structure this translates into a "genome size" of 100 bits. Estimations for the number of evaluations (one cycle of Fourier recycling/recombination for one phase set) needed to obtain convergence of the "gene pool" revealed that the attempted procedure would require several orders of magnitude more computing time than is practically available [28]. Thus the genetic algorithm idea reached an impasse and was put aside.

2.3 Automatic Fourier recycling & topology search

However, surprisingly enough, simple Fourier recycling of random starting phases without the genetic algorithm recombination revealed the correct structure for several test cases of medium complexity. Once this fact had been discovered, the development proceeded in this direction. The idea was to let the “recycling engine” run and to devise some way of filtering out the correct solution.

With only a few exceptions, zeolite framework structures can be described as 4-connected and 3-dimensional nets. Since interatomic distances are very well known for the elements found in zeolites, the automatic Fourier recycling was first improved to take individual minimum interatomic distances into account in the interpretation step, in the hope that the rate of occurrence of “detectable” correct solutions could be improved. Then a second feature, an exhaustive search for 4-connected, 3-dimensional topologies (i.e. a search for a set of T-sites where each T-site has exactly four neighboring T-sites), was added.

The algorithm employed is known in computer science as a *backtracking algorithm* [29]. The adaptation developed in this work takes the first ~50 largest peaks in the asymmetric unit of each electron density map produced during the Fourier recycling process. In a systematic and non-redundant way, the algorithm culls all subsets of these ~50 peaks which build up a tetrahedral framework.

Again surprisingly, the recycling/topology-search combination produces hundreds or even thousands of topologies for a given unit cell and space group, so a way of identifying and sorting the topologies (i.e. of recognizing equivalent solutions) had to be found.

2.4 Coordination sequences

This was where coordination sequences[30] (CS) came into play. CS’s are integer sequences, in which the first term is the conventional coordination number (i.e. the number of atoms bonded to a certain pivot atom). The second term is the number of additional atoms bonded to all atoms counted in the previous step, and so on. In effect, CS’s give a “finger print” of a specific framework topology and provide a very elegant method of circumventing origin choice ambiguities and atom coordinate deviations in the comparison of crystal structures. By means of CS’s, the huge number of topologies obtained could be sorted. Examination of several test cases revealed that the most frequently occurring topology is, in general, the correct one.

3 The FOCUS method

The automatic Fourier recycling / topology search / topology classification and sorting algorithms were combined to form the core of a program system that was given the name FOCUS. The FOCUS method can be viewed as a tool that can be added to the set of conventional structure determination techniques. It is itself a combination and adaptation of classical methods.

3.1 The FOCUS environment

Figure 3-1 shows a flowchart of the complete structure determination procedure, and indicates where FOCUS is applied.

First, a data set is collected in the usual way on a high resolution powder diffractometer. The resulting powder pattern is analyzed using a peak finding program in order to obtain a set of peak positions (i.e. a list of 2θ's or their corresponding d-values). These are then input to an automatic indexing program to determine the unit cell parameters [31]. At this point, the critical question is whether the sample is one pure phase or a mixture of two or more phases. If the latter proves to be the case, purification methods have to be found, since otherwise the data is of very limited value for structure solution.

The next step, the determination of possible space groups, deserves special attention. While space group determination with single crystal data is generally straight forward, the extent of reflection overlap in a powder pattern, whether due to sample quality or to structure complexity, has a significant potential to obscure the symmetry and thereby severely hamper the solution process. The structure solutions of both VPI-10, where an unfortunate combination of unit cell parameters introduced an ambiguity that allowed ten possible space groups (see section 4.8) and VPI-9 (see section 4.7), illustrate the importance of this step. An appreciable amount of time has to be spent on space group determination.

Through the implementation of Le Bail's ingenious technique for the extraction of integrated intensities from powder patterns [11] in various Rietveld programs (e.g. [32, 33, 34]), the extraction process has recently become relatively easy to conduct and is now almost a routine task. The extracted intensities are normalized by means of a Wilson plot. However, for zeolites and zeolite like materials, it is very common for the Wilson plot to diverge significantly from the ideal straight line. Figure 3-2 shows the Wilson plot ("○") with the calculated structure factors of the clathrasil **DOH**, a typical example, to a resolution of 1.3 Å ≈ 0.15 $(\sin\theta/\lambda)^2$. The "." line is the fit computed by the Xtal 3.2 [35] GENEV module, while the "+" line is computed with the overall temperature factor that was used in the

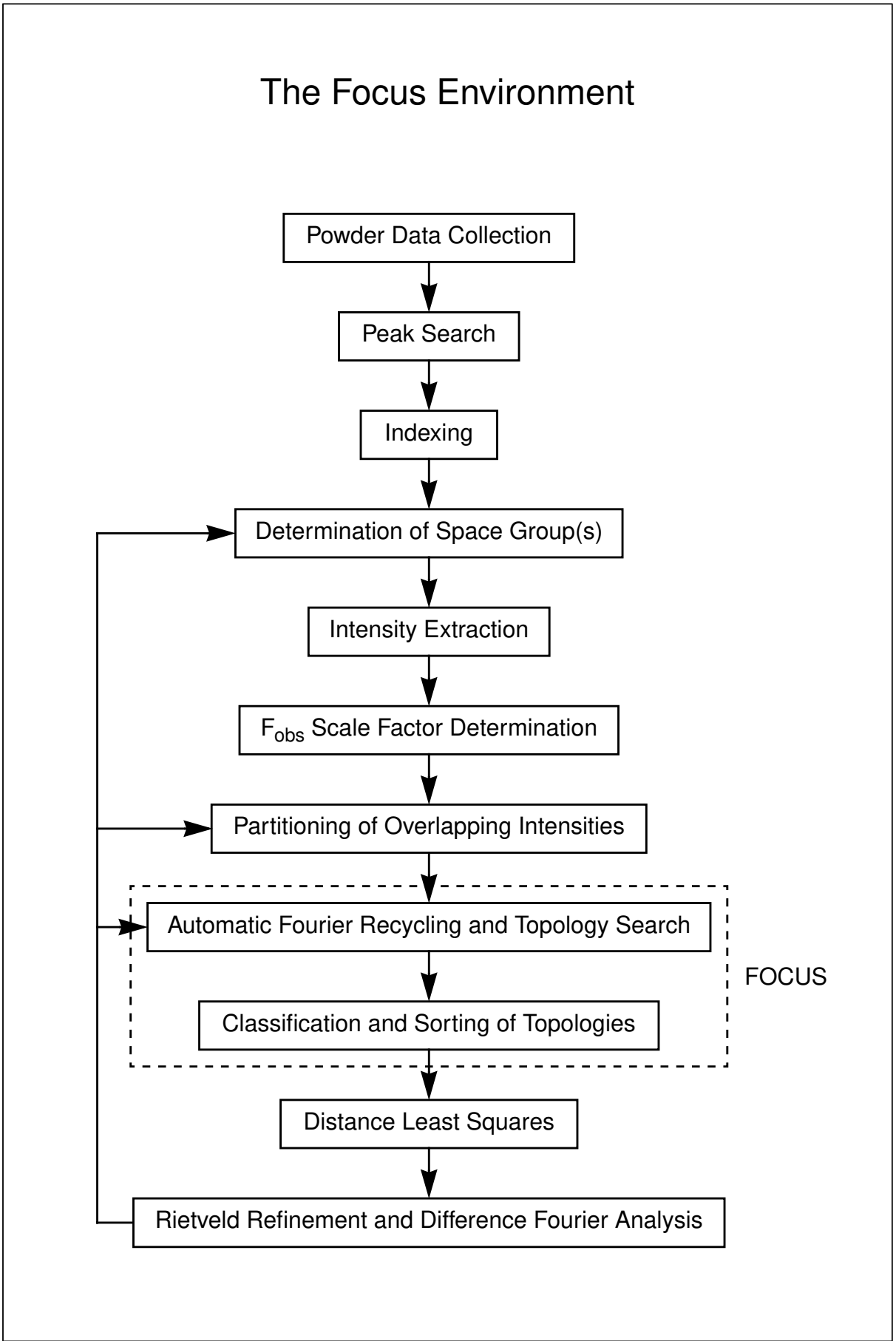

The Focus Environment
Powder Data Collection
Peak Search
Indexing
Determination of Space Group(s)
Intensity Extraction
F_{obs} Scale Factor Determination
Partitioning of Overlapping Intensities
Automatic Fourier Recycling and Topology Search
Classification and Sorting of Topologies
FOCUS
Distance Least Squares
Rietveld Refinement and Difference Fourier Analysis

Figure 3-1

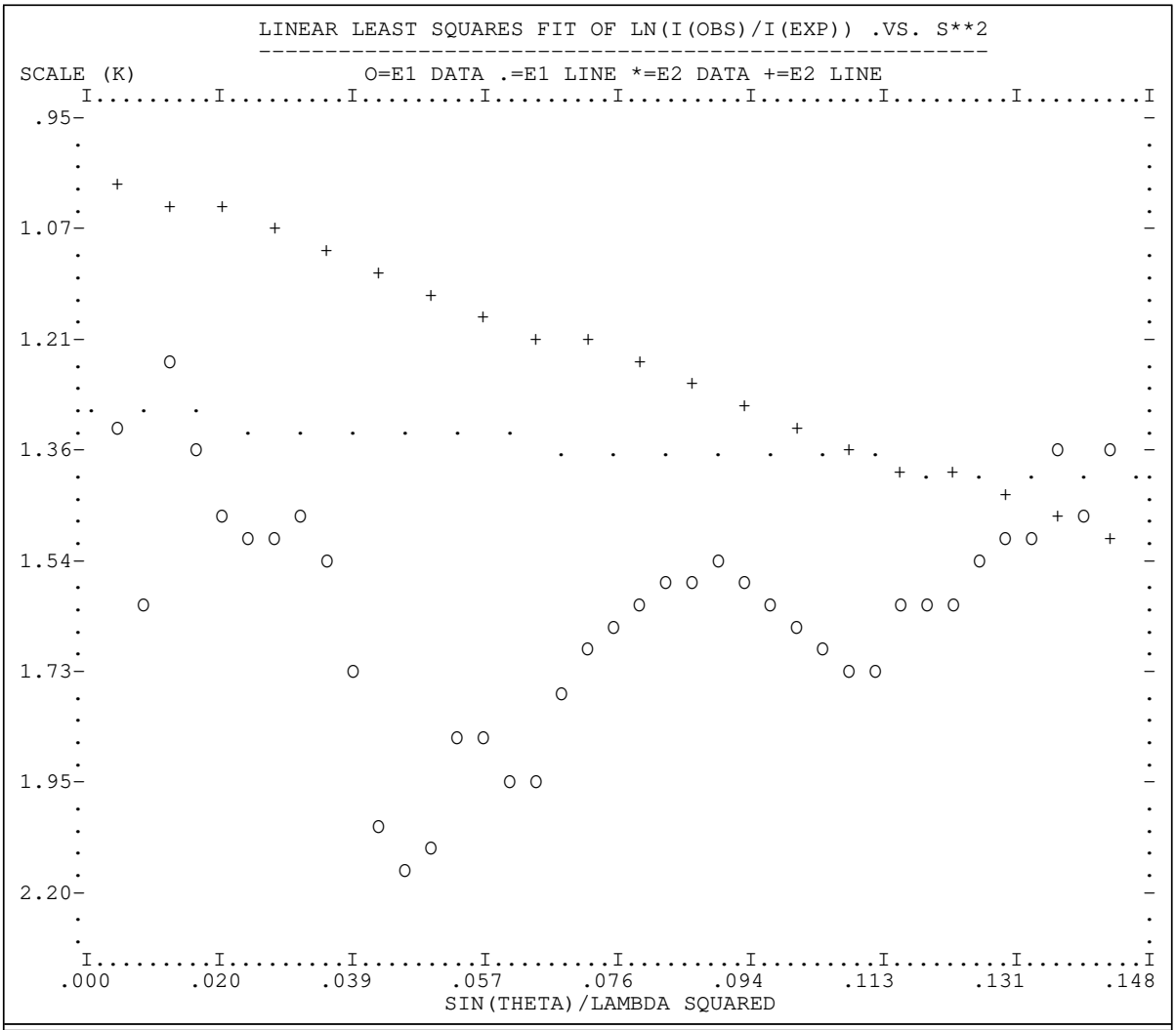


Figure 3-2: Wilson plot with calculated structure factors of **DOH**

computation of the structure factors, and a scale factor of one. Based on experience gathered while working on test cases, a pragmatic approach was used for unknown structures: the overall temperature factor was held fixed at $U_{overall} = 0.025$ Å^2, and the straight line shifted parallel to the y-axis so that the intersection with the observed data is at about 1.3 Å (usually at the right edge of the plot). While this simple procedure gives satisfactory results in most cases, it should be mentioned that Estermann [36] has recently presented a more elaborate and promising approach for the normalization of diffraction data from structures, which significantly violate the random atom expectation, on which the Wilson plot is based. However, this new approach was not applied here.

After scaling, the extracted intensities need further processing. The minimum treatment is the equipartitioning of overlapping intensities. That means that a sensible “overlap factor”

(*of*) is chosen, typically 0.3, and the intensities of all groups of reflections with

$$2\theta_2 - 2\theta_1 < \frac{FWHM_1 + FWHM_2}{2}\, of \qquad (3\text{-}1)$$

are averaged (2θ = reflection position in the powder pattern, *FWHM* = full width at half maximum). According to Estermann [37], averaging of the N_g Fourier magnitudes $|F|^2_{\vec{H}}$ in a particular overlap group is carried out using

$$|F|^2_{\vec{H},\,i} = \frac{\sum_{j=1}^{N_g} m_{\vec{H},\,j} |F|^2_{\vec{H},\,j}}{N_g \cdot m_{\vec{H},\,i}} \qquad i = 1, 2, \ldots, N_g$$

(which results in equal $m \cdot |F|^2$, m = reflection multiplicity), but the gain compared to

$$|F|^2_{\vec{H},\,i} = \frac{\sum_{j=1}^{N_g} m_{\vec{H},\,j} |F|^2_{\vec{H},\,j}}{\sum_{j=1}^{N_g} m_{\vec{H},\,j}} \qquad i = 1, 2, \ldots, N_g$$

(which results in equal $|F|^2$) is only marginal.

For more sophisticated partitioning of overlapping intensities, David [14, 15] has suggested two approaches. The first is based on intensity statistics (“squaring method”), and the second approach is founded on a maximum-entropy formalism (“maximum-entropy Patterson method”). For the same purpose, Jansen et al. [16] have developed the DOREES procedure. The redistribution of overlapping intensities is based on five E-value relations (two triplet relationships, two quartet relationships, and one relationship based on the Patterson function). Another approach, the FIPS (“Fast Iterative Patterson Squaring”) method, was developed by Estermann et al. [17]. In an iterative process, the overlapping intensities are redistributed according to the partitioning found for the magnitudes obtained by Fourier transforming squared Patterson maps.

In cases where a solution attempt with equipartitioned data is unsuccessful, the application of these methods could be helpful. However, only the FIPS method was applied in this investigation (see page 63).

At this stage, the normalized and partitioned data are input to the FOCUS procedure. To complete the overall picture before going into details, it is sufficient to know that FOCUS produces a list of topologies, ranked by frequency of occurrence, which can be interpreted as likelihood of correctness. The topologies are given as fractional coordinates of the “node

positions" (usually tetrahedrally coordinated positions, occupied with e.g. Si, Al, P).

For the most likely topologies, bridging oxygen atoms are inserted at the center of all node-node connections, and the resulting completed framework is subjected to a distance least-squares refinement with the DLS-76 program [38]. After careful inspection of the DLS-76-residuals and the refined bond lengths and angles, the most promising structure can be selected as starting model for a conventional Rietveld refinement with difference Fourier analysis to find missing atoms (i.e. non-framework atoms). In cases where the refinement does not converge, other reasonable structures from the list - if present - can be tried, or parts of the whole procedure can be repeated. For example, a different space group could be selected, the partitioning of overlapping reflections could be varied, or the parameters for the FOCUS procedure could be changed.

3.2 Automatic Fourier recycling

3.2.1 Overview

Fourier recycling can be started in two ways: either with starting phases or with a starting model, and both approaches are in active use. For example, the first Fourier map can be generated using phases from a promising direct methods solution, and used to complete or correct the model. For zeolite structures, on the other hand, model building has frequently been the key to successful structure determination. In this case, the partial model is used to compute a phase set, which, in turn, is used together with experimentally determined Fourier magnitudes ($|F|$'s) to calculate an electron density map. Typically, the preliminary stages of a single crystal or Rietveld refinement involve the generation of a sequence of difference Fourier maps and the feed-back of an improved model.

In this study, mainly the first approach, i.e. the use of starting phases, has been employed. The Fourier recycling can be initialized either with random starting phases, or with phases from some other source. An electron density map is generated by a Fourier transform and subjected to a peak search algorithm. If random starting phases are used, the resulting peaklist can also be viewed as "random starting model". In other words, starting with random phases or with a random model is essentially equivalent. Since it is technically easier to set up a random phase set than a random model, only the first possibility was used.

3.2.2 The automatic Fourier recycling loop

3.2.2.1 Prerequisites

The automatic Fourier recycling is initialized by:

- selecting a subset of reflections for active use.
- defining structural properties, namely approximate unit cell contents and a minimum distance for each pair of atom types.
- defining technical parameters like grid spacings for the electron density map or maximum number of peaks in the electron density peaklist.

For the selection of the subset of reflections to be used in the recycling, the reflections (hkl, normalized and partitioned Fourier magnitudes) are sorted in descending order with respect to magnitude times multiplicity. Two selection procedures are possible: a) a prescribed number of the strongest reflections is selected, or b) the sum of all magnitudes – weighted by the multiplicities – is taken to be 100%, and the strongest reflections are selected from the sorted list until a prescribed percentage of the total sum is accumulated.

> It should be mentioned that a more involved selection procedure was used when the genetic algorithm idea was still being pursued. First, the semi-invariant vectors and moduli of the space group were determined. The reflection list was sorted by two criteria: the potentially origin defining (“non-semi-invariant”) reflections at the top, the semi-invariant at the bottom, and secondly, by magnitude in descending order. From this list the strongest set of origin defining reflections was selected for the subset of actively used reflections. Only then was the subset completed with either procedure a) or b). In order to make a population of phase sets ready for the crossover procedure of the genetic algorithm, the individual phase sets have to be maintained on the same origin. That means, the origin defining reflections have to have the same phase angles. However, while this can be achieved easily for centrosymmetric space groups, major difficulties arise for acentric space groups. Moreover, it turned out that the structure determination in “random starting phases” mode is more efficient when no phases are held fixed, that means the phases angles of all active reflections are randomly set for each trial set. This observation is not really understood.

The description of the approximate unit cell contents is simply a list of expected atom types, the number of atoms per unit cell for each type, an isotropic displacement factor and an occupancy factor. In addition, structural information can be supplied by defining whether a certain type is expected to be a framework node, an atom bridging two framework nodes, or a general type. Further structural information is given in the form of minimum distances for pairs of atom types. For an example, refer to page 30.

3.2.2.2 Initialization of a new trial and Fourier transform

The automatic Fourier recycling loop is illustrated in Fig. 3-3. A single trial is initialized by assigning starting phases to the selected reflections. The next step is a Fourier transform of magnitudes and phases to produce an electron density map.

> Two Fourier transform algorithms were investigated: a radix 2 fast Fourier transform (FFT) and the Beevers-Lipson algorithm. The O(n log n) FFT algorithm is by far the superior one in cases where the resolution in direct space is equal to the resolution in reciprocal space. The number of grid points has to be a power of two, and the whole cell must be transformed. However, in practice, the O(n^2) Beevers-Lipson algorithm turned out to be more efficient. In powder work, the resolution used in direct space is usually three to four times the resolution in reciprocal space. The number of grid points resulting from the choice of (usually) 1/3 Å resolution is, in general, not a power of 2, and depending on space group symmetry, only fractions of the unit cell have to be computed via the Fourier transform. More diverse FFT algorithms for crystallographic applications have been presented in the literature [39, 40], but these were not used in this investigation.

3.2.2.3 Peak search

The first processing step of the electron density map is a 27 point peak search in the asymmetric unit of the unit cell: a “pivot” or “central” point is marked as a peak if its electron density value is higher than that of all of its 26 nearest neighbors (Fig. 3-4). A histogram of the peak heights found is maintained throughout the search. After all grid points in the asymmetric unit have been scanned, the histogram is used to determine the height cut off, such that a preset maximum number of peaks is not exceeded.

3.2.2.4 Peak interpolation

Since the exact peak maxima do not generally coincide with a grid point, the positions of the peak maxima are determined, or refined, with a formalism found in [41] (pp. 35-37). A peak which is not centered at the origin and is not necessarily spherical is modeled by

$$\rho_{calc}(xyz) = \exp(a + bx + cy + dz + ex^2 + fy^2 + gz^2 + hyz + kzx + lxy) \qquad (3\text{-}2)$$

where $\rho_{calc}(xyz)$ is the electron density at the point with fractional coordinates xyz. The ten coefficients $a,b,c,\ldots,l$ are determined by least-squares so as to minimize $\sum(\ln\rho_{obs}-\ln\rho_{calc})^2$, summed over all points used, usually the central point and its 18 nearest neighbors (Fig. 3-4). However, ρ_{obs} with values below a preset threshold (arbitrarily fixed at 10^{-6} electrons/Å^3) are omitted. If less than a definable number of points have a value which is sufficiently high, no interpolation is carried out and the central grid point position is retained.

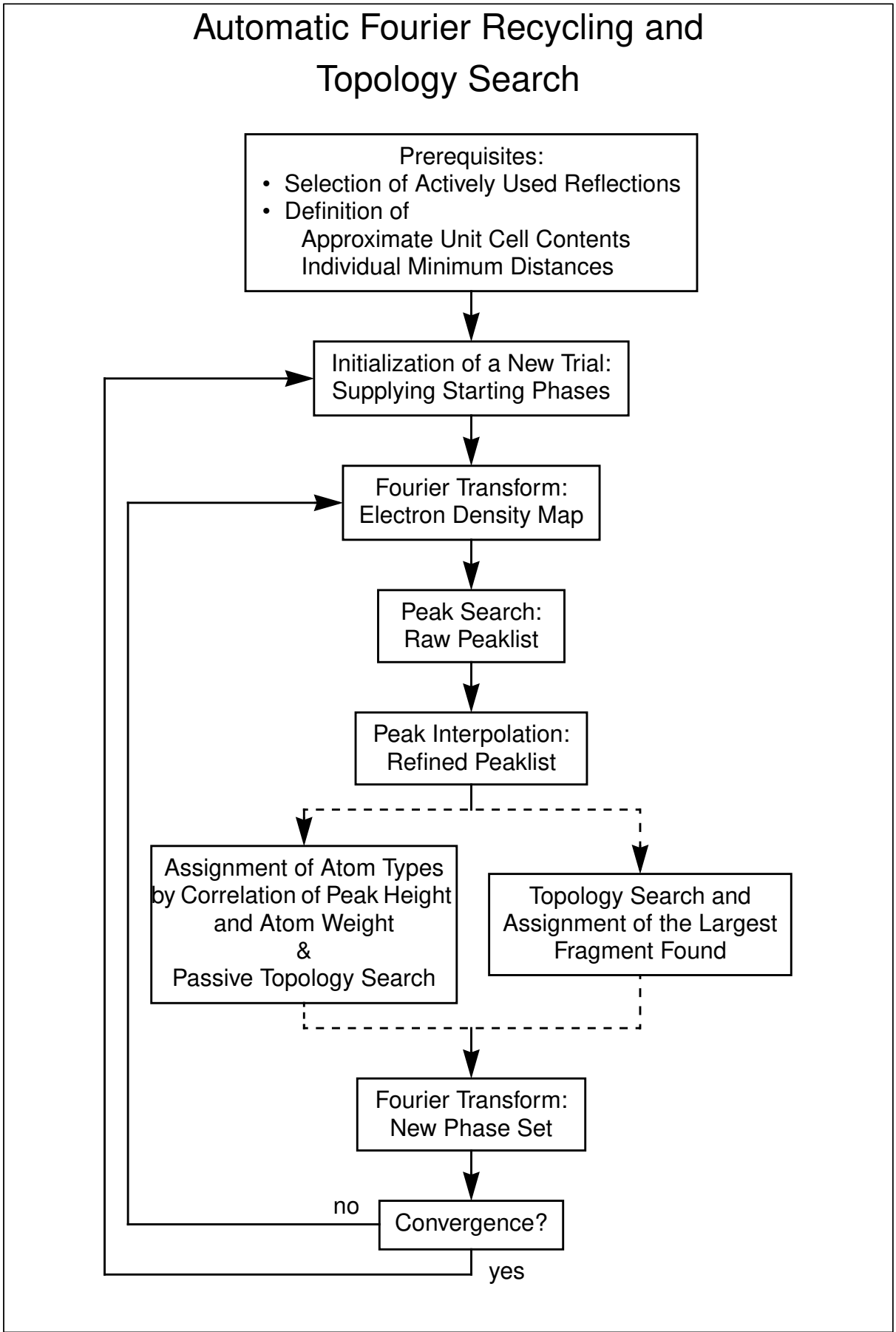
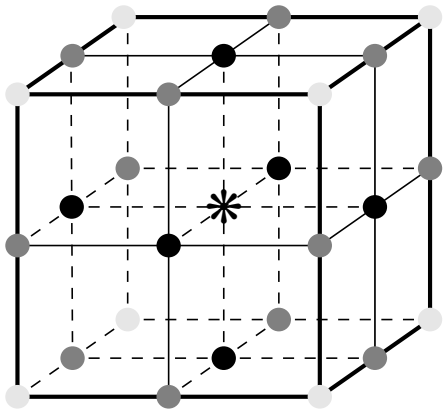

Automatic Fourier Recycling and Topology Search
Prerequisites:
• Selection of Actively Used Reflections
• Definition of
Approximate Unit Cell Contents
Individual Minimum Distances
Initialization of a New Trial:
Supplying Starting Phases
Fourier Transform:
Electron Density Map
Peak Search:
Raw Peaklist
Peak Interpolation:
Refined Peaklist
Assignment of Atom Types
by Correlation of Peak Height
and Atom Weight
&
Passive Topology Search
Topology Search and
Assignment of the Largest
Fragment Found
Fourier Transform:
New Phase Set
no
Convergence?
yes

Figure 3-3



✱	Pivot or central point
● ● ●	Points used in peak search
● ●	Points used for interpolation

Figure 3-4: Grid point usage

Upon successful determination of the coefficients $a,b,c,\ldots,l$ the position of the maximum is determined by finding that xyz_{max} for which $\partial(\ln\rho_{calc})/\partial x = \partial(\ln\rho_{calc})/\partial y = \partial(\ln\rho_{calc})/\partial z = 0$. To account for numerical inaccuracies, the "shift vector" – pointing from the central grid point to the interpolated position of the maximum – is projected onto a plane when the grid point is on a special position with two degrees of freedom (e.g. a mirror plane), or an axis when the special position has one degree of freedom (e.g. a two-fold rotation axis).

After a peak position has been refined, the shortest distance to all symmetrically equivalent positions (self-distance) is computed. To do this, it is necessary to generate all equivalent positions in 27 unit cells (i.e. the "center" unit cell with fractional coordinates $0 \le x,y,z < 1$, and the 26 surrounding unit cells). If the self-distance is smaller than a prescribed minimum distance (e.g. for a position too close to a mirror plane), the peak is moved onto the symmetry element which is responsible for the close contact. After the shift, the self-distance calculation is repeated. Under certain conditions, the peak position will be corrected more than once.

In the next processing step, the list of interpolated peak positions is sorted in descending order by one of these criteria:

(a) the peak height found at the central grid point.

(b) the peak height $\rho_{calc}(xyz_{max})$.

(c) the analytical integral $\int_V \rho_{calc}(xyz)dV$ (see page 121).

Experience has shown that the peak shapes in the electron density maps produced by the automatic recycling procedure are often very distorted and poorly modeled by Eq. (3-2), and frequently introduce numerical instabilities. Therefore the most simple approach – use of the peak height found at the central grid point – turned out to give the best results.

The last treatment of the refined peaklist is to set an "N-marker" for each entry which can satisfy the node atom requirements (see page 13 and page 30).

3.2.2.5 Construction of a structural model

At this point there are two alternatives:

(a) Assignment of atom types by correlation of peak height and atomic number

The outer assignment loop steps over the defined atom types, which are sorted in descending order by means of atomic number. The inner loop steps over the unassigned entries of the refined peaklist, trying to find a position for the pivot atom type. The pivot atom type is assigned to a previously unassigned entry if (i) the N-marker is set for atom types of class "node", (ii) the multiplicity of the entry is not greater than the number left for assignment for the pivot atom type, and (iii) the prescribed minimum distances to all assigned atoms are not violated. The inner loop is terminated when the prescribed number of atoms per unit cell for the pivot atom type is assigned, or the end of the refined peaklist is reached.

> This assignment algorithm is about as primitive as it can be. Several enhancements are possible, but were not put into practice. For example, maximum distances could also be considered, or coordination numbers could be tracked, allowing for example no more than four bridging oxygen atoms in a sphere of a given radius around a node position. Also, an algorithm could be employed which tries to find the "best" assignment for a given atom type rather than the first possible. However, to a certain extent this is also achieved by alternative (b) below.

Independently, an exhaustive topology search (see section 3.3) among the first 50…60 highest peaks in the asymmetric unit is performed and these are written to a file, when found.

(b) Topology search and assignment of the largest fragment found

An exhaustive topology search similar to the one of alternative (a) above is used to find the largest framework fragment that can be built from a subset of the peaks in the refined peaklist with the N-marker set. The selection criterion is the total number of node-node bonds in the fragment divided by the number of active node positions. Of fragments with equal number of bonds and node positions, the one with the greatest sum of peak-heights is selected.

At the end of the topology search, atom types of class "node" are assigned to the

fragment positions with an algorithm similar to that of alternative (a): the outer loop steps over the atom types of class "node" – again sorted in descending order of atomic number – and the inner loop searches for an unassigned fragment position. However, distances do not need to be checked, because the topology search has already taken care of these.

3.2.2.6 Fourier Transform and convergence test

The recycling loop is closed by a straight forward Fourier transform (see for example [42], "Calculation of the structure factor") of the structural model constructed through one of these processes, and generates a new phase set. By means of a convergence test, which is based on the F-weighted ratio (see page 119) of phase changes, the decision is made as to whether the new phase set is used to calculate a new electron density map, or, in the case of convergence, a new trial is initialized by supplying new starting phases.

3.3 Topology search

The topology search is an application of the well known backtracking algorithm (see for example [29]) and operates on the refined peaklist. To make the topology search efficient, it was divided into two stages: the preparation of a list of potential node-node bonds ("bondlist") for each entry of the refined peaklist, and the actual backtracking which then operates on these bondlists.

3.3.1 Creation of the bondlists

For the creation of the bondlists, a *minimum node distance* (ND_{min}) and a *maximum node distance* (ND_{max}) is prescribed. Values typically used for SiO_2 frameworks were ND_{min} = 2.6 Å and ND_{max} = 3.6 Å, which allows for a tolerance of 0.5Å around the "ideal" node distance ND_{ideal} = 3.1 Å.

In a first scan through the refined peaklist, entries are marked as "Inactive" if the N-marker is not set, or the self-distance is less than ND_{min}. In the second scan, potential node-node bonds with distances in the range ND_{min} through ND_{max} are tabulated for each peak. If the distance between two nodes is less than ND_{min}, or if two peaks form more than the *maximum number of node-node bonds* (NN_{max}), they cannot be present together in the type of framework sought, and an "Exclusive" marker is set. In the next scan, all entries with less than the *minimum number of node-node bonds* (NN_{min}) are eliminated by setting the "Inactive" marker. Of course, the number of bondlist entries of peaks which had potential node-node bonds to those just eliminated is thereby reduced. Therefore, the last scan has to be repeated until no further changes are necessary. Finally, the refined peaklist is resorted by means of the

number of active bondlists per entry, and the bondlists themselves are also sorted such that the order is optimized for the backtracking. Tab. 3-1 gives an example for the final bondlists of a refined peaklist. For example, position number one in the refined peaklist has four active bondlists, and one “Exclusive” marker, which indicates that positions one and four in the peaklist cannot occur together in a framework. The asterisk behind a distance signals that this bond is symmetrically equivalent to the previous bond (see page 119). In addition to the distances, the bond vectors (in Cartesian coordinates), pointing from the pivot peak to the corresponding bonded peaks, are also stored for use in the actual backtracking procedure.

No. in refined peaklist	No. of active bondlists or marker	Bondlists	
		Bond to No. in refined peaklist	Distance(s) [Å] or marker
0	5	0	3.3114
		1	3.3015
		2	3.4472 3.0204
		3	3.4349
		4	3.2393
1	4	1	3.1951 2.8726
		0	3.3015
		3	3.4216
		5	3.5401 3.2220
		4	Exclusive
2	3	0	3.4472 3.4472* 3.0204 3.0204*
		4	2.9356
		5	3.3422
3	3	0	3.4349 3.4349*
		1	3.4216 3.4216*
		4	3.4111
4	3	0	3.2393 3.2393*
		2	2.9356
		3	3.4111
		1	Exclusive
5	2	1	3.5401 3.5401* 3.2220 3.2220*
		2	3.3422
6	Inactive		
7	Inactive		

Table 3-1: Final bondlists of a refined peaklist

3.3.2 The backtracking procedure

The first level of the backtracking procedure consists of an outer loop which steps over the active peaklist entries. Each pivot entry is used as "seed node" to initialize a set of "present" framework positions ("F-set"). On the next level, a *connectivity completion procedure* (CCP), which loops all possibilities for the construction of NN_{min} through NN_{max} bonds for the pivot entry, is called. In these constructions, refined peaklist entries with indices less than the index of the pivot entry have to be omitted in order to avoid redundancy. For each possible bond configuration, a test, which checks its geometrical validity (see page 120), is carried out. If the geometry proves to be acceptable, the positions which are newly bonded to the pivot position are added to the F-set. Then the enlarged F-set is searched for the first entry which is not already a pivot-element (in a previous level) and the CCP is *recursively* called with this entry as new pivot element. If all elements of the F-set have NN_{min} through NN_{max} bonds, a framework topology which meets the prescribed criteria has been found and it is written to a file.

Two basic types of backtracking algorithms are known: the algorithm which terminates as soon as a solution has been found, and the alternative algorithm which searches for all possible solutions and writes a protocol. The implementation discussed here is of the second kind. This means that the only condition on which the recursive CCP returns to the previous level is, that the possibilities for the construction of the desired connectivities for a given pivot position are depleted.

3.3.3 Selecting truly 3-dimensional frameworks

Experience revealed that another geometry filter is necessary to reduce the number of obviously useless frameworks produced by the search procedure. Very frequently, heavily distorted "layer structures" appeared. In an attempt to suppress all but truly 3-dimensional frameworks, a simple algorithm to test whether or not a path from an arbitrary starting node in the unit cell to all other nodes in the same cell exists was introduced.

The procedure operates on the full set of nodes in 27 unit cells (a "center" unit cell with fractional coordinates $0 \leq x,y,z < 1$, and the 26 surrounding unit cells). To start, an arbitrary node in the center cell is selected. In the next step, all nodes in the unit cell which are bonded to the starting node are marked as "reached", and the starting node itself is marked with "all bonds followed". In the subsequent steps, nodes previously marked as "reached" are sought, and processed in the same manner as the starting node. However, only previously unmarked nodes are set to "reached". This process continues until no more nodes in the 27 unit cells are

“reached”.

The final step is to check whether there are “unmarked” nodes left in the center cell, or if there are surrounding cells with all nodes unmarked (i.e. no nodes reached). If so, the framework is suppressed since it is not truly 3-dimensional.

3.3.4 Modified topology search: “two color” frameworks

There are a large number of zeolite frameworks with two types of strictly alternating node atoms, for example Si-Al, Al-P, Ga-P or Zn-P. While the node-node distances of pure silicon frameworks are always such that the (four) nodes bonded through bridging oxygen are also the next (four) neighboring nodes, this is not always true for other types of node atom pairs. For example, the gallophosphate ULM-5 [43] has one gallium in the asymmetric unit which is bonded to four phosphorous atoms through oxygen, and to another gallium, again through oxygen, at a distance smaller than the largest Ga-P distance. This is illustrated in Fig. 3-5. By ignoring this Ga-Ga bonded oxygen and also the four fluorine atoms per asymmetric

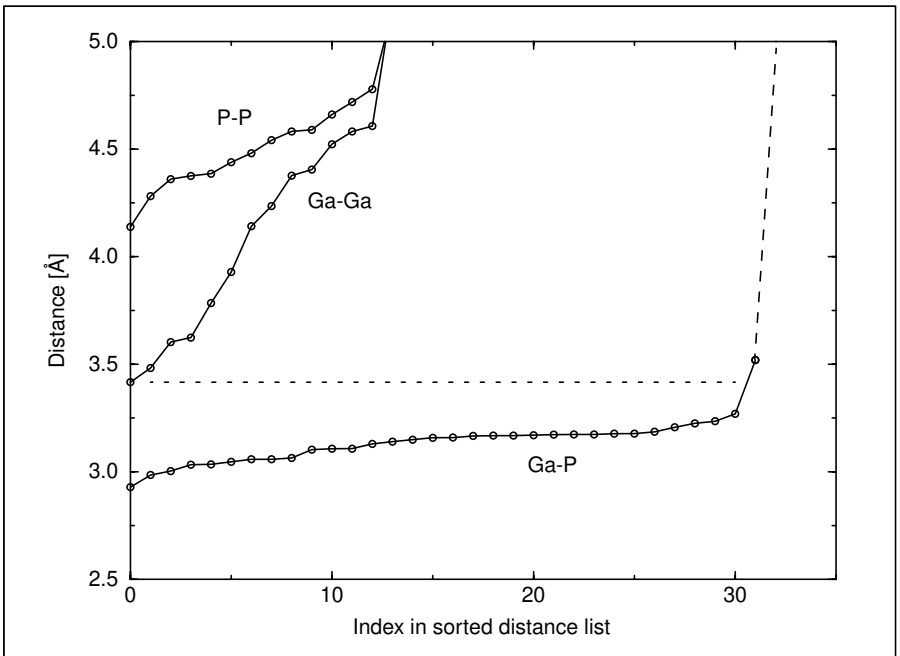


Figure 3-5: ULM-5 node distances

unit, ULM-5 can still be viewed as tetrahedral framework with strict alternation of Ga and P. However, since the smallest Ga-Ga distance of the special gallium is smaller than the largest Ga-P distance, the topology search will not recover this framework.

To overcome this problem, the search algorithm was modified for frameworks with strictly alternating occupation of the nodes: a “color”, say white, is assigned to the seed node, which is set in the outer loop. In the CCP, all positions which are connected to the pivot

position are assigned the “opposite color”, say black. Node distances smaller than ND_{min} are still not allowed, but bonds are created only between positions of different color.

This simple modification is sufficient to recover the tetrahedral topology of ULM-5 (given the correct peak positions). Furthermore, this modification also acts as a filter which allows only strictly alternating topologies to be accepted, and thereby reduces the number of non-feasible topologies that have to be investigated in the subsequent steps.

3.4 Sorting of topologies

A fast and efficient way of classifying and sorting the frameworks produced by the backtracking procedure, based on the evaluation of the site multiplicities, loop configurations (LC) and coordination sequences (CS), was developed for the next stage. While the multiplicities are available immediately, because they are needed in several of the preceding steps, the determination of LC’s and CS’s is more involved.

3.4.1 Determination of a CS: a node counting algorithm

The notion of CS was formally introduced by Brunner & Laves [44] in order to investigate the topological identity of frameworks and of atomic positions within a framework. The CS is a number sequence in which the k-th term is the number of atoms in “shell” k that are bonded to atoms in “shell” k-1. Shell 0 consists of a single atom, and the number of atoms in the first shell is the conventional coordination number.

The CS determination algorithm used here can be described as a *node counting algorithm* or a *coordination shell algorithm*. The algorithm is started by selecting an *initial node* ($k = 0$). In the next step, all nodes bonded to the initial node are determined ($k = 1$). For $k \geq 2$, all characteristics of the algorithm become evident: those nodes, which are bonded to the “new nodes of the previous step (k-1)”, but have not been counted before, are counted.

> This means that three sets of nodes for three topological distances (three coordination shells) have to be maintained: the *middle* (k-1) nodes, whose bonds are followed to determine the *next* (k) nodes, and the *back* (k-2) nodes, to know which of the nodes bonded to the middle nodes have already been counted. The innermost shells with $k < k_{next}$-2 are not needed and can be deleted. In this way, the memory required grows quadratically with k, while other algorithms presented in the literature [45] have a cubic growth rate. This algorithm was suggested by G.O. Brunner (member of our research group).

3.4.2 Determination of a LC: modification of the node counting algorithm

The term LC as used here follows a definition in [46] (where the term “Maschensymbol” is used) and is a generalization of the LC as defined in [4]. The LC of a framework node N_i with NN_i *node-node bonds* is understood as a set of $\binom{NN_i}{2}$ (binomial coefficient) pairs of integer numbers. Each pair characterizes the angle described by node N_i in the center and two bonded nodes. Fig. 3-6 gives an illustration of the six angles found for a node which is coordinated by four neighboring nodes.

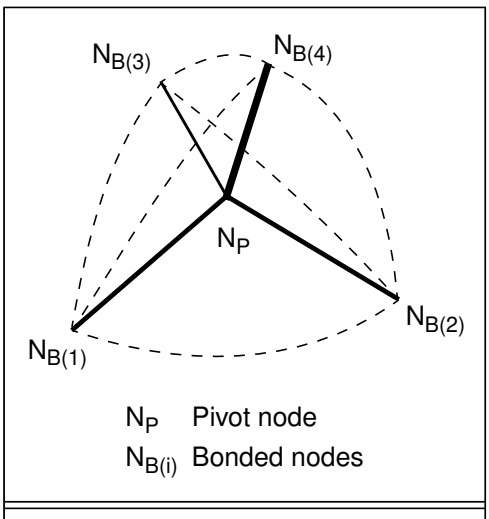


Figure 3-6: Six tetrahedral angles

The first integer of a pair is the number of nodes in the shortest loop which contains the corresponding angle. The second integer gives the number of loops with that number of nodes. For example, the loop configuration “4 1 4 1 5 1 5 1 6 1 7 2” says that two (of the six) angles are each part of single loops with four nodes, two angles are each part of single loops with five nodes, one angle is part of a loop with six nodes, and one angle is part of two distinct loops each with seven nodes.

The LC determination algorithm is very similar to the CS algorithm. The modified node counting algorithm is surrounded by an outer loop which steps over $NN_i - 1$ bonded nodes. Let j be the index in the list of bonded nodes (the first entry has index 0), such that $N_{b(j)}$ is the pivot node in this loop. Each pivot node is taken as the *initial node* ($k = 0$), and the algorithm works its way through the coordination shells until all *target nodes* $N_{b(j+1)} \ldots N_{b(NN_i-1)}$ are visited. The crucial modification of the CS algorithm is that bonds to the center node N_i are never followed.

Each time a target node is hit, $k + 2$ gives the number of nodes in the corresponding loop. If the target was not hit before, this number is recorded and the counter for the number of

loops is set to one. If the target was hit before in the same shell (that means with the same loop size) the counter is advanced by one.

After all integer pairs are obtained, they are sorted in ascending order to give the final LC for the node N_i.

Strictly speaking, a definition of the term loop has not been given, but follows immediately from the design of the LC algorithm. For clarification it should be mentioned that *loop* as used here is not equivalent to what is "commonly" referred to as *ring*. Fig. 3-7 illustrates the situation for the two 7-membered loops of the node **T5** in the topology **DDR** [4].

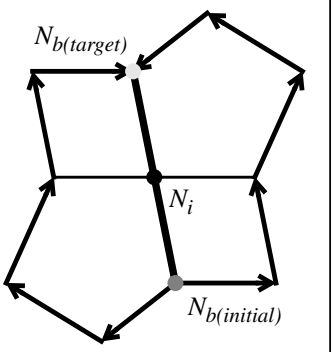


Figure 3-7

Commonly, a ring is thought to be a linear sequence of nodes, with the last node connected to the first one. Each node has exactly two connections to two other nodes in the ring. The loops in Fig. 3-7 do not comply with this concept, since one of the nodes in each of the two loops has an additional third connection to the center node N_i. However, various "exact" definitions of the term *ring* have been presented in the literature, some of which come close to *loop* as used here. For a discussion refer to [47].

3.4.3 Combined evaluation of multiplicities, LC's and CS's

A characteristic "fingerprint" of a structure is obtained by constructing a sequence of integers for each node in the asymmetric unit, by merging site multiplicity, LC and CS as illustrated in Fig. 3-8 (which is for node **T5** in the topology **DDR** [4]). The LC consists of

Multiplicity	Loop Configuration	Coordination Sequence
18	4 1 4 1 5 1 5 1 6 1 7 2	4 10 21 37 62 94 124 158 196 252

Figure 3-8

$\binom{4}{2}$ = 6 pairs of integers, and the CS is computed up to the 10^{th} member. Altogether one 4-connected node position is described by 23 integer numbers.

Two frameworks – as produced by the search algorithm – are considered to be equivalent if the sets of lexically sorted integer sequences are equal.

However, it has to be mentioned that Fischer [48] has derived four pairs of distinct sphere packings which cannot be distinguished by comparing the integer sequences. On the other hand, these examples look unrealistic for crystal structures, and no example is known where two crystal structures cannot be distinguished.

4 Applications

The FOCUS procedure has been applied to six test cases of different complexity and to three previously unknown structures. Characteristic data for the nine structures is summarized in Tab. 4-1. In all cases, the full procedure outlined in Fig. 3-1 was followed. The whole-profile intensity extraction was carried out on measured data using the GSAS program suite in “Le Bail extraction mode” with “CW Peak profile type no. 2” [32]. The refined profile parameters were used to prepare the overview of the overlap situation shown in Fig. 4-1. The overlap factor (eq. 3-1) used is 0.3, which means that reflections which are less than about 30% of their FWHM apart are put into the same overlap group. The plot shows how the ratio of overlapping and non-overlapping reflections develops with increasing resolution. For example, down to a d-spacing of 5.0 Å all reflection of EMC-2 are single, at a resolution of 3.0 Å, about 24% of all reflections overlap, and finally at 1.3 Å, the degree of overlap has reached 83%.

Name	Formula	Space group	Unit cell	Volume	Overlap*
Dodecasil-1H	$Si_{34}\ O_{68}$	P 6/m m m (No. 191)	a = 13.798 Å c = 11.211 Å	1848 Å^3	15 %
NU-3	$[Si_{54}\ O_{108}]\ \cdot(C_{10}H_{15}NH_2)_6$	R $\bar{3}$ m (No. 166)	a = 13.184 Å c = 22.221 Å	3345 Å^3	42 %
RUB-17	$K_4\ Na_{12}\ [Si_{28}\ Zn_8\ O_{72}]\ \cdot 18\ H_2O$	C m (No. 8)	a = 7.239 Å b = 40.562 Å c = 7.309 Å β = 91.84°	2145 Å^3	52 %
SAPO-40	$[(Si,Al,P)_{32}\ O_{64}]\ \cdot 2((CH_3CH_2CH_2)_4NOH)$	P m m n (No. 59)	a = 22.041 Å b = 13.698 Å c = 7.122 Å	2150 Å^3	64 %
Zeolite-A	$Na_{96}\ [Al_{96}\ Si_{96}\ O_{384}]\ \cdot 150\ H_2O$	F m $\bar{3}$ c (No. 226)	a = 24.558 Å	14811 Å^3	67 %
EMC-2	$Na_{11}\ [(Si,Al)_{96}\ O_{192}]\ \cdot 6\ H_20$	P 6_3/m m c (No. 59)	a = 17.378 Å c = 28.344 Å	7413 Å^3	83 %
VPI-9	$(NH_4^+)_{24}\ [Si_{44}\ Zn_{12}\ O_{112}]\ \cdot 50H_20$ †	P 4_2/n c m (No. 138)	a = 9.895 Å c = 36.872 Å	3610 Å^3	47 %
VPI-10	$(NH_4^+)_{16}\ [Si_{28}\ Zn_8\ O_{72}]\ \cdot 28\ H_20$ †	I 2 m m (No. 44)	a = 12.599 Å b = 21.810 Å c = 7.022 Å	1930 Å^3	80 %
B2	$K_4\ Na_4\ [Si_{16}\ Be_4\ O_{40}]\ \cdot\ 16\ H_2O$ †	P 2_1 m a (No. 26)	a = 13.173 Å b = 7.126 Å c = 12.678 Å	1190 Å^3	34 %

Table 4-1: Summary of characteristic data of the structures presented.
The first six are test cases and the last three novel structures.

*N(overlap)/N(total)*100 at a resolution of 1.3 Å
†estimated formula

In all cases, data up to a resolution of 1.3 Å were used (indicated by the dashed line in Fig. 4-1). In the following sections, the test cases are presented in the order of increasing overlap.

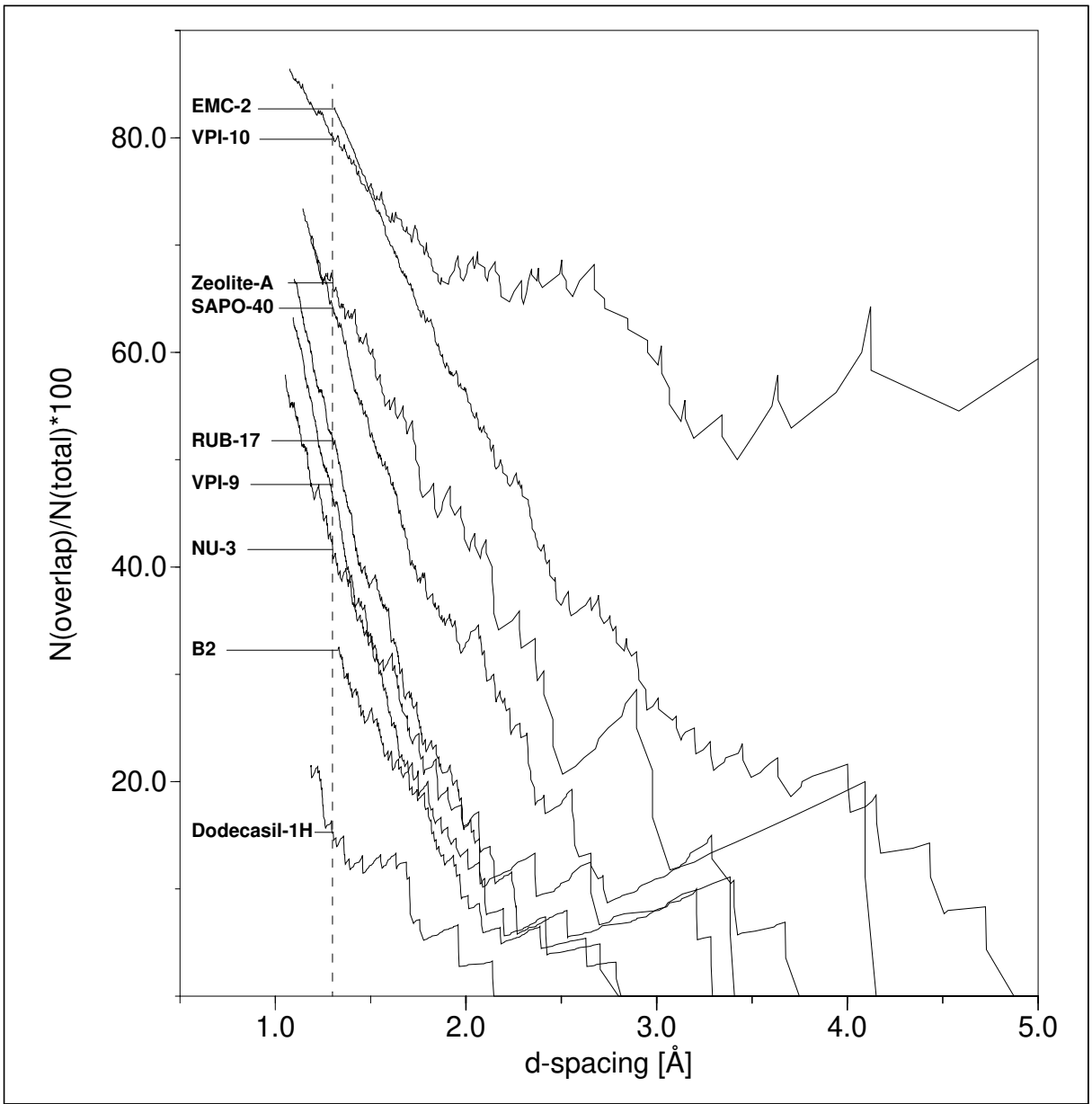


Figure 4-1: Overview of the overlap situation

4.1 The test structure Dodecasil-1H (**DOH**)

4.1.1 Preparation

Gerke & Gies [49] solved the structure by means of single crystal measurements and direct methods. Fig. 4-2 shows the **DOH** topology.

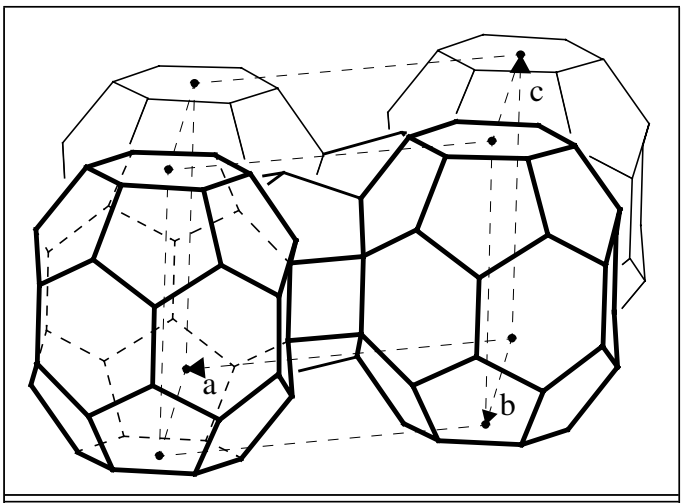


Figure 4-2: The framework of **DOH**
(O atoms have been omitted for clarity)

The polycrystalline Dodecasil-1H sample used for this investigation was available from a previous study [50]. To eliminate the organic template (Quinuclidine = $C_7H_{13}N$), the sample was calcined for six days at a temperature of 850 ˚C. The powder profile was collected on a STOE Stadi-P diffractometer with strictly monochromatic Cu-$K_{\alpha 1}$ radiation and a linear position sensitive detector (PSD), which covers approximately 6° 2θ. The sample was filled into a 0.3 mm capillary, which was rotated during the measurement. Initial lattice constants were obtained by running the PEAKFIND [51] program, and using the peak positions as input for the POWDER (indexing) program of Taupin [52]. After manual determination of the background intensity, GSAS was used to extract integrated intensities up to a resolution of 1.19 Å. Tab. 4-2 shows the relevant data for Dodecasil-1H and the final values for the eight refined parameters (all GSAS [32] parameters are defined in the GSAS user manual). The final plot of the profile fit is shown in Fig. 4-3. Due to pronounced peak shape asymmetry in the low angle region, the fit is not fully satisfactory (see inset in Fig. 4-3). However, the integrated intensities are only slightly affected by this phenomenon, and no further action was considered to be necessary. The overall temperature factor was held fixed at $U_{iso} = 0.02$ Å^2, and the Xtal 3.2 GENEV module was used to determine the scaling factor K = 0.08 (see page 8).

Formula		
	Si_{34} O_{68}	
Number of node atoms in the asymmetric unit		
	Observed space group	4
	Highest topological symmetry (P 6/m m m)	4
Data collection		
	STOE Stadi-P, linear PSD	
	Rotating 0.3 mm capillary	
	Cu-$K_{\alpha 1}$ (1.5406 Å) radiation	
	2θ range 5 - 80°, step size 0.01°	
Intensity extraction		
	Space group	P 6/m m m (No. 191)
	Unit cell	a = 13.798 c = 11.211 Å
	GU	90.6
	GV	-52.3
	GW	15.3
	LX	5.274
	LY	6.036
	asym	0.4008
	R_p	0.0813
	R_{wp}	0.1381

Table 4-2: Selected data for Dodecasil-1H

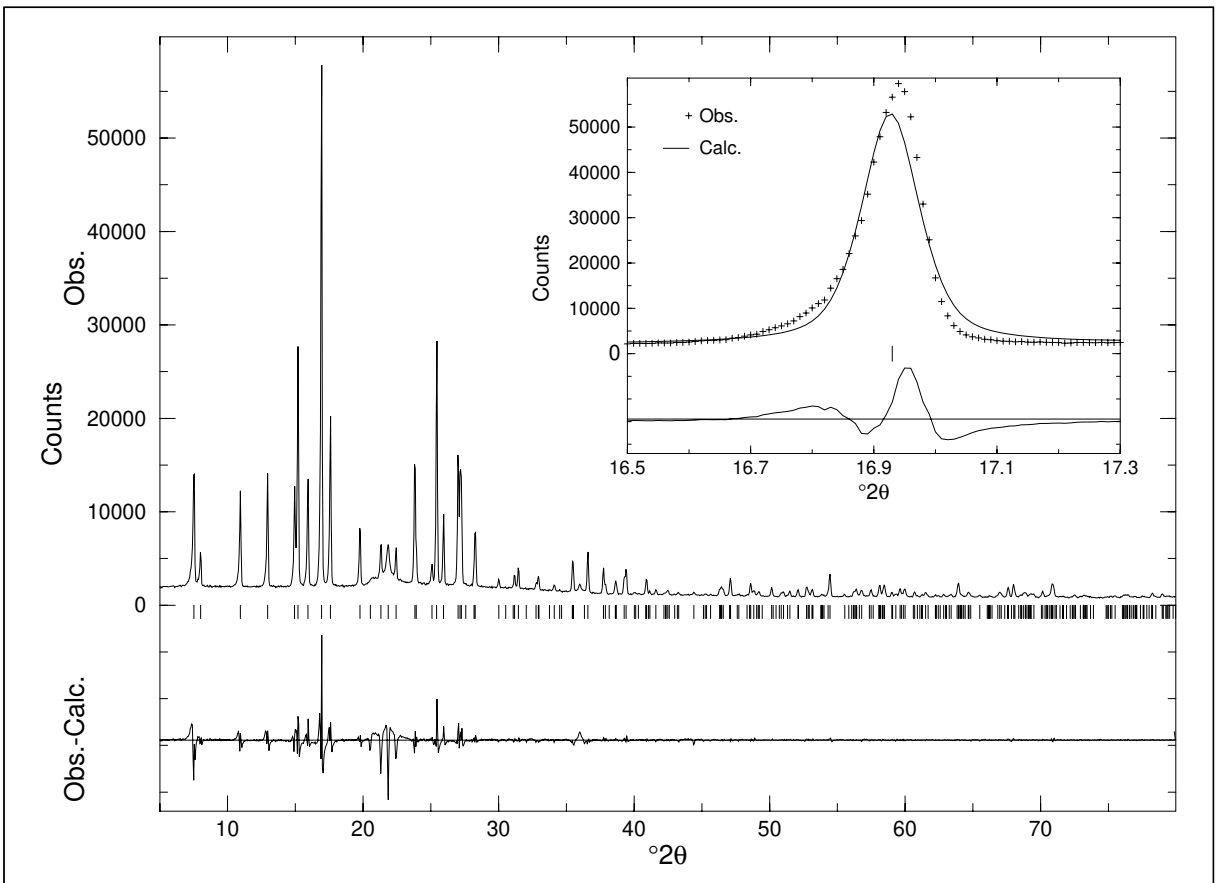


Figure 4-3: GSAS whole-profile intensity extraction for Dodecasil-1H

4.1.2 The FOCUS input file

The main parts of the input file for the FOCUS run are shown in Fig. 4-4. The general information supplied at the beginning defines the space group and lattice constants as refined with GSAS.

The next two `AtomType` lines define the cell contents of the structure to be solved, as determined by a chemical analysis or estimated by other means. The first item after the keyword `AtomType` is either "`+`" or "`–`". All atoms specified with an AtomType line are used in the calculation of F000 (the Fourier magnitude at the origin of reciprocal space), but only atoms with the "`+`" marker are considered in the atom and/or the framework fragment recycling procedures. The next item is a "class label" `Node`, `NodeBridge`, or "`*`", where the latter is for non-framework atoms (see page 13). After the class label, an "atom label" and the number of atoms of this type per unit cell are supplied. Also possible - but not used here – is the definition of the occupancy factor to be used in the recycling (preset to 1.0), the isotropic temperature factor (preset to 0.035), and a "scattering factor label" (derived from the preceding atom label).

> For example, the line
> ```
> AtomType  -  *  Ow  20  1.25  0.05  O
> ```
> describes an oxygen with an occupancy of 1.25 and an isotropic temperature factor of 0.05 Å^2, which is a commonly used approximation for water molecules in zeolite channels. The scattering factor used is that of oxygen, and 20 water per unit cell are expected. However, experience has shown that recycling extra framework atoms is not efficient, and for the calculation of F000 it would be sufficient to supply one `AtomType` line for oxygen and one for hydrogen using the default occupancy and temperature factors.
> The scattering factor labels and the corresponding scattering curves are defined in [53].

The next block of five lines is related to the atom recycling procedure. The `Chemistry MinDistance` lines define the individual minimum distances for each pair of atom types which are used in the atom recycling procedure. Following `Chemistry MinDistance` are two pairs of "class label" and "atom label" as defined on `AtomType` lines, and the minimum distance for this pair of atom types in the same units as the lattice constants, usually Å.

> Remark: since there is a "-" on the `AtomType` line for `NodeBridge O`, this atom type is not used in the atom recycling procedure. Therefore it would be sufficient to supply only the first Chemistry MinDistance line.

`MaxPotentialAtoms` gives the maximum number of peaks which are considered in the assignment algorithm. For example, with the value in Fig. 4-4, if the algorithm tries to assign a silicon atom to one of the peaks in the asymmetric unit, but is not able to find a valid position among the peaks in the asymmetric unit which generate the 46 highest peaks in the unit cell,

```
Title Dodecasil-1H calc. 850 deg.C Stoe-PSD/ETH GSAS extraction
SpaceGroup  P 6/m m m
UnitCell  13.7977 11.2105

AtomType  +  Node        Si  34
AtomType  -  NodeBridge  O   68

Chemistry  MinDistance  Node        Si  Node        Si  2.6
Chemistry  MinDistance  Node        Si  NodeBridge  O   1.4
Chemistry  MinDistance  NodeBridge  O   NodeBridge  O   2.3
MaxPotentialAtoms  46
MaxRecycledAtoms  34

FwSearchMethod  FwTracking
MaxPeaksFwSearch  680
MaxPeaksFwFragmentSearch  408
MinNodeDistance  2.6
MaxNodeDistance  3.6
MinSymNodes  0
MaxSymNodes  48
NodeType  4  *  -6 -3 -1 4 6
MinLoopSize  3
MaxLoopSize  24
EvenLoopSizesOnly  Off
Check3DimConnectivity  On
IdealT_NodeDistance  3.1
CheckTetrahedralGeometry  Hard

RandomInitialization  Time
FeedBackCycles  1 1 1 1 1 1
FeedBackBreakIf  PhaseDiff < 5.00 % and DeltaR < 1.00 %

Grid_xyz  42  42  32
eDensityCutOff  1 %
MinPfI  17
CatchDistance  0.5
eD_PeaksSortElement  Grid_eD

Lambda  CuA1
FobsMin_d  1.3
FobsScale  0.08
SigmaCutOff  0
OverlapFactor  0.15
OverlapAction  EqualMF2
ReflectionUsage  75 %
Grid_hkl  +11  22  18

#  h   k   l       Fobs      Sigma     FWHM
   1   0   0      817.68        *   0.11672
   0   0   1      726.02        *   0.11636
   1   0   1      644.98        *   0.11439
   1   1   0     1139.74        *   0.11322
End
```

Figure 4-4: A sample input for the Dodecasil-1H test structure

the silicon is not assigned at all. `MaxRecycledAtoms` prescribes the maximum number of atoms in the unit cell that are actually assigned and is forced to be smaller or equal to `MaxPotentialAtoms`.

The following block of 14 lines specifies the parameters for the framework and framework fragment search procedure. `FwSearchMethod` is either `FwTracking` or `AltFwTracking`, which are simple backtracking and "colored" backtracking, respectively (see page 21). When atoms are recycled and only complete frameworks are sought, `MaxPeaksFwSearch` defines the maximum number of peaks in the unit cell that are used in the backtracking procedure.

In framework fragment recycling mode, `MaxPeaksFwFragmentSearch` determines the maximum number of peaks. Since the fragment search is significantly slower than the search for complete frameworks only, it is sometimes necessary to set `MaxPeaksFwFragmentSearch` to a smaller value than `MaxPeaksFwSearch` in order to retain reasonable computing times.

`MinNodeDistance` and `MaxNodeDistance` establish the lower and upper limits for the node-node distances which are used in the preparation of the lists of potential node-node bonds (see page 18). In this case, a tolerance of 0.5 Å around the "ideal" distance of 3.1 Å is set.

`MinSymNodes` and `MaxSymNodes` set the lower and upper limits for the number of framework nodes per unit cell. While `MinSymNodes` just prevents frameworks with too low a density from being evaluated and printed, `MaxSymNodes` cuts complete branches of the search tree. On the one hand, this can reduce the computing time for frameworks with a well-established low density, but on the other, one has to be careful not to prescribe a value that is too small.

The `NodeType` line defines the number of bonds for a given node type, the maximum number of nodes of this type in the asymmetric unit and a list of the symmetry elements which can not be occupied by a node of this type. In Fig. 4-4, only one node type with tetrahedral connectivity is defined. The asterisk "`*`" specifies that an unlimited number of nodes in the asymmetric unit can be of this type. The following numbers "`-6 -3 -1 4 6`" specify that this node type cannot be on a six- or threefold rotoinversion axis, an inversion center, or a four- or sixfold rotation axis.

Supplying a value greater than three for `MinLoopSize` has two consequences: when atoms are recycled and only complete frameworks are sought, frameworks which have loops with less than `MinLoopSize` members are rejected (that just means they are not printed). In framework fragment recycling mode, the fragments which are candidates for the "largest

fragment" for recycling are checked for `MinLoopSize`. Unfortunately, the present implementation of the loop size test is very time consuming. The time spent for the fragment search increases by roughly 40%. In this example, `MinLoopSize` was therefore kept at its default value of three, although four is perhaps more appropriate for high silica frameworks. (However, the structure of the high silica ZSM-18 (**MEI**) does contain 3-rings).

`MaxLoopSize` is less critical than `MinLoopSize` and just specifies the maximum loop size up to which the LC algorithm advances (see page 23). The default value of 24 is sufficient for all known zeolite topologies. For loops with more than `MaxLoopSize` members, a "0" is printed. Cases where smaller values would result in a speed gain for the price of having some zeros in the LC are hardly imaginable.

In this example, `EvenLoopSizesOnly` is switched `Off`. This means, all loop sizes greater than or equal to `MinLoopSize` are allowed. The `EvenLoopSizesOnly` option was introduced for the search for frameworks where a strict alternation of two atom types is expected. In these cases, only even loop sizes are possible. `EvenLoopSizesOnly` provides an alternative to `AltFwTracking`, especially for aluminum phosphates. For further discussion refer to the SAPO-40 test case on page 51.

> It has to be noted that in framework fragment search mode the impact of `EvenLoopSizesOnly` on the computing time requirements is similar to setting `MinLoopSize` to a value greater than three. However, since loop sizes have to be computed only once per framework or framework fragment, `MinLoopSize` greater than three does not result in more time consumption if `EvenLoopSizesOnly` is switched `On`.

`Check3DimConnectivity` is followed by one of the keywords `On` or `Off`. If `On`, a filter procedure is called for each framework topology found (see page 20). Only 3-dimensionally connected frameworks can pass this filter, layer or chain structures are rejected. `IdealT_NodeDistance` specifies the "ideal" node-node distance for four-connected nodes. This is the basic value for the geometrical tests described in chapter 7.3, which are further specified by the `CheckTetrahedralGeometry` keyword, which is followed by `Off`, `Normal`, or `Hard`. For high silica and Si-Al frameworks like Dodecasil-1H, the `Hard` test is appropriate.

The next block with three input lines describes the initialization and development of the "trials". The keyword `RandomInitialization` is used to define the "seed" value for the portable pseudo random number generator [54], which is used to generate the starting phases. The special value `Time` tells FOCUS to use the machine time for the automatic determination of the seed value, which is then printed on the output file. This integer value – like any positive integer value – can be resupplied with `RandomInitialization` in order to rerun FOCUS with

different output options or for testing or debugging purposes.

The `FeedBackCycles` keyword is followed by an arbitrarily long sequence of nonnegative integers (including zero). The first integer specifies the number of times the atom recycling procedure is to be used in one trial, the second integer is for the number of framework fragment recycling loops, the third again for atom recycling, and so on. In Fig. 4-4, six cycles with alternation of atom and framework fragment recycling are requested. However, as the next keyword `FeedBackBreakIf` indicates, the recycling is prematurely terminated if both the phase set and the R_F residual value have converged. Another special situation is, when no fragment which can be recycled is found. In this case, a trial continues with atom recycling (but the cycle is still counted as framework fragment recycling cycle).

The next block concerns the layout of the electron density map and the characteristics of the peak search and refinement. In Fig. 4-4, the grid for the electron density map is defined such that a resolution of about 1/3 Å is achieved. One has also to take care that all symmetry elements pass through grid points. In its present form, FOCUS does not automatically generate an appropriate grid, but it does refuse to work with grid sizes that do not conform to this requirement. For example, in space group $P\bar{1}$ the grid sizes for all directions have to be a multiple of two, in order to have all inversion centers laying on a grid point. In the present case of space group P6/*mmm*, the grid size in the z-direction has to be a multiple of two, and the grid sizes for the x- and y-direction have to be a multiple of six.

The `eDensityCutOff` value specifies the lower cut-off value for the peak search in the electron density maps. This specification can either be an absolute value, e.g. `eDensityCutOff 1.0`, or relative to the maximum value of the whole map, as is in Fig. 4-4. The overall maximum value of `MaxPotentialAtoms`, `MaxPeaksFwSearch`, and `MaxPeaksFwFragmentSearch` is the maximum number of peaks in the unit cell which are put on the peaklist by the peaksearch procedure. However, if there are less than this number of peaks with a maximum peak height above the value set by `eDensityCutOff`, the list will contain fewer peaks. The next three keywords, `MinPfI`, `CatchDistance`, and `eD_PeaksSortElement` determine the behavior of the peaklist refinement procedures. `MinPfI` (“minimum number of points for interpolation”) defines the minimum number of grid points with a positive electron density value surrounding a grid peak position. If the actual number is fewer than `MinPfI`, no interpolation for the peak position is carried out and the coordinates of the grid point are retained (see page 14). `CatchDistance` is the minimum distance a peak has to have to all of its symmetrically equivalent peaks (self-distance). For self-distances smaller

than `CatchDistance`, a procedure is activated, which moves the peak onto the symmetry element which is responsible for the close contact.

> Depending on the Wyckoff position, this process is repeated up to three times (for example, with space group P*mmm* and a position close to all three mirror planes, after a first move onto one mirror plane there are two degrees of freedom left, the second move leaves only one degree of freedom along the intersection of the two mirror planes, and only with the third move is the final position found).
> In cases, where no peak position with a self-distance below `CatchDistance` can be found, FOCUS terminates and issues an error message. This should only happen when `CatchDistance` is unreasonably high compared to the lattice constants defined with `UnitCell`. Peaks with a self-distance larger than CatchDistance but smaller than the smallest interatomic distance set for the atom recycling or framework fragment recycling procedures (see below) are effectively dead. However, at present no extra provision has been made to eliminate such peaks from the list.

After all peak positions have been refined, the peaklist is sorted according to `eD_PeakSortElement`, which can be specified as `Grid_eD`, `Maximum`, or `Integral` (see page 16).

The last block specifies treatment and usage of the extracted intensities. First of all, the wavelength used in the diffraction experiment is specified with `Lambda`, followed by either a decimal value for the wavelength (in the same units as the values supplied with `UnitCell`) or one of the codes for the internally stored wavelengths (which are in Å units). (For the list of available keywords see page 122.) `FobsMin_d` sets the minimum d-spacing for the reflections to be used. `FobsScale` defines the scale factor, which was determined with the Xtal GENEV module. `SigmaCutOff` is set to zero in this example, because the GSAS REFLIST command does not produce standard deviations for the extracted intensities. If standard deviations are available, reflections with an intensity smaller than `SigmaCutOff` times their standard deviation can be excluded.

The `OverlapFactor` together with the individual FWHM for each reflection is used to determine the overlap groups, which are then processed according to `OverlapAction`, which is one of `NoAction`, `EqualF2`, or `EqualMF2` (see page 10). (To comply with the convention adopted by the Xtal module DIVIDE, the input `OverlapFactor` is *off*/2 of eq. 3-1).

`ReflectionUsage` specifies the number of reflections that are actually used. This can be absolute, for example `ReflectionUsage 80` will select the 80 highest reflections, or it can be relative, as in Fig. 4-4. In the latter case, reflections are selected in descending order of (equipartitioned) intensity times multiplicity ($M \cdot F$) until the prescribed percentage of the total sum of $M \cdot F$ over all input reflections is accumulated.

At the moment, one also has to prescribe grid sizes for reciprocal space with `Grid_hkl`, such that there is a grid point for all selected reflections and their symmetry equivalents. This is due to the development history of the program, when tests with fast Fourier transform algorithms were made. Through restructuring of the transform procedures, the necessity for the specification of the grid in reciprocal space could be eliminated.

> FOCUS always ignores anomalous dispersion effects and treats the Fourier magnitudes in reciprocal space as centro-symmetric. Therefore only one half of the reciprocal space needs to be transformed. The direction which is to be halved (usually the direction with the most grid points) can be specified by a plus sign in front of the corresponding grid size. In Fig. 4-4, the grid is halved in the x-direction, ranging from 0 to 10. The grid in the y-direction ranges from -11 to 10 and in the z-direction from -9 to 8.

The last part of the input file is a listing of the extracted Fourier magnitudes. The data are given as reflection indices hkl, observed relative Fourier magnitude, the estimated standard deviation of the Fourier magnitude and the FWHM as derived from the refined profile parameters. As mentioned before, GSAS does not produce estimated standard deviations. Therefore asterisks are supplied instead.

4.1.3 Tests with varying reflection usage and recycling strategy

Early tests revealed that the Dodecasil-1H framework is easily recovered using FOCUS. Therefore it is possible to get quick responses on the effect of parameter changes. However, it has to be mentioned that this test structure is a very “simple” case for FOCUS and is recovered with almost any reasonable input parameters. The results which follow are typical, but not readily transferable to other structures.

Fig. 4-5 shows the histograms of absolute frequency of occurrence of unique topologies for a series of test runs. The input files were derived from the sample input in Fig. 4-4. In the left column of Fig. 4-5, `ReflectionUsage` was set to 50%, which results in the active use of 55 of the 214 reflections. The second input line which was varied is `FeedBackCycles`. The recycling sequences used are printed in the upper right corner of each histogram. The same recycling strategies were used in the middle and right column, but with `ReflectionUsage` 75% and 100% respectively.

For each histogram of Fig. 4-5, “FT” indicates the number of Fourier transforms which were made in 100 trials. Furthermore the computing time in minutes is given.

In all cases, the most frequently occurring framework is the **DOH** topology. Two conclusions can be drawn from the histograms: the choice of `ReflectionUsage` is not critical for the success rate, and recycling with framework fragment search is superior to atom

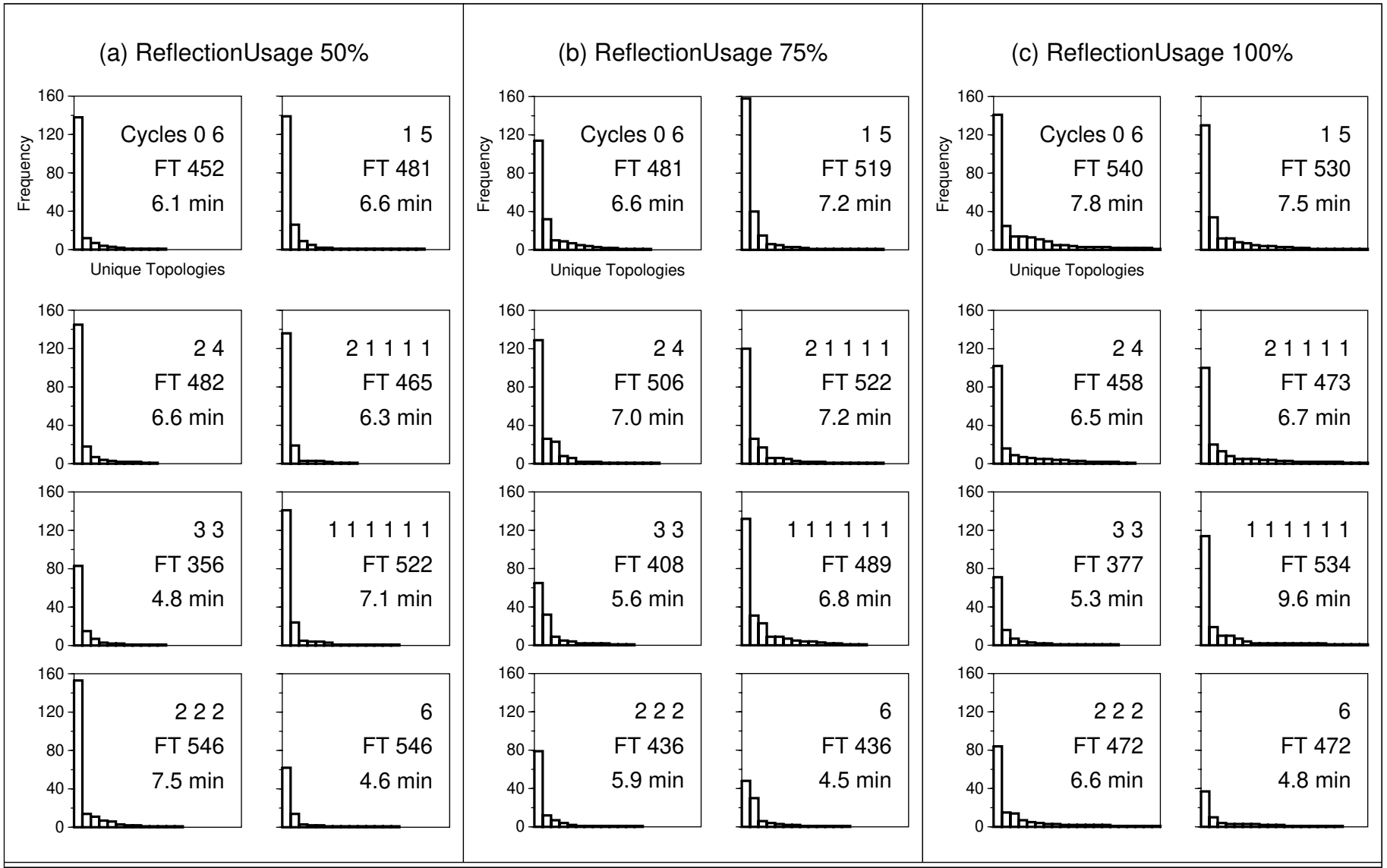


Figure 4-5: Influence of reflection usage and recycling strategy for Dodecasil-1H per 100 trials

recycling, and alternation of atom recycling and fragment recycling is even better. The exact alternation strategy does not appear to be very important, but long periods of atom recycling like `FeedBackCycles 3 3` are less productive. In this example, the fragment recycling alone works very well, so `FeedBackCycles 1 5` or `2 4` also yield high success rates. However, this is not transferable to more complex cases.

4.2 The test structure NU-3 (LEV)

The origin of the 1-aminoadamantane (ADAM) NU-3 sample, the collection of the synchrotron dataset, and the structure is described in [55]. Fig. 4-6 shows the **LEV** topology of NU-3.

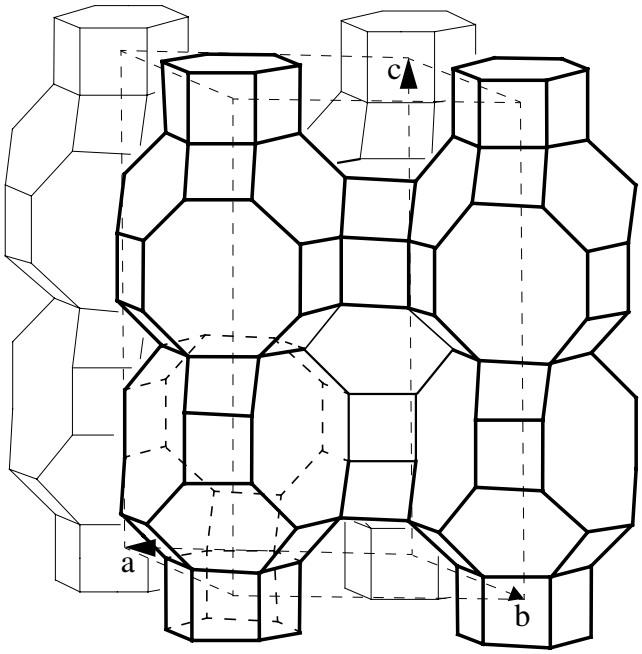


Figure 4-6: The framework of NU-3
(O atoms have been omitted for clarity)

Using GSAS, integrated intensities were extracted up to a resolution of 1.05 Å. Tab. 4-3 shows the relevant data for ADAM NU-3 and the final values for the seven refined parameters. The final plot of the profile fit is shown in Fig. 4-7.

In the last preparation step, the scaling factor was determined by running the Xtal GENEV module. The main parts of the FOCUS input are shown in Fig. 4-8. Since NU-3 is a high silica framework like **DOH**, the parameters for atom recycling and topology search are very similar to those of Fig. 4-4. However, a different test series of six runs was carried out. Based on the template of Fig. 4-8, four input lines were varied to give the input for “Run A” to “Run F”. The construction of the test series is given in Tab. 4-4. Three recycling strategies (rows in Tab. 4-4) were investigated, with and without recycling of oxygen in atom recycling mode (columns in Tab. 4-4).

The evaluation of the FOCUS results is done with the histograms of Fig. 4-9 and the data compiled in Tab. 4-5. The heading row of Tab. 4-5 is in turn explained in Tab. 4-6.

Formula		
	$[Si_{54}\ O_{108}]\ \cdot(C_{10}H_{15}NH_2)_6$	
Number of node atoms in the asymmetric unit		
	Observed space group	2
	Highest topological symmetry (R $\overline{3}$ m)	2
Data collection		
	Station 9.1, SRS, Daresbury, U.K.	
	Rotating 0.5 mm capillary	
	Wavelength 1.5388 Å	
	2θ range 8 - 92°, step size 0.01°	
Intensity extraction		
	Space group	R $\overline{3}$ m (No. 166)
	Unit cell	a = 13.184 c = 22.221 Å
	GU	240.8
	GV	-89.8
	GW	13.2
	LX	2.154
	asym	-0.0715
	R_p	0.0617
	R_{wp}	0.0975

Table 4-3: Selected data for ADAM NU-3

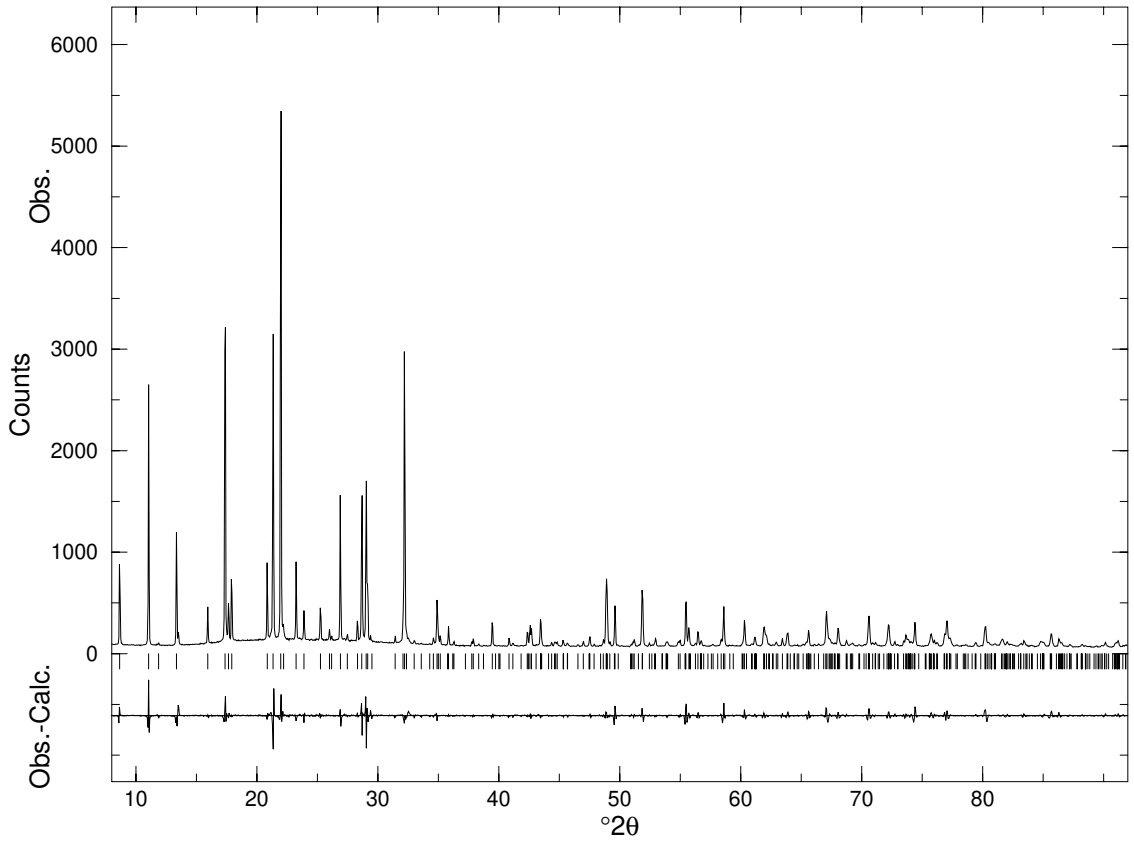


Figure 4-7: GSAS whole-profile intensity extraction for NU-3

```
Title NU-3, Daresbury Data, GSAS extraction
SpaceGroup  R -3 m
UnitCell  13.1835 22.2207

AtomType  +  Node        Si   54
AtomType  -  NodeBridge  O   108
AtomType  -  *           N     6
AtomType  -  *           C    60
AtomType  -  *           H   102

Chemistry  MinDistance  Node        Si  Node        Si  2.6
Chemistry  MinDistance  Node        Si  NodeBridge  O   1.4
Chemistry  MinDistance  NodeBridge  O   NodeBridge  O   2.3
MaxPotentialAtoms  64
MaxRecycledAtoms  54

FwSearchMethod  FwTracking
MaxPeaksFwSearch  960
MaxPeaksFwFragmentSearch  640
MinNodeDistance  2.6
MaxNodeDistance  3.6
MinSymNodes  0
MaxSymNodes  64
NodeType  4  *  -6 -3 -1 4 6
MinLoopSize  4
MaxLoopSize  24
EvenLoopSizesOnly  Off
Check3DimConnectivity  On
IdealT_NodeDistance  3.1
CheckTetrahedralGeometry  Normal

RandomInitialization  Time
FeedBackCycles  6
FeedBackBreakIf  PhaseDiff < 5.00 % and DeltaR < 1.00 %

Grid_xyz  42  42  66
eDensityCutOff  1 %
MinPfI  17
CatchDistance  0.5
eD_PeaksSortElement  Grid_eD

Lambda  1.53388
FobsMin_d  1.3
FobsScale  0.2
SigmaCutOff  0
OverlapFactor  0.15
OverlapAction  EqualMF2
ReflectionUsage  75 %
Grid_hkl  22  22  +17

#  h   k   l       Fobs  Sigma  FWHM
   1   0   1     216.54 *   0.07781
   1   0  -2     461.95 *   0.07348
End
```

Figure 4-8: A sample input for the ADAM NU-3 test structure

FeedBackCycles	AtomType – NodeBridge O 108 MaxPotentialAtoms 64 MaxRecycledAtoms 54	AtomType + NodeBridge O 108 MaxPotentialAtoms 162 MaxRecycledAtoms 162
6	Run A	Run B
0 6	Run C	Run D
1 1 1 1 1 1 1 1 1 1	Run E	Run F

Table 4-4: Construction of a test series

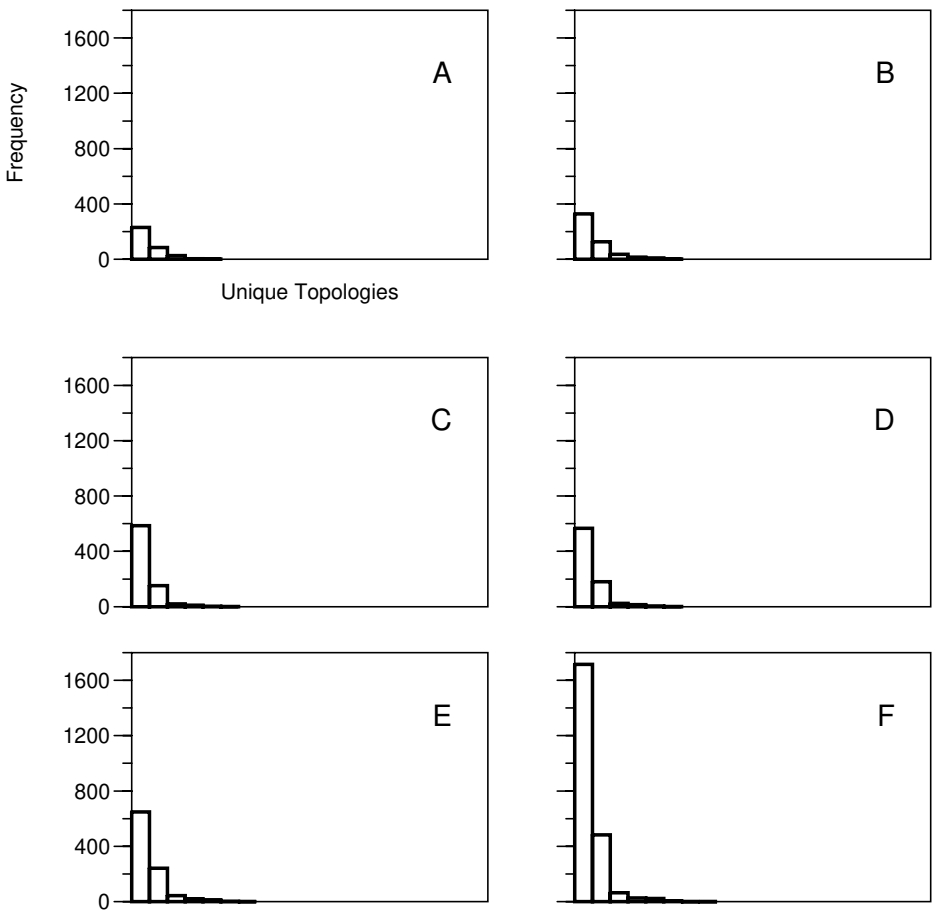


Figure 4-9: FOCUS Results for NU-3 (**LEV**) – 500 trials per run

For convenience and for compatibility with equivalent tables for the structures in the chapters to follow, the first five columns of Tab. 4-5 repeat the information given in a different arrangement in Tab. 4-4.

The main observation is that for all six runs, the most frequently occurring framework is the **LEV** topology, i.e. the correct solution. Secondly, in terms of the number of trials, pure atom recycling is less efficient than pure framework fragment recycling, which in turn is less efficient than alternating atom/fragment recycling. Moreover, atom recycling with oxygen, in general, is better than recycling of node atoms alone. The most efficient technique is found in

Run	NBO	MPA	MRA	FBC	FT	Fw	UFw	RFw	t/min	%tFwS
A	-	64	54	6	1577	350	5	7507	50	16
B	+	162	162	6	2130	519	6	10563	69	16
C	-	64	54	0 6	2327	775	6	6248	83	24
D	+	162	162	0 6	2397	795	6	6777	85	25
E	-	64	54	1 1 1 1 1 1 1 1 1 1	3395	971	7	13550	114	20
F	+	162	162	1 1 1 1 1 1 1 1 1 1	5362	2319	8	23144	186	23

Table 4-5: FOCUS Results for NU-3 (**LEV**) – 500 trials per run (108 actively used reflections)

Run	Label for run in test series
NBO	`AtomType ?` **`N`**`ode`**`B`**`ridge` **`O`** `...`
MPA	**`M`**`ax`**`P`**`otential`**`A`**`toms ?`
MRA	**`M`**`ax`**`R`**`ecycled`**`A`**`toms ?`
FBC	**`F`**`eed`**`B`**`ack`**`C`**`ycles ?`
FT	Number of **F**ourier **T**ransforms computed during the run
Fw	Total number of **F**rame**w**orks found
UFw	Number of **U**nique **F**rame**w**orks found
RFw	Total number of **R**ejected **F**rame**w**orks = Sum of number of frameworks + which were rejected by the geometry filter + with loop sizes < `MinLoopSize` + with odd loop sizes (only if the `EvenLoopSize` option is `On`)
t/min	Total computing **t**ime in **min**utes (MIPS R4400 CPU,150 MHz clock rate)
%tFwS	**%** of computing **t**ime spent for **F**rame**w**ork **S**earch

Table 4-6: Legend for headings in Tab. 4-5

"Run F", with alternation of recycling modes and use of oxygen in the atom recycling mode. This is also true, if the histograms are renormalized (Fig. 4-10) by dividing the frequency of occurrence of a unique topology by the total number of Fourier transforms (**FT** in Tab. 4-5). This means, the efficiency is measured relative to the number of Fourier transforms rather than relative to the number of trials. The advantage of technique F is now less pronounced, but still clear.

Another interesting aspect is the investigation of the second most frequently occurring topology (HB2 – "histogram bar 2"), which is the same for all six runs. As can be seen in Fig. 4-11, the projections of the NU-3 framework and the DLS refined HB2 framework along [001] are indistinguishable. However, as is shown in Fig. 4-12, the similarities between the measured powder profile of NU-3 and the simulated profile of HB2 are not very obvious. Yet, the influence of the intensities on the outcome of a FOCUS run can be demonstrated by repeating "Run F" with the calculated and equipartitioned Fourier magnitudes of HB2, rather than the magnitudes extracted from the NU-3 powder profile. In the resulting histogram

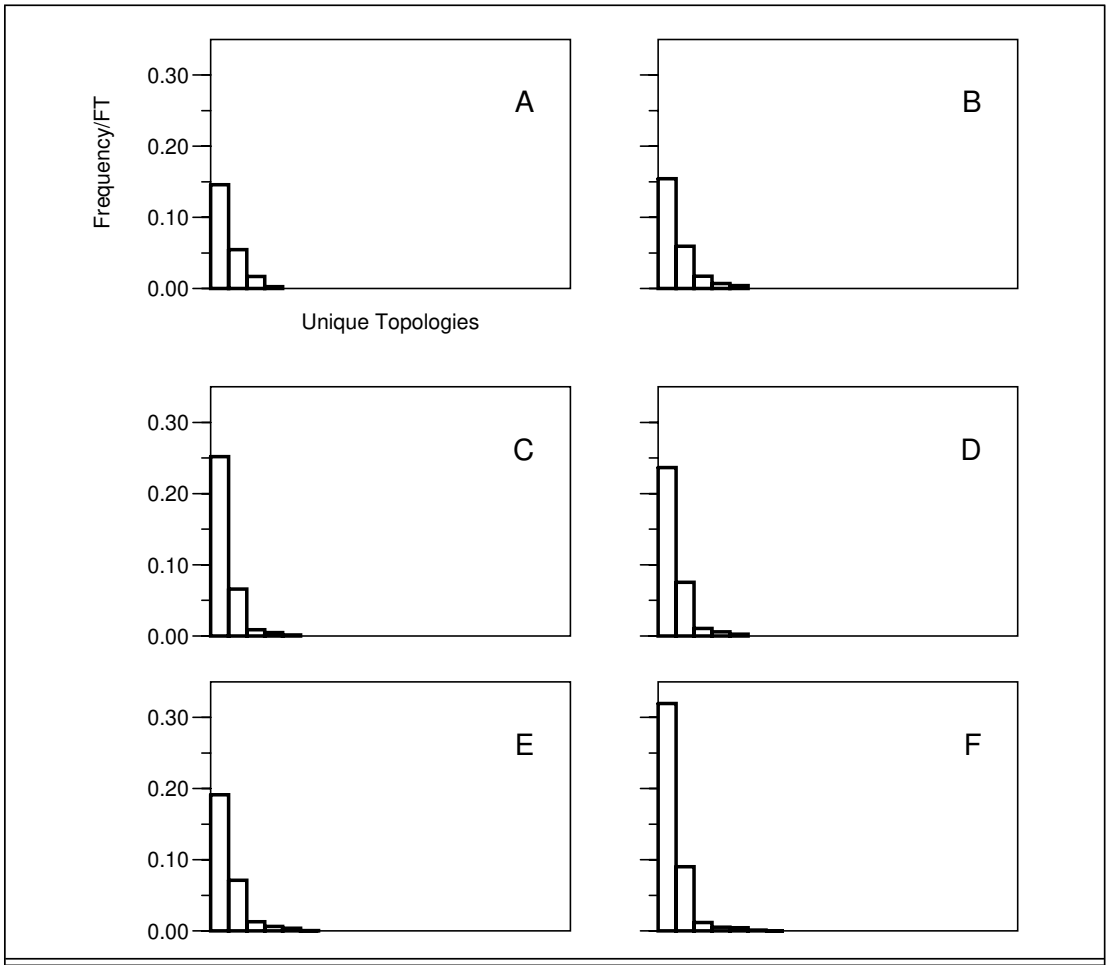


Figure 4-10: FOCUS Results for NU-3 (**LEV**) – Frequency per Fourier transform in 500 trials

(Fig. 4-13(a)), the first bar now corresponds to HB2, while the second one corresponds to the **LEV** topology. However, the difference between the first and the second bar is not as large as in Fig. 4-9(F). To check for the influence of the lattice constants on the preference of one topology over the other, the lattice constants of HB2 were refined with the DLS program. The lattice constants changed from a = 13.184 and c = 22.221 Å to a = 12.709 and c = 23.268 Å, and the DLS residual dropped from 0.0122 to 0.0048. (For comparison: the DLS residual for the **LEV** topology drops from 0.0061 to 0.0030 upon refinement of the lattice constants.) The final DLS refined coordinates of HB2 are listed in Tab. 4-7.

The procedure described above was then repeated with the new unit cell, and the results shown in Fig. 4-13(b). The difference between the first and the second bar in the histogram is even slightly less pronounced than in 4-13(a). Because the number of positional parameters is exactly equal for the two topologies, it is unclear why the preference for HB2 with HB2 Fourier magnitudes is less strong than the preference for NU-3 with NU-3 magnitudes. Nonetheless, the general rule of correspondence between the most frequently occurring topology and the correct solution still holds.

Space group	R $\bar{3}$ m		
Unit cell	a = 12.70897 c = 23.26756 Å		
Si1	0.00000	0.24441	0.00000
Si2	0.99636	0.24562	0.22068

Table 4-7: DLS coordinates of HB2 node atoms

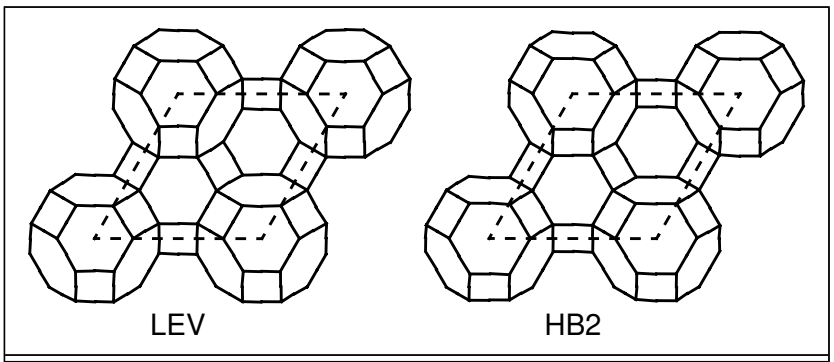


Figure 4-11: Projections of **LEV** and HB2 topologies along [001]

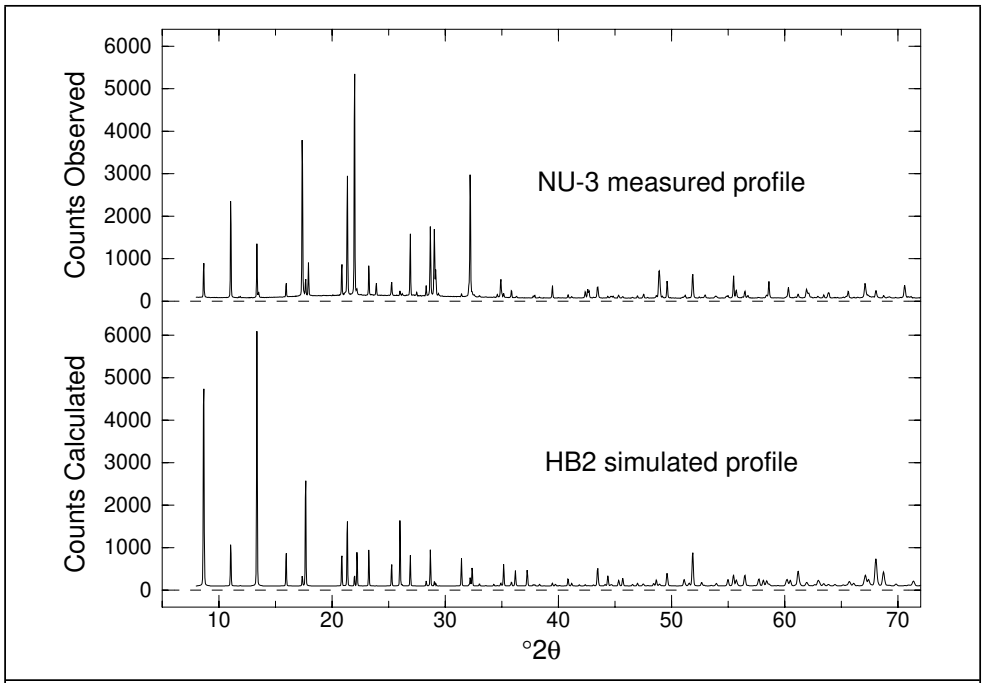


Figure 4-12: Comparison of NU-3 measured powder profile and HB2 calculated profile

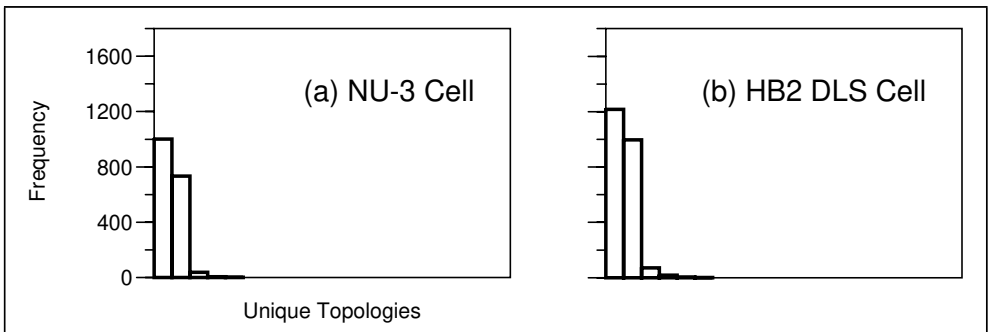


Figure 4-13: Result of 500 FOCUS trials with recycling technique F
and calculated and equipartitioned Fourier magnitudes of HB2

4.3 The test structure RUB-17 (RSN)

The powder profile for the test calculations with RUB-17 was kindly provided by C. Röhrig. Röhrig et al. [19] have described the synthesis procedure, the X-ray measurement on a laboratory diffractometer, and the structure solution, and have discussed the relationship between RUB-17 and other zincosilicate and beryllosilicate zeolites. Fig. 4-14 shows the topology of the RUB-17 structure.

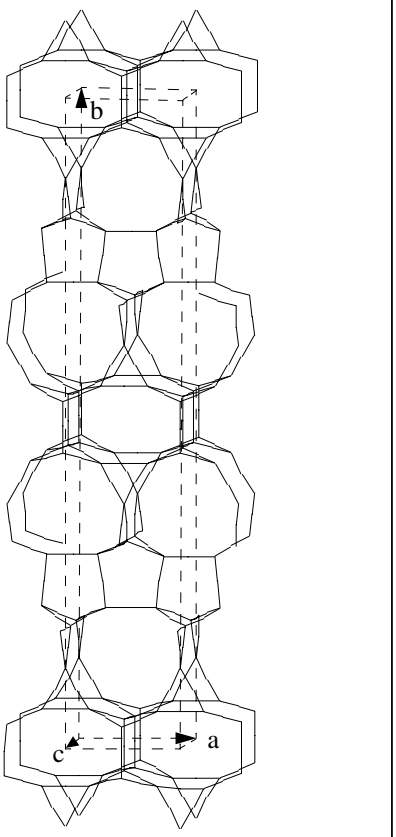


Figure 4-14: The framework of RUB-17
(O atoms have been omitted for clarity)

Using GSAS, integrated intensities were extracted up to a resolution of 1.1 Å. Tab. 4-8 shows the relevant data for RUB-17 and the final values for the 11 refined parameters. The final plot of the profile fit is shown in Fig. 4-15. After the determination of the intensity scaling factor with help of GENEV, the FOCUS input template of Fig. 4-16 was prepared. In the same manner as for the NU-3 test case, a test series of six runs was derived from the input template (see also Tab. 4-4). For the three runs without oxygen recycling, `MaxPotentialAtoms` and `MaxRecycledAtoms` were set to 72 and 36, respectively. In the other three runs with oxygen recycling, both parameters were set to 108.

The results of the FOCUS test runs are documented in Fig. 4-17 and Tab. 4-9 (for the

Formula		
	$K_4\ Na_{12}\ [Si_{28}\ Zn_8\ O_{72}]\ \cdot 18\ H_2O$	
Number of node atoms in the asymmetric unit		
	Observed space group	9
	Highest topological symmetry (C 2/m)	5
Data collection		
	Siemens D5000, linear PSD	
	Rotating 0.3 mm capillary	
	Cu-$K_{\alpha 1}$ (1.5406 Å) radiation	
	2θ range 7 - 87°, step size 0.01°	
Intensity extraction		
	Space group	C m (No. 8)
	Unit cell	a = 7.239 b = 40.562 c = 7.309 Å β = 91.84°
	Zero	-1.50 (⇒ -0.015° 2θ)
	GU	94.9
	GV	-33.5
	GW	6.7
	LX	1.509
	LY	5.404
	asym	0.2672
	R_p	0.0704
	R_{wp}	0.0900

Table 4-8: Selected data for RUB-17

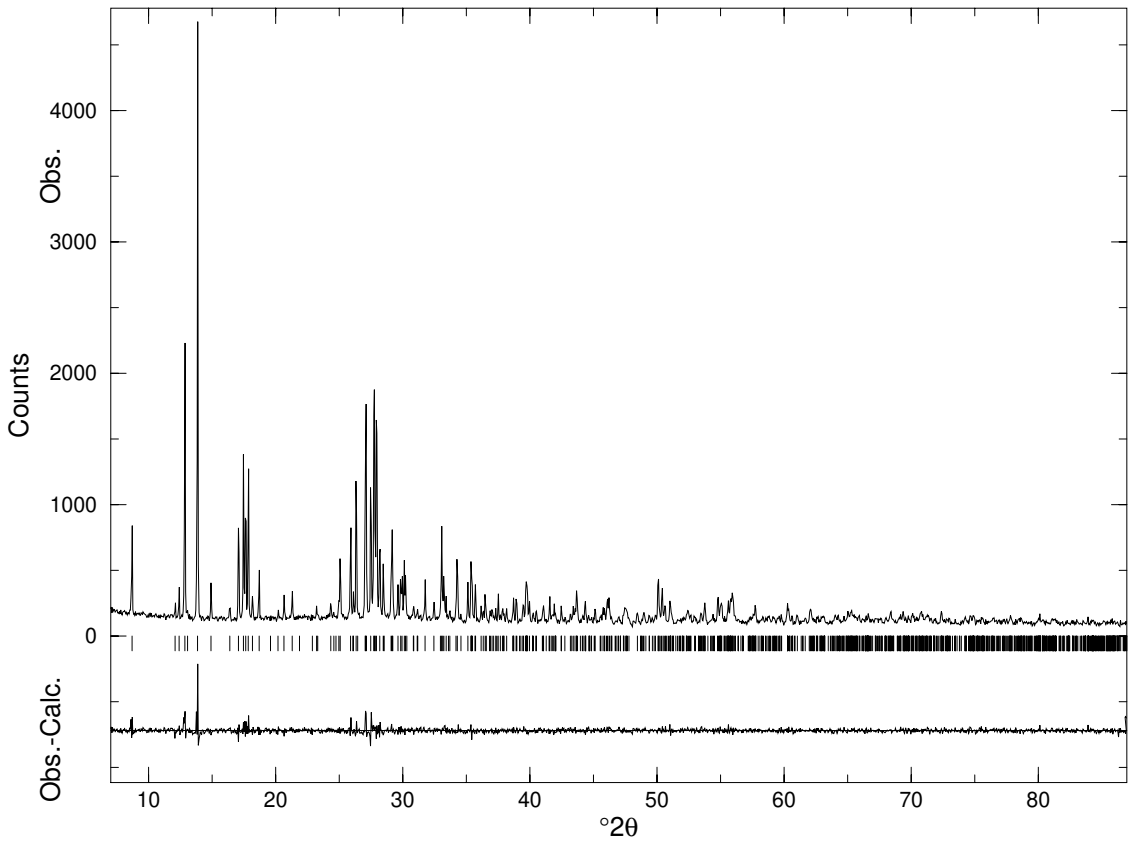


Figure 4-15: GSAS whole-profile intensity extraction for RUB-17

```
Title RUB-17 C. Rohrig  Siemens D5000  GSAS extraction
SpaceGroup  C 1 m 1
UnitCell  7.2392 40.5617 7.3086  90 91.844 90

AtomType  +  Node        Zn   8  *  0.01
AtomType  +  Node        Si  28  *  0.01
AtomType  -  NodeBridge  O   72  *  0.02
AtomType  -  *           K    4
AtomType  -  *           Na  12
AtomType  -  *           H   36
AtomType  -  *           O   18

Chemistry  MinDistance  Node        Zn  Node        Zn  4.0
Chemistry  MinDistance  Node        Zn  Node        Si  2.5
Chemistry  MinDistance  Node        Zn  NodeBridge  O   1.7
Chemistry  MinDistance  Node        Si  Node        Si  2.5
Chemistry  MinDistance  Node        Si  NodeBridge  O   1.5
Chemistry  MinDistance  NodeBridge  O   NodeBridge  O   2.3
MaxPotentialAtoms  72
MaxRecycledAtoms  36

FwSearchMethod  FwTracking
MaxPeaksFwSearch  200
MaxPeaksFwFragmentSearch  80
MinNodeDistance  2.3
MaxNodeDistance  3.7
MinSymNodes  0
MaxSymNodes  44
NodeType  4  *  -6 -3 -1 4 6
MinLoopSize  3
MaxLoopSize  24
EvenLoopSizesOnly  Off
Check3DimConnectivity  On
IdealT_NodeDistance  3.15
CheckTetrahedralGeometry  Normal

RandomInitialization  Time
FeedBackCycles  6
FeedBackBreakIf  PhaseDiff < 5.00 % and DeltaR < 1.00 %

Grid_xyz  22  120  22
eDensityCutOff  1 %
MinPfI  17
CatchDistance  0.5
eD_PeaksSortElement  Grid_eD

Lambda  CuA1
FobsMin_d  1.3
FobsScale  0.35
SigmaCutOff  0
OverlapFactor  0.15
OverlapAction  EqualMF2
ReflectionUsage  75 %
Grid_hkl  12  +31  12
```

Figure 4-16: A sample input for the RUB-17 test structure

explanation of the headings in Tab. 4-9 refer to Tab. 4-6). As in the previous test cases, the most frequently occurring unique topology is the correct solution in all six runs. Also familiar, the success rate of the recycling strategy with alternating atom and framework fragment recycling is superior to pure atom or fragment recycling. However, in this case, the advantage of alternating recycling modes is much more striking than before. This is also true for the discrimination of most frequently and second most frequently occurring topology.

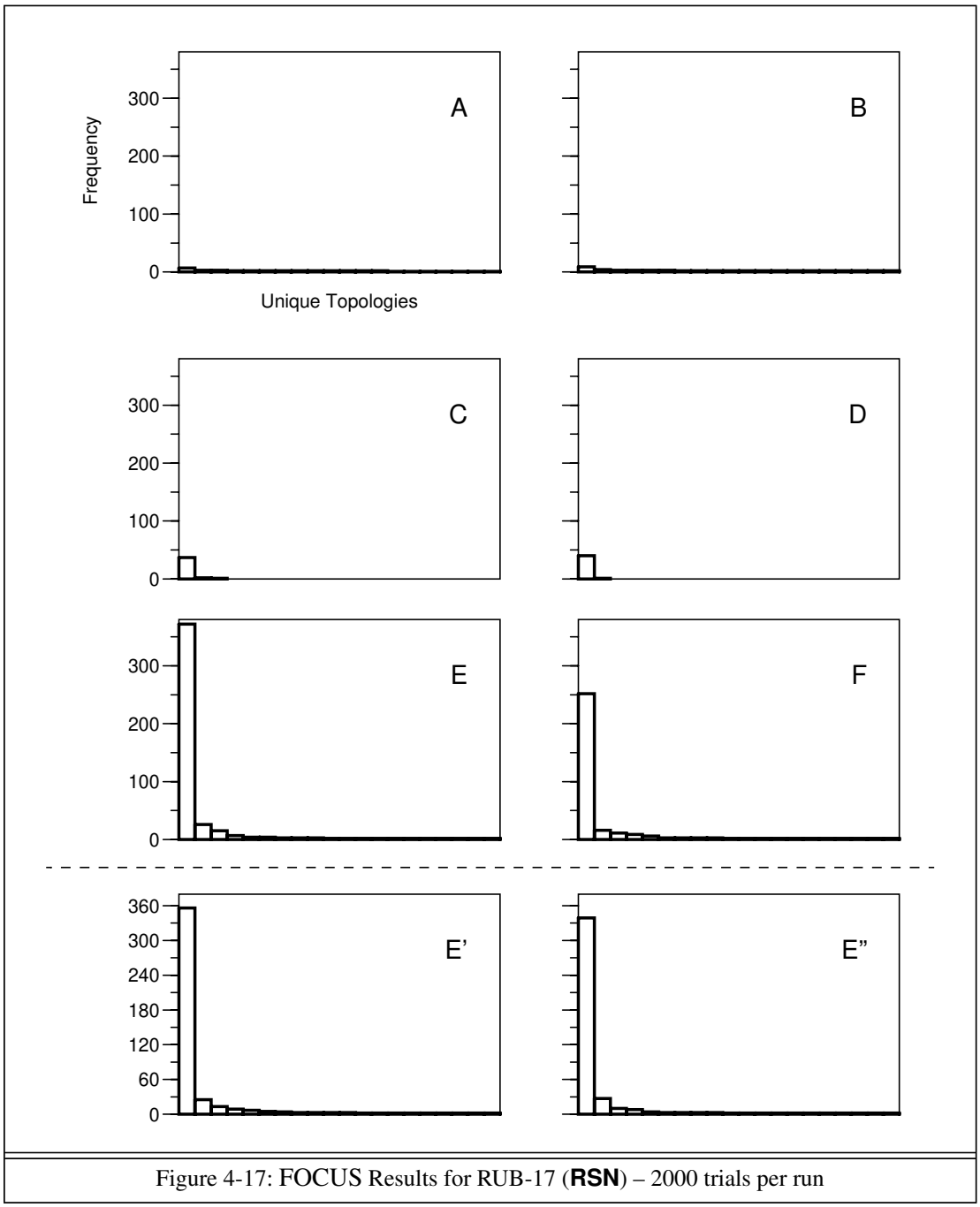


Figure 4-17: FOCUS Results for RUB-17 (**RSN**) – 2000 trials per run

Run	NBO	MPA	MRA	FBC	FT	Fw	UFw	RFw	t/min	%tFwS
A	-	72	36	6	7130	218	198	87561	[b]767	92
B	+	108	108	6	9013	335	299	117264	[b]1040	92
C	-	72	36	0 6	11119	40	3	4075	[b]165	43
D	+	108	108	0 6	11351	41	2	4014	[b]169	43
E	-	72	36	1 1 1 1 1 1 1 1 1 1	20019	800	350	121242	[b]1209	86
F	+	108	108	1 1 1 1 1 1 1 1 1 1	21716	708	390	132634	[b]1381	86
E’	-	44	36	1 1 1 1 1 1 1 1 1 1	20016	798	359	121561	[o]1097	86
E”	-	108	36	1 1 1 1 1 1 1 1 1 1	20171	805	388	123699	[o]1123	86

Table 4-9: FOCUS Results for RUB-17 (**RSN**) – 2000 trials per run (309 actively used reflections)

[b] time normalized by dividing by 0.7
[o] time normalized by dividing by 0.7

In view of the previous examples, it is somewhat surprising that technique E without oxygen recycling outperforms technique F with use of oxygen. One reason could be the use of a higher value for `MaxPotentialAtoms` relative to `MaxRecycledAtoms`. Alternatively one could argue that oxygen in a structure with heavy atoms like zinc has less influence, and that its use in the automatic recycling is therefore less advantageous. An additional factor leading to the negative effect on the success rate could simply be the fact that node atoms are potentially assigned to much weaker peaks, since `MaxPotentialAtoms` has a larger value (108 to account for oxygen, compared to 72), instead of being skipped.

To get a better idea, Run E was repeated twice with `MaxPotentialAtoms` set to 44 and 108, respectively. The resulting histograms E’ and E” are also shown in Fig. 4-17. As can be seen, the change of `MaxPotentialAtoms` has only a very minor effect on the success rate. Therefore, the main reason for the lower success rate of run F must be seen in the use of oxygen. For a structure as complex as RUB-17 and/or in presence of heavy atoms like zinc it seems to be advantageous to recycle the node atoms alone.

4.4 The test structure SAPO-40 (**AFR**)

The synthesis of the SAPO-40 sample, the collection of data on a STOE laboratory diffractometer, and the structure solution from powder data with a combination of "Fast Iterative Patterson Squaring" and direct methods is described in [37, 56]. Fig. 4-18 shows the **AFR** topology.

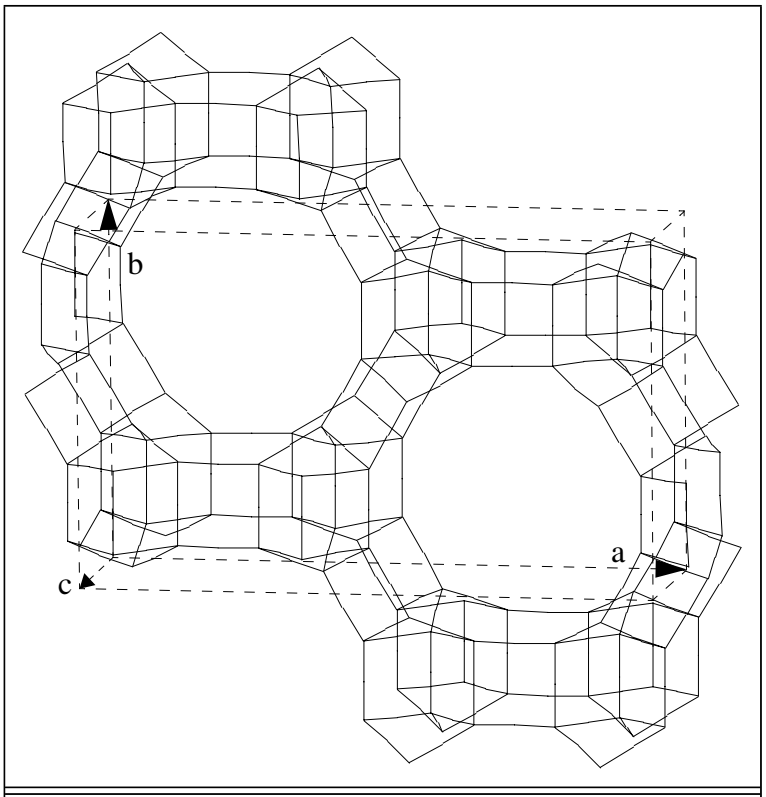


Figure 4-18: The framework of SAPO-40
(O atoms have been omitted for clarity)

Using GSAS, integrated intensities were extracted up to a resolution of 1.19 Å. Tab. 4-10 shows the relevant data for SAPO-40 and the final values for the 10 refined parameters. The final plot of the profile fit is shown in Fig. 4-19. After determination of the intensity scaling factor with help of GENEV, the FOCUS input template of Fig. 4-20 was prepared.

Since the scattering powers of Si, Al, and P are only slightly different, only Si was used in the recycling. This is, in general, a proper approach for aluminum phosphates. Only after the structure is known, can one introduce the strict Al-P alternation, which in many cases reduces the symmetry. One example is SAPO-40, where the introduction of Al-P alternation is only possible in a cell with twice as large a repeat unit in the c-direction, and the new space group P*ccn* [57]. Another example is the calcined form of $AlPO_4$-18 (**AEI**) [58]. The powder profile can be indexed on the orthorhombic space group C*mcm*, which corresponds to the topological symmetry of **AEI**. But with Al-P ordering, the highest possible space group is monoclinic C2/*c*

Formula

$[(Si,Al,P)_{32}\ O_{64}]\ \cdot 2((CH_3CH_2CH_2)_4NOH)$

Number of node atoms in the asymmetric unit

Observed space group	4
Highest topological symmetry (P m m n)	4

Data collection

STOE Stadi-P, linear PSD
Rotating 0.5 mm capillary
Cu-$K_{\alpha 1}$ (1.5406 Å) radiation
2θ range 5 - 80°, step size 0.02°

Intensity extraction

Space group	P m m n (No. 59)
Unit cell	a = 22.041 b = 13.698 c = 7.122 Å
Zero	1.06 (⇒ 0.0106° 2θ)
GU	105.6
GV	-0.1
GW	1.4
LX	5.683
LY	0.898
asym	2.6033
R_p	0.1073
R_{wp}	0.1802

Table 4-10: Selected data for SAPO-40

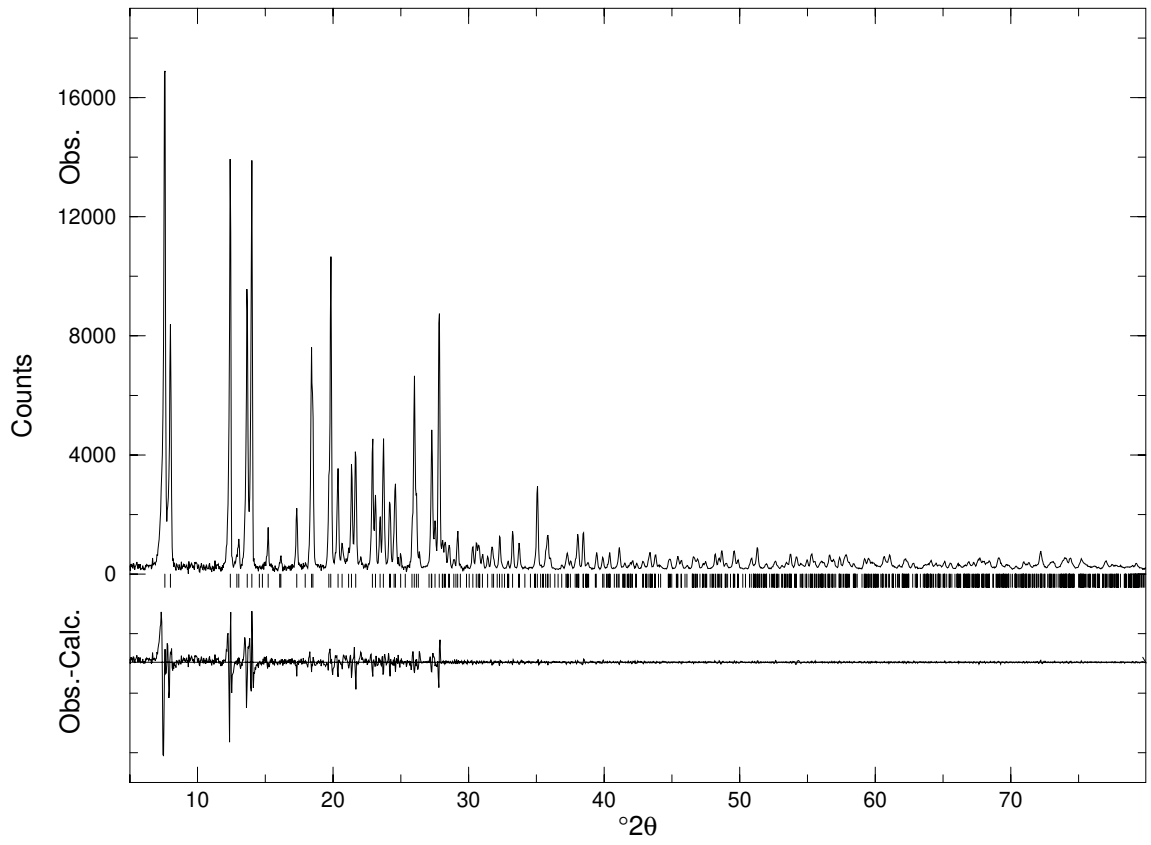


Figure 4-19: GSAS whole-profile intensity extraction for SAPO-40

```
Title SAPO-40, 0.5 mm capillary, STOE STADI-P, lin. PSD
SpaceGroup  P m m n Z
UnitCell  22.0412 13.6979 7.1222

AtomType  +  Node        Si  32
AtomType  -  NodeBridge  O   64
AtomType  -  *           C   24
AtomType  -  *           N    2
AtomType  -  *           O    2
AtomType  -  *           H   58

Chemistry  MinDistance  Node        Si  Node        Si  2.6
Chemistry  MinDistance  Node        Si  NodeBridge  O   1.4
Chemistry  MinDistance  NodeBridge  O   NodeBridge  O   2.3
MaxPotentialAtoms  48
MaxRecycledAtoms  32

FwSearchMethod  FwTracking
MaxPeaksFwSearch  320
MaxPeaksFwFragmentSearch  180
MinNodeDistance  2.6
MaxNodeDistance  3.6
MinSymNodes  0
MaxSymNodes  40
NodeType  4  *  -6 -3 -1 4 6
MinLoopSize  4
MaxLoopSize  24
EvenLoopSizesOnly  Off
Check3DimConnectivity  On
IdealT_NodeDistance  3.1
CheckTetrahedralGeometry  Normal

RandomInitialization  Time
FeedBackCycles  6
FeedBackBreakIf  PhaseDiff < 5.00 % and DeltaR < 1.00 %

Grid_xyz  68  44  24
eDensityCutOff  1 %
MinPfI  17
CatchDistance  0.5
eD_PeaksSortElement  Grid_eD

Lambda  CuA1
FobsMin_d  1.3
FobsScale  0.11
SigmaCutOff  0
OverlapFactor  0.15
OverlapAction  EqualMF2
ReflectionUsage  75 %
Grid_hkl  +17  22  12

#  h   k   l       Fobs  Sigma  FWHM
   1   1   0     1317.57 *   0.07884
End
```

Figure 4-20: A sample input for the SAPO-40 test structure

with a β angle very close to 90°. Therefore the use of the `AltFwTracking` framework search method is not recommended for aluminum phosphates. Instead, FOCUS offers the `EvenLoopSizesOnly` option, which takes care of the fact that only even loop sizes are possible for structures with strictly alternating occupation of the node positions.

Two test series of six runs were derived from the input template (see also Tab. 4-4). The first series (A_1-F_1) was with `EvenLoopSizesOnly Off`, the second series (A_2-F_2) made use of this filter. For the runs without oxygen recycling, `MaxPotentialAtoms` and `MaxRecycledAtoms` were set to 48 and 32, respectively. In the other runs with oxygen recycling, both parameters were set to 96.

The results of the runs are summarized in the histograms of Figs. 4-21 and 4-22 and Tab. 4-11. For runs A_i and B_i, the `EvenLoopSizesOnly` mechanism acts as a pure filter which prevents frameworks with odd loop sizes from being printed, and there is no interaction with the recycling process. However, in runs C_i-F_i, possible candidates for the "largest framework fragment" are also rejected, with the intention of giving the recycling algorithm a stronger bias towards the production of the correct topology.

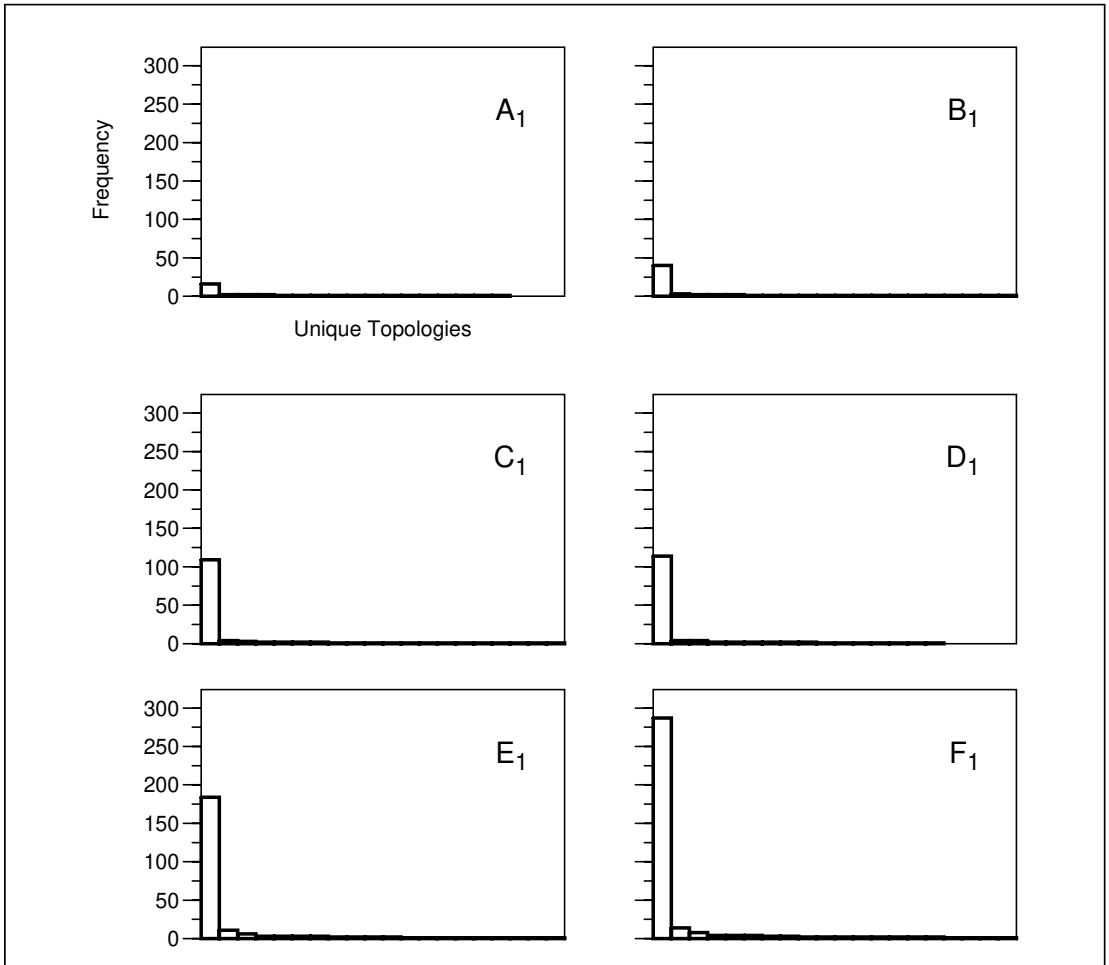


Figure 4-21: FOCUS Results for SAPO-40 (**AFR**) - `EvenLoopSizesOnly Off` - 1000 trials per run

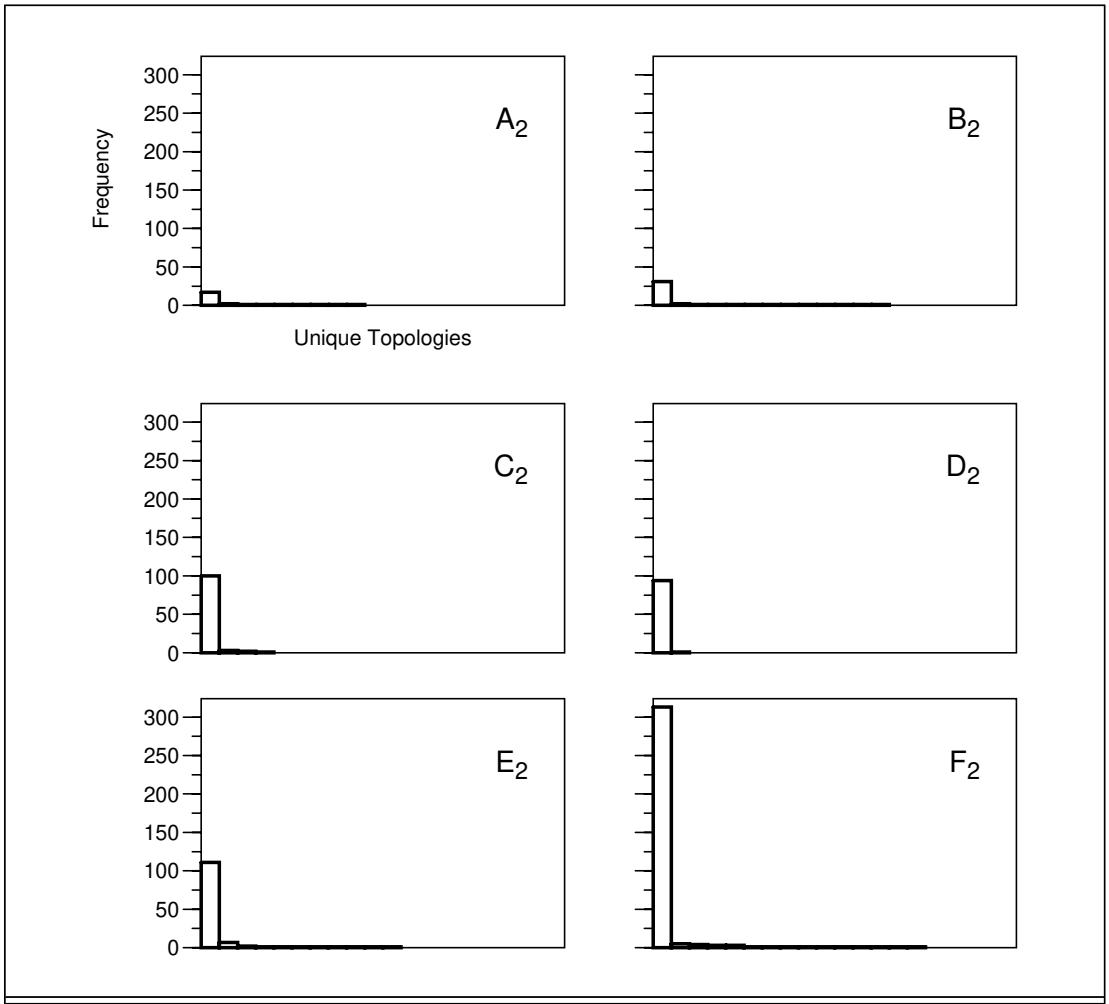


Figure 4-22: FOCUS Results for SAPO-40 (**AFR**) - `EvenLoopSizesOnly On` - 1000 trials per run

Run	NBO	MPA	MRA	FBC	FT	Fw	UFw	RFw	t/min	%tFwS
A_1	-	48	32	6	3689	35	17	23058	[g]47	36
B_1	+	96	96	6	4617	70	26	31245	[g]61	34
C_1	-	48	32	0 6	5837	137	20	2966	[g]1604	97
D_1	+	96	96	0 6	5918	141	16	3048	1470	96
E_1	-	48	32	1 1 1 1 1 1 1 1 1 1	9456	247	37	36778	1314	92
F_1	+	96	96	1 1 1 1 1 1 1 1 1 1	10994	371	44	41346	1643	93
A_2	-	48	32	6	3695	26	9	23333	[g]51	35
B_2	+	96	96	6	4680	44	13	32217	[b]49	37
C_2	-	48	32	0 6	5721	106	4	1094	[g]1858	97
D_2	+	96	96	0 6	5718	95	2	1114	[b]1343	97
E_2	-	48	32	1 1 1 1 1 1 1 1 1 1	9067	128	11	31871	[o]1064	95
F_2	+	96	96	1 1 1 1 1 1 1 1 1 1	10975	338	15	37756	[o]1582	96

Table 4-11: FOCUS Results for SAPO-40 (**AFR**) – 1000 trials per run (314 actively used reflections)

[g] time normalized by dividing by 1.1
[b] time normalized by dividing by 0.7
[o] time normalized by dividing by 0.7

A somewhat unexpected result is revealed in a comparison of the histograms for C_1-E_1 and C_2-E_2. In these runs, the effect of the `EvenLoopSizesOnly` option is a slight reduction of the success rate, rather than an increase. Only run F_2 is doing slightly better than F_1, and the discrimination of the most frequently and second most frequently occurring topology is even more pronounced. Altogether, except from reducing the output of useless topologies, employment of the `EvenLoopSizesOnly` option does not seem to have a real advantage. In the case of SAPO-40, the correct framework is produced either way, and, with use of recycling technique F, also with high efficiency.

4.5 The test structure Zeolite-A (LTA)

The origin of the zeolite-A sample, the collection of the dataset on a laboratory diffractometer, and the results of a Rietveld refinement are given in [59]. Fig. 1-1 on page 2 shows the **LTA** topology of zeolite-A.

Using GSAS, integrated intensities were extracted up to a resolution of 1.15 Å. Tab. 4-12 shows the relevant data for zeolite-A and the final values for the seven refined parameters. The final plot of the profile fit is shown in Fig. 4-23. After determination of the intensity scaling factor with help of GENEV, the FOCUS input template of Fig. 4-24 was prepared. As for the previous examples, a test series of six runs was derived from the input template (see also Tab. 4-4). For the three runs without oxygen recycling, `MaxPotentialAtoms` and `MaxRecycledAtoms` were set to 288 and 192, respectively. In the other three runs with oxygen recycling, both parameters were set to 576.

Formula	
$Na_{96}\,[Al_{96}\,Si_{96}\,O_{384}]\cdot 150\,H_2O$	
Number of node atoms in the asymmetric unit	
Observed space group	2
Highest topological symmetry (P m $\bar{3}$ m)	1
Data collection	
Enraf-Nonius PDS 120, linear PSD	
Rotating 0.3 mm capillary	
Cu-$K_{\alpha 1}$ (1.5406 Å) radiation	
2θ range 5 - 80°, step size 0.03°	
Intensity extraction	
Space group	F m $\bar{3}$ c (No. 226)
Unit cell	a = 24.558 Å
GU	138.3
GV	-58.2
GW	35.7
LX	4.950
LY	20.641
asym	1.7878
R_p	0.0310
R_{wp}	0.0485

Table 4-12: Selected data for zeolite-A

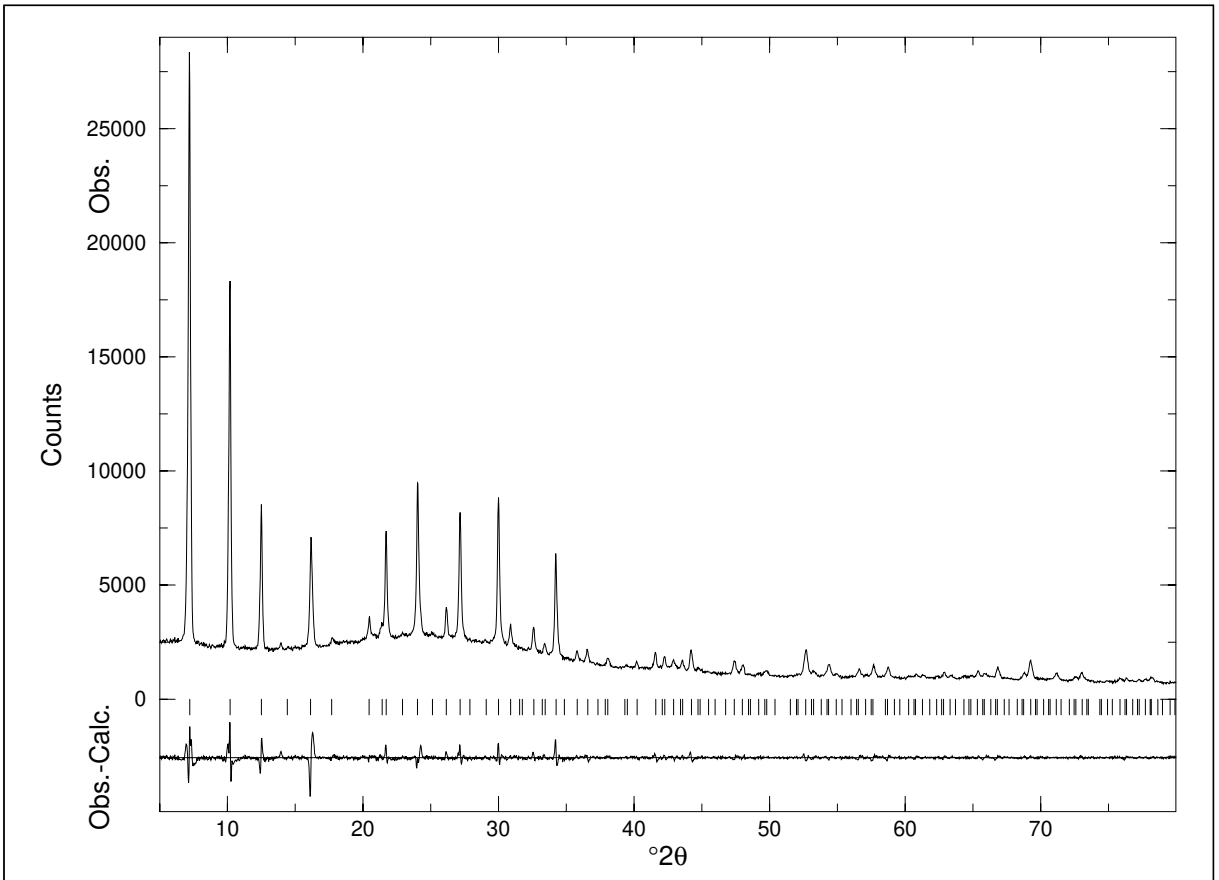


Figure 4-23: GSAS whole-profile intensity extraction for zeolite-A

Although integrated intensities can be extracted from the profile in the higher space group Pm$\bar{3}$m with a = 12.3 Å equally well, the larger unit cell, which describes the alternation of Si and Al, was used in this example. In the smaller cell, where Si and Al are not distinguished, the solution of the structure is almost trivial, despite the high ratio of overlapping reflections. For example, with the small cell and recycling technique F, FOCUS produced a total of 708 topologies in 300 trials, and 695 of these were the correct solution. The total computing time for the 300 trials was 12 minutes. This means that a correct solution was produced nearly every second.

With the larger cell, zeolite-A offers a good example for the employment of the `AltFwTracking` framework search mode, since the framework nodes are occupied by Si and Al in strict alternation. However, even with the larger cell with a reflection overlap at a resolution of 1.3 Å of 67%, the zeolite-A topology was easily recovered in all six test runs. As can be seen from Tab. 4-13, only one unique topology – the correct solution – was produced. Again, as in all of the previous examples except RUB-17, recycling technique F is the most successful. In the present case, the success rate of run F is three to more than eight times higher than that of the other runs.

```
Title Zeolite-A, Enraf-Nonius PDS 120, GSAS extraction
SpaceGroup  F m -3 c
UnitCell  24.5583

AtomType  +  Node        Si   96
AtomType  +  Node        Al   96
AtomType  -  NodeBridge  O   384
AtomType  -  *           Na   96
AtomType  -  *           O   150
AtomType  -  *           H   300

Chemistry  MinDistance  Node        Si  Node        Si  2.6
Chemistry  MinDistance  Node        Si  Node        Al  2.6
Chemistry  MinDistance  Node        Si  NodeBridge  O   1.4
Chemistry  MinDistance  Node        Al  NodeBridge  Al  2.6
Chemistry  MinDistance  Node        Al  NodeBridge  O   1.4
Chemistry  MinDistance  NodeBridge  O   NodeBridge  O   2.3
MaxPotentialAtoms  288
MaxRecycledAtoms  192

FwSearchMethod  AltFwTracking
MaxPeaksFwSearch  9600
MaxPeaksFwFragmentSearch  4800
MinNodeDistance  2.6
MaxNodeDistance  3.6
MinSymNodes  0
MaxSymNodes  288
NodeType  4  *  -6 -3 -1 4 6
MinLoopSize  3
MaxLoopSize  24
EvenLoopSizesOnly  Off
Check3DimConnectivity  On
IdealT_NodeDistance  3.1
CheckTetrahedralGeometry  Hard

RandomInitialization  Time
FeedBackCycles  6
FeedBackBreakIf  PhaseDiff < 5.00 % and DeltaR < 1.00 %

Grid_xyz  76  76  76
eDensityCutOff  1 %
MinPfI  17
CatchDistance  0.5
eD_PeaksSortElement  Grid_eD

Lambda  CuA1
FobsMin_d  1.3
FobsScale  0.35
SigmaCutOff  0
OverlapFactor  0.15
OverlapAction  EqualMF2
ReflectionUsage  75 %
Grid_hkl  +19  38  38
```

Figure 4-24: A sample input for the zeolite-A test structure

Run	NBO	MPA	MRA	FBC	FT	Fw	UFw	RFw	t/min	%tFwS
A	-	288	192	6	791	122	1	14	121	7
B	+	576	576	6	1279	158	1	19	194	7
C	-	288	192	0 6	1230	326	1	36	191	9
D	+	576	576	0 6	1222	307	1	28	203	8
E	-	288	192	1 1 1 1 1 1 1 1 1 1	1581	277	1	37	242	8
F	+	576	576	1 1 1 1 1 1 1 1 1 1	3022	1023	1	80	508	17

Table 4-13: FOCUS Results for zeolite-A (**LTA**) – 300 trials per run (80 actively used reflections)

At this point, a remark of technical nature regarding the computing times might be appropriate. At present, FOCUS computes the Fourier transform for a whole cell and does not make use of symmetry in this step. Therefore the computing time needed for the tests with the larger unit cell is significantly longer than for the tests working with the smaller unit cell. With technique F, for example, the time required increases from 12 to 508 minutes. By the introduction of Fourier transform procedures which make use of symmetry, the computing time requirements could be lowered significantly. (Related to the Fourier transform routines, see also the remark on page 36.)

4.6 The test structure EMC-2 (EMT)

The origin of the ammonium exchanged and calcined EMC-2 sample, the collection of the dataset on a synchrotron beamline, and the results of a Rietveld refinement are given in [60]. Fig. 4-25 shows the **EMT** topology of EMC-2.

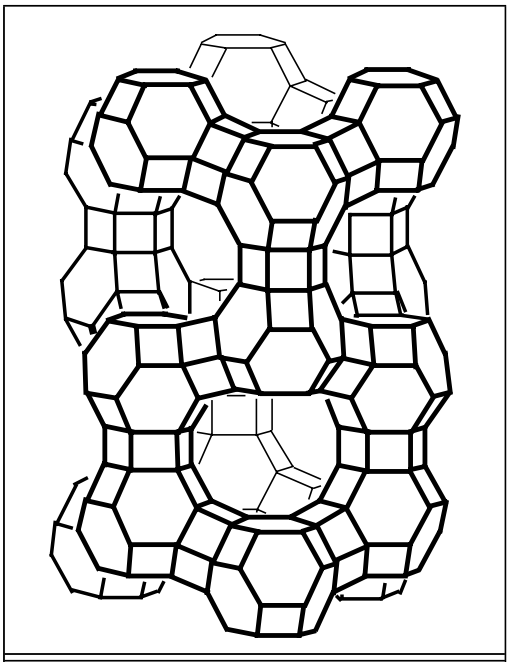


Figure 4-25: The framework of EMT
(O atoms have been omitted for clarity)

In this case, the experience gained in the extraction step deserves a more extensive discussion. To start a Le Bail extraction, GSAS requires the input of a dummy model. This model is used to calculate the initial partitioning, which is then modified according to the Le Bail method in the subsequent refinement cycles. Reflections, which are apportioned exactly zero intensity in the initialization step, cannot gain intensity in the refinement, even if they are not overlapping. Therefore, in all cases except for EMC-2, one or two dummy atoms were put onto random – but not special – positions.

The first intensity extraction for EMC-2 was started with an existing GSAS file containing the correct model. Subsequent FOCUS runs making use of these intensities were extremely successful. Due to an oversight, the equipartitioning step had been switched off (`OverlapAction NoAction`). After correcting this, the success rate dropped significantly, so apparently the initial partitioning with the correct model was retained in further refinement steps with GSAS and not redistributed as assumed. On the one hand, this observation is very helpful for the initialization of a Rietveld refinement. This means, it is a good approach to supply the model to be refined, to proceed with a Le Bail extraction in order to fit the zero

correction and profile parameters, and to continue with the model refinement only after these parameters have converged satisfactorily. On the other hand, for testing a structure solution program, using the correct model to initialize partitioning is a hidden form of cheating. Therefore, the extraction was repeated with a dummy model of two atoms in random (general) positions. Surprisingly, with these intensities, FOCUS did not give any solution at all. Obviously, the overlap factor of 0.3 used for the equipartitioning is too conservative to eliminate all of the bias which is introduced by the dummy model. In the previous test cases with less overlap, FOCUS was still able to recover the correct model, but with 83 % overlap at 1.3 Å resolution, and a structure as complex as EMC-2, the limit of the method was reached.

Increasing the overlap factor in order to eliminate the model bias was not an option, because this would mean that nearly all reflections would overlap, making the problem even worse. So, an attempt was made to eliminate any model bias in the initial partitioning. For this purpose, a small program was written to modify the GSAS "direct access" file which stores the reflection data. After setting up the input files, but before starting the intensity refinement, all "calculated" intensities were set to "1.0". Then the intensity refinement was conducted as usual, and the results are summarized in Tab. 4-14. Fig. 4-26 shows the final plot of the profile.

Formula	
Na_{11} $[(Si,Al)_{96}$ $O_{192}]$ ·6 H_20	
Number of node atoms in the asymmetric unit	
Observed space group	4
Highest topological symmetry (P 63/m m c)	4
Data collection	
HASYLAB beamline B2	
Rotating 1.0 mm capillary	
Wavelength 1.38981 Å	
2θ range 4.6 - 60°, step size 0.01°	
Intensity extraction	
Space group	P 63/m m c (No. 59)
Unit cell	a = 17.378 c = 28.344 Å
GU	140.1
GV	-57.4
GW	6.8
LX	1.469
LY	36.086
asym	0.2600
R_p	0.0871
R_{wp}	0.1064

Table 4-14: Selected data for EMC-2

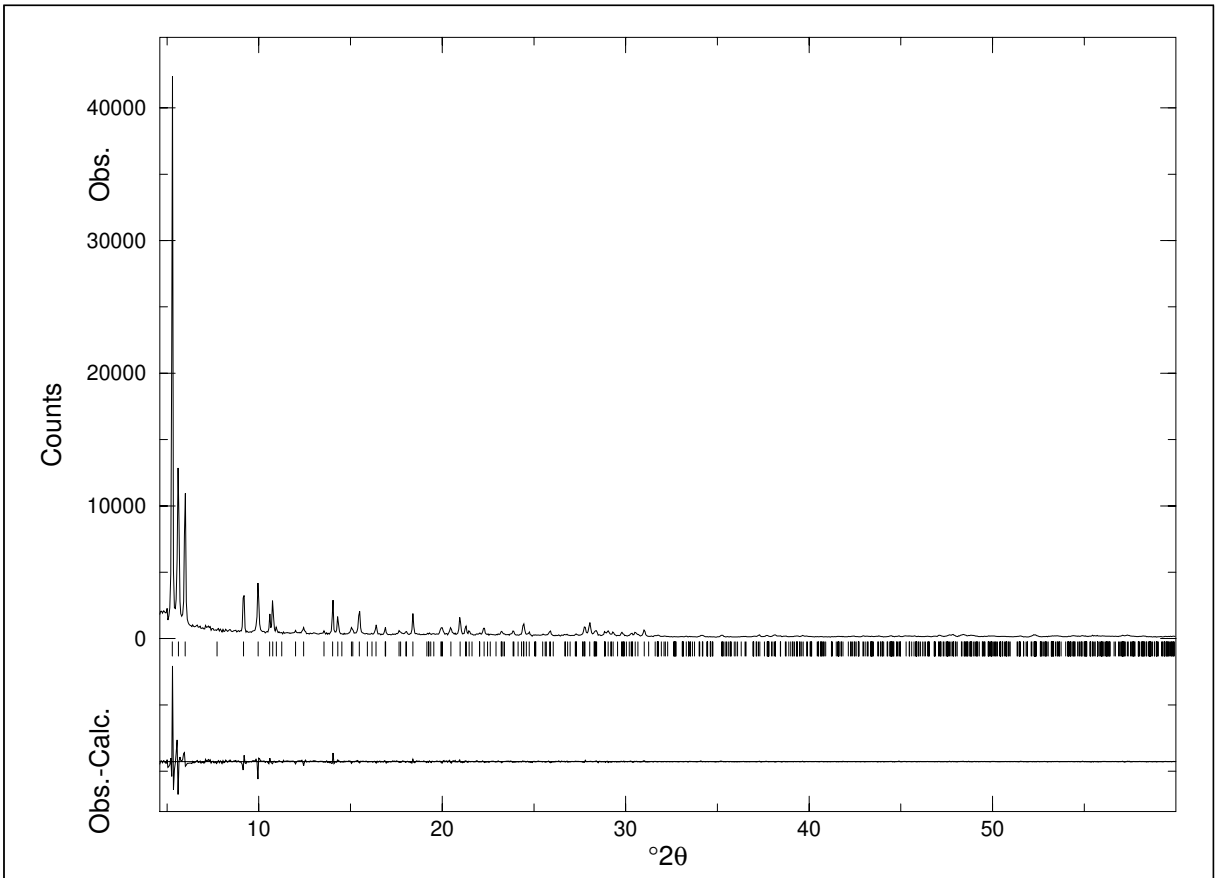


Figure 4-26: GSAS whole-profile intensity extraction for EMC-2

Two test series of six runs were derived from the input template shown in Fig. 4-27 (see also Tab. 4-4). For the runs without oxygen recycling, `MaxPotentialAtoms` and `MaxRecycledAtoms` were set to 144 and 96, respectively. In the other runs with oxygen recycling, both parameters were set to 288. In the first series, equipartitioned intensities as extracted with GSAS were used. Fig. 4-28 and the upper half of Tab. 4-15 summarize the results of the first series. With pure atom recycling, no framework at all was produced in 1000 trials. However, the careful extraction starting with an unbiased partitioning was rewarded by the successful recovery of the correct topology using framework fragment recycling and the technique with alternation of atom and fragment recycling. Again, technique F performed best, with a clear discrimination of the two most frequently occurring topologies.

Compared with the other examples, the absolute success rate in the first test series is poor. With the intention of improving the success rate, the extracted intensities were redistributed using the FIPS method [17, 37]. The best FIPS `BOOSTR` parameter was found by testing all values from two to twenty with a step of two. The best result was obtained with `BOOSTR 4` and four partitioning cycles. To avoid destroying the FIPS improved partitioning, `OverlapFactor` in the FOCUS input was set to zero, and `OverlapAction` was set to `NoAction`.

```
Title EMC-2 calc., HASYLAB B2, 1.00 mm cap., GSAS extraction
SpaceGroup  P 63/m m c
UnitCell  17.3783 28.3439

AtomType  +  Node        Si   96
AtomType  -  NodeBridge  O   192
AtomType  -  *           O     6
AtomType  -  *           Na   11

Chemistry  MinDistance  Node        Si  Node        Si  2.6
Chemistry  MinDistance  Node        Si  NodeBridge  O   1.4
Chemistry  MinDistance  NodeBridge  O   NodeBridge  O   2.3
MaxPotentialAtoms  144
MaxRecycledAtoms  96

FwSearchMethod  FwTracking
MaxPeaksFwSearch  960
MaxPeaksFwFragmentSearch  480
MinNodeDistance  2.6
MaxNodeDistance  3.6
MinSymNodes  0
MaxSymNodes  144
NodeType  4  *  -6 -3 -1 4 6
MinLoopSize  4
MaxLoopSize  24
EvenLoopSizesOnly  Off
Check3DimConnectivity  On
IdealT_NodeDistance  3.1
CheckTetrahedralGeometry  Normal

RandomInitialization  Time
FeedBackCycles  6
FeedBackBreakIf  PhaseDiff < 5.00 % and DeltaR < 1.00 %

Grid_xyz  54  54  84
eDensityCutOff  1 %
MinPfI  17
CatchDistance  0.5
eD_PeaksSortElement  Grid_eD

Lambda  1.38981
FobsMin_d  1.3
FobsScale  0.35
SigmaCutOff  0
OverlapFactor  0.15
OverlapAction  EqualMF2
ReflectionUsage  75 %
Grid_hkl  28  28  +22

#  h   k   l       Fobs  Sigma  FWHM
   1   0   0     1382.75 *   0.06829
   0   0   2     1596.87 *   0.06831
End
```

Figure 4-27: A sample input for the EMC-2 test structure

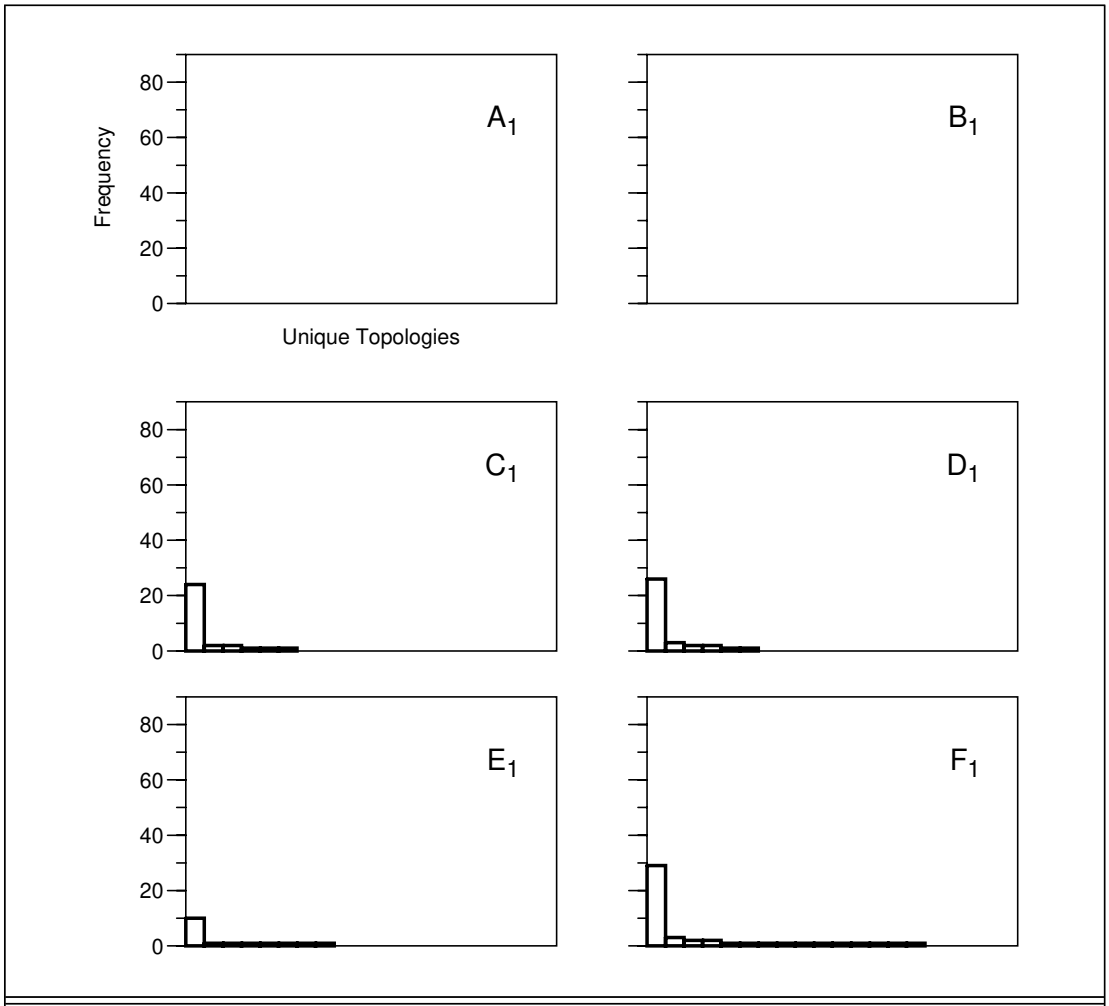


Figure 4-28: FOCUS Results for EMC-2 (**EMT**) – Not using FIPS – 1000 trials per run

The effect of the FIPS treatment can be seen by comparing the histograms of the second series in Fig. 4-29 with those of the first series in Fig. 4-28, and by comparing the upper and lower halfs of Tab. 4-15. With FIPS, frameworks are produced with all techniques A_2-F_2. A_2 is a special case, because none of the four frameworks found in 1000 trials is the correct solution. In the histogram this is indicated by the missing first bar. B_2 is also a special case, because the first two bars in the histograms are of equal height. Consequently, the correct solution is only one of the two most frequent topologies. But with techniques C_2 to F_2 the general rule that the most frequent topology is the correct solution is found to be valid. Moreover, the success rate is roughly three to four times higher than for the first series. For runs C_2 to E_2, the discrimination of the two most frequent topologies is less clear than for the first series, but with technique F_2, which is generally to be preferred, the discrimination is again very clear and only slightly weaker than for F_1 (9.7:1 vs. 7.1:1).

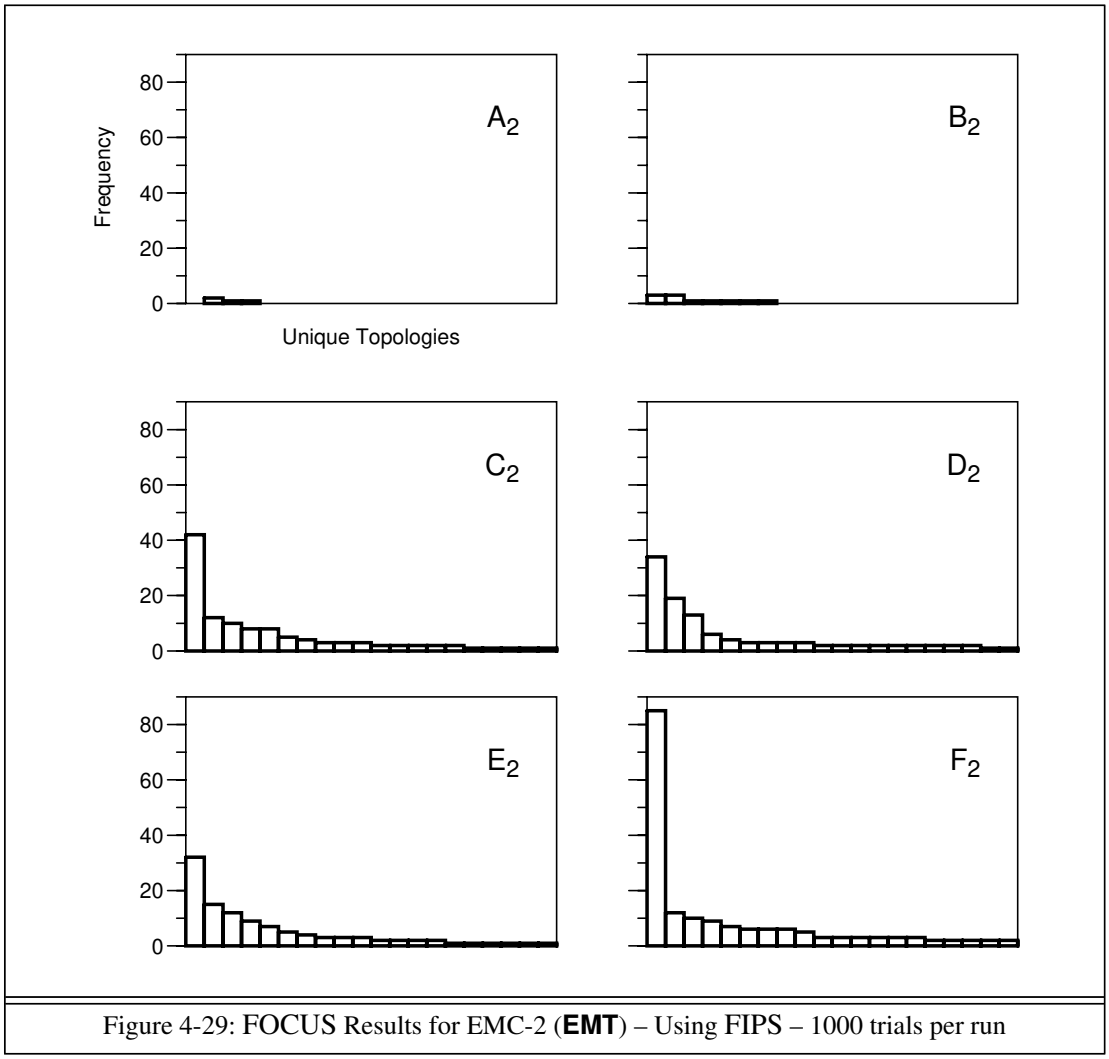


Figure 4-29: FOCUS Results for EMC-2 (**EMT**) – Using FIPS – 1000 trials per run

Run	NBO	MPA	MRA	FBC	FT	Fw	UFw	RFw	t/min	%tFwS
A_1	-	144	96	6	3769	0	0	11746	[o]209	9
B_1	+	288	288	6	4532	0	0	15706	[o]276	11
C_1	-	144	96	0 6	5639	31	6	3127	[g]593	52
D_1	+	288	288	0 6	5697	35	6	3142	[b]690	55
E_1	-	144	96	1 1 1 1 1 1 1 1 1 1	9832	17	8	24178	[o]812	40
F_1	+	288	288	1 1 1 1 1 1 1 1 1 1	10938	47	15	28560	[o]1132	52
A_2	-	144	96	6	3720	4	3	14145	[o]216	10
B_2	+	288	288	6	4571	11	7	19841	[o]278	9
C_2	-	144	96	0 6	5884	118	25	4776	[o]1173	74
D_2	+	288	288	0 6	5857	109	21	4642	[o]1218	75
E_2	-	144	96	1 1 1 1 1 1 1 1 1 1	9673	131	44	31146	[o]1300	44
F_2	+	288	288	1 1 1 1 1 1 1 1 1 1	10966	211	53	38212	[o]1826	68

Table 4-15: FOCUS Results for EMC-2 (**EMT**) – 1000 trials per run (370 actively used reflections)

[o] time normalized by dividing by 0.7
[g] time normalized by dividing by 1.1
[b] time normalized by dividing by 0.7

4.7 The structure of the zincosilicate zeolite VPI-9

4.7.1 Preparation and application of FOCUS

Two different as-synthesized VPI-9 samples were kindly supplied by the group of M.E. Davis at CalTech, U.S.A.

Results of previous investigations on the first sample, synthesized by M.J. Annen (MJA2-19) are published in [21, 22, 23]. The chemical formula for VPI-9, based on chemical analysis and NMR data, is given in [23] as $Rb_{24}Zn_{12}Si_{48}O_{120}\cdot wH_2O$. With as-synthesized MJA2-19, three datasets at three different wavelengths were collected on the synchrotron beamline X7A at the NSLS in Brookhaven (see page 5). The lattice constants of about a = 9.9 and c = 37.0 Å given in [22] allowed for a reasonable, but not a completely satisfactory whole-profile intensity extraction. Some small and broader peaks remained unindexed. The extinction symbol derived from the profiles was P-c-. This means that the space groups to be considered were $P4_2/mcm$ (No. 132), $P\bar{4}c2$ (No. 116), and $P4_2cm$ (No. 101). After a long series of unsuccessful attempts to determine the structure, the profiles were reinvestigated. It was found that the profiles could be indexed better in the space group $P4_12_12$ (No. 92) [or alternatively in the enantiomorphic space group $P4_32_12$ (No. 96)] with lattice constants of about a = 9.9 Å and c = 74 Å. However, the solution of a structure with such a large cell was considered to be beyond the capacity of all available methods, including FOCUS.

In another attempt to solve the structure, an ion exchange experiment that would eliminate the strong scatterer Rb was performed. A 1M NH_4Cl solution was prepared and its pH adjusted to 8 by adding NH_4OH. A small amount of MJA2-19 was added to the solution. After three days it was washed with distilled water and then dried at 80 °C. A new powder pattern was measured on a laboratory diffractometer. Although the peak positions changed only slightly, the differences in the intensities were significant. Nevertheless, the extinction symbol derived from the new profile was again P-c-, and subsequent attempts to solve the structure were unsuccessful.

Because it was suspected that an impurity was present in MJA2-19, a second VPI-9 sample (MY027) was synthesized by M. Yoshikawa. In the synthesis, potassium was present in addition to rubidium. To minimize the contribution of these non-framework atoms, MY027 was subjected to a carefully conducted ammonium exchange experiment right from the beginning. The effect of this treatment can be seen in Fig. 4-30, which compares the profiles of as-synthesized and NH_4^+-exchanged MY027. Both profiles were collected on the Swiss-Norwegian beamline (SNBL) at the ESRF in Grenoble, but with slightly different

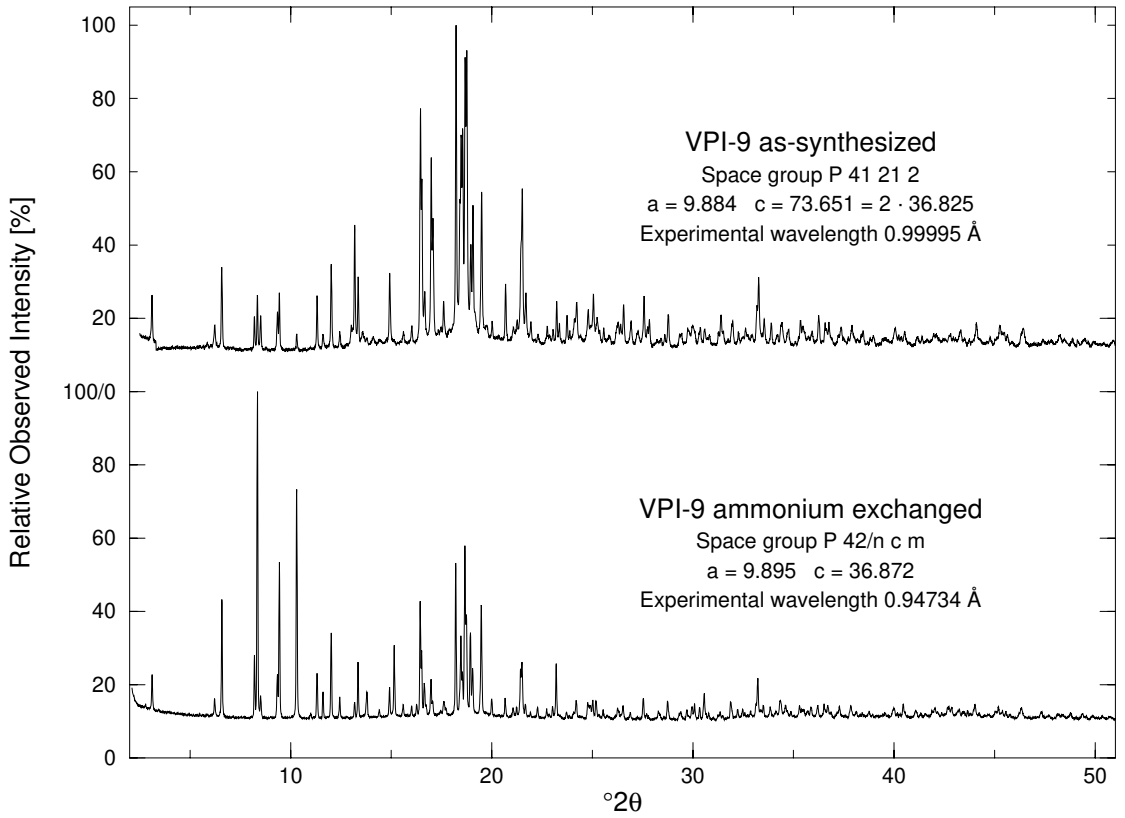


Figure 4-30: VPI-9 (MY027) ESRF profiles normalized to a wavelength of 1.0 Å

wavelengths. For comparison purposes, the profiles in Fig. 4-30 have been normalized to a wavelength of 1.0 Å.

As for MJA2-19, the shifts in the peak positions are only minor, but the intensity distribution is very different. Examination of the systematic absences gave a new and most important result: the appropriate extinction symbol was found to be P*nc*- rather than P-*c*-. With these absences, the new space group, $P4_2/ncm$, was indicated.

From this point on, the structure determination closely resembled that of the test cases. Tab. 4-16 summarizes the results of the intensity extraction with GSAS, and Fig. 4-30 shows the final profile fit. After determination of the intensity scaling factor, the FOCUS input template in Fig. 4-32 was prepared. The Si/Zn ratio in the input file reflects the results of a new chemical analysis of as-synthesized MY027 provided by M. Davis. No chemical analysis was made for the exchanged MY027 sample. Instead, the NH_4^+ and H_2O content was only roughly approximated by `AtomType - * O 74`.

Estimated formula

$(NH_4^+)_{24}$ $[Si_{44}$ Zn_{12} $O_{112}]$ $\cdot 50H_20$

Data collection

SNBL at the ESRF in Grenoble, Debye-Scherrer Geometry
Rotating 0.5 mm capillary
Wavelength 0.94734 Å
2θ range 2 - 55°, step size 0.01°

Intensity extraction

Space group	P 42/n c m (No. 138)
Unit cell	a = 9.895 c = 36.872 Å
Zero	0.46 (⇒ 0.0046° 2θ)
GU	56.6
GV	-14.0
GW	3.6
LX	0.888
LY	10.282
R_p	0.0154
R_{wp}	0.0217

Table 4-16: Selected data for VPI-9 (MY027)

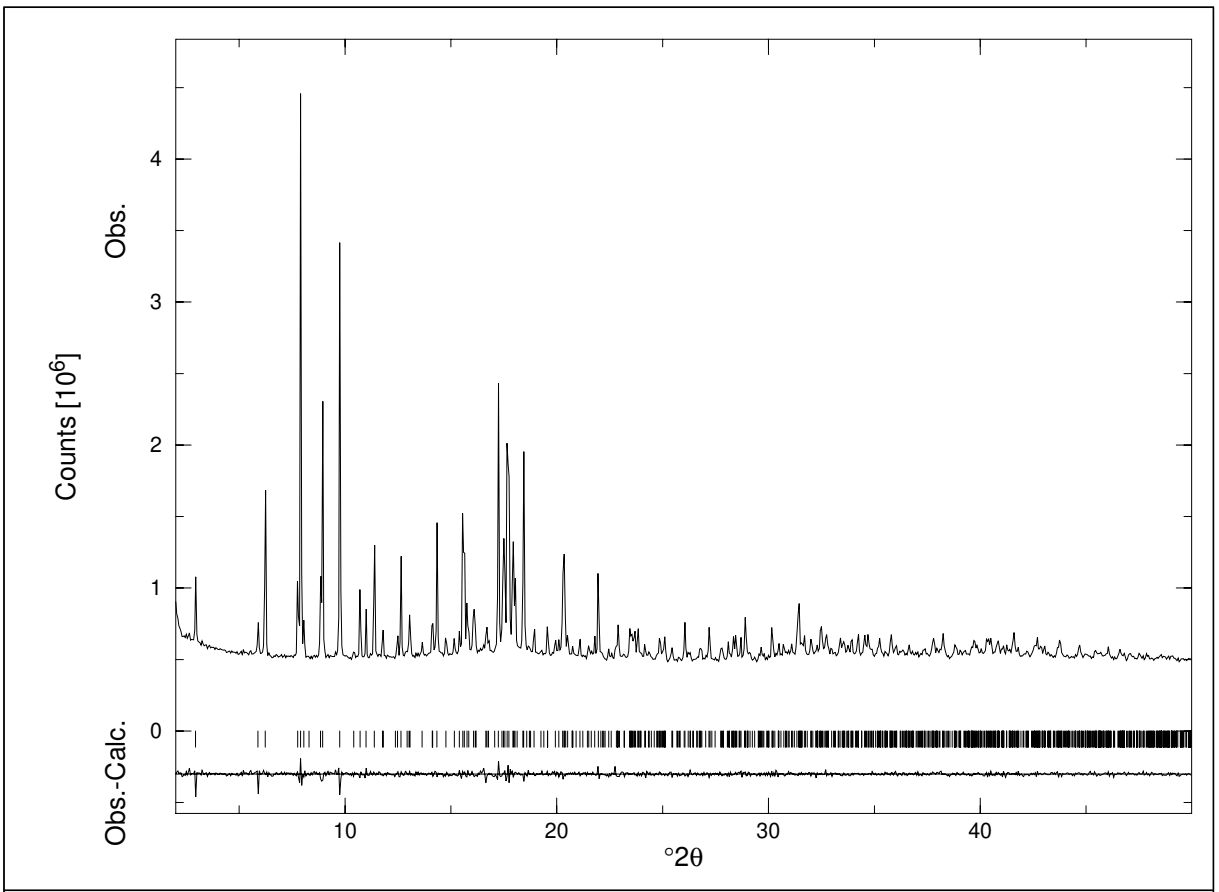


Figure 4-31: GSAS whole-profile intensity extraction for ammonium exchanged VPI-9

```
Title MY027 NH4+ my027_95_03_15.esrf GSAS Extraction 9503311635
SpaceGroup  P 42/n c m Z
UnitCell  9.89461 36.8715

AtomType  +  Node        Zn   12  *  0.01
AtomType  +  Node        Si   44  *  0.01
AtomType  -  NodeBridge  O   112  *  0.02
AtomType  -  *           O    74

Chemistry  MinDistance  Node        Zn  Node        Zn  4
Chemistry  MinDistance  Node        Zn  Node        Si  2.9
Chemistry  MinDistance  Node        Zn  NodeBridge  O   1.7
Chemistry  MinDistance  Node        Si  Node        Si  2.9
Chemistry  MinDistance  Node        Si  NodeBridge  O   1.5
Chemistry  MinDistance  NodeBridge  O   NodeBridge  O   2.3
MaxPotentialAtoms  72
MaxRecycledAtoms  56

FwSearchMethod  FwTracking
MaxPeaksFwSearch  480
MaxPeaksFwFragmentSearch  480
MinNodeDistance  2.3
MaxNodeDistance  3.7
MinSymNodes  0
MaxSymNodes  64
NodeType  4  *  -6 -3 -1 4 6
MinLoopSize  3
MaxLoopSize  24
EvenLoopSizesOnly  Off
Check3DimConnectivity  On
IdealT_NodeDistance  3.15
CheckTetrahedralGeometry  Normal

RandomInitialization  Time
FeedBackCycles  6
FeedBackBreakIf  PhaseDiff < 5.00 % and DeltaR < 1.00 %

Grid_xyz  32  32  112
eDensityCutOff  1 %
MinPfI  17
CatchDistance  0.5
eD_PeaksSortElement  Grid_eD

Lambda  0.94734
FobsMin_d  1.3
FobsScale  0.016
SigmaCutOff  0
OverlapFactor  0.15
OverlapAction  EqualMF2
ReflectionUsage  75 %
Grid_hkl  16  16  +29
```

Figure 4-32: A sample input for VPI-9 (MY027)

The original structure solution was obtained with atom recycling without the use of oxygen (technique A), because framework fragment recycling was not available at that time. However, for comparison purposes, again a series of six runs (see Tab. 4-4) was derived from the input template in Fig. 4-32. For the three runs without oxygen recycling, `MaxPotentialAtoms` and `MaxRecycledAtoms` were set to 72 and 56, respectively. In the other three runs with oxygen recycling, both parameters were set to 168.

The histograms resulting from the FOCUS runs are shown in Fig. 4-33. Tab. 4-17 summarizes the characteristic data. As observed in all cases before, alternation of atom and framework fragment recycling yields the best success rate. Also, with the exception of run B, the most frequently occurring topology is the same in all runs. With technique B, the first two histogram bars are of equal height, and the topology in question is represented by one of them. In all cases the use of oxygen lowered the success rate. This was only observed for the RUB-17 test case. Not only is the degree of overlap for VPI-9 and for RUB-17 very similar, but the chemistry and complexity of the two structures are also comparable. Thus the results of the VPI-9 runs give further to support to the interpretation of the RUB-17 test results on page 50.

4.7.2 Verification of the structure: preliminary Rietveld refinement

For simplicity, the most frequently occurring topology found in the FOCUS runs will now be referred to as the "VPI-9 framework".

The first step after the evaluation of the histograms was the preparation of an input file for DLS-76 with help of KRIBER [61]. No attempt was made to derive the zinc distribution in the framework from the FOCUS output. For the DLS refinement, silicon atoms were assigned to all framework positions. Initial positions for the bridging oxygen atoms were automatically generated by KRIBER. The DLS calculation converged to a residual value of 0.0102 (the DLS prescriptions used are listed in section 7.6). This value is at the high end of the scale for plausible structures, but still acceptable. For comparison, the DLS residual for the Lovdarite (**LOV**) topology assuming a pure silicon composition is 0.0117. Lovdarite is a beryllosilicate with 3-rings.

Fig. 4-34 shows the DLS refined VPI-9 framework viewed parallel [110]. Seven node atoms in the asymmetric unit – which generate 60 atoms in the unit cell – form a framework of 3, 4, 5, and 8-rings, and a complex 3-dimensional channel system. Space group $P4_2/ncm$ is the highest topological symmetry of this framework. This means that the number of node atoms per asymmetric unit cannot be reduced by the introduction of further symmetry operations.

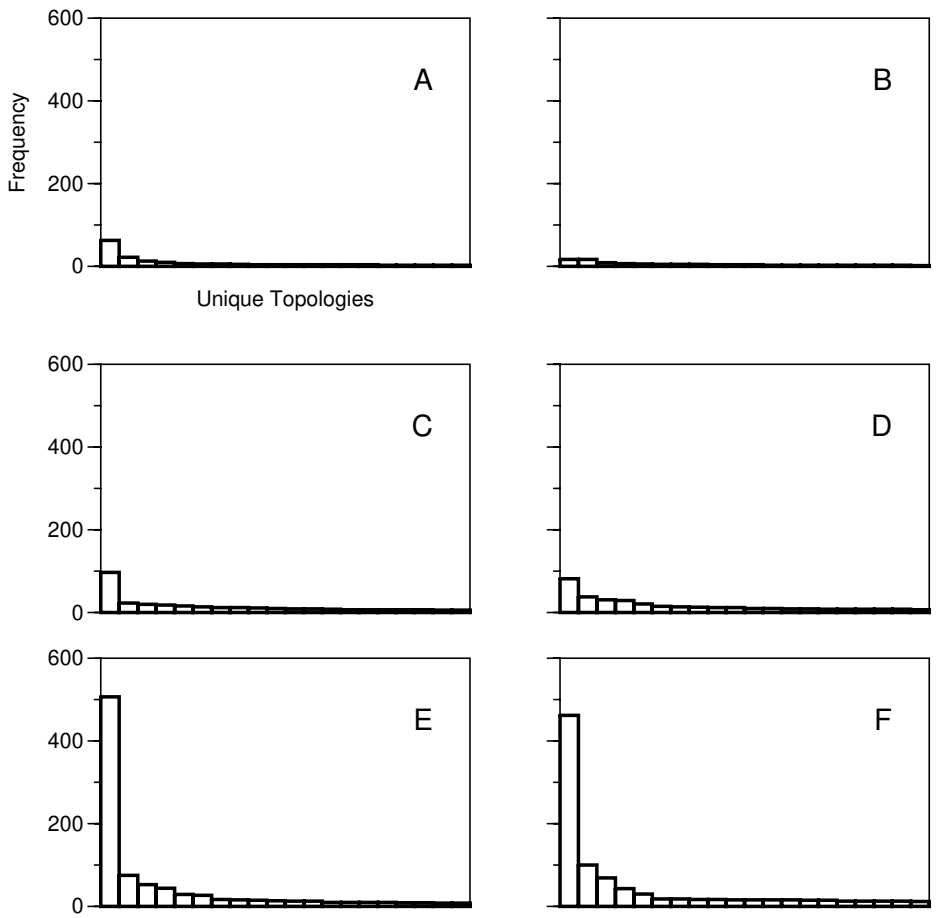


Figure 4-33: FOCUS Results for VPI-9 – 2000 trials per run

Run	NBO	MPA	MRA	FBC	FT	Fw	UFw	RFw	t/min	%tFwS
A	-	72	56	6	8339	378	193	114844	229	28
B	+	168	168	6	9193	333	208	131694	224	15
C	-	72	56	0 6	12368	960	396	114021	368	33
D	+	168	168	0 6	12487	990	389	114387	374	33
E	-	72	56	1 1 1 1 1 1 1 1 1 1	21637	1783	546	260235	629	31
F	+	168	168	1 1 1 1 1 1 1 1 1 1	21743	2096	642	270107	653	32

Table 4-17: FOCUS Results for VPI-9 – 2000 trials per run (258 actively used reflections)

The framework can be subdivided into two layer-like building units (LLBU's) – A and B in Fig. 4-34 – which are generated by four and two node atoms of the asymmetric unit, respectively, and are connected by one additional node atom. The LLBU's A and B are repeated in different perspectives in Fig. 4-35. The simpler LLBU B is a two dimensional undulating net of 4- and 8-rings, which is also known as "4.8 net". LLBU B occurs in several

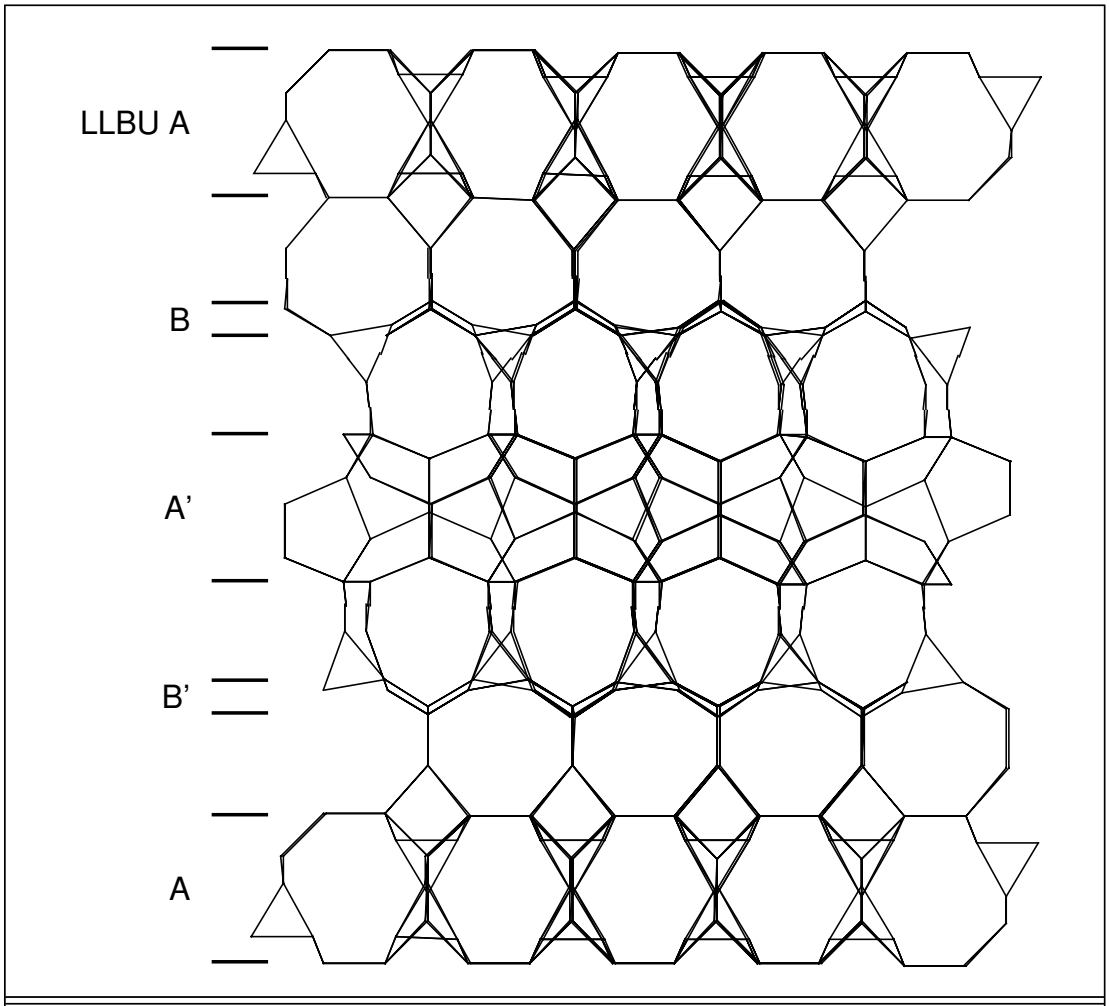


Figure 4-34: The framework of VPI-9 viewed parallel [110] (O atoms have been omitted for clarity)

other zeolites, among them the zincosilicates RUB-17 and VPI-7 [62].

The basic building unit of LLBU A itself is a polyhedron consisting of a 3-ring with three bent 5-rings attached to it. The polyhedra share 3-ring faces on one side, and 5-ring edges on the other side, to form infinite chains running parallel [110] or [1$\bar{1}$0]. Neighboring chains in the same plane run parallel, and are shifted by half a chain period length with respect to one another. The chains are linked through bridging oxygens. – Neither LLBU A, nor the basic polyhedron has been observed in other zeolites before.

The DLS refined model was used to initialize a Rietveld refinement for ammonium exchanged MY027. Fig. 4-36 shows the initial plot of the observed and the difference (observed - calculated) profile. To say the least, this plot was very discouraging, and various alternative structures were considered before the refinement was finally continued. In order to keep the framework reasonable, soft-constraints were introduced. Zinc positions were identified through refinement of the occupancy factors for the node atoms (all silicon to start with). Then the model was further improved by including oxygen atoms as approximations for

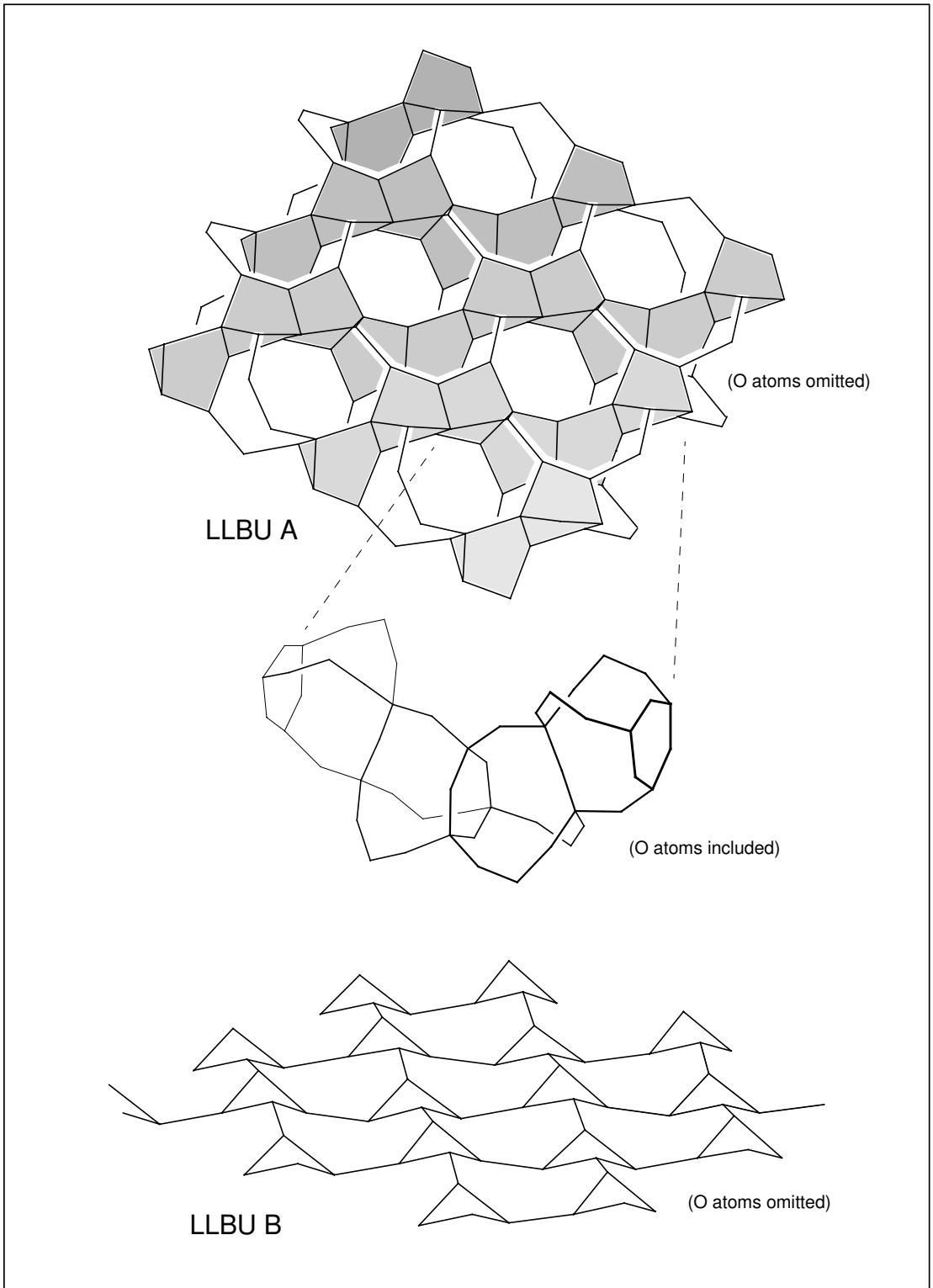


Figure 4-35: VPI-9 layer-like building units

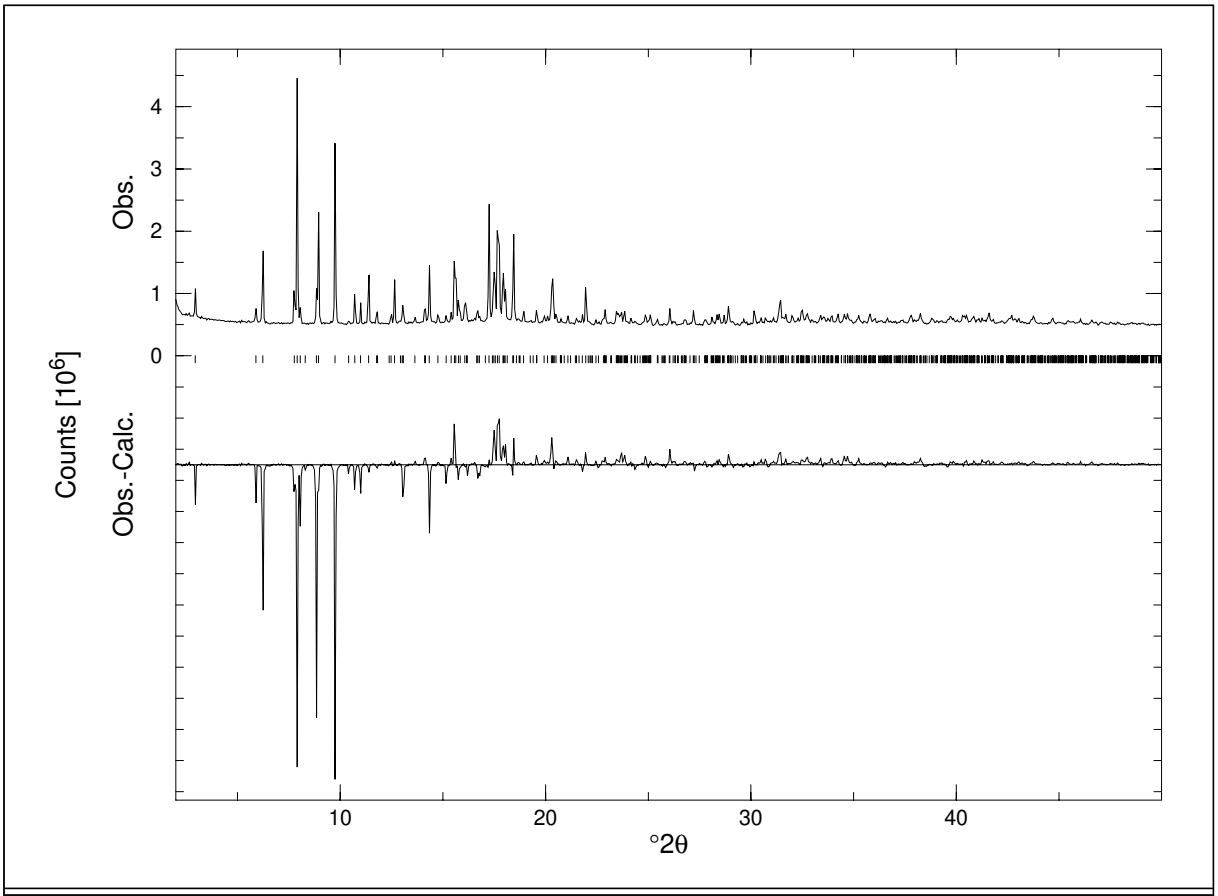


Figure 4-36: VPI-9 (MY027) Initial Rietveld plot with DLS model

ammonium ions and water molecules at positions gleaned from a series of difference Fourier maps. The refinement was taken to a point where the correctness of the framework became evident. Fig. 4-37 shows the plot of the observed and difference profiles at that stage. Tab. 4-18 summarizes the results. Minimum, maximum and average bond distances and angles of the framework atoms are listed in Tab. 4-19. For the Ow positions, Tab. 4-20, lists the distances to the six nearest neighbors.

A remark is necessary regarding the estimated standard deviations (ESD's) listed in Tab. 4-18. Since the Durbin-Watson d-value (DWd) [63] is not in the appropriate range (ca. 2), all ESD's are too low, presumably at least one order of magnitude. The correct approach to obtain reliable ESD's is to take only every *n*'th profile point in the refinement, where *n* has to be chosen such that DWd is in the range of two, as indicated in Tab. 4-18. However, only with *n* = 33, could DWd be raised to 2.069, leaving 160 profile points to fit 98 parameters. This is of course unacceptable. Furthermore, the development of DWd with *n* was very unstable for *n* greater than 20. Since the objective of the present refinement was to verify the framework topology rather than to obtain exact structural parameters, the ESD's were not considered further. The data in Tab. 4-18 are presented for completeness, but have to be considered in

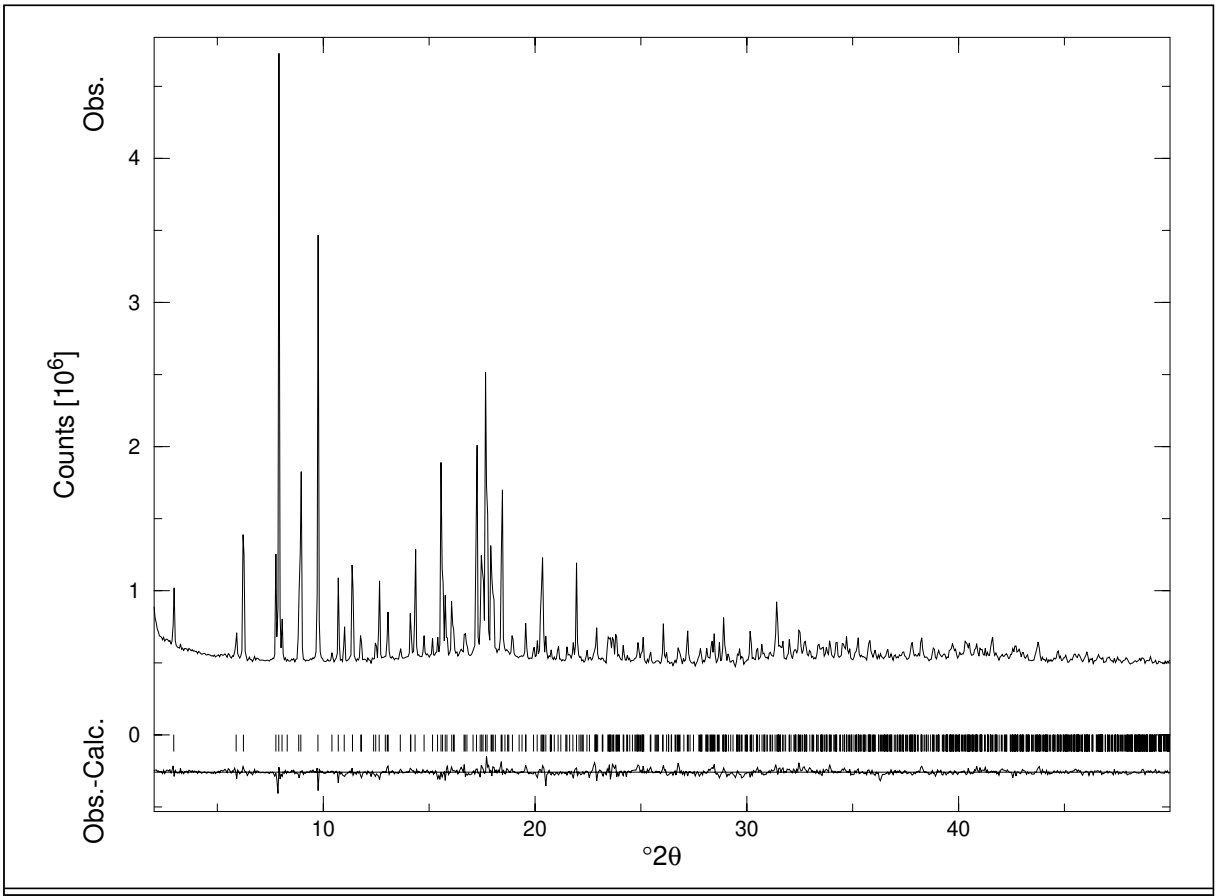


Figure 4-37: GSAS Rietveld refinement for ammonium exchanged VPI-9

conjunction with this reservation.

Meanwhile, L.B. McCusker (in our research group) has carried out a careful Rietveld refinement of as-synthesized MY027. With space group and lattice constants as indicated in Fig. 4-30 (i.e. the larger unit cell), and a chemical composition of

$$Rb_{44}K_4[Si_{96}Zn_{24}O_{240}] \cdot 48H_20,$$

the refinement converged to $R_F = 0.069$ and $R_{wp} = 0.147$. The assignment of the zinc positions of the preliminary refinement was confirmed. However, in contrast to the data listed in Tab. 4-18, the occupancy factors for all silicon atoms were found to be 1.0.

On the basis of these data, a proposal for the assignment of a new structure type code was submitted to the Structure Commission of the International Zeolite Association, and they have now approved the code **VNI** for the VPI-9 topology.

Space group	P 4_2/n c m (No. 138) [Origin choice 2]
Unit cell	a = 9.89460(8) c = 36.8715(4) Å
Soft-constraints	62
Zero	0.46 ($\Rightarrow$ 0.0046° 2θ)
GU	22.9
GV	-0.1
GW	2.7
LX	1.432
LY	3.183
Background	fixed + 6 parameters for function #2
R_p	0.0214
R_{wp}	0.0292
R(F**2)	0.1536
X^2	3.029
DWd	0.542 (No serial correlation in fit at 90% confidence for 1.911 < DWd < 2.089)

Number of structural parameters 86 (48 positional, 26 displacement, 12 occupancy)

	x	y	z	U [$Å^2.10^2$]	Occupancy
Si1	0.1012(5)	0.6947(5)	0.90419(18)	0.16(30)	1.178(9)
Zn2	0.25	0.25	0.49297(18)	0.04(23)	1.0
Si3	0.1031(7)	0.1031(7)	0.27089(26)	1.96(55)	1.191(17)
Si4	0.0365(6)	0.0365(6)	0.45889(17)	0.35(43)	1.185(11)
Si5	0.1403(5)	0.1403(5)	0.06549(26)	1.94(48)	1.187(14)
Si6	0.1107(7)	0.1107(7)	0.73491(24)	3.52(55)	1.248(18)
Zn7	0.9783(3)	0.9783(3)	0.34102(12)	0.10(14)	1.0
O1	0.0851(12)	0.6019(9)	0.93994(27)	3.4(6)	1.0
O2	0.25	0.75	0.90955(57)	3.7(9)	1.0
O3	0.0065(10)	0.8203(10)	0.91742(37)	5.2(8)	1.0
O4	0.0731(15)	0.6264(12)	0.86580(29)	5.5(7)	1.0
O5	0.1531(7)	0.1531(7)	0.46083(51)	0.0(9)	1.0
O6	0.1458(8)	0.3542(8)	0.52100(31)	0.5(9)	1.0
O7	0.0628(10)	0.2533(7)	0.25457(32)	1.2(6)	1.0
O8	0.1006(13)	0.1006(13)	0.31514(31)	2.5(10)	1.0
O9	-0.0004(8)	-0.0004(8)	0.25028(37)	3.5(10)	1.0
O10	0.0	0.0	0.5	0.2(11)	1.0
O11	0.25	0.25	0.08233(54)	0.5(13)	1.0
O12	0.1127(12)	0.1127(12)	0.69059(30)	0.9(9)	1.0
Ow1	0.25	0.75	0.32260(64)	14.8(21)	1.016(32)
Ow2	0.1051(12)	0.1051(12)	0.16503(48)	4.4(11)	1.315(29)
Ow3	0.6442(7)	0.1442(7)	0.5	2.8(9)	1.402(24)
Ow4	0.3754(13)	0.1246(13)	0.40821(75)	16.8(18)	1.327(34)
Ow5	0.25	0.75	0.75	1.5(12)	1.274(38)
Ow6	0.4142(44)	0.0858(44)	0.3336(14)	36.2(47)	1.023(70)
Ow7	0.3113(52)	0.1887(52)	0.3474(19)	0.5(42)	0.313(47)

Table 4-18: Results of the preliminary Rietveld refinement for VPI-9 (MY027)

Si - O distances		Si - O - Si angles		O - Si - O angles	
Minimum	1.583	Minimum	132.8	Minimum	97.6
Maximum	1.624	Maximum	180.0	Maximum	119.3
Average	1.656	Average	151.0	Average	109.2
Zn - O distances		**Zn - O - Si angles**		**O - Zn - O angles**	
Minimum	1.670	Minimum	120.4	Minimum	97.7
Maximum	1.960	Maximum	147.9	Maximum	113.4
Average	1.775	Average	133.8	Average	109.4

Table 4-19: VPI-9 framework distances [Å] and angles [°] of the preliminary refinement

4.8 The structure of the zincosilicate zeolite VPI-10

4.8.1 Preparation and application of FOCUS

The synthesis and chemical composition of VPI-10 is very similar to VPI-9. Like VPI-9, VPI-10 had already been characterized by Annen et al. [21, 22]. The sample used for structure determination was synthesized by M. Yoshikawa, in the same series as MY027 (VPI-9). In view of the experience gathered with VPI-9, the ammonium exchange experiment described on page 67 was repeated with VPI-10, prior to any structure determination attempt. The effect of this treatment can be seen by comparing the observed profiles shown in Fig. 4-38 and Fig. 4-39, which were collected on the same laboratory diffractometer with Cu-$K_{\alpha 1}$ radiation.

Since the lattice constants were not known, indexing of the as-synthesized sample was attempted with TREOR [64] as well as with POWDER [52]. At first a solution with the hexagonal metric of a = 12.64 and c = 3.53 Å was found by TREOR. However, nine lines remained unindexed. These were not very strong, but clearly present. After some difficulties, and careful selection of peak positions to be used for indexing, POWDER produced a body-centred orthorhombic unit cell with a = 12.64, b = 21.92, and c = 7.07 Å. A relation to the hexagonal cell can be established by transforming the hexagonal cell to an orthorhombic C-centred cell:

$$a_{orth} = a_{hex} = b_{hex} = 12.64 \text{ Å}$$
$$b_{orth} = b_{hex} \cdot 2 \cdot \cos(30°) = 21.89 \text{ Å}$$

However, the cell found by POWDER is not C-centred, but I-centred. Furthermore, the c direction is doubled. This means that although the ratio of the orthorhombic lattice constants a and b is that of a pseudo-hexagonal metric, the lattice itself is clearly not pseudo-hexagonal. The nearly perfect pseudo-hexagonal ratio of a and b must be seen as pure chance, caused by a structural peculiarity.

Besides making indexing difficult, the pseudo-hexagonal ratio of a and b caused further trouble in the determination of the space group. For an I-centred orthorhombic space group, there are only three systematic absences to look for: 0kl:k,l=2n, h0l:h,l=2n, and hk0:h,k=2n. The first two of them could be ruled out immediately, but as a consequence of the pseudo-hexagonal ratio of a and b, it was not clear whether or not the third reflection condition was violated. For example, the 110 reflection at position 8.0598° in the profile overlaps with 020 at 8.0690°, or 130 at 13.9885° overlaps with 200 at 14.0046°. Up to high reflection angles, the reflections in question always overlap with some non-absent reflection. So it was impossible to decide whether extinction symbol I--- or I--(ab) was appropriate. I--- is

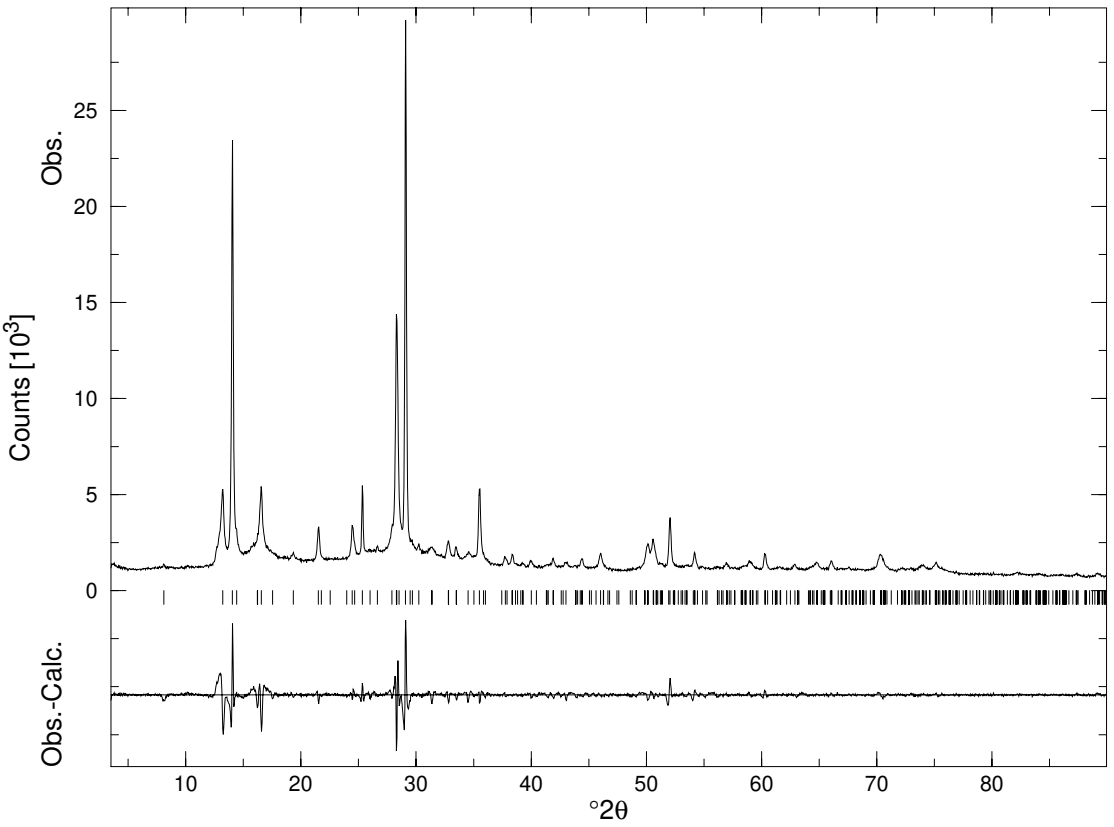


Figure 4-38: GSAS whole-profile intensity extraction for ammonium exchanged VPI-10 – Space group I222

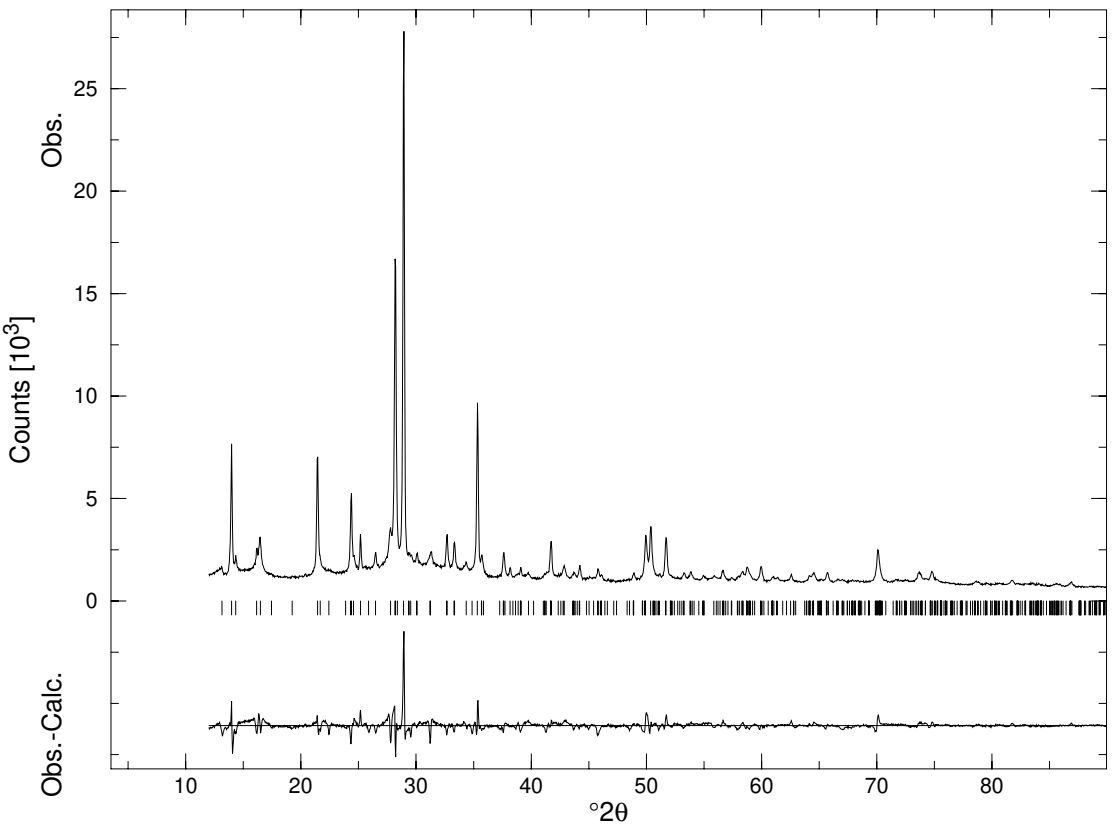


Figure 4-39: Preliminary GSAS Rietveld refinement for as-made VPI-10

consistent with four different space groups with a total of six different settings, and *I*--(*ab*) with two different space groups with a total of four different setting. Altogether, ten different possibilities had to be considered.

For structure determination with the ammonium exchanged sample, one intensity extraction with GSAS was carried out in space group I222 [extinction symbol *I*---], a second one in space group I*m*2*a* [extinction symbol *I*--(*ab*)]. In the second extraction, all parameters were held fixed, and only the intensity distribution was refined. Both extractions converged to exactly the same residual values. The final profile plots are also indistinguishable, so only one plot is shown in Fig. 4-38. The chemical formula estimated from chemical analysis data provided by M.E. Davis for as-synthesized VPI-10, and the results of the intensity extraction are listed in Tab. 4-21. No chemical analysis was made on the NH^{4+}-exchanged material. Instead, the NH_4^+ and H_2O content was only roughly approximated.

Estimated formula		
	$(NH_4^+)_{16}$ [Si_{28} Zn_8 O_{72}] ·28 H_20	
Data collection		
	STOE Stadi-P, linear PSD	
	Rotating 0.3 mm capillary	
	Cu-$K_{\alpha 1}$ (1.5406 Å) radiation	
	2θ range 3.5 - 90°, step size 0.02°	
Intensity extraction		
	Space group	I 2 2 2 (No. 23) & I m 2 a (No. 46)
	Unit cell	a = 12.599 b = 21.810 c = 7.022 Å
	GU	1561.8
	GV	-709.9
	GW	92.0
	LX	12.067
	asym	-0.0017
	R_p	0.0576
	R_{wp}	0.0938

Table 4-21: Selected data for VPI-10

Fig. 4-40 shows the main parts of a sample FOCUS input file. The original structure determination was made before framework fragment recycling was introduced (this means, Fig. 4-40 is an idealized remake of the original input file). Therefore, the recycling technique employed was atom recycling with use of oxygen. The same input template was used to set up input files for six space groups and settings with the intensities extracted in the space group I222, and for four space groups and settings with the intensities extracted in the space group I*m*2*a*. Each FOCUS run produced hundreds to more than a thousand unique topologies. Since the absolute success rates varied significantly, the histograms in Fig. 4-41 were normalized by

```
Title VPI-10 NH4+ exch. 0.3 cap. STOE lin. PSD GSAS extraction
SpaceGroup  I 2 2 2
UnitCell  12.5995 21.8104 7.0223

AtomType  +  Node        Zn    8  *  0.01  Zn2+
AtomType  +  Node        Si   28  *  0.01  Si4+
AtomType  +  NodeBridge  O    72  *  0.02  O2-
AtomType  -  *           O    28
AtomType  -  *           N    16
AtomType  -  *           H   120

Chemistry  MinDistance  Node        Zn  Node        Zn  4
Chemistry  MinDistance  Node        Zn  Node        Si  2.6
Chemistry  MinDistance  Node        Zn  NodeBridge  O   1.4
Chemistry  MinDistance  Node        Si  Node        Si  2.6
Chemistry  MinDistance  Node        Si  NodeBridge  O   1.4
Chemistry  MinDistance  NodeBridge  O   NodeBridge  O   2.3
MaxPotentialAtoms  108
MaxRecycledAtoms  108

FwSearchMethod  FwTracking
MaxPeaksFwSearch  216
MaxPeaksFwFragmentSearch  216
MinNodeDistance  2.3
MaxNodeDistance  3.7
MinSymNodes  0
MaxSymNodes  40
NodeType  4  *  -6 -3 -1 4 6
MinLoopSize  3
MaxLoopSize  24
EvenLoopSizesOnly  Off
Check3DimConnectivity  On
IdealT_NodeDistance  3.15
CheckTetrahedralGeometry  Normal

RandomInitialization  Time
FeedBackCycles  6
FeedBackBreakIf  PhaseDiff < 5.00 % and DeltaR < 1.00 %

Grid_xyz  40  68  24
eDensityCutOff  1 %
MinPfI  17
CatchDistance  0.5
eD_PeaksSortElement  Grid_eD

Lambda  CuA1
FobsMin_d  1.3
FobsScale  0.13
SigmaCutOff  0
OverlapFactor  0.15
OverlapAction  EqualF2
ReflectionUsage  90 %
Grid_hkl  20  +17  12
```

Figure 4-40: A sample FOCUS input for VPI-10

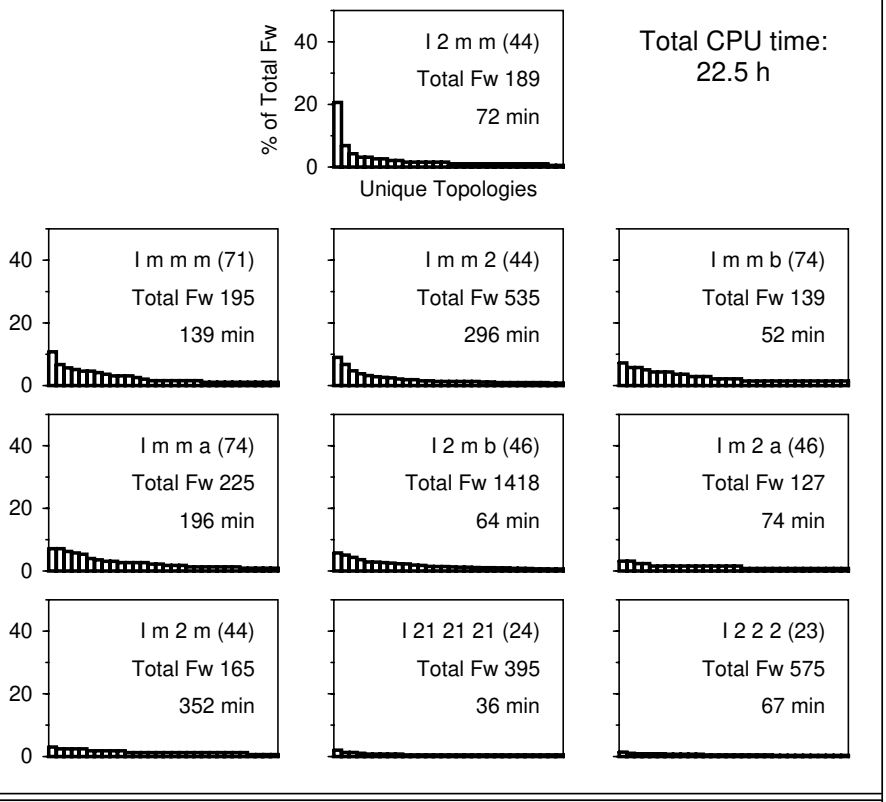


Figure 4-41: VPI-10: atom recycling in 10 space groups

dividing the absolute frequency of occurrence of a unique topology by the total number of topologies found in that run. In the histograms of Fig. 4-41, the upper line designates the space group used, the second line shows the total number of topologies found, and the third line shows the computing time in minutes used for the run.

Only one of the ten histograms shows a clear discrimination of the two most frequently occurring topologies: the histogram for space group I2*mm*. The DLS residual computed for the pure silicon framework of the most frequently occurring topology in I2*mm* was 0.0068 (as usual, oxygen atoms were inserted with help of KRIBER). There was another topology with a DLS residual of 0.0067 in space group I*mm*2, which also occurred in I222 and I*mmm*, but the relative frequencies of occurrence were only 0.37%, 0.17%, and 1.54%, respectively. Another reason not to consider this topology further is based on a geometrical argument: the topology contains extremely improbable elliptical 8-rings.

Investigation of the topology which is represented by the second bar in the histogram for space group I2*mm* (DLS residual 0.0080) revealed that the projections of the two most frequently occurring topologies along the *c* axis are identical. This was also observed for the NU-3 test case, and was therefore taken as another argument to support the correctness of the solution in this space group.

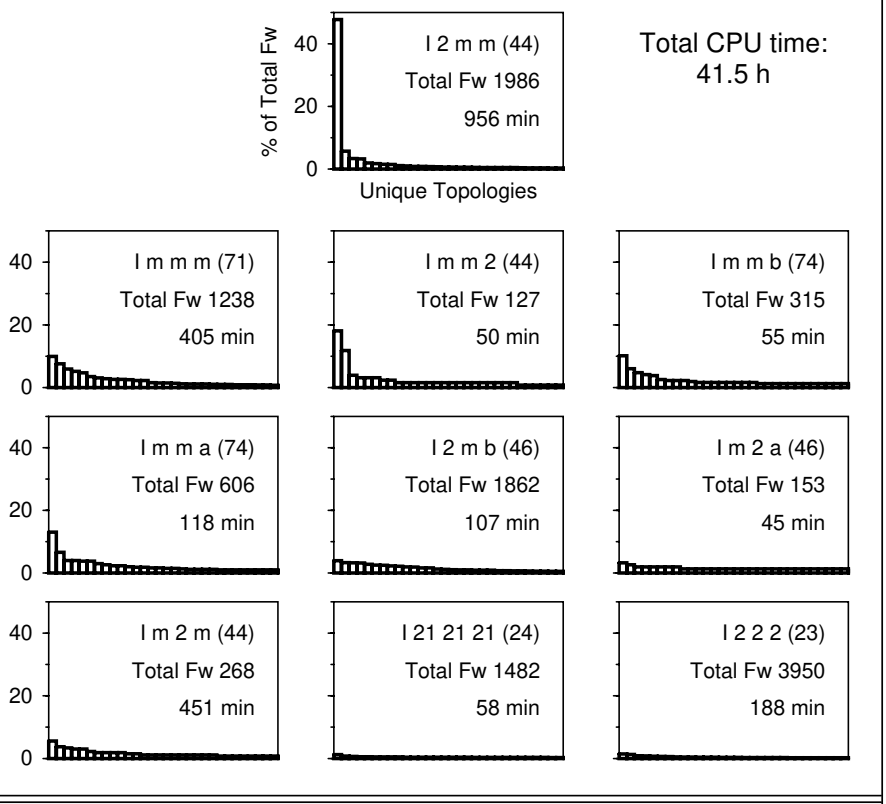


Figure 4-42: VPI-10: framework fragment recycling in 10 space groups

Before proceeding further, the results of subsequent runs with the improved version of FOCUS are presented, in order to support the correctness of the structure to be proposed. First, the ten runs of the original structure solution were repeated with framework fragment recycling, and the histograms in Fig. 4-42 prepared in the same manner – and on the same scales – as in Fig. 4-41. As can be seen, the overall outcome is very similar. Again, a clear discrimination of the two most frequently occurring topologies is only found for space group I2*mm*. In addition, the discrimination is significantly better than in Fig. 4-41.

Finally, for comparison, the test series of six runs as presented for most of the previous cases (see also Tab. 4-4) was also set up for VPI-10 in space group I2*mm*. `ReflectionUsage` was set to 75%, and `MaxPeaksFwSearch` to 320. For the three runs without oxygen recycling, `MaxPotentialAtoms` and `MaxRecycledAtoms` were set to 48 and 36, respectively. In the other three runs with oxygen recycling, both parameters were set to 108. The results of the runs are shown in the histograms of Fig. 4-43. Tab. 4-22 summarizes the relevant data. In all cases, the most frequently occurring topology is the same, and will be referred to as the "VPI-10 framework" in the following presentation of the structure. In all respects, the results from the previous cases are repeated very nicely for VPI-10. In contrast to the results found for the chemically similar structures RUB-17 and VPI-9, the use of oxygen consistently enhances the

success rates. However, the advantage or disadvantage of using oxygen does not appear to be a very important factor for any of the examples studied.

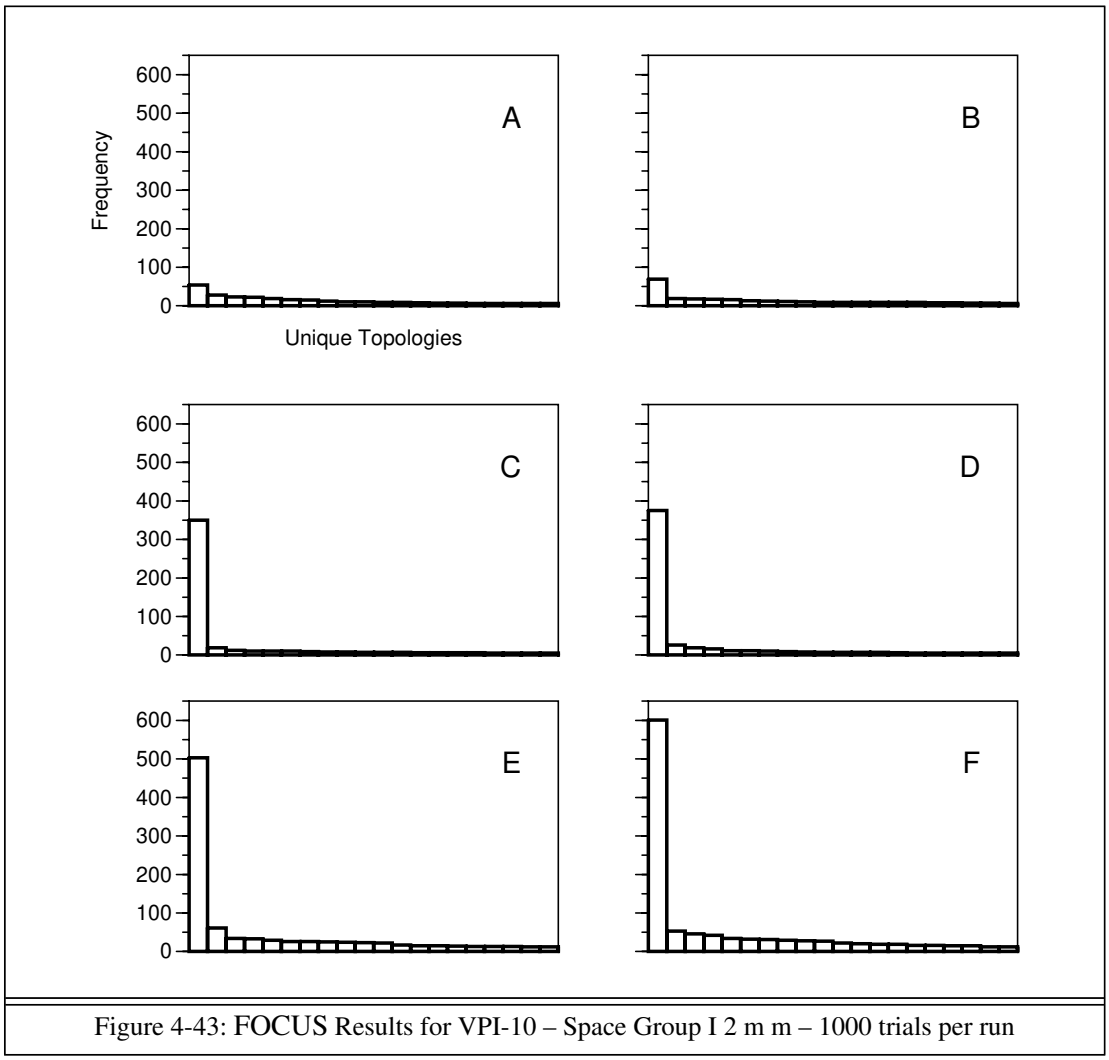


Figure 4-43: FOCUS Results for VPI-10 – Space Group I 2 m m – 1000 trials per run

Run	NBO	MPA	MRA	FBC	FT	Fw	UFw	RFw	t/min	%tFwS
A	-	48	36	6	4111	868	362	266259	73	40
B	+	108	108	6	4462	788	339	256626	79	37
C	-	48	36	0 6	6338	697	144	55537	272	75
D	+	108	108	0 6	6314	696	121	54556	248	73
E	-	48	36	1 1 1 1 1 1 1 1 1 1	10939	1848	450	370229	310	62
F	+	108	108	1 1 1 1 1 1 1 1 1 1	11000	2020	445	367367	314	62

Table 4-22: FOCUS Results for VPI-10 – 1000 trials per run (153 actively used reflections)

4.8.2 Verification of the structure: preliminary Rietveld refinement

Fig. 4-44 shows the DLS refined VPI-10 framework viewed approximately parallel [001]. Seven node atoms in the asymmetric unit – which generate 30 atoms in the unit cell – form a framework of 3, 4, 8, and 9-rings, and a complex 3-dimensional channel system. The projection of the framework along [001] is almost perfectly trigonal, giving a good explanation for the pseudo-hexagonal ratio of the lattice constants *a* and *b*. The space group I2*mm* is the highest topological symmetry of this framework, so, the number of node atoms per asymmetric unit cannot be reduced by the introduction of further symmetry operations.

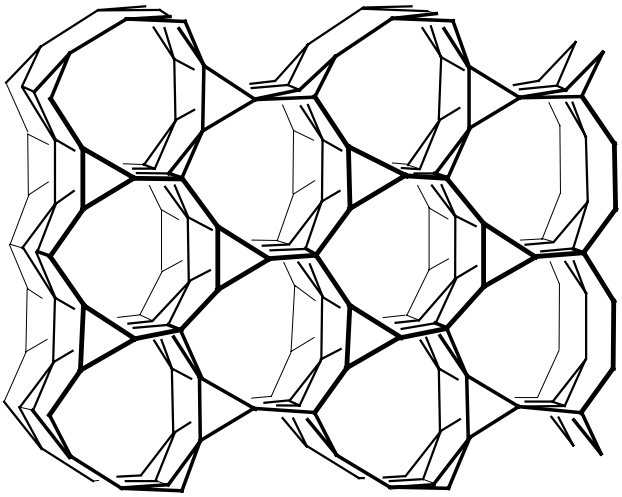


Figure 4-44: The framework of VPI-10
(O atoms have been omitted for clarity)

The DLS refined model was used to initialize a Rietveld refinement for as-synthesized VPI-10. To give an impression of the difficulties involved in the verification of the structure, Fig. 4-45 shows the initial plot of the observed and the difference profile. In order to keep the framework reasonable, soft-constraints were introduced. One zinc position was tentatively assigned after refinement of the occupancy factors for the node atoms (all silicon to start with). However, based on NMR measurements provided by M.E. Davis, it was assumed that zinc is not fully ordered in the framework. Therefore, the other six positions were assigned Si scattering factors and their occupancy factors refined. The model was improved by adding K atoms (to approximate Rb, K and H_2O) at positions gleaned from a series of difference Fourier maps. After each addition to the model, positions and occupancy factors were refined. This way the refinement was taken to a point where the correctness of the framework became evident. Fig. 4-38 shows the plot of the observed and difference profiles at the point, where the preliminary refinement was stopped. Tab. 4-23 summarizes the results. The essentials of the

comment on page 75 regarding the ESD's also apply to Tab. 4-23. The fit in Fig. 4-38 strongly supports the assumption that the proposed topology is correct. However, the limited quality of the refinement is underlined by the interatomic distances shown in Tab. 4-24 and Tab. 4-25. Especially the very short distance between K1 and O12 makes it necessary to await the results of a careful refinement before the topology can be accepted without reservation. Recently, a synchrotron dataset was collected for as-synthesized VPI-10, and that refinement is in progress.

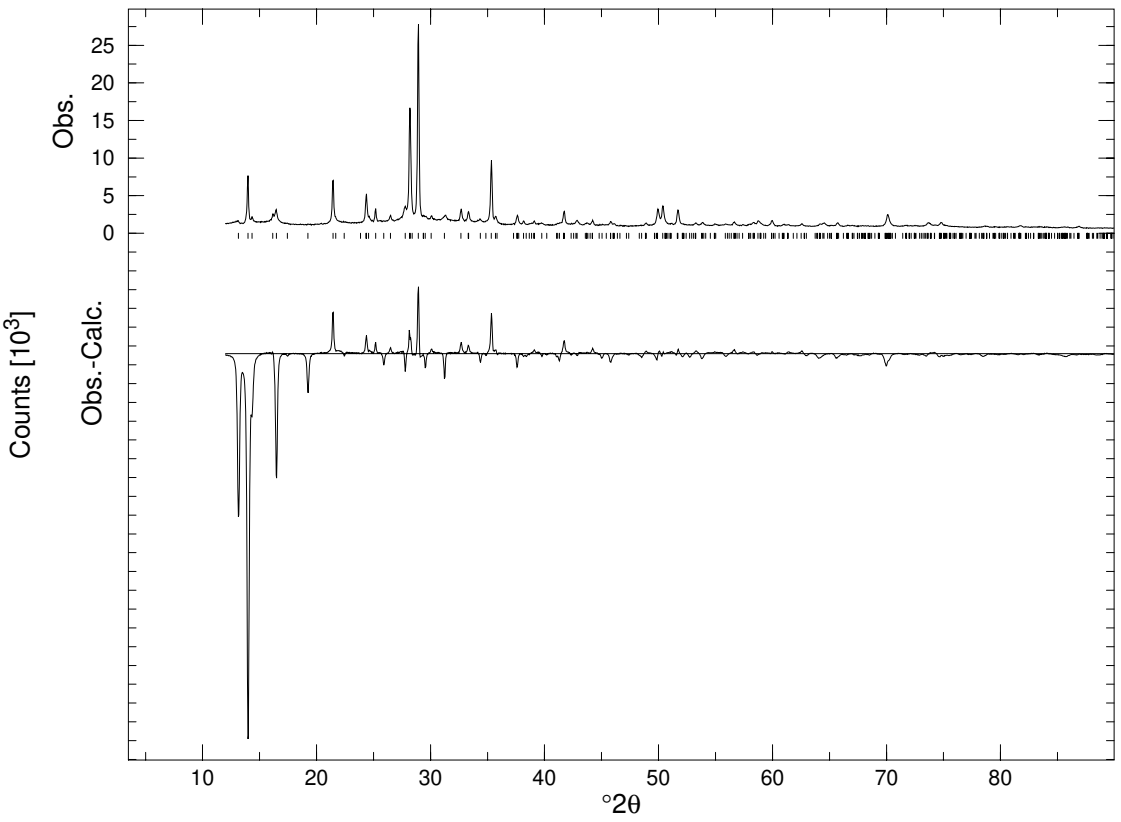


Figure 4-45: VPI-10 Initial Rietveld plot with DLS model

Space group	I 2 m m (No. 44)
Unit cell	a = 12.6498(11) b = 21.8995(19) c = 7.06534(19) Å
Soft-constraints	62
GU	1137.3
GV	-553.1
GW	84.1
LX	8.984
Background	fixed + 6 parameters for function #2
R_p	0.0697
R_{wp}	0.0948
R(F**2)	0.1299
χ^2	9.181
DWd	0.157 (No serial correlation in fit at 90% confidence for 1.911 < DWd < 2.089)

Number of structural parameters 90 (52 positional, 25 displacement, 13 occupancy)

	x	y	z	U [Å^2·10^2]	Occupancy
Si1	0.7854(28)	0.0	0.5	4.0(24)	1.83(17)
Si2	0.0145(31)	0.3963(14)	0.0	1.9(23)	0.93(10)
Si3	0.0931(30)	0.5	0.2685(34)	0.3(20)	1.21(10)
Zn4	0.0087(24)	0.2601(11)	0.0	3.5(12)	1.02(4)
Si5	0.9908(26)	0.0653(12)	0.5	4.2(19)	1.86(15)
Si6	0.7904(24)	0.3346(13)	0.0	1.6(22)	1.39(13)
Si7	0.1012(23)	0.1567(12)	0.2235(34)	1.2(15)	1.01(7)
O1	0.7128(33)	0.0	0.305(4)	2.0(46)	1.0
O2	0.8592(27)	0.9384(14)	0.5	1.8(50)	1.0
O3	0.0582(42)	0.3279(13)	0.0	0.9(37)	1.0
O4	0.0441(43)	0.4380(13)	0.180(4)	4.8(27)	1.0
O5	0.8833(31)	0.3902(21)	0.0	0.7(42)	1.0
O6	0.0526(60)	0.5	0.5	8.1(58)	1.0
O7	0.0542(44)	0.2212(17)	0.814(4)	3.6(27)	1.0
O8	0.8766(28)	0.2766(19)	0.0	0.8(30)	1.0
O9	0.0410(57)	0.0	0.5	0.1(69)	1.0
O10	0.0165(37)	0.1106(20)	0.323(4)	2.3(23)	1.0
O11	0.7238(27)	0.3508(26)	0.802(4)	1.5(31)	1.0
O12	0.0856(65)	0.1104(23)	0.0	5.1(41)	1.0
K1	0.4338(41)	0.4072(26)	0.5	38.0(47)	1.95(17)
K2	0.2531(26)	0.3466(12)	0.2593(31)	3.0(9)	1.27(5)
K3	0.5199(21)	0.2291(10)	0.0	6.8(15)	1.96(8)
K4	0.3734(23)	0.1520(16)	0.0	8.0(20)	1.37(9)
K5	0.5489(35)	0.5	0.5	4.4(26)	1.42(14)
K6	0.2529(53)	0.0	0.7636(33)	17.0(25)	1.91(12)

Table 4-23: Results of the preliminary Rietveld refinement for VPI-10

Si - O distances		Si - O - Si angles		O - Si - O angles	
Minimum	1.555	Minimum	113.6	Minimum	87.7
Maximum	1.887	Maximum	137.1	Maximum	126.3
Average	1.648	Average	150.5	Average	109.1
Zn - O distances		**Zn - O - Si angles**		**O - Zn - O angles**	
Minimum	1.611	Minimum	136.9	Minimum	100.7
Maximum	1.710	Maximum	142.8	Maximum	116.4
Average	1.665	Average	138.9	Average	109.5

Table 4-24: VPI-10 framework distances [Å] and angles [°] of the preliminary refinement

K1	O12	1.958	K4	K3	2.507
	K5	2.500		O11	2.853
	O10	2.541		O11	2.853
	O10	2.541		O5	3.654
	SI7	2.989		O5	3.654
	SI7	2.989		O4	3.696
K2	O2	3.033	K5	O12	2.462
	O3	3.099		O12	2.462
	K1	3.143		K1	2.500
	O4	3.363		K1	2.500
	K2	3.401		K6	3.074
	O1	3.428		K6	3.074
K3	K4	2.507	K6	K5	3.074
	O7	2.509		O9	3.264
	O7	2.509		O11	3.321
	SI7	3.336		O11	3.321
	SI7	3.336		K6	3.341
	ZN4	3.543		O5	3.460

Table 4-25: VPI-10 K distances [Å]
of preliminary refinement

4.9 The proposed structure of the beryllosilicate B2

4.9.1 Preparation and structure determination

The B2 sample was obtained by S. Ueda et al. [65] in a series of syntheses conducted with the aim of producing zeolitic beryllosilicates. There are no publications dealing with B2 directly, but in [65] (title: “Synthetic Lovdarite”) the synthesis is described:

> SYNTHESIS. Reaction mixtures of the composition $5.25Na_2O{\cdot}1.75K_2O{\cdot}0.50TEA_2O{\cdot}0.30BeO{\cdot}(8.0\text{-}18.0)\ SiO_2{\cdot}200H_2O$ were prepared using 10M aqueous solutions of sodium and potassium hydroxide, 10% aqueous solution of tetraethylammonium hydroxide, beryllium carbonate, colloidal silica sol and water. The reactant mixtures were heated to 200°C for 3 to 12 days in autoclaves with teflon linings. Organic bases such as TMAOH or TEAOH, though not incorporated in the zeolitic product, appear necessary to obtain the pure phase. In addition to lovdarite two unidentified phases were also encountered in the course of the synthesis experiments.

B2 is one of the “two unidentified phases”, and a sample was given to Prof. W.M. Meier (at that time head of this research group) for structure determination in 1985. Over the intervening years, several members of this group have worked on the problem. In 1988, a synchrotron dataset was collected at the HASYLAB in Hamburg by Ch. Baerlocher, and the profile could be indexed on an orthorhombic cell with a = 13.17, b = 7.126, and c = 12.68 Å. Investigation of the systematic absences suggested two possible extinction symbols, P*n*-*a* or P--*a*. The systematic absence due to the *n*-glide was violated by only one small peak, which could also be the trace of an impurity phase. Based on these observations, and expecting 3-rings, G.O. Brunner devised a model, which is referred to as “B2 model A” in the following. A Rietveld refinement was attempted by R.M. Kirchner and Ch. Baerlocher, but was not rewarded with success. Another attempt to determine the structure was made by M.A. Estermann in 1992, with a combination of the FIPS method [17] and conventional direct methods, but the structure of B2 remained unsolved.

For the first attempt to solve the structure with FOCUS, the HASYLAB profile was reinvestigated, and the previous indexing and possible extinction symbols confirmed. The manually fixed background was also redetermined, and – ignoring the small violation of the *n*-glide – the intensity extraction was carried out in the space group P*nma*. Next, the test series of six runs (see also Tab. 4-4) was derived from the FOCUS input template in Fig. 4-46. The prescribed cell content is Ch. Baerlocher’s estimate based on a chemical analysis of S. Ueda. There are additional water molecules in the channel system, but, because of their minor contribution and for the sake of simplicity, these were ignored. For the three runs without

```
Title B2 Hasylab 1988, GSAS extraction
SpaceGroup  P n m a
UnitCell  13.1731 7.1256 12.6777

AtomType  +  Node        Si  16  *  0.01
AtomType  +  Node        Be   4  *  0.01
AtomType  -  NodeBridge  O   40  *  0.02
AtomType  -  *           K    4
AtomType  -  *           Na   4

Chemistry  MinDistance  Node        Si  Node        Si  2.6
Chemistry  MinDistance  Node        Si  Node        Be  2.6
Chemistry  MinDistance  Node        Si  NodeBridge  O   1.4
Chemistry  MinDistance  Node        Be  Node        Be  2.6
Chemistry  MinDistance  Node        Be  NodeBridge  O   1.4
Chemistry  MinDistance  NodeBridge  O   NodeBridge  O   2.3
MaxPotentialAtoms  24
MaxRecycledAtoms  20

FwSearchMethod  FwTracking
MaxPeaksFwSearch  400
MaxPeaksFwFragmentSearch  320
MinNodeDistance  2.6
MaxNodeDistance  3.6
MinSymNodes  0
MaxSymNodes  24
NodeType  4  *  -6 -3 -1 4 6
MinLoopSize  3
MaxLoopSize  24
EvenLoopSizesOnly  Off
Check3DimConnectivity  On
IdealT_NodeDistance  3.1
CheckTetrahedralGeometry  Normal

RandomInitialization  Time
FeedBackCycles  6
FeedBackBreakIf  PhaseDiff < 5.00 % and DeltaR < 1.00 %

Grid_xyz  40  24  40
eDensityCutOff  1 %
MinPfI  17
CatchDistance  0.5
eD_PeaksSortElement  Grid_eD

Lambda  1.4030
FobsMin_d  0
FobsScale  0.15
SigmaCutOff  0
OverlapFactor  0.15
OverlapAction  EqualMF2
ReflectionUsage  75 %
Grid_hkl  +10  12  20
```

Figure 4-46: A sample FOCUS input for B2

oxygen recycling, `MaxPotentialAtoms` and `MaxRecycledAtoms` were set to 24 and 20, respectively. In the other three runs with oxygen recycling, both parameters were set to 80 (admittedly due to an error, 60 would have been the "ideal" value).

As can be seen in Fig. 4-47 and the corresponding Tab. 4-26, the absolute success rate of the FOCUS runs was surprisingly high. With recycling technique F, 500 trials were sufficient to accumulate nearly 2500 topologies of the most frequently occurring type. This topology, now referred to as "B2 model B", had not been considered before. Another surprise was the observation that in all six runs, the second most frequently occurring topology is model A mentioned above.

Since the histograms of Fig. 4-47 were considered to give a clear indication as to the correct B2 structure, and the DLS residual value for the pure silicon framework converged satisfactorily to 0.0065, no other space groups were investigated at this point. Ch. Baerlocher tried to refine model B with the Rietveld technique. However, although considerable effort was put into the refinement, model B could not be confirmed.

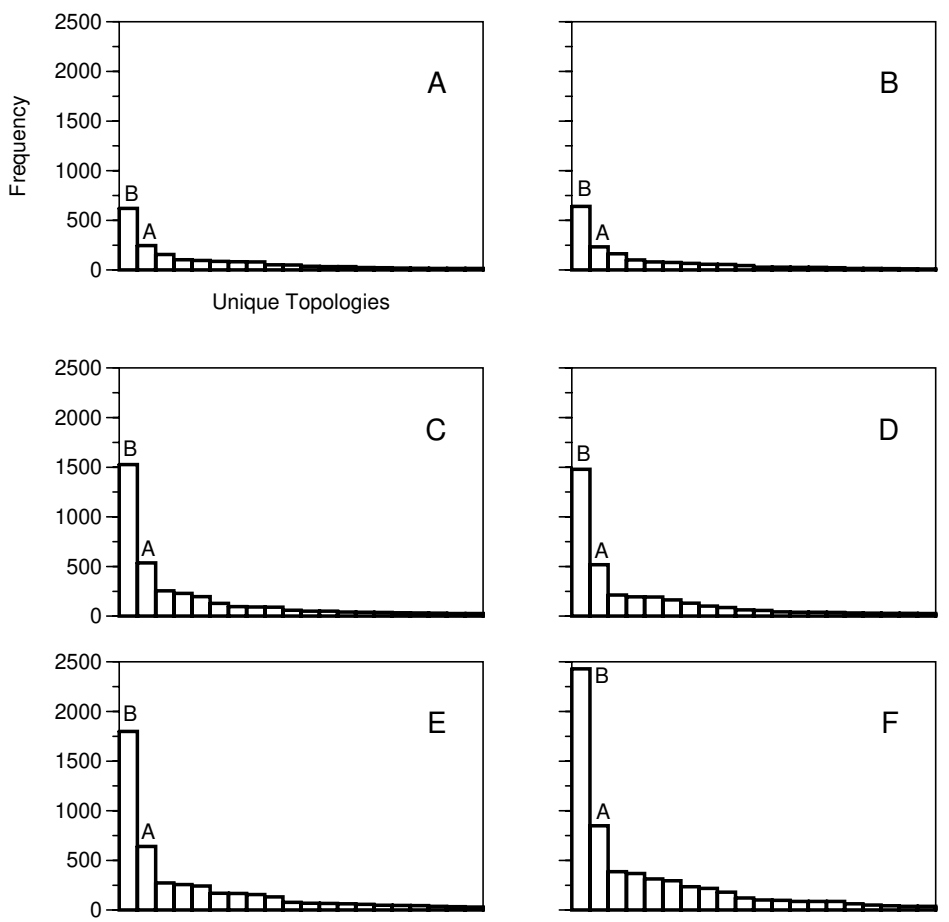


Figure 4-47: FOCUS Results for B2 – Space Group Pnma – 500 trials per run

Run	NBO	MPA	MRA	FBC	FT	Fw	UFw	RFw	t/min	%tFwS
A	-	24	20	6	1904	1914	45	98452	17	47
B	+	80	80	6	2198	1807	42	96524	18	44
C	-	24	20	0 6	2970	3783	47	61260	35	63
D	+	80	80	0 6	2965	3720	47	61069	34	59
E	-	24	20	1 1 1 1 1 1 1 1 1 1	3984	4662	54	142374	40	55
F	+	80	80	1 1 1 1 1 1 1 1 1 1	5247	6442	55	175623	60	62

Table 4-26: FOCUS Results for B2 – Space Group Pnma – 500 trials per run (155 actively used reflections)

In the hope of finding alternative models, the intensity extraction was repeated in the space group P2_1*ma*. The resulting plot, which is very similar to the plot for the previous extraction in space group P*nma*, is shown in Fig. 4-48. Tab. 4-27 lists the relevant results of the extraction.

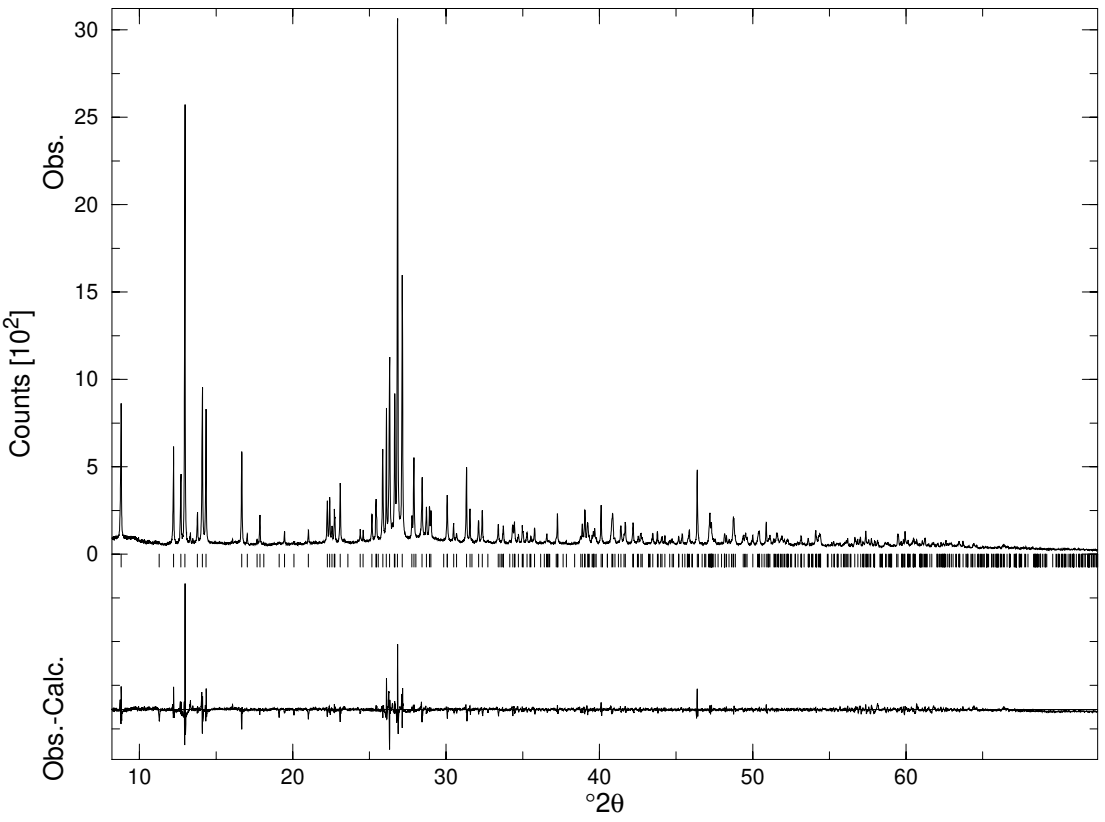


Figure 4-48: GSAS whole-profile intensity extraction for B2 – Space group P2_1ma

In contrast to the intensities extracted in the space group P*nma*, the partitioning of the overlapping intensities extracted in the lower space group P2_1*ma* was enhanced with FIPS as explained on page 63. FOCUS input files for four space groups (P*n*2_1*a*, P2_1*ma*, P*m*2*a*, P*mma*) were generated by combining the template file in Fig. 4-46 with the appropriate extracted

Estimated formula	
$K_4\ Na_4\ [Si_{16}\ Be_4\ O_{40}] \cdot 16\ H_2O$	
Data collection	
HASYLAB beamline B2	
Rotating 1.0 mm capillary	
Wavelength 1.4030 Å	
2θ range 8.2 - 50°, step size 0.01°	
2θ range 50 - 90°, step size 0.02°	
Intensity extraction	
Space group	P 2_1 m a (No. 26)
Unit cell	a = 13.173 b = 7.126 c = 12.678 Å
GU	62.6
GV	-18.4
GW	1.7
LX	4.008
asym	0.5694
R_p	0.0826
R_{wp}	0.1116

Table 4-27: Selected data for B2

intensities. Since either recycling technique E or F had proved to be the most efficient in all cases studied, only these two were applied. Fig. 4-49 shows the histograms accumulated in the eight runs, and also the histograms E and F of the previous runs with space group P*nma*. The histograms are given on two scales. The first two columns are on absolute scales, subdivided in two groups, one for each extinction symbol. The histograms in the two right columns show the fraction (in percent) of a unique topology with respect to the total number of topologies found in a run.

With space group P$n2_1a$, the results for space group P*nma* are essentially repeated. The most frequently occurring topology is model B, and model A is represented by the fourth histogram bar with both recycling techniques. However, no convincing discrimination between the first histogram bars is observed. The same is true for the histograms for space groups P*mma* and P*m*2*a*. Interestingly enough, model A is also found in three of the four runs, but with negligible frequency. The only reasonable discrimination in the space groups with extinction symbol P--*a* is observed for space group P2_1ma – although not a really striking one on either scale, absolute or relative. Closer investigation revealed that the most frequently occurring model is a new type, referred to as “B2 model C” in the following. In the second place is model B, and model A occurs in the fourth or fifth place. To get a better overview, the standard test series of six runs was completed for space group P2_1ma by setting up the runs with pure atom and pure framework fragment recycling. As can be seen in Fig. 4-50, the results obtained with techniques E and F are essentially repeated, but with pure atom recycling,

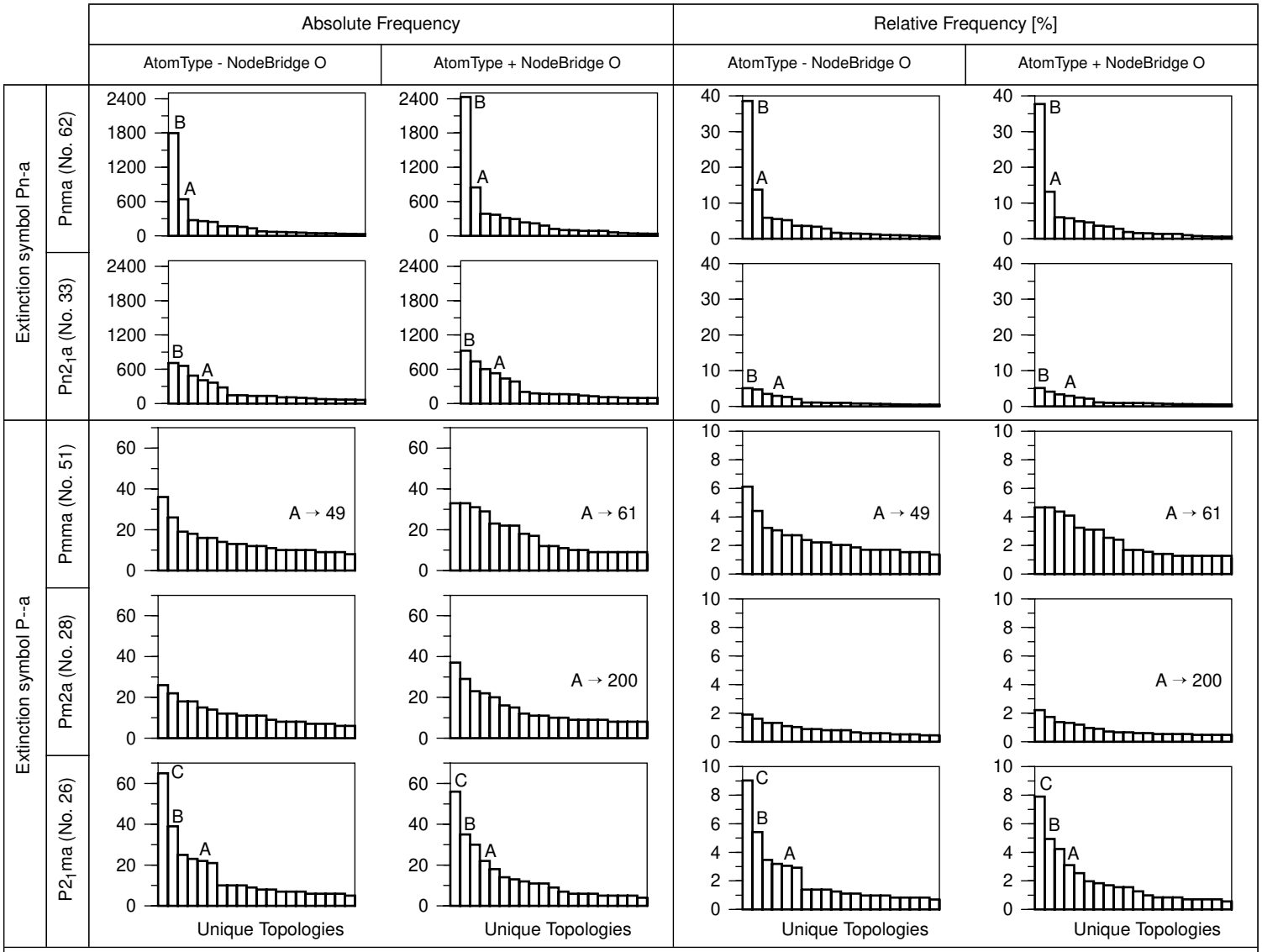


Figure 4-49: FOCUS Results for B2 in five space groups

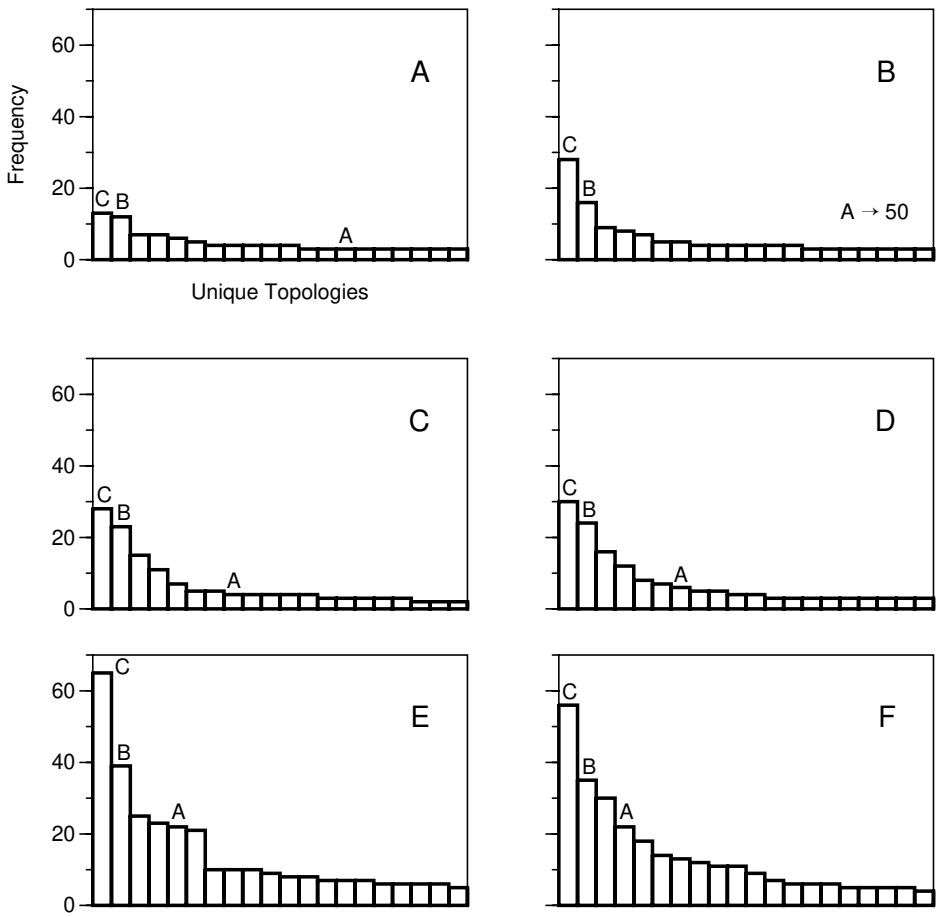


Figure 4-50: FOCUS Results for B2 – Space Group $P2_1ma$ – 500 trials per run

Run	NBO	MPA	MRA	FBC	FT	Fw	UFw	RFw	t/min	%tFwS
A	-	24	20	6	1850	290	178	95621	[o]48	83
B	+	80	80	6	2270	343	194	105084	[o]53	81
C	-	24	20	0 6	3262	286	142	32363	[g]1456	99
D	+	80	80	0 6	3263	296	138	32508	[b]1728	99
E	-	24	20	1 1 1 1 1 1 1 1 1 1	4966	720	296	156076	[o]1266	98
F	+	80	80	1 1 1 1 1 1 1 1 1 1	5500	709	311	164851	[o]1097	98

Table 4-28: FOCUS Results for B2 – Space Group $P2_1ma$ – 500 trials per run (152 actively used reflections)

[o] time normalized by dividing by 0.7
[g] time normalized by dividing by 1.1
[b] time normalized by dividing by 0.7

model A occurs much less frequently. The statistics of the test series are shown in Tab. 4-28.

The first test for the validity of model C was a DLS run with all nodes occupied by silicon, and oxygen automatically inserted with KRIBER. At first, the model only converged to a DLS residual value of 0.0095. But when the DLS run was repeated with the “random starting coordinates” option set (but with the same connectivity), the model finally converged to a

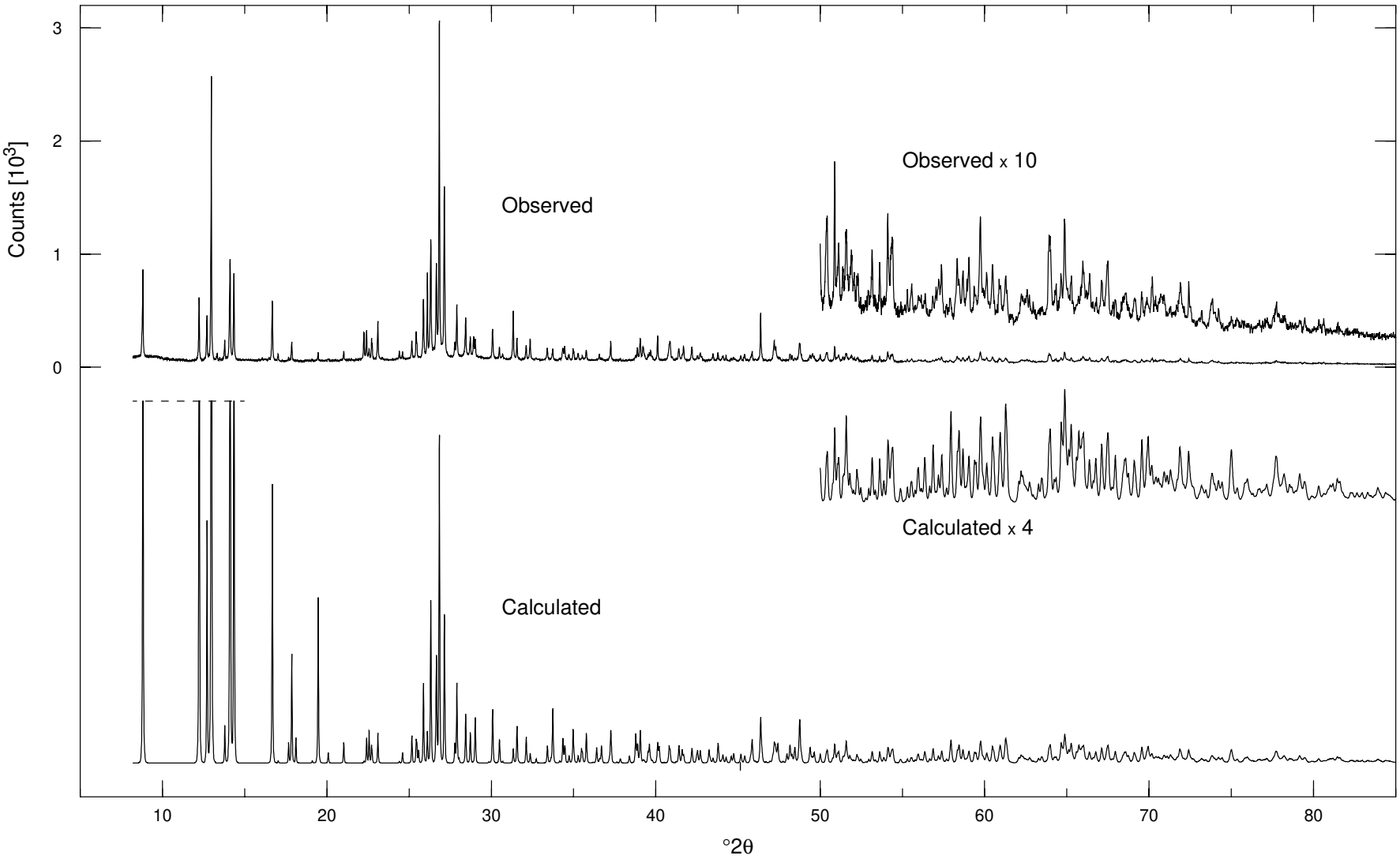


Figure 4-51: Qualitative comparison of observed and calculated (B2 model C, DLS coordinates) profiles

residual value of 0.0069, indicating that the topology as picked from the FOCUS output was heavily distorted. Fig. 4-51 shows a qualitative comparison of the observed profile and the profile calculated with the pure silicon DLS model. The low angle reflections, which are heavily affected by the missing potassium and sodium cations and the missing water molecules, are much too high and were cut at the dashed line. The scaling factor for the calculated profile was chosen such that the reflections in the range from twenty to forty degrees are best approximated. For the high angle region, another scaling factor is necessary. This is also due to the incomplete structure, and because of the more or less arbitrary choice of temperature factors.

The two profiles were considered to have enough similarities to encourage the initialization of a Rietveld refinement, which was done by Ch. Baerlocher. Indeed, the residuals obtained after the first steps are significantly better than ever obtained with models A or B. The refinement is still in progress, and the results have to be awaited. Of course, with powder data, a successful Rietveld refinement is still the only way to get a definitive answer as to whether or not a structure has been successfully determined.

4.9.2 Review of the three models suggested for B2

Looking at the histograms of Fig. 4-49 the question arises, why model A occurs under almost all circumstances, and why model B is found very often together with model C in space group $P2_1ma$. Surprisingly the answer was easily found: the three models are very similar. With the large tolerances prescribed for the node-node distances in the FOCUS input files (±0.5 Å), slight shifts of some node positions are sufficient to convert any model to one of the other two.

Tables 4-29 - 4-31 show the DLS refined coordinates of the three models. For the sake of completeness, the coordinates of model A in its space group of highest topological symmetry are also listed. Figures 4-52 - 4-54 show the wire-frame plots of the three models, and Fig. 4-55 the relationship between the three models. The wire-frame plot (which includes oxygen) shows a part of the DLS refined model C, which can be seen as the “compromise model”. Through the operation indicated by “C→A” model C is transformed to model A. The operation involves breaking of two bonds (indicated by the scissors) and forming two new ones (shown by the dashed lines). Focusing on the node atoms only, it can be seen that only a small shift of the center of the “spiro-5” unit to the right is needed for the transformation.

The term “spiro-5” unit is defined in the Atlas of Zeolite Structure Types [4] as a “secondary building unit” of five node atoms. This unit can also be described as two 3-rings

Space Group	P 63/m m c (No. 194) [topological symmetry]		
Unit cell	a = 7.273 b = 13.987 Å [refined with DLS]		
DLS residual	0.0064		
SI1	0.00000	0.00000	0.38389
SI2	0.19639	0.39279	0.25000
O1	0.75654	0.87827	0.34492
O2	0.00000	0.00000	0.50000
O3	0.45387	0.54613	0.25000

Space Group	P n m a (No. 62)		
Unit cell	a = 13.1731 b = 7.1256 c = 12.6777 Å		
DLS residual	0.0071		
SI1	0.21619	0.54110	0.35570
SI2	0.10968	0.25000	0.21958
SI3	0.20760	0.75000	0.55756
SI4	0.12354	0.75000	0.78020
O1	0.19742	0.56572	0.48190
O2	0.32271	0.43548	0.32986
O3	0.22438	0.75000	0.30403
O4	0.12478	0.43412	0.29396
O5	0.99300	0.25000	0.17726
O6	0.18220	0.25000	0.11572
O7	0.12351	0.75000	0.65178

Table 4-29: DLS coordinates of B2 model A (G.O. Brunner)

Space Group	P n m a (No. 62) [topological symmetry]		
Unit cell	a = 13.1731 b = 7.1256 c = 12.6777 Å		
DLS residual	0.0066		
SI1	0.21364	0.54118	0.11755
SI2	0.14560	0.75000	0.92935
SI3	0.14123	0.75000	0.69435
SI4	0.05225	0.25000	0.18462
O1	0.12424	0.43534	0.18292
O2	0.18190	0.56434	0.99411
O3	0.22660	0.75000	0.16815
O4	0.32228	0.43473	0.12912
O5	0.19348	0.75000	0.81087
O6	0.02239	0.75000	0.91786
O7	0.01850	0.75000	0.70997

Table 4-30: DLS coordinates of B2 model B

sharing one node atom, rotated 90° with respect to one another. The operation “C→A” transforms a spiro-5 unit to another unit, which is arbitrarily called “spiro-7” unit in the following. As can be seen from the inset in Fig. 4-55, a spiro-7 unit can be viewed as one 3-ring and one 4-ring, but this time with no shared nodes. Model C contains both types of spiro units. For example, the unit to the left of “C→B” in Fig. 4-55 is a spiro-7 unit. Now, since the

Space Group	$P\ 2_1\ m\ a$ (No. 26) [topological symmetry]	
Unit cell	a = 13.1731 b = 7.1256 c = 12.6777 Å	
DLS residual	0.0069	
SI1	0.12500	0.20727 0.12381
SI2	0.47797	0.50000 0.20204
SI3	0.28838	0.50000 0.07456
SI4	0.03743	0.29061 0.60521
SI5	0.13074	0.00000 0.32200
SI6	0.20474	0.00000 0.54989
SI7	0.43386	0.00000 0.53594
SI8	0.98708	0.50000 0.03348
O1	0.01861	0.31500 0.10097
O2	0.14687	0.18619 0.24951
O3	0.11559	0.00000 0.07070
O4	0.21864	0.31297 0.06554
O5	0.53178	0.50000 0.08629
O6	0.50761	0.68528 0.27083
O7	0.35570	0.50000 0.18222
O8	0.36359	0.50000 0.97266
O9	0.14638	0.18539 0.59360
O10	0.05007	0.50000 0.55513
O11	0.94922	0.18432 0.53821
O12	1.01541	0.00000 0.36771
O13	0.20862	0.00000 0.42154
O14	0.32196	0.00000 0.59053

Table 4-31: DLS coordinates of B2 model C

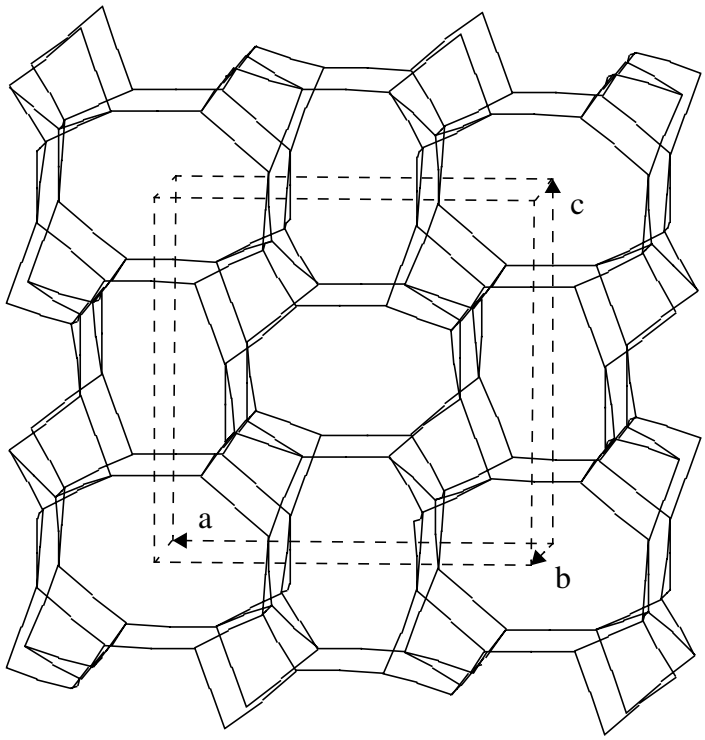


Figure 4-52: The topology of B2 model A
(O atoms have been omitted for clarity)

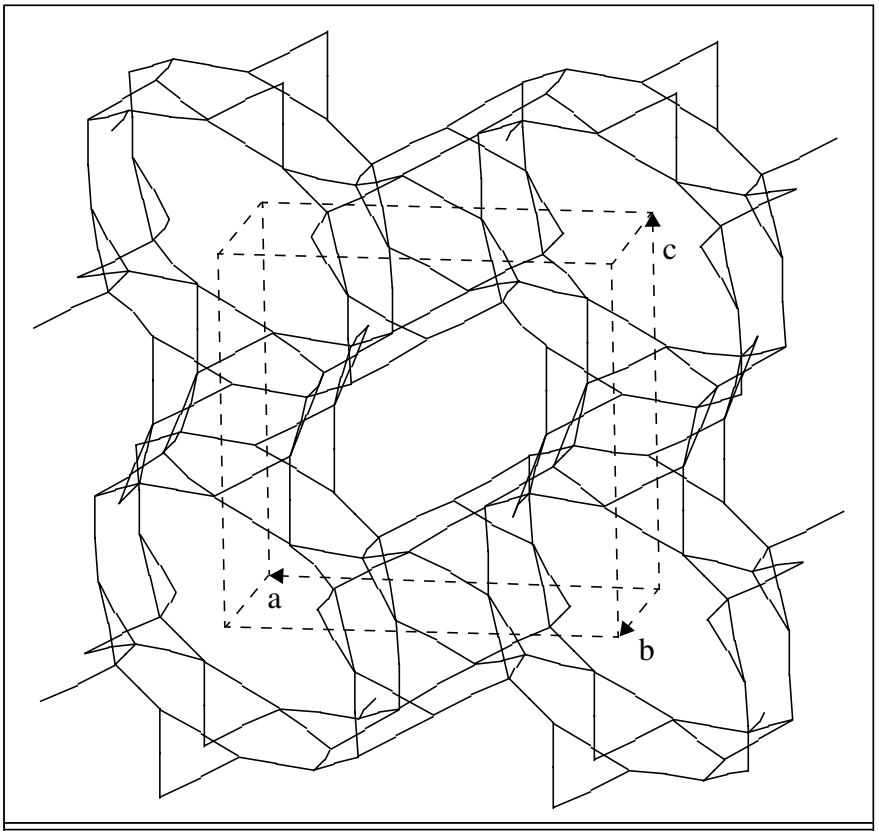


Figure 4-53: The topology of B2 model B
(O atoms have been omitted for clarity)

effect of the operation "C→A" is to transform *all* (due to symmetry, as indicated by the light dashed lines) spiro-5 units of model C to spiro-7 units, model A can be described completely with one spiro-7 unit and the space group. In contrast, the effect of operation "C→B" is to transform *all* spiro-7 units of model C to spiro-5 units. Model B contains only spiro-5 units. Focusing on the node atoms, the operation involves again only a small shift of one node, the center node of the spiro-7 unit.

With this background, the histograms of Fig. 4-49 can be reinterpreted. Model A occurs in all five space groups, because it is closely related to the (presumably) correct solution, model C, and because all five space groups are subgroups of $P6_3/mmc$, which is the topological symmetry of model A. Model B is the most frequently occurring model in space groups P*nma* and $Pn2_1a$, also due to its close relation to model C. Model C occurs only in space group $P2_1ma$, because this is the highest topological symmetry of model C, and none of the other four space groups is a subgroup of $P2_1ma$. Finally, the discrimination of the most frequently occurring topologies in space group $P2_1ma$ is not very pronounced, because both model B and model A are generated from model C by a slight shift of a single node position. Generating the histograms, with models A, B, and C counted as only one unique topology, would results in a very good discrimination.

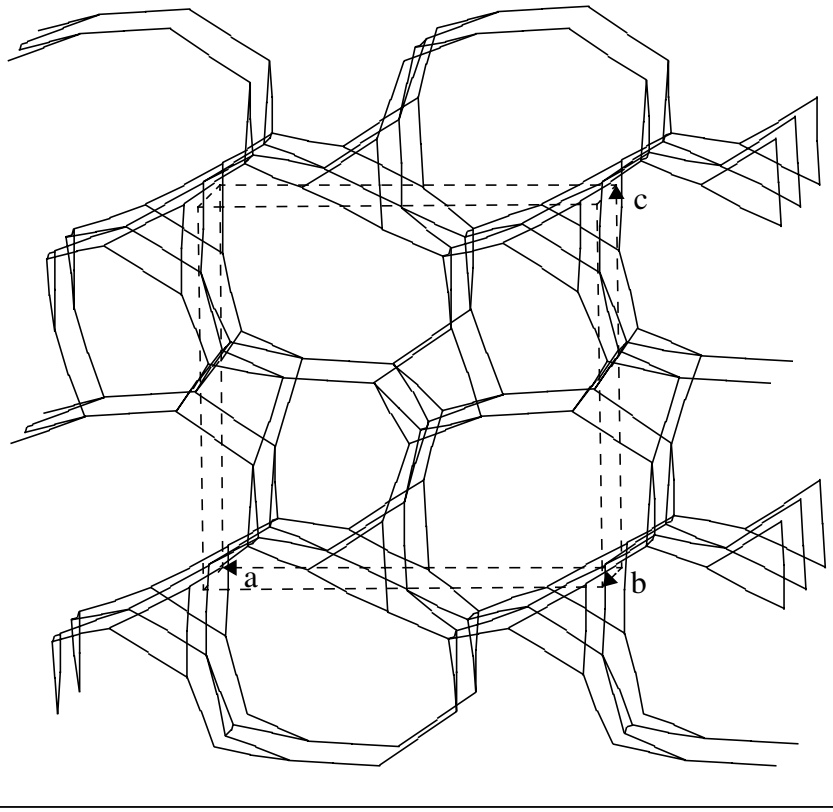


Figure 4-54: The topology of B2 model C
(O atoms have been omitted for clarity)

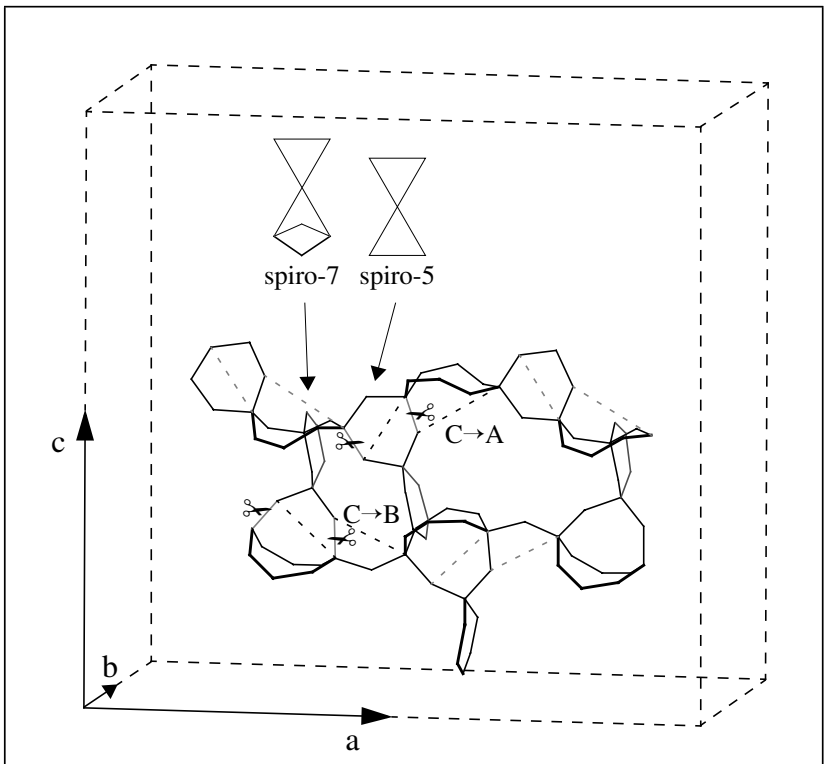


Figure 4-55: B2 – Relation of model C to models A and B

5 Further aspects

5.1 Review VPI-9 anomalous scattering

As mentioned on page 5, three synchrotron data sets were collected for VPI-9, with the intention of exploiting anomalous dispersion effects for structure determination. Now that the structure is solved and refined, the prospect of solving VPI-9 following the suggestions of W. Prandl [25] can be evaluated. For this purpose, the three correlation plots in Fig. 5-1 were prepared. In all plots, the horizontal axis shows the structure amplitudes F_{cal} down to a d-spacing of 1.3 Å, as computed (XTAL FC module) with the refined coordinates of L.B. McCusker (see page 76), and with anomalous dispersion corrections for a wavelength of $\lambda = 1.0198$ Å. The structure amplitudes ordered along the vertical axis in the first plot (top of Fig. 5-1) were calculated with anomalous dispersion corrections for $\lambda = 0.8164$ Å, and for the third plot (bottom) with $\lambda = 1.2831$ Å. (Tab. 5-1 shows the applied anomalous absorption corrections, which were determined with the GSAS FPRIME module.) In the middle plot, the structure amplitudes for $\lambda = 1.0198$ Å were equipartitioned with FOCUS method `EqualMF2`. `OverlapFactor` was set to 0.15, and FWHM values were generated with the function $FWHM = U + V \cdot \tan(\theta)$, where θ is the diffraction angle. U and V were chosen such that the FWHM is 0.05 at 3° 2θ, and 0.1 at 45° 2θ. The resulting FWHM simulate a very good synchrotron measurement, and are slightly better than the refined FWHM. From the plots it can be seen immediately that the effect of the equipartitioning alone – ignoring all other systematic errors – introduces deviations which are at least as severe as the effects due to anomalous disperson, even in comparison with the Zn-edge plot, with an f' of -11.5 electrons for zinc. In practice, the "noise" overlaying the anomalous dispersion effects is certainly a lot worse. One major source of error is the difficulty in establishing reliable scaling factors for combining the profiles. Further sources are the errors in the estimation of the background intensity, peakshape misfits, and inexact absorption corrections. Therefore, it seems very unlikely that anomalous dispersion effects could be used to determine the structure of VPI-9, or other structures of similar or higher complexity.

In the introduction of [66] W. Prandl wrote: "Because of the high resolution of the X-ray diffractometers that are available now at many synchrotron sources, powder diffraction has become nearly equivalent to single-crystal methods: …". This statement is not true for any of the synchrotron measurements presented in this work. Furthermore, if this statement were true, there would be no point in doing two or three measurements at different wavelengths (given, there *are* anomalously scattering atoms). Direct methods, especially with the new tangent

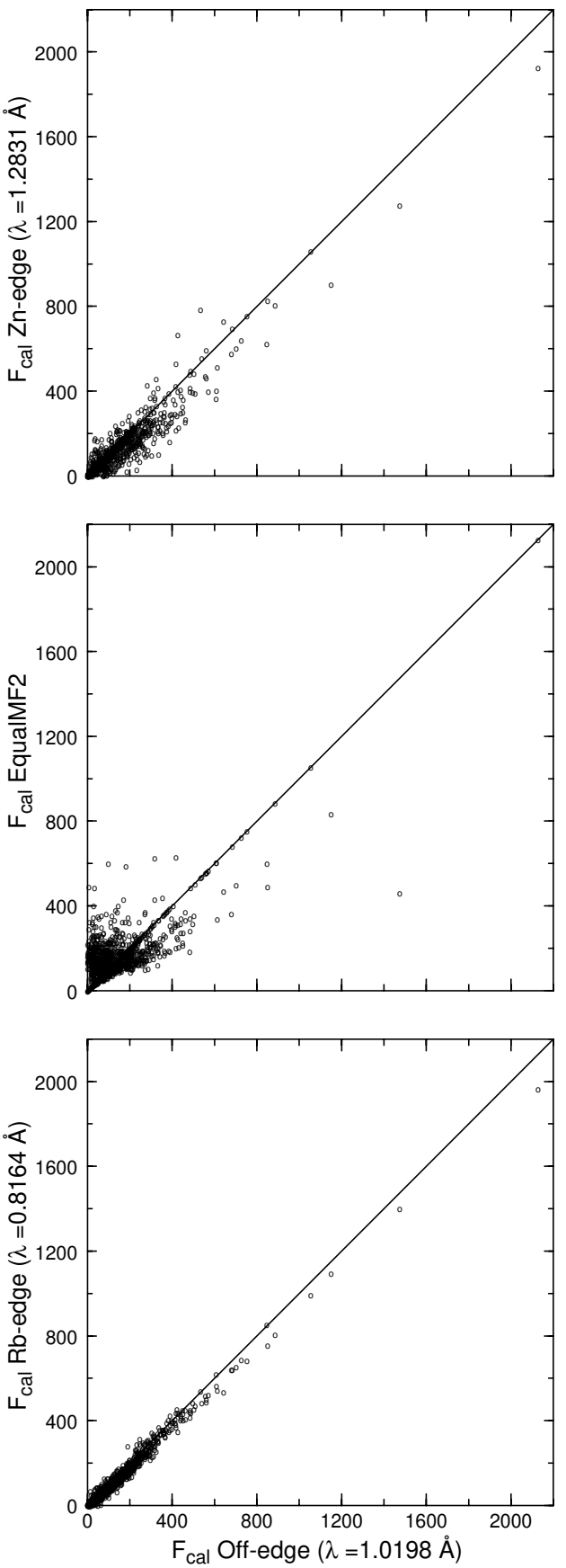


Figure 5-1: Calculated anomalous dispersion effects compared to equipartitioning effect for VPI-9

	Rb-edge (λ = 0.8164)		Off-edge (λ = 1.0198)		Zn-edge (λ = 1.2831)	
Element	f'	f"	f'	f"	f'	f"
Si	0.094	0.094	0.137	0.147	0.193	0.231
Zn	0.100	1.823	-0.470	2.647	-11.519	0.486
Rb	-6.574	0.509	-1.436	0.765	-0.877	1.160
K	0.220	0.328	0.288	0.500	0.348	0.766
O	0.012	0.008	0.021	0.013	0.033	0.022

Table 5-1: GSAS FPRIME anomalous dispersion corrections

formula presented by Rius et al. [27] are about as likely to reveal a structure, as is the experimentally cumbersome procedure of Prandl.

5.2 Searching for non-tetrahedral node connectivities

All the examples presented in the previous section have one feature in common: a 3-dimensional 4-connected network of nodes was sought. To a certain degree, this is a consequence of the main idea which inspired the design of FOCUS: the integration of structural knowledge into the solution process. However, as FOCUS has been described, the specialization is extreme. To demonstrate the consequences of relaxing the structural assumptions, two further examples can be given.

5.2.1 Searching for interrupted frameworks

A FOCUS input file was composed to determine the structure of Roggianite (**-RON**) [67, 68]. The dash preceding the structure type code is used for “interrupted frameworks” (frameworks that are not fully 4-connected, but have one or more nodes in the asymmetric unit which are connected to only three neighboring nodes). These three connected node atoms are still tetrahedrally coordinated, but one of the four bonded oxygens is a terminal hydroxyl group pointing into a cage or channel rather than a node-bridging oxygen.

The relevant data for this test can be taken from the FOCUS input in Fig. 5-2. The intensities used were calculated from the coordinates published in [67]. FWHM values were generated with the function FWHM = U + V · tan(θ), where θ is the diffraction angle. U and V were chosen such that the FWHM is 0.1 at 6° 2θ, and 0.2 at 70° 2θ and the intensities equipartitioned. The main difference with respect to the other input examples is the `NodeType 3 1 …` line, which allows for frameworks with one 3-connected node in the asymmetric unit.

100 trials were calculated in about 20 minutes. The histogram in Fig. 5-3 shows a quite clear discrimination of the most frequently occurring topologies, and the first histogram bar

```
Title Code -RON
SpaceGroup  I 4/m c m
UnitCell  18.33 9.16

AtomType  +  Node        Si  32
AtomType  +  Node        Al  16
AtomType  +  NodeBridge  O   88
AtomType  -  *           O   32
AtomType  -  *           Be   8
AtomType  -  *           H   16

Chemistry  MinDistance  Node        Si  Node        Si  2.6
Chemistry  MinDistance  Node        Si  Node        Al  2.6
Chemistry  MinDistance  Node        Si  NodeBridge  O   1.4
Chemistry  MinDistance  Node        Al  Node        Al  2.6
Chemistry  MinDistance  Node        Al  NodeBridge  O   1.4
Chemistry  MinDistance  NodeBridge  O   NodeBridge  O   2.3
MaxPotentialAtoms  136
MaxRecycledAtoms  136

FwSearchMethod  FwTracking
MaxPeaksFwSearch  320
MaxPeaksFwFragmentSearch  640
MinNodeDistance  2.6
MaxNodeDistance  3.6
MinSymNodes  0
MaxSymNodes  64
NodeType  3  1  -6 -3 -1 4 6
NodeType  4  *  -6 -3 -1 4 6
MinLoopSize  3
MaxLoopSize  24
EvenLoopSizesOnly  Off
Check3DimConnectivity  On
IdealT_NodeDistance  3.1
CheckTetrahedralGeometry  Normal

RandomInitialization  Time
FeedBackCycles  1 1 1 1 1 1 1 1 1 1
FeedBackBreakIf  PhaseDiff < 5.00 % and DeltaR < 1.00 %

Grid_xyz  52  52  28
eDensityCutOff  1 %
MinPfI  17
CatchDistance  0.5
eD_PeaksSortElement  Grid_eD

Lambda  CuA1
FobsMin_d  1.3
FobsScale  1
SigmaCutOff  0
OverlapFactor  0.15
OverlapAction  EqualMF2
ReflectionUsage  75 %
Grid_hkl  +15  30  14
GenerateFWHM  6 0.1  70 0.2
```

Figure 5-2: FOCUS input for Roggianite

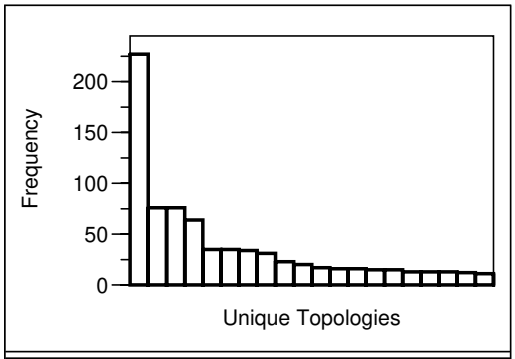


Figure 5-3: FOCUS histogram for **-RON**

represents the **-RON** topology. No more tests have been carried out, but, based on this example, it can be assumed that in general a search for an interrupted framework is as likely to give a solution, as is a search for a fully 4-connected framework of the same complexity. In this simple example, the time spent for the framework search increased by a factor of 2.4, compared with that required for the search for a fully 4-connected net under similar conditions.

5.2.2 Searching for 3, 4, and 6-fold connectivities

For testing purposes, FOCUS offers the possibility of bypassing the Fourier transform and peak search procedures, and adding arbitrary peak positions to the (refined) peaklist. The backtracking procedure then works with the externally supplied positions. A simple test was made with a gallophosphate structure solved from single crystal data by A. Chippindale et al. [69]. Fig. 5-4 shows the corresponding FOCUS input, which now also contains the coordinates for this gallophosphate. The three `NodeType` lines reflect the different node connectivities that are present in the structure. To obtain a reference point for the search times, two of the three `NodeType` lines were deleted, in order to restrict the search to fully 4-connected frameworks. Then a test series was generated, with increasing numbers of positions on the peaklist, starting with the nine node atoms only, and ending up with the whole structure (56 positions). The plot of the computing times is shown by the circles in Fig. 5-5. Next, the two deleted `NodeType` lines were reintroduced, and the test series repeated. The resulting computing times are marked by the triangles in Fig. 5-4. It turned out that in both test series the time required for the topology search increases approximately exponentially. To show the similarity of the two curves, the times of the first series are plotted again, multiplied by a factor of 29. This means that, as a consequence of allowing connectivities other than four, the time to search a peaklist of the same size increases by a factor of nearly 30. Since the run times for complex structures are currently better measured in days rather than in hours, this factor increases the computing time from one day to one month.

```
Title [Me2NH(CH2)2NHMe2][Ga4P5O20H].H2O (A.M. Chippindale)
SpaceGroup  P 21 21 21
UnitCell  9.574 14.000 17.435
MaxNodeDistance  3.8
NodeType  3  2  -6 -3 -1 4 6
NodeType  4  *  -6 -3 -1 4 6
NodeType  6  1
CheckTetrahedralGeometry  Normal
GA1       0.22463   0.29395   0.69846
GA2       0.45118   0.50175   0.45624
GA3       0.24554   0.20719   0.47185
GA4       0.46107   0.51150   0.92710
P1        0.25470   0.34939   0.87405
P2        0.22190   0.42018   0.55107
P3        0.48320   0.21287   0.61963
P4        0.76340   0.52288   0.51133
P5        0.00220   0.18191   0.60985
O1        0.19030   0.33900   0.79420
O2        0.40120   0.24630   0.69270
O3        0.08960   0.20400   0.68300
O4        0.19610   0.39700   0.63590
O5        0.36700   0.46730   0.54660
O6        0.13740   0.38730   0.92600
O7        0.61840   0.55300   0.48300
O8        0.93370   0.08190   0.62430
O9        0.20890   0.33850   0.49760
O10       0.60290   0.28250   0.60940
O11       0.38600   0.20270   0.55270
O12       0.26900   0.07030   0.44590
O13       0.89250   0.25940   0.60210
O14       0.09500   0.16550   0.54100
O15       0.37490   0.42100   0.86790
O16       0.11130   0.49310   0.52440
O17       0.54110   0.11220   0.63880
O18       0.64300   0.47550   0.93990
O19       0.30780   0.25640   0.90480
O20       0.81210   0.59970   0.56850
O99       0.08550   0.26820   0.06360
N1        0.79060   0.56060   0.21740
N2        0.70920   0.78170   0.33390
C1        0.81440   0.63920   0.27400
C2        0.68290   0.69140   0.28880
C3        0.92630   0.50990   0.20190
C4        0.68500   0.49250   0.24250
C5        0.58360   0.80750   0.37790
C6        0.76000   0.86070   0.28590
H19       0.75960   0.59070   0.16840
H29       0.78640   0.76870   0.37140
H11       0.88570   0.68500   0.25350
H12       0.84880   0.61160   0.32350
H21       0.63780   0.70950   0.23840
H22       0.61580   0.64970   0.31790
H31       0.90400   0.45660   0.16370
H32       0.99390   0.55360   0.18040
H33       0.95880   0.48000   0.25050
H51       0.66960   0.44360   0.20160
H52       0.70610   0.46500   0.29140
H53       0.59060   0.53050   0.24670
H71       0.60450   0.86630   0.40830
H72       0.56110   0.75370   0.41400
H73       0.50510   0.81850   0.34240
H91       0.77530   0.91820   0.32090
H92       0.84870   0.84300   0.26160
H93       0.68720   0.87730   0.24760
End
```

Figure 5-4: FOCUS input for the test with the gallophosphate

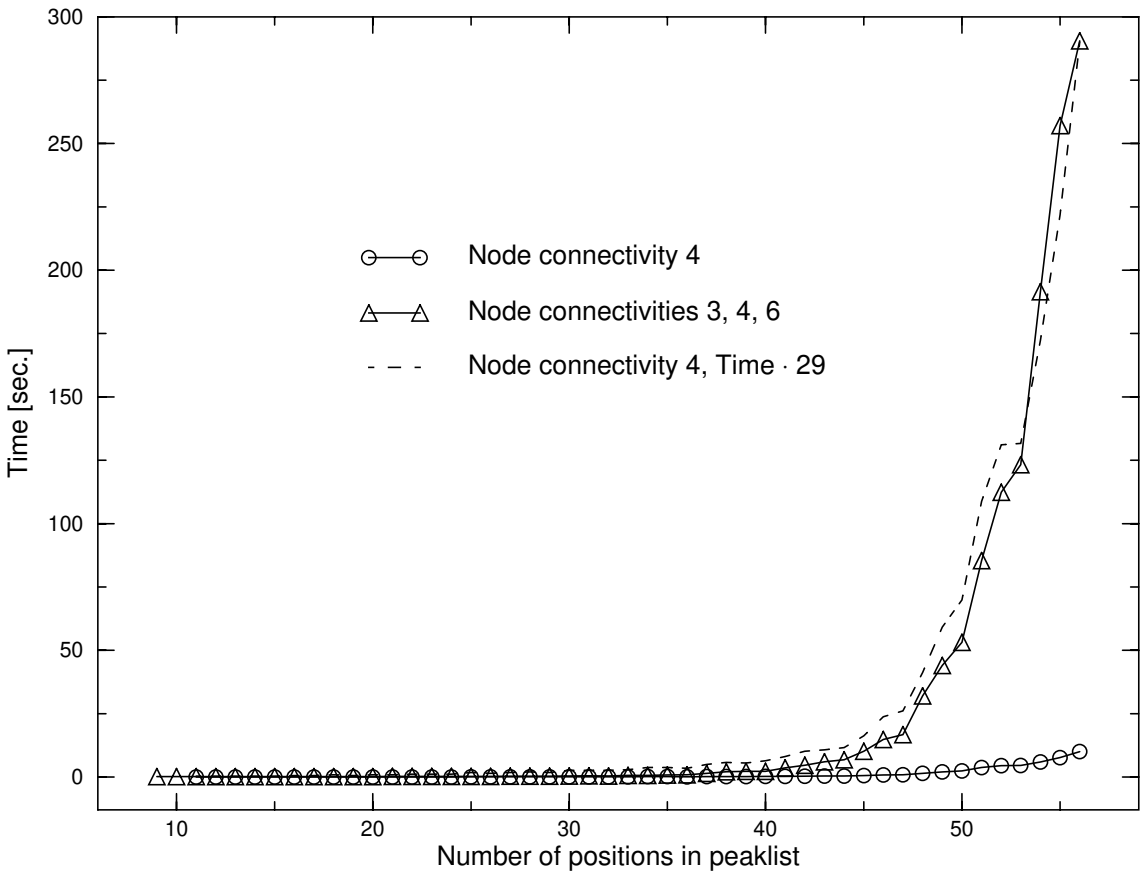


Figure 5-5: Time behavior of the backtracking algorithm with growing peaklist

One way of overcoming such overwhelming time requirements is, of course, to work with a smaller peaklist. To get an estimate of the possibilities, the test with the gallophosphate was extended. First, the calculated (but not equipartitioned) Fourier magnitudes were used (as in the previous example of Roggianite). `MaxPeaksFwSearch` and `MaxPeaksFwFragment-Search` were set to 100 and 60, respectively. Since the only Wyckoff position in the space group $P2_12_12_1$ has a multiplicity four, there are at most 25 positions for the framework search, and just 15 positions for the framework fragment search to work on. The test was terminated after 989 trials, which took about 25 hours of computing time. The time spent for the framework and framework fragment search was 94% of the total time. With the ideal intensities, the correct topology was generated 80 times, and only five other topologies occurred (once each). Then, the overlapping intensities were equipartitioned, but everything else kept unchanged. With these intensities, 1943 trials were computed in about 43 hours, but only three topologies were found at all, and the correct solution occurred only once. Of course, the cause of the poor success rate can be sought in the treatment of overlapping reflections. On the other hand, it is probable that larger values for `MaxPeaksFwSearch` and `MaxPeaksFw-FragmentSearch` would increase the success rate considerably. Unfortunately, the price to be

paid for this way of recovering from poor intensities is a profound increase in computing time.

Another weak point that should not be forgotten is that the possible node connectivities have to be prescribed *before* the structure is solved. In the gallophosphate example, it would be sensible to expect more than two three-fold node connectivities, since phosphorus was also observed to be bonded to terminal hydroxyl groups in related structures [e.g. cloverite (**-CLO**)]. In addition, it is not easy to decide beforehand, how many gallium atoms are six- or maybe five-connected to other nodes. Allowing for these uncertainties requires an even further increase in computing time. To solve a structure like this gallophosphate from powder data using FOCUS would certainly require tremendous computing capabilities (at the time this is written) and effort. However, the "massive parallel" computers, equipped with several thousand processors, that are currently emerging would be very well suited for the algorithmic approach adopted by FOCUS, and might render attempts to determine structures of this complexity level successful.

5.3 Backtracking on a grid

Another technique, which was also employed in the case of B2 in the space group P*nma*, is to dispense with Fourier recycling completely, and to take only the very basic information from the powder diagram (i.e. unit cell parameters and space group). A grid is generated, and all grid points in the asymmetric unit are put into the peaklist, which is then passed to the topology search procedure. In theory, such a topology search should produce all topologies which are possible within the framework prescriptions. Of course, compromises have to be made in order to keep computing times "reasonable". Fig. 5-6 shows the FOCUS input file used for B2. Besides the `NodeType` line(s), the critical input parameters are `MinNodeDistance`, `MaxNodeDistance`, and `Grid_xyz`. The number of grid points determines directly the number of positions in the asymmetric unit. As shown before, the computing time grows approximately exponentially with the number of positions, so the grid used is quite coarse, with grid spacings of 0.659, 0.445, and 0.634 Å for the *x*, *y*, and *z* directions, respectively, and the tolerances prescribed with `MinNodeDistance` and `MaxNodeDistance` have to be chosen appropriately. However, not much effort has been spent in optimizing the balance of these five input parameters, so the values in Fig. 5-6 are no more than an educated guess.

FOCUS generated 901 grid points in the asymmetric unit. After the preparation of the list of potential node-node bonds, only 473 were left, because 428 positions were incompatible with the framework prescriptions. The number of seed nodes (see page 20) to process was further reduced to 238 by skipping entries which were connected to a previous seed node by an

```
Title B2 lattice constants
SpaceGroup  P n m a
UnitCell  13.1731 7.1256 12.6777

FwSearchMethod  FwTracking
MinNodeDistance  2.5
MaxNodeDistance  3.7
MinSymNodes  0
MaxSymNodes  *
NodeType  4  *  -6 -3 -1 4 6
MinLoopSize  3
MaxLoopSize  24
EvenLoopSizesOnly  Off
Check3DimConnectivity  On
IdealT_NodeDistance  3.1
CheckTetrahedralGeometry  Normal

Grid_xyz  20  16  20

End
```

Figure 5-6: FOCUS input file for B2 grid search

allowed origin shift. The actual backtracking runs were distributed on several machines of different kinds. At that time, FOCUS did not produce time protocols, so the exact normalized computing time could not be calculated. With the available machines, the computation was finished after about one week, and the rough estimate for a single processor yields times between four and six weeks. A total of 49296 topologies were produced, and an uncounted number of topologies were rejected by the geometry filter (estimated five million). After sorting, there were 364 unique topologies. Among these were “B2 model A” and “B2 model B” (see section 4.9), but also the topologies **ABW**, **ATS**, **BIK**, **CAN**, **JBW**, **PHI**, **TON**, the newly approved **AHT** (AlPO4-H2 [70, 71]), and the topology of the dense silica polymorph tridymite.

A comparison of the topologies produced by the grid search with those produced by Fourier recycling and topology search revealed that the grid search missed some, as might be expected. However, working with a finer grid is out of question with the computer technology currently available. Furthermore, to give only one example, doing the grid search for B2 in the space group $P2_1ma$, which is now assumed to be the correct one, would be virtually impossible, even with the grid sizes shown in Fig. 5-6, because the number of grid points would be approximately doubled (1800 in the asymmetric unit, 1000 compatible with framework prescriptions, 500 seed nodes). This means that the complexity limit of the grid search is probably reached before that of the combination of Fourier recycling and topology

search. In view of this it can be stated that the active use of the intensity data – regardless of the overlap problem – is still the best approach.

In the literature, various methods for a systematic derivation of zeolite topologies can be found: e.g. Smith's enumerations of 4-connected nets (e.g. [72, 73]), the Brunner approach [74], the Bennett-Schomaker method [8], Treacy's combinatorial method [75], the Akporiaye-Price formalism [76], and its extension by Shannon [77]. However, none of these approaches are generalized or fully automatic and all require problem specific adaptations or extensions (like Shannon's extension to the Akporiaye-Price formalism). Although the grid search as outlined here is no more than a by-product of the FOCUS method, it is – to the author's knowledge – the only truly automatic procedure which requires only very basic prescriptions. The need for involved tailoring of the input or algorithm to fit the problem is effectively eliminated by the introduction of the brute-force backtracking algorithm.

5.4 (Mis)using FOCUS for the generation of hypothetical topologies

Closely related to the previous section is the idea that FOCUS could be used to generate hypothetical topologies. For example, the 364 topologies found with the grid search for B2 were processed in two DLS refinement steps. First, of course, oxygens were inserted at the center of all node-node connections, and then the atomic coordinates were refined within the prescribed unit cell. In the second step, the node and oxygen coordinates obtained in the first step were refined together with the lattice constants. In the latter step, the refinement of 18 structures diverged completely, and these were discarded. Fig. 5-7 shows the histograms of the final DLS residuals obtained with and without refinement of lattice constants. After the first refinement step, just five topologies were found with a DLS residual below 0.01 (indicated by the dashed lines in Fig. 5-7, which can be viewed as an upper limit for structures likely to exist in reality. After the second step, the distribution of the residuals has become much better: now there are 37 structures with residuals below 0.005, and a total of 163 below 0.01. Not all of these structures are sensible. For example, sometimes two oxygens, which are not bonded to the same node, are too close to each other after refinement. In the DLS refinement, three types of distances are refined: node-oxygen bond distances, oxygen-oxygen distances for pairs of oxygens bonded to the same node atom (node-oxygen-node angles) and node-node distances for pairs of nodes bonded to the same oxygen (oxygen-node-oxygen angles). DLS is virtually blind to all other geometrical features (unless they are also prescribed). To discard structures with a low DLS residual, but unreasonable distances, a post-processor which reevaluates the geometry would be useful. Unfortunately, no such program was available, so the structures

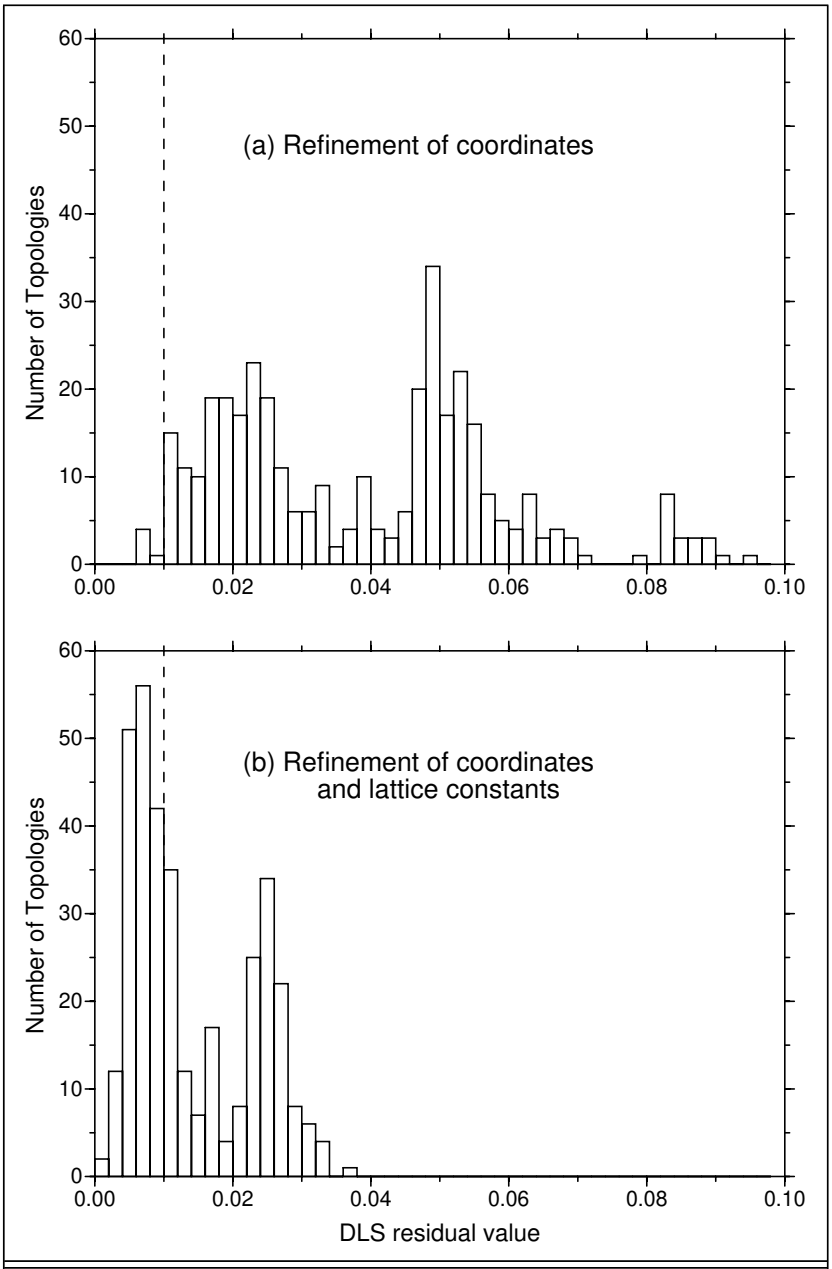


Figure 5-7: DLS residuals for topologies produced by FOCUS

were investigated by eye.

To get an estimate of the predictive potential of the DLS residuals, Tab. 5-2 lists the values obtained for the known topologies after the treatment just explained. Except for **ATS**, all residuals are below 0.01, as expected. Reinvestigation of the **ATS** topology revealed that the DLS residual drops to 0.0081 if the lattice constants are refined in the first step. An even better residual (0.0026) was obtained, when the oxygen positions refined in the first step were ignored, and new positions at the center of all node-node connections were calculated. This means that the high residual after the "standard" treatment reveals a weakness of the method for obtaining the residuals, but the residual itself seems to be very predictive, given that a

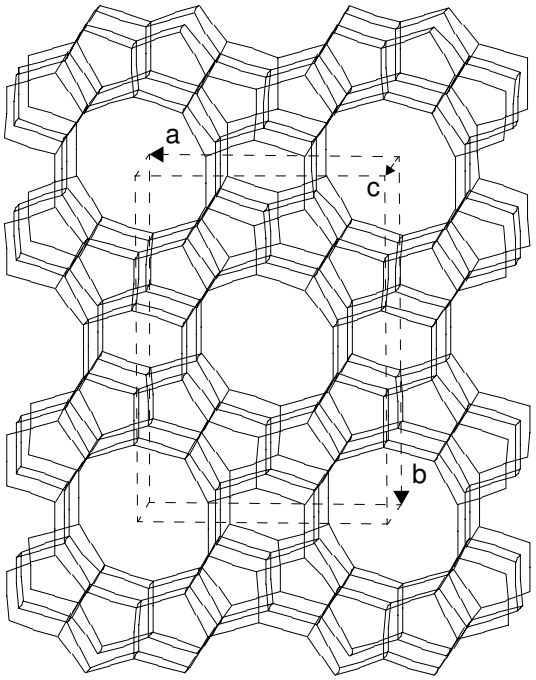
Topology	DLS residual
JBW	0.0039
PHI	0.0041
TON	0.0047
CAN	0.0049
Tridimyte	0.0049
ABW	0.0051
BIK	0.0054
AHT	0.0071
ATS	0.0167 (0.0026)

Table 5-2: DLS residuals for the known structures found with the B2 grid search (after refinement of lattice constants)

proper method is applied to get the lowest value possible. To be certain that no low-residual topology is missed, it seems to be necessary to try various DLS refinement strategies. However, no other strategies were developed in the course of this work.

A grid search similar to that for B2 was set up with the lattice constants of VPI-10 in the space group I2*mm* (but could not be completed in the time available). To give an interesting example, Fig. 5-8 shows the DLS coordinates, a wire-frame plot, and the three projections of a novel topology found in the course of this grid search with a DLS residual of 0.0050. This topology is interesting, because the projection along [001] is also a projection of the **MFI** topology, and the projection along [010] resembles a projection of the **AHT** topology [70, 71].

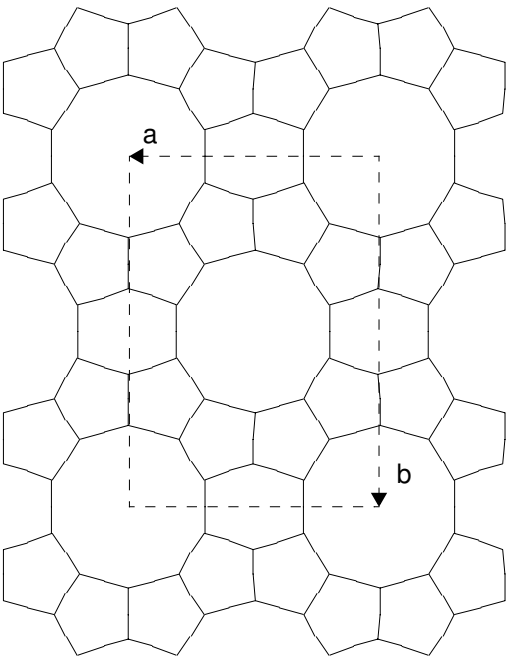
Another source of potentially low-residual topologies is found in the output of the normal FOCUS runs (Fourier recycling and topology search). In most cases, a huge number of "incorrect" topologies are produced along with the desired solution. These topologies can, of course, be processed in the same way as the ones produced by the grid search. For example, this was done for the output of one test run for EMC-2 (see section 4.6). Fig. 5-9 shows the DLS coordinates and the super-cage of a topology with a DLS residual of 0.0051. As in EMC-2, the super-cage is accessible via 12-rings, and is circumscribed by an 18-ring. However, in EMC-2, the super-cages are connected to form a system of straight 12-ring channels along [001], while the arrangement of the cages in the new topology generates a complex system of undulating channels. Therefore, different chemical and catalytical properties can be expected. The structure is beautiful, geometrically plausible, and if it could be synthesized, maybe also useful.

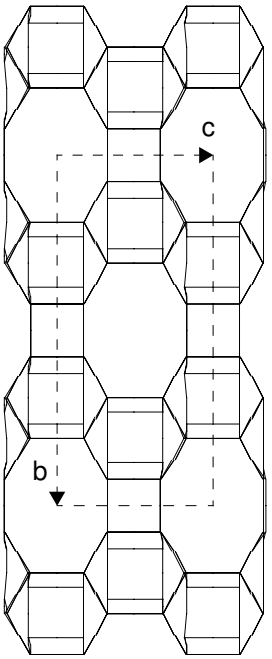


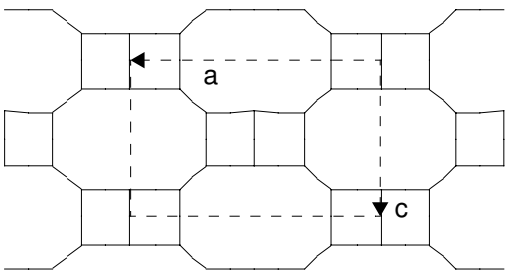
Topological space group: I m m m (71)
Unit cell: a = 14.78 b = 20.55 c = 9.16 Å
DLS refined node coordinates:

T1	0.19607	0.18941	0.67656
T2	0.50000	0.73197	0.17646
T3	0.50000	0.87685	0.17649
T4	0.30580	0.07575	0.83010

DLS residual in I m m m: 0.0058
I 2 m m: 0.0050







O atoms have been omitted for clarity

Figure 5-8: Example topology found with grid search (VPI-10 lattice constants, space group I2*mm*)

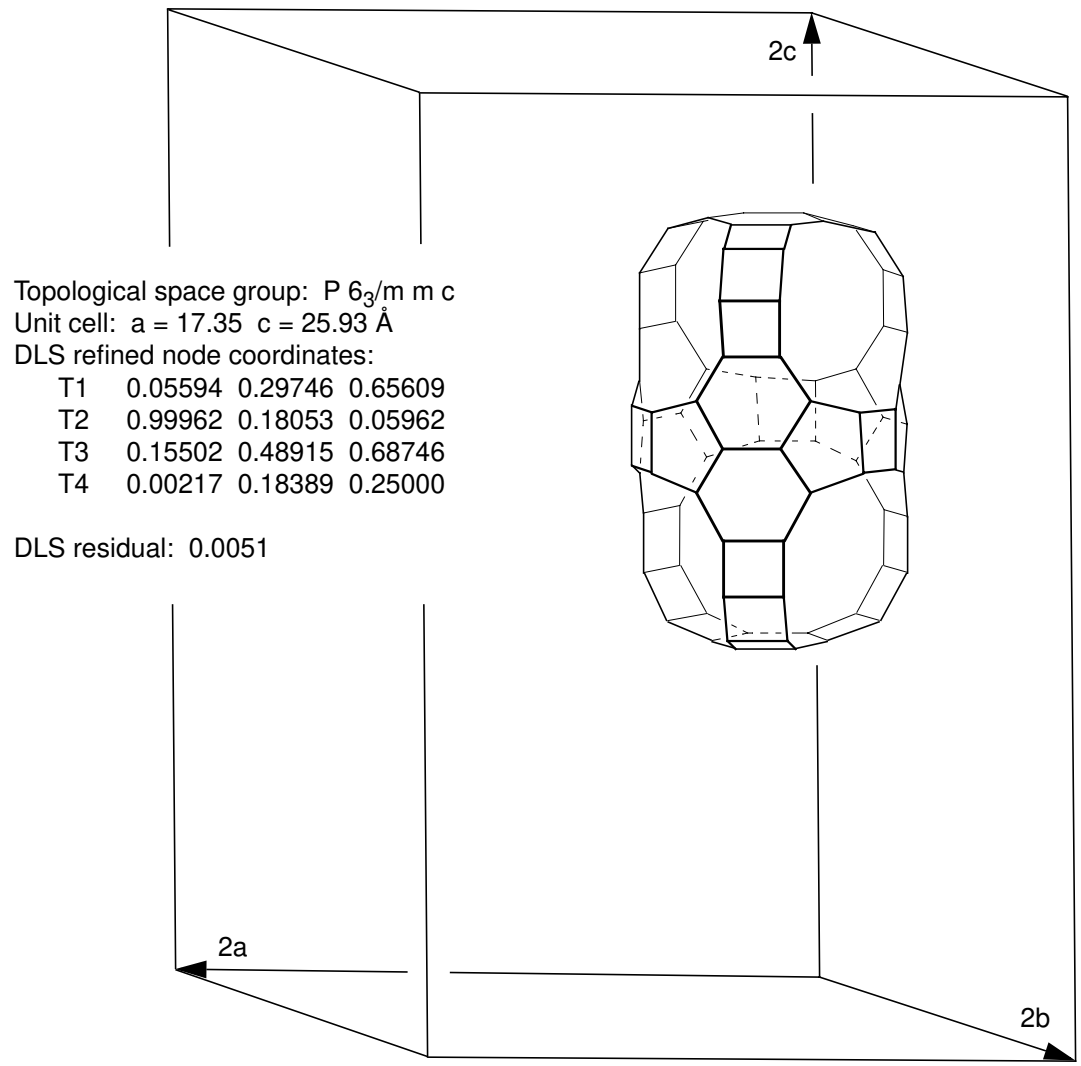


Figure 5-9: Super-cage of a topology found in EMC-2 test run (O atoms have been omitted for clarity)

5.5 Possible developments of FOCUS algorithm

The two short examples in section 5.2 have shown that the computing time requirements of FOCUS grow very rapidly with the number of different possible connectivity types. This means that for a more general structure solution algorithm, it is unlikely that an exhaustive search like the topology search of FOCUS is practicable. Maybe an approach for a more generalized solution mechanism could be derived from the way the presumably most advanced algorithmic game strategies (i.e. chess programs) work. An exhaustive search for the best chess moves – which means playing all games possible – will probably remain impossible forever. However, *searching only to a certain depth*, combined with carefully compiled databases for “standard situations” turns out to show a most impressive strength. To build an analogy between chess programs and structure determination, a “standard situation” could be

identified as a known structural fragment, for example a ring or a chain; the chess rules find their counterparts in the chemical and geometrical restrictions; and a "move" is the selection of a structural model for the automatic Fourier recycling. In view of this, FOCUS could be characterized as playing a very simple game, but it is hoped that some of the experience gathered in the development can contribute to a more generalized mechanism.

FOCUS represents a purely algorithmic, heavily computer-based method. A source of information which was left untapped (except for a few tentative tests not reported here), is the exploitation of the statistics of the Fourier magnitudes, which is the foundation of direct methods. In 1989, Rius et al. [27] derived a new tangent formula, and, in contrast to that used in conventional direct methods, this has recently been shown to be applicable to low resolution (with respect to d-spacings) data [78]. A combination of the "Fourier refinement" (recycling) of FOCUS and phase refinement with the new tangent formula offers tantalizing possibilities. An interesting aspect related to this is the fact that the proposed combination resembles recent developments in direct methods aiming at the determination of larger structures (e.g. "small proteins") from single crystal data. In the "Shake-and-Bake" procedure presented by De Titta et al. [79] and Weeks et al. [80], phase refinement ("shake") alternates with Fourier refinement "bake". Similarly, Sheldrick & Gould [81] have presented a procedure with alternating phase refinement and "peaklist optimization" (which they classify as "half baked" with reference to the Shake-and-Bake procedure). However, the powder specific difference between these procedures and the proposed combination of FOCUS and the Rius' tangent formula, is a stronger enforcement of a prescribed class of structures at the Fourier refinement stage, and a significantly weaker demand for high resolution at the phase refinement stage.

At present, FOCUS only recycles phases derived from the automatically constructed models. However, it would also be possible to derive a new partitioning of overlapping intensities from the models. As has been shown in the EMC-2 test case, intensities play a vital role for the success rate. Obviously, those models that are in best agreement with the intensities have the highest chance of reproducing themselves. Of course, the correct model has no more chance of being randomly created than any other model, but once parts of it are present in the electron density map, the automatic Fourier recycling is likely to enforce it, while incorrect models are more likely to disintegrate. It is an open and highly interesting question, whether repartitioning of overlapping intensities during the recycling process would help to enforce the correct model, or whether it is more likely to "dilute" the already fragile intensity information extractable from a powder pattern.

6 Conclusions

The aim of this project was to contribute to the further development of structure determination from powder data. An approach, which has often been discussed in the past, was adopted: well-known structural properties were integrated into the structure solution process. Structural information, such as the types and numbers of atoms present, the expected connectivity types, and interatomic distances and bond-angles has been exploited

As a member of a research group that specializes in zeolite structure analysis, the new FOCUS method was developed for the integration of zeolite specific information into the structure solution process. FOCUS makes extensive use of modern computer technology, and many substeps make use of well established techniques, such as the conversion of powder data to a pseudo single crystal data set. The conventional treatment of the pseudo single crystal data is replaced, or enhanced, by a combination of automatic Fourier recycling and a topology search. Finally, the usefulness of the FOCUS procedure has been demonstrated by its successful application in the structure determination of two complex novel zeolite structures, where only powder data were available. Furthermore, a structure proposal for a novel beryllosilicate zeolite has been found, but has not yet been confirmed by a Rietveld refinement.

Experience gathered during the course of this project shows that the methodologically attractive approach of using chemical and geometrical knowledge can compensate for some of the information lost as a result of the overlap problem. At the same time, there is an intrinsic disadvantage: any method based on assumptions of certain structural properties is also limited to materials which conform to these assumptions. Unlike direct methods, which only make assumptions valid for all X-ray diffraction experiments, the consideration of more specific structural information also introduces a certain specialization. However, from the outset it has been foreseen that the basic idea – the integration of structural assumptions into the solution process – should also be applicable to other classes of materials. Two short examples have been presented (section 5.2), which show the consequences of relaxing the structural assumptions to allow solution attempts for non-4-connected frameworks. It was found that the computing time requirements of FOCUS grow very rapidly with the number of different possible connectivity types. Suggestions for further developments to overcome this problem were outlined in section 5.5, and it is hoped that some of the experience gathered in the development of FOCUS contributes to the evolution of a more generalized mechanism.

7 Appendix

7.1 *F*-weighted phase changes

Let $F_{\vec{H}}$ be the Fourier magnitude of a reciprocal lattice point $\vec{H}$, and $\varphi_{1,\vec{H}}$ and $\varphi_{2,\vec{H}}$ two phase angles (in degrees) for the same lattice point. With $\Delta\varphi_p = \left|\varphi_{1,\vec{H}} - \varphi_{2,\vec{H}}\right| \bmod 360$, the smallest angle $\Delta\varphi_s$ between $\varphi_{1,\vec{H}}$ and $\varphi_{2,\vec{H}}$ is defined as:

$$\Delta\varphi_s = \begin{cases} \Delta\varphi_p & \text{for } \Delta\varphi_p \le 180° \\ 360 - \Delta\varphi_p & \text{for } \Delta\varphi_p > 180° \end{cases}$$

The *F*-weighted phase difference (or change) $\Delta\Phi_w$ of two phase sets with *n* elements $\varphi_{1,\vec{H}(j)}$ and $\varphi_{2,\vec{H}(j)}$, and a corresponding set of Fourier magnitudes with elements $F_{\vec{H}(j)}$, is then defined as:

$$\Delta\Phi_w = \frac{\sum_{j=1}^{n} m_{\vec{H}(j)} \cdot F_{\vec{H}(j)} \cdot \frac{\Delta\varphi_{s(j)}}{180}}{\sum_{j=1}^{n} m_{\vec{H}(j)} \cdot F_{\vec{H}(j)}}$$

$m_{\vec{H}}$ is the number of symmetry equivalent copies (multiplicity) of the reciprocal lattice point $\vec{H}$.

As $\Delta\Phi_w$ is defined, it can take values from zero (all phase angles equal) to one (all phase angles rotated 180° with respect to each other).

7.2 Symmetrically equivalent bonds

Let X_1 and X_2 be two positions in the unit cell, and let $B_{1\text{-}2}$ be a bond between X_1 and X_2. If there is a symmetry operation which maps X_1 onto itself and X_2 onto $X_{2'} \neq X_2$, then there is a bond $B_{1\text{-}2'}$ which is *symmetrically equivalent* to bond $B_{1\text{-}2}$.

Fig. 7-1 illustrates the situation for an example where X_1 is laying on a mirror plane.

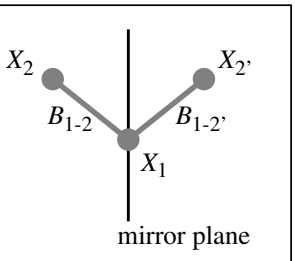


Figure 7-1

7.3 Geometrical evaluation of tetrahedral node connectivities

(a) Test of the six tetrahedral angles (see Fig. 3-6 on page 23):

A tetrahedron is rejected, if there is an angle larger than 175°, or more than two angles are smaller than 75°.

(b) Test of the four “sub-tetrahedra” defined by each (i) the pivot node and (ii) three bonded nodes (see Fig. 3-6 on page 23):

For this, four groups of three tetrahedral angles (of the main tetrahedron) are evaluated:

Group	Angle 1	Angle 2	Angle 3
1	$N_{B(1)}$ - N_P - $N_{B(2)}$	$N_{B(1)}$ - N_P - $N_{B(3)}$	$N_{B(1)}$ - N_P - $N_{B(4)}$
2	$N_{B(1)}$ - N_P - $N_{B(2)}$	$N_{B(2)}$ - N_P - $N_{B(3)}$	$N_{B(2)}$ - N_P - $N_{B(4)}$
3	$N_{B(1)}$ - N_P - $N_{B(3)}$	$N_{B(2)}$ - N_P - $N_{B(3)}$	$N_{B(3)}$ - N_P - $N_{B(4)}$
4	$N_{B(1)}$ - N_P - $N_{B(4)}$	$N_{B(2)}$ - N_P - $N_{B(4)}$	$N_{B(3)}$ - N_P - $N_{B(4)}$

Table 7-1: Groups of tetrahedral angles used for the “sub-tetrahedra” test

A tetrahedron is rejected, if in one of the four groups more than one out of the three angles is smaller than 60°.

(c) Test of the volume:

With the assumption that all six tetrahedral angles have the ideal value of arccos(-1/3) = 109.47°, `IdealT_NodeDistance` is used to compute the ideal tetrahedral volume. A tetrahedron is rejected, if its volume is less than 20% of the ideal volume.

(d) Test of the distortion (I):

Idea: at least one bond has to be approximately perpendicular to the plane defined by the other three bonds. However, since the tolerance is set to ±60°, this is only a weak filter.

(e) Test of distortion (II):

A tetrahedron is rejected, if the pivot node (see Fig. 3-6 on page 23) is not inside the tetrahedron. (The tetrahedron is confined by four planes. Each plane is defined by one of the four subsets of three nodes which can be constructed from the four bonded nodes.)

With `CheckTetrahedralGeometry Normal`, tests (a) - (d) are carried out.

With `CheckTetrahedralGeometry Hard`, tests (a) - (e) are carried out. (There are some aluminophosphate zeolites which do not pass this test with the published node atom coordinates.)

7.4 Analytic integration of electron density peaks

This formalism is due to M. Schwarz [82].

The solution of the improper integral

$$\Phi = \int_V \rho_{calc}(xyz)dV$$

with

$$\rho_{calc}(xyz) = \exp(a + bx + cy + dz + ex^2 + fy^2 + gz^2 + hyz + kzx + lxy)$$

(Eq. (3-2)) is sought.

We define

$$A = \begin{pmatrix} e & l/2 & k/2 \\ l/2 & f & h/2 \\ k/2 & h/2 & g \end{pmatrix} \quad \text{and} \quad \vec{a} = \begin{pmatrix} a \\ b \\ c \end{pmatrix} .$$

Since A is a symmetric matrix, all three of its eigenvalues, λ_1, λ_2 and λ_3, are real. Now it has to be established that all three eigenvalues are negative. If even one eigenvalue is zero or positive, the integral diverges ($\Phi = \infty$). If all of the eigenvalues are negative, the integral is finite and the analytic formula is

$$\Phi = V \cdot \pi^{3/2} \frac{\exp\left(a - \frac{1}{4}(\vec{a}A^{-1}\vec{a})\right)}{\sqrt{|\det A|}} ,$$

where V is the volume of the unit cell.

G.W. Stewart [83] suggested the technique which is used to test if the three eigenvalues are all negative. This can be done by performing Gaussian elimination on A. Specifically, if e, f and g are not all negative then the eigenvalues are not all negative. Now the matrix

$$\begin{pmatrix} f' & H' \\ * & g' \end{pmatrix} = \begin{pmatrix} f - \frac{(l/2)^2}{e} & h/2 - \frac{l/2 \cdot k/2}{e} \\ * & g - \frac{(k/2)^2}{e} \end{pmatrix}$$

is formed. If f' and g' are not negative, then the eigenvalues are not all negative. Finally $g'' = g' - H'^2/f'$ is formed. If g'' is not negative then the eigenvalues are not all negative. Otherwise they all are negative. This test is due to Lagrange.

7.5 Internally stored X-ray wavelengths [Å] and their keywords

$K_{\alpha 1}$		$K_{\alpha 2}$		Average $K_{\alpha 1}$, $K_{\alpha 2}$	
CrA1	2.28970	CrA2	2.29361	Cr	2.2909
FeA1	1.93604	FeA2	1.93998	Fe	1.9373
CuA1	1.54056	CuA2	1.54439	Cu	1.5418
MoA1	0.70930	MoA2	0.71359	Mo	0.7107
AgA1	0.55941	AgA2	0.56380	Ag	0.5608

Table 7-2

7.6 DLS prescriptions used

	KRIBER input		DLS input	
Constraint	**Prescription**	**ESD**	**Distance**	**Weight**
Si - O distance	1.628	0.01	1.628	2
Si - O - Si angle	145	8	3.105	0.229
O - Si - O angle	109	2	2.651	0.406

Table 7-3: DLS prescriptions used

For the DLS [38] input, Si-O-Si angles are converted to Si-Si distances, and O-Si-O angles are converted to O-O distances by KRIBER [61].

All values have to be in the same units as the lattice constants supplied. In this work, the unit Å was used in all cases.

7.7 FOCUS source code

The source code of FOCUS was written in ANSI-C. The basic part of the program has a length of ca. 19700 lines (477 kBytes). For use in FOCUS, a space group library was developed which has a length of ca. 10500 lines (257 kBytes). The program is portable and has been successfully compiled and tested with various Unix operating systems (ranging from Linux to Cray Unicos), and also the OpenVMS operating system.

All input and output files are in ASCII format. Two simple accompanying tools were developed for the evaluation of the output. The first tool culls the coordination sequences from the output file, and the second tool prepares the histogram and performs a database lookup for the coordination sequences. The first tool is also used to extract the crystallographic coordinates of the topologies from the FOCUS output. The format of these structure data is readable by the KRIBER [61].

8 References

[1] L.B. Sand, F.A. Mumpton; Natural Zeolites; Pergamon Press, Oxford 1978

[2] G.T. Kerr; Synthetic Zeolites; Scientific American (1989), **261**, 82-87

[3] Zeolites and Related Microporous Materials: State of the Art 1994, Proceedings of the 10th International Zeolite Conference, Garmisch-Partenkirchen, Germany, July 17-22, 1994; Elsevier, Amsterdam-London-New York-Tokyo 1994

[4] W.M. Meier, D.H. Olson; Atlas of Zeolite Structure Types; Third Revised Edition; Butterworth-Heinemann, London 1992

[5] H.M. Rietveld; A Profile Refinement Method for Nuclear and Magnetic Structures; J. Appl. Cryst. (1969), **2**, 65

[6] Ch. Baerlocher; Restraints and constraints in Rietveld refinement; In: R.A. Young; The Rietveld Method; IUCr Monographs on Crystallography 5; International Union of Crystallography, Oxford University Press (1993), 186-196

[7] R.L. Snyder; Analytical profile fitting of X-ray powder diffraction profiles in Rietveld analysis; In: R.A. Young; The Rietveld Method; IUCr Monographs on Crystallography 5; International Union of Crystallography, Oxford University Press (1993), 111-131

[8] J.M. Bennett; Determining the Structure of Molecular Sieve Materials Using High-Resolution Powder Data; In: W.H. Flank, T.E. Whyte (Eds.); Perspectives in Molecular Sieve Science; ACS Symp. Ser. 368 (1988), 162-176

[9] G.S. Pawley; Unit-Cell Refinement From Powder Diffraction Scans; J. Appl. Cryst. (1981), **14**, 357-361

[10] D.S. Sivia; W.I.F. David; A Bayesian Approach to Extracting Structure-Factor Amplitudes from Powder Diffraction Data; Acta Cryst. (1994), **A50**, 703-714

[11] A. Le Bail, H. Duroy, J.L. Fourquet; Ab-Initio Structure Determination of $LiSbWO_6$ by X-Ray Powder Diffraction; Mat. Res. Bull. (1988), **23**, 447-452

[12] M.M. Woolfson; Direct Methods - from Birth to Maturity; Acta Cryst. (1987), **A43**, 593-612

[13] A. Le Bail; A compilation of publications dealing with ab initio structure determination from powder diffraction data, March 1995; Universite du Maine, 72017 Le Mans Cedex, France; ftp://aviion.univ-lemans.fr/pub/fluorlab/iniref.html

[14] W.I.F. David; The probabilistic determination of intensities of completely overlapping reflections in powder diffraction patterns; J. Appl. Cryst. (1987), **20**, 316-319

[15] W.I.F. David; Extending the power of powder diffraction for structure determination; Nature (1990), **346**, 731-734

[16] J. Jansen, R. Peschar, H. Schenk; On the Determination of Accurate Intensities from Powder Diffraction Data. II. Estimation of Intensities of Overlapping Reflections; J. Appl. Cryst. (1992), **25**, 237-243

[17] M.A. Estermann, V. Gramlich; Improved Treatment of Severely or Exactly Overlapping Bragg Reflections for the Application of Direct Methods to Powder Data; J. Appl. Cryst. (1993), **26**, 396-404

[18] G. Cascarano, L. Favia, C. Giacovazzo; Sirpow.91 - a direct-methods package optimized for powder data; J. Appl. Cryst. (1992), **25**, 310-317

[19] C. Rohrig, H. Gies; A New Zincosilicate Zeolite with Nine-Ring Channels; Angewandte Chemie – International Edition in English (1995), **34**, 63-65

[20] B. Marler, A. Grünewald-Lüke, H. Gies; Decasils, a new order-disorder family of microporous silicas; Zeolites (1995), **15**, 388-399

[21] M.J. Annen; Ph.D. thesis: Synthesis and characterization of novel molecular sieves containing three-membered rings; Virginia Polytechnic Institute and State University, Chemical Engineering, Blacksburg (1992)

[22] M.J. Annen, M.E. Davis; Raman and ^{29}Si MAS NMR spectroscopy of framework materials containing three-membered rings; Microporous Materials (1993), **1**, 57-65

[23] M.A. Camblor, M.E. Davis; ^{29}Si MAS NMR Spectroscopy of Tectozincosilicates; Journal of Physical Chemistry (1994), 98, 13151-13156

[24] R.W. Grosse-Kunstleve, L.B. McCusker, Ch. Baerlocher; Anomalous dispersion effects on the powder diffraction pattern of the rubidium zincosilicate molecular sieve VPI-9; Activity Report 1993 (BNL 52415), BNL National Synchrotron Light Source, B-49

[25] W. Prandl; Phase Determination and Patterson Maps from Multiwavelength Powder Data; Acta Cryst. (1990), **A46**, 988-992

[26] D.E. Goldberg; Genetic algorithms in search, optimization, and machine learning; Addison-Wesley (1989)

[27] J. Rius, C. Miravitlles; A Geometrically Constrained Phase Refinement Function: the Derivation of a New Modified Tangent Formula; Acta Cryst. (1989), A45, 490-494

[28] D.E. Goldberg, P. Segrest; Finite Markov chain analysis of genetic algorithms; In: John J. Grefenstette, Ed.; Genetic algorithms and their applications: proceedings of the Second International Conference on Genetic Algorithms, July 28-31, 1987, at the Massachusetts Institute of Technology, Cambridge, MA; Lawrence Erlbaum (1987)

[29] N. Wirth; Algorithms and data structures; Prentice-Hall (1986)

[30] G.O. Brunner; The Properties of Coordination Sequences and Conclusions Regarding the Lowest Possible Density of Zeolites; J. Solid State Chem. (1979), **29**, 41-45

[31] D. Louër; Automatic Indexing: Procedures and Applications; In: Accuracy in Powder Diffraction II: Proceedings of the International Conference May 26-29, 1992; E. Prince, J.K. Stalick (Eds.); NIST SP 846, 92-104

[32] GSAS - General Structure Analysis System; A.C. Larson, R.B. von Dreele; Los Alamos National Laboratory, NM 87545, USA, 1985-1995

[33] Ch. Baerlocher, A. Hepp; XRS-82 – X-ray Rietveld System; Laboratory of Crystallography, ETH Zurich, Switzerland, 1982-1995

[34] FullProf ; J. Rodriguez-Carvajal; Institut Laue-Langevin, Grenoble, France, 1993

[35] S.R. Hall, H.D. Flack, J.M. Stewart (Eds.); Xtal 3.2 User manual; Universities of Western Australia and Maryland (1992)

[36] M.A. Estermann; Optimised Wilson normalisation of integrated Bragg intensities from powder diffraction data lacking atomic resolution; presented at: IUCr Workshop "Structure determination from Powder Data", Wadham College, Oxford, U.K., July 1995

[37] M.A. Estermann; Beiträge zur Verbesserung der *ab initio* Strukturanalyse von Pulverdaten; Ph.D. thesis, ETH Zurich, Switzerland, 1991

[38] Ch. Baerlocher, A. Hepp, W.M. Meier; DLS-76 – A Program for the Simulation of Crystal Structures by Geometric Refinement; Institute of Crystallography and Petrography, ETH Zurich, Switzerland, 1977

[39] L.F. Ten Eyck; Crystallographic Fast Fourier Transforms; Acta Cryst. (1973), **A29**, 183-191

[40] L.F. Ten Eyck; Efficient Structure-Factor Calculation for Large Molecules by the Fast Fourier Transform; Acta Cryst. (1977), **A33**, 486-492

[41] J.S. Rollet (Ed.); Computing Methods in Crystallography; Pergamon Press (1965)

[42] C. Giacovazzo (Ed.); Fundamentals of Crystallography; International Union of Crystallography; Oxford University Press (1992)

[43] T. Loiseau, G. Férey; Oxyfluorinated Microporous Compounds: VII. Synthesis and Crystal Structure of ULM-5, a New Fluorinated Gallophosphate $Ga_{16}(PO_4)_{14}(HPO_4)_2(OH)_2F_7$, $[H_3N(CH_2)_6NH_3]_4$, 6 H_2O with 16-Membered Rings and Both Bonding and Encapsulated F^-; J. Solid State Chem. (1994), **111**, 403-415

[44] G.O. Brunner, F. Laves; Zum Problem der Koordinationszahl.; Wiss. Z. Techn. Univers. Dresden (1971), **20**, 387-390

[45] C.P. Herrero; Coordination Sequences of Zeolites Revisited: Asymptotic Behaviour for Large Distances; J. Chem. Soc. Faraday Trans. (1994), **90**, 2597-2599

[46] W. Fischer; Existenzbedingungen homogener Kugelpackungen zu kubischen Gitterkomplexen mit weniger als drei Freiheitsgraden; Z. Krist. (1973), **138**, 129-146

[47] L. Stixrude, M.S.T. Bukowinski; Rings, topology, and the density of tectosilicates; American Mineralogist (1990), **75**, 1159-1169

[48] W. Fischer; Existenzbedingungen homogener Kugelpackungen zu kubischen Gitterkomplexen mit drei Freiheitsgraden; Z. Krist. (1974), **140**, 50-74

[49] H. Gerke, H. Gies; Studies on clathrasils. IV, Crystal structure of dodecasil 1H, a synthetic clathrate compound of silica; Z. Krist. (1984), 166, 11-22

[50] R.W. Grosse-Kunstleve; Diplom-Arbeit, Institut für Mineralogie, Ruhr-Universität Bochum (1992)

[51] D. Alexander; PEAKFIND – A program for the determination of peak positions in powder profiles; ICI, U.K. (1973)

[52] D. Taupin; Enhancements in Powder-Pattern Indexing; J. Appl. Cryst. (1989), **22**, 455-459

[53] International Tables for Crystallography - Volume C - Mathematical, Physical and Chemical Tables; Edited by A.J.C. Wilson; Kluwer Academic Publishers, Dordrecht/ Boston/London 1992

[54] G. Marsaglia, A. Zaman; Toward a Universal Random Number Generator; Florida State University Report: FSU-SCRI-87-50 (1987) – Remark: this is the well-known “ranmar” portable pseudo-random number generator.

[55] L.B. McCusker; Zeolite Structure Analysis Using Powder Diffraction Data; Materials Science Forum Vol. 133-136 (1993) pp. 423-434

[56] M.A. Estermann, L.B. McCusker, Ch. Baerlocher; *Ab initio* structure determination from severely overlapping powder diffraction data; J. Appl. Cryst. (1992), **25**, 539-543

[57] L.B. McCusker, Ch. Baerlocher; A re-examination of the structure of SAPO-40; Microporous Materials, in press

[58] A. Simmen, L.B. McCusker, Ch. Baerlocher, W.M. Meier; The structure determination and Rietveld refinement of the aluminophosphate $AlPO_4$-18; Zeolites (1991), **11**, 654-661

[59] C. Deroche, B. Marler, H. Gies, G.T. Kokotailo, P.U. Pennartz; The structure of partially hydrated zeolite NaA drived from Rietveld refinement; Z. Krist. (1992), Suppl. No. **5**, 46

[60] Ch. Baerlocher, L.B. McCusker, R. Chiapetta; Location of the 18-crown-6 template in EMC-2 (EMT), Rietveld refinement of the calcined and as-synthesized forms; Microporous Materials (1994), **2**, 269-280

[61] R. Bialek; KRIBER - an interactive PASCAL program to calculate distances and angles, to generate input files for the programs DLS-76 and XRS-82, and to calculate coordination sequences and loop configurations; Institut für Kristallografie und Petrografie, ETH Zurich, Switzerland (1991)

[62] C. Röhrig, H. Gies, B. Marler; Rietveld refinement of the crystal structure of the synthetic porous zincosilicate VPI-7; Zeolites (1994), 14, 498-503

[63] R.J. Hill, H.D. Flack; The Use of the Durbin-Watson d Statistic in Rietveld Analysis; J. Appl. Cryst. (1987), **20**, 356-361

[64] P.-E. Werner, L. Erikson, M. Westdahl; TREOR, a semi-exhaustive trial-and-error powder indexing program for all symmetries; J. Appl. Cryst. (1985), **18**, 367-370

[65] S. Ueda, M. Koizumi, Ch. Baerlocher, L.B. McCusker, W.M. Meier; Synthetic Lovdarite; Preprints of Poster Papers, The 7th International Zeolite Conference (1986), p. 23

[66] W. Prandl; Phase Determination from X-ray Powder Diffraction Data. II. Partial Patterson Maps and the Localization of Anomalously Scattering Atoms; Acta. Cryst. (1994), **A50**, 52-55

[67] G. Giuseppetti, F. Mazzi, C. Tadini, E. Galli; The revised crystal structure of Roggianite – Ca_2 (Be $(OH)_2$ Al_2 Si_4 O_{13}) < 2.5 H_2O; Neues Jahrbuch für Mineralogie - Monatshefte (1991). **7**, 307-314

[68] IZA Structure Commission Report; Zeolites (1994), **14**, 389-392

[69] A.M. Chippindale, R.I. Walton, C. Turner; Synthesis and structure of a novel open-framework gallium phosphate – $[Me_2NH(CH_2)_2NHMe_2]^{2+}[Ga_4P_5O_{20}H]^{2-}\cdot H_2O$; J. Chem. Soc., Chem. Commun. (1995), **12**, 1261-1262

[70] H.-X. Li, M.E. Davis, J.B. Higgins, R.M. Dessau; Aluminophosphate Molecular Sieves Comprised Of Hydrated Triple Crankshaft Chains. J. Chem. Soc., Chem. Commun. (1993), **4**, 403-405

[71] G.J. Kennedy, J.B. Higgins, C.F. Ridenour, H.-X. Li, and others; Multifield Magic-Angle Spinning and Double-Rotation Nuclear Magnetic Resonance Studies of a Hydrated Aluminophosphate Molecular Sieve - AlPO4-H2; Solid State Nuclear Magnetic Resonance (1995), **4**, 173-178

[72] J.V. Smith; Topochemistry of Zeolites and Related Materials. 1. Topology and Geometry; Chem. Rev. (1988), **88**, 149-182

[73] J.V. Smith; Towards a comprehensive mathematical theory for the topology and geometry of microporous materials; In: P.A. Jacobs, R.A. van Santen (Eds.); Zeolites: Facts, Figures, Future; Elsevier Science Publishers B.V., Amsterdam (1989)

[74] G.O. Brunner; Criteria for the evaluation of hypothetical zeolite frameworks; Zeolites (1990), **10**, 612-614

[75] M.M.J. Treacy; S. Rao, I. Rivin; A combinatorial method for generating new zeolite frameworks; In: R. von Ballmoos, J.B. Higgins, M.M.J. Treacy (Eds.); Proceedings from the Ninth International Zeolite Conference, Montreal (1992), 381-388

[76] D.E. Akporiaye, G.D. Price; Systematic enumeration of zeolite frameworks; Zeolites (1989), **9**, 23-32

[77] M.D. Shannon; Method of solution and structure of 3 novel high-silica medium-pore zeolites with multi-dimensional channel system; In: R. von Ballmoos, J.B. Higgins, M.M.J. Treacy (Eds.); Proceedings from the Ninth International Zeolite Conference, Montreal (1992), 389-398

[78] J. Rius, S. Vortmann, H. Gies; Solution of zeolite precursors from low-resolution X-ray powder diffraction data; Poster abstract presented at: IUCr Workshop “Structure determination from Powder Data”, Wadham College, Oxford, U.K., July 1995

[79] G.T. DeTitta, C.M. Weeks, P. Thuman, R. Miller, H.A. Hauptman; Structure solution by minimal function phase refinement and Fourier filtering: theoretical basis; Acta Cryst. (1994) **A50**, 203-210

[80] C.M. Weeks, G.T. DeTitta, H.A. Hauptman, P. Thuman, R. Miller; Structure solution by minimal function phase refinement and Fourier filtering: implementation and applications, Acta Cryst. (1994) **A50**, 210-220

[81] G.M. Sheldrick, R.O. Gould; Structure Solution by Iterative Peaklist Optimization and Tangent Expansion in Space Group P1; Acta Cryst. (1995), **B51**, 423-431

[82] M. Schwarz; Department of Mathematics, ETH Zurich; Private communication.

[83] G.W. Stewart; Department of Computer Science, University of Maryland; Private communication.

Acknowledgments

The investigations described in this thesis were carried out in the Laboratory of Crystallography at the Swiss Federal Institute of Technology, in collaboration with the Department of Chemistry.

I would like to express my thanks to Prof. Dr. M. Dobler for agreeing to supervise this project and for the freedom granted in the course of the research, and to Prof. Dr. W. Steurer for allowing me to work in his laboratory and for his willingness to become a referee of this thesis.

I also thank C. Röhrig and C.A. Deroche of the Ruhr-Universität Bochum for providing me with the RUB-17 and Zeolite-A datasets. I particularly acknowledge Prof. Dr. M.E. Davis of Caltech for providing me with some inspiring ideas, and for sending us the VPI-9 and VPI-10 samples and chemical analyses. I am indebted to M. Schwarz of ETH Zurich and G.W. Stewart of the University of Maryland for their help in the calculation of the analytical integral for electron density peaks.

Furthermore, I must express my deepest thanks to “the anonymous Internet user” – all those people who put so much effort into bringing this incredible medium alive. On more than one occasion, the immediate availability of information and the fast exchange of know-how through news groups made the impossible possible.

I would like to thank Mrs. E. Spengler for her successful efforts to make the administrative burden during my sojourn in Zurich nearly imperceptible, and for acting as the soul of the laboratory.

The help of G.O. Brunner, and the discussions regarding the development of the coordination sequence and loop configuration algorithms are very much appreciated. Furthermore, I want to express my sincere thanks to all members of our laboratory, for contributing to the very friendly working environment and for their constant willingness to participate in discussions, especially to the members of the zeolite group, J. Onate, Dr. C. Lengauer and Dr. A. Meden. But, of course, above all, I am deeply indebted to Dr. Ch. Baerlocher and Dr. L.B. McCusker for giving me the feeling of having “scientific parents”, open to all ideas, always supporting, and always present when I needed help – in almost any respect.

My last acknowledgment goes again to Dr. L.B. McCusker, for her constant patience in correcting and proof-reading the entire manuscript.

This work was supported by funds from the Swiss National Science Foundation.

Curriculum Vitae

Birth
Warendorf, Germany, April 3, 1964

Citizenship
German

Marital Status
Married

Education in Germany

1970 - 1974	Grundschule, Sassenberg
1974 - 1983	Gymnasium, Warendorf
1983	Abitur
1983 - 1985	Social service (substitute for military service)
1987 - 1988	Studies of architecture, Universität Dortmund
1988 - 1992	Studies of mineralogy, Ruhr-Universität Bochum
1992	Diplom in mineralogy with emphasis on crystallography

I started the work on this thesis in November 1992.