

PAULING FILE verifies / reveals 12 principles in materials science supporting four cornerstones given by Nature

Pierre VILLARS^{1*}, Shuichi IWATA²

¹ *Material Phases Data System (MPDS), Unterschwanden 6, CH-6354 Vitznau, Switzerland*

² *The University of Tokyo, 5-1-5 Kashiwanoha, Kashiwa City, Chiba, 277-8563, Japan*

* *Corresponding author. E-mail: villars.mpds@bluewin.ch*

Received December 3, 2013; accepted December 25, 2013; available on-line August 30, 2014

Here we outline four cornerstones given by Nature, which on the one hand are responsible for the fact that we are confronted with infinitely many chemical element combinations, on the other hand provide a systematic framework of restraints. To support these four cornerstones, statistical plots and materials-overview – governing factor maps were used. This led to the establishment of twelve principles, which can be used in materials design as restraints. The principles were verified or derived with the help of the PAULING FILE, the most comprehensive inorganic materials database, consisting of three interconnected parts: structure-diffraction, constitution, and intrinsic physical properties. The principles have general validity and provide a possibility to develop efficient experimentation and calculation exploration strategies.

First cornerstone: Infinitely many chemical element combinations

i) Number of potential chemical element combinations; ii) Number of potential compounds per chemical system and their stoichiometric ratios; iii) Number of potential prototypes per potential compound.

Second cornerstone: Core principle that defines compound formation

1) Compound-formation map principle; 2) Number of chemical elements – atomic-environment type (AET) correlation principle; 3) Active concentration range principle; 4) Stoichiometric ratio condition principle.

Third cornerstone: Core principle that defines ordering of chemical elements within a structure type

5) Simplicity principle; 6) Symmetry principle; 7) Atomic-environment type principle; 8) Chemical element ordering principle.

Fourth cornerstone: Core principle that links the position of chemical elements within a structure type ↔ in the Periodic System

9) Prototype – Periodic System correlation (chemistry) principle; 10) Structure stability map principle; 11) Generalized AET stability map principle; 12) Complete solid solution stability map principle.

Database / Governing factors / Data mining / Inorganic compounds / Intermetallic compounds / Materials informatics

1. Introduction

Confronted with the explosion of computing power, as well as of materials data, Gray proposed in 2009 the *Fourth Paradigm of Science: Data-Intensive Discovery through Data Exploration* (eScience) [1], which means electronically unify experiment, theory and computation. The Executive Office of the President of the United States, National Science and Technology Council, launched mid-2011 the whitepaper *Materials Genome Initiative for Global Competitiveness* [2], having as major aim to shorten the time between discovery of advanced materials and their industrial application by at least a factor two. Reflecting these new trends, many ideas have been

proposed to explore new dimensions, trying to derive knowledge from a collection of data. To show a clear direction for such trends, we want to draw a roadmap in this work by taking advantage of scientific data, namely in the case of this publication we select scientific data on inorganic compounds (materials).

Three key developments during the past decades have opened up unprecedented opportunities, namely: 1) The power of high-speed computers has reached incredible levels, and their price has decreased. 2) Computational materials science, algorithmic developments, and sophisticated software systems for simulation are advancing at increasing rates. In particular, so-called first-principles methods are now established that can predict the prototype (crystal

structure) and intrinsic physical properties of inorganic compounds by the only use of fundamental constants. The computed results are often comparable with experimental data. In a similar way, the CALPHAD approach has reached maturity, and is able to calculate phase diagrams that can be compared with experimentally determined phase diagrams. 3) High bandwidth communication at extremely low cost has revolutionized global collaboration and knowledge exchange.

Despite the enormous importance of inorganic compounds for our industrialized society in areas such as housing, energy, transportation, civil engineering, communication, and health, mankind's knowledge of inorganic compounds is astoundingly sparse. For example, only about 16% of all possible ternary chemical systems have been at least partly characterized. In the case of inorganic compounds containing four or more chemical elements, this fraction drops to 0.6% or less [3]. In fact, the remarkable accomplishments in the development of advanced technology such as aircraft engines, computer processors, magnetic recording devices, or chemical catalysts, rely on the optimization of physical properties of inorganic compounds.

In this publication we focus on inorganic substances. On the one hand it is worth noticing that the number of experimentally investigated inorganic compounds is very low. But on the other hand the ability to calculate crystallographic data and intrinsic physical properties for inorganic compounds from first-principles methods, and phase diagrams with CALPHAD methods, is rapidly growing. These two facts motivated us to explore new ways to build up a Materials Genome, we call it World Materials, database system consisting of two major inter-linked data parts: The experimentally determined data part, to be used as reference, and the (to be) calculated (using first-principles and CALPHAD mass-calculations) data part.

It has to be highlighted that engineering materials are typically multi-phase materials, which are in addition affected by defects, interfaces, and microstructure. Nevertheless the fundamental basis of all these engineering materials is made from the individual inorganic compounds.

Neither the Fourth Paradigm of Science [1] nor the Materials Genome Initiative [2] can be realized without the integration of restraints obtained by "Materials data exploration searching for principles (governing factors) with the aim to formulate restraints", as shown in this work. The infinite number of potential chemical element combinations forces us to develop approaches that are able to reduce this infinite number to a practicable number of the most probable potential inorganic compounds, to be theoretically and experimentally investigated. In addition it is essential to note that the realization of World Materials requires two pre-conditions:

i) The first requirement is the introduction of the distinct phases concept to link different kinds of materials data.

This concept was introduced for the PAULING FILE materials database system [3-5], and is implemented in its derived products [6-9]. A phase is defined by the chemical system and the structural prototype and has been given a unique name by a representative chemical formula and, when relevant, a specification. As the linkage of different groups of data was considered as most important, the PAULING FILE was designed as a phase-oriented database, using a fully relational database system. This was achieved by the creation of a distinct phases table, as well as the required internal links. In practice this means that each chemical system has been evaluated and the distinct phases identified based on available information. Finally every database entry has been linked to such a distinct phase. This is also a requirement for linking materials databases created by different teams.

ii) The second requirement is the existence of a comprehensive, critically evaluated materials database system of experimentally determined single-phase materials data, to be used as reference.

The PAULING FILE was launched 20 years ago and represents meanwhile the sole and therefore world-largest data collection of its kind with over 500'000 data sets (subdivided into structure-diffraction, constitution, intrinsic physical properties) from over 150'000 scientific publications, covering in principle all crystalline inorganic compounds with no C-H bonds. It is now becoming feasible to use it as a starting reference. PAULING FILE data have been carefully checked, and fully standardized [10-12].

2. Exploration of materials data searching for principles (governing factors) with the aim to formulate restraints

One of the most challenging tasks in materials science is the design of new inorganic compounds with beforehand-defined intrinsic physical and/or chemical properties. In order to reach this objective, two different approaches are, in general, explored [4].

1. The first approach is to simulate the motion of the atoms in the inorganic compounds, as well as their electronic interactions, as close to reality as possible by using quantum-mechanical calculations. In theory there is no input other than the laws of quantum mechanics and the atomic number (AN) of the involved chemical elements. Using these calculations the intrinsic properties of inorganic compounds can be understood from first principles (theoretical and simulation branch of science, representing the Second and Third Paradigm of Science [1]).

2. The second approach is based on a pragmatic level. It is a fact that most of our current knowledge in

chemistry and materials science has been collected empirically, searching for patterns, rules and principles among experimental results (empirical branch of science, representing the First Paradigm of Science [1]).

For both approaches, to make predictions it is necessary to start from experimental variables such as the selection of the chemical elements to be combined, their concentrations, temperature, and pressure. For the first approach we use, in addition, the laws of quantum mechanics, including the atomic number (AN), whereas for the second approach we use tabulated elemental-property parameters (and/or expressions), such as the atomic number (AN), periodic number (PN), atomic size ($SZ_a(AN,PN)$) and atomic reactivity ($RE_a(AN,PN)$) [13].

For both approaches the key-position is the structural prototype of the inorganic compound, which represents our 'window' to view the electronic interactions of the atoms within a specific compound. Presently over 32'000 different prototypes have been experimentally established for inorganic compounds. In other words, Nature realizes more than 32'000 different geometrical arrangements of atoms, only for compounds without C-H bonds.

Recently several strong patterns in materials-overviews – elemental-property parameters maps [13-16], considering thousands of data sets of different chemical systems / inorganic compounds (summarizing the content of thousands of scientific publications), have been published. This proves that the underlying quantum mechanical laws can be parameterized by the use of elemental-property parameters (and/or expressions) of the constituent chemical elements. The search for optimal parameters led to relatively simple maps with well-defined stability domains, giving excellent overviews of experimental data for known inorganic compounds. As a direct consequence the maps give some prediction ability.

Before initiating the PAULING FILE project, one of us reviewed the world literature, focusing on intermetallics and alloys, in context with the topic

“Factors Governing Crystal Structures” [17], and came up with nine principles. The validity of these principles was tested on a fair number of compounds in well-defined groups of experimentally determined data sets and showed accuracy in the range of 90-100%. It was concluded, at that time, that we can rely on predictions based on those nine principles with considerable confidence. Applying the nine principles will both reduce the number of systems (or samples) to be investigated and remarkably increase the success rate of finding new compounds.

Now, 20 years later, having access to five times more experimental facts with a coverage extended from intermetallics and alloys to include other inorganic compounds such as *e.g.* halides and ceramics, the situation has significantly changed. The PAULING FILE contains structural information for over 140'000 inorganic phases (of which about 50% contain oxygen), compared to Pearson's Handbook of Crystallographic Data for Intermetallic Phases [18], which covered about 28'000 intermetallics and alloys (including sulfides and selenides).

In this work we sometimes focus on the 1'000 most populous prototypes and their representatives. This covers 70% of the entries in the 2013/14 edition of Pearson's Crystal Data (PCD-2013/14 [6]), which contains crystal structure data from the PAULING FILE. Tables 1 and 2 give some details about the content of PCD-2013/14, and a brief comparison with the competing product ICSD-2013/14 from FIZ/NIST [19]. By comparing the given numbers, two facts become obvious:

- 1) PCD-2013/14 [6] contains 60% more entries than ICSD-2013/14 [19], and therefore covers the world literature more comprehensively.
- 2) PCD-2013/14 [6] contains 31'883 prototypes, *i.e.* five times more than ICSD-2013/14 [19] with 6'357 prototypes. The PCD-2013/14 data have been carefully checked [10], consequently standardized with the program STRUCTURE TIDY [12] and the prototypes assigned according to the principles defined in [11].

Table 1 Number of distinct chemical systems, distinct phases, and database entries in PCD-2013/14 [6] and number of database entries in ICSD-2013/14 [19] (grouped into unaries, binaries, ternaries, four or more chemical elements).

NUMBER OF CHEMICAL ELEMENTS	CHEMICAL SYSTEMS PCD-2013/14	DISTINCT PHASES PCD-2013/14	DATABASE ENTRIES PCD-2013/14	DATABASE ENTRIES ICSD-2013/14
Unaries	95	424	2,727	1,716
Binaries	2'553	17'672	48'775	30'968
Ternaries	17'131	54'673	95'596	60'640
>3 elements	33'714	68'919	95'567	56'730
Total	53'493	141'688	242'665	150'054

Table 2 Number of database entries in PCD-2013/14 [6] compared with ICSD-2013/14 [19].

	PCD-2013/14	ICSD-2013/14
Number of prototypes	31'883	6'357
Number of database entries	242'665	150'054
- refined atom coordinates	154'847	122'300
- assigned atom coordinates	72'088	27'754
- no atom coordinates	15'730	–

The main aim of the present work was on the one hand to test the nine principles proposed in [17] (based on about 20% of the by now available facts) and, if necessary, to modify them, and on the other hand to discover additional principles. Finally we ended up with in total twelve principles, derived either from statistical plots or from materials-overview – elemental-property parameters maps. Each of them revealed governing factors, which could be used to formulate restraints.

3. Nature defines cornerstones providing a marvelously rich, but still very rigid systematic framework of restraints

Below we outline four fundamental cornerstones defined by Nature. The first cornerstone (3.1.) is responsible for the fact that we are confronted with infinitely many chemical element combinations. The second to fourth cornerstones (3.2.-3.4.) provide a very rigid systematic framework of restraints. Each of the latter is supported by four principles, *i.e.* in total twelve principles. Since the cornerstones reflect underlying natural laws, they have general validity. For comparison purposes, we will briefly outline the nine principles proposed earlier [17] and, when relevant, their extensions, and describe the newly discovered three principles.

3.1. First cornerstone: Infinitely many chemical element combinations

The experimental data in the PAULING FILE as included in SpringerMaterials [9], shows that the materials knowledge rate, by comparing the number of potential chemical element combinations with the number of chemical systems where we have at least partial information, is low. This is shown in Fig. 1 for unary to quinary systems, and it is clearly seen that we have robust knowledge only for unary and binary systems. For ternary systems we are still at the very beginning and for higher systems we have close to no knowledge. This is surprising as, since 1980, materials scientists are mainly working on ternaries and higher systems. Going to higher-order systems leads to an astronomically large number of potential chemical systems to be considered. Fig. 2a shows that the average number of inorganic phases per binary chemical system is 6.9, for ternaries about 3.2 and for quinary and higher systems about 2 (over 200

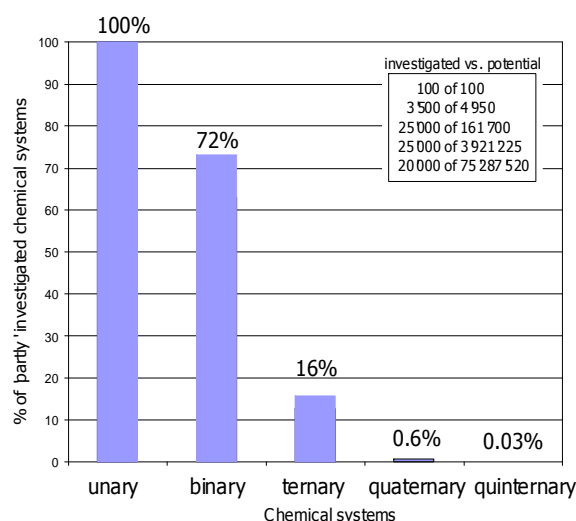


Fig. 1 Percentage of partly or fully investigated chemical systems grouped into unary, binary, ternary, quaternary, and quinary systems. The inset shows the actual numbers of investigated *versus* potential systems. The numbers are based on data of SpringerMaterials/PAULING FILE [9].

common stoichiometric ratios occur). In the following we will use the expression *basic prototype* for a prototype that has no mixed occupancy. Basic prototypes and their representatives are here referred to as daltonide compounds (no significant homogeneity range), which may not always be true. Experimentally investigated systems reveal up to maximum 30 daltonide compounds per system. Fig. 2b shows for 17'083 basic (no mixed site occupancy) ternary compounds 490 different stoichiometric ratios in a ternary concentration triangle, limiting the selection to the 1'000 most populous prototypes and their representatives. Focusing on potential inorganic compounds, going to higher-order systems leads in practice, with an average of about 2 inorganic compounds per system, to an infinitely large variety of potential inorganic compounds to be considered.

Fig. 3 shows that the number of prototypes per literature year has increased approximately linearly from 100 prototypes per year in 1960 to 1'100 prototypes per year in 2010. Nature has so far realized

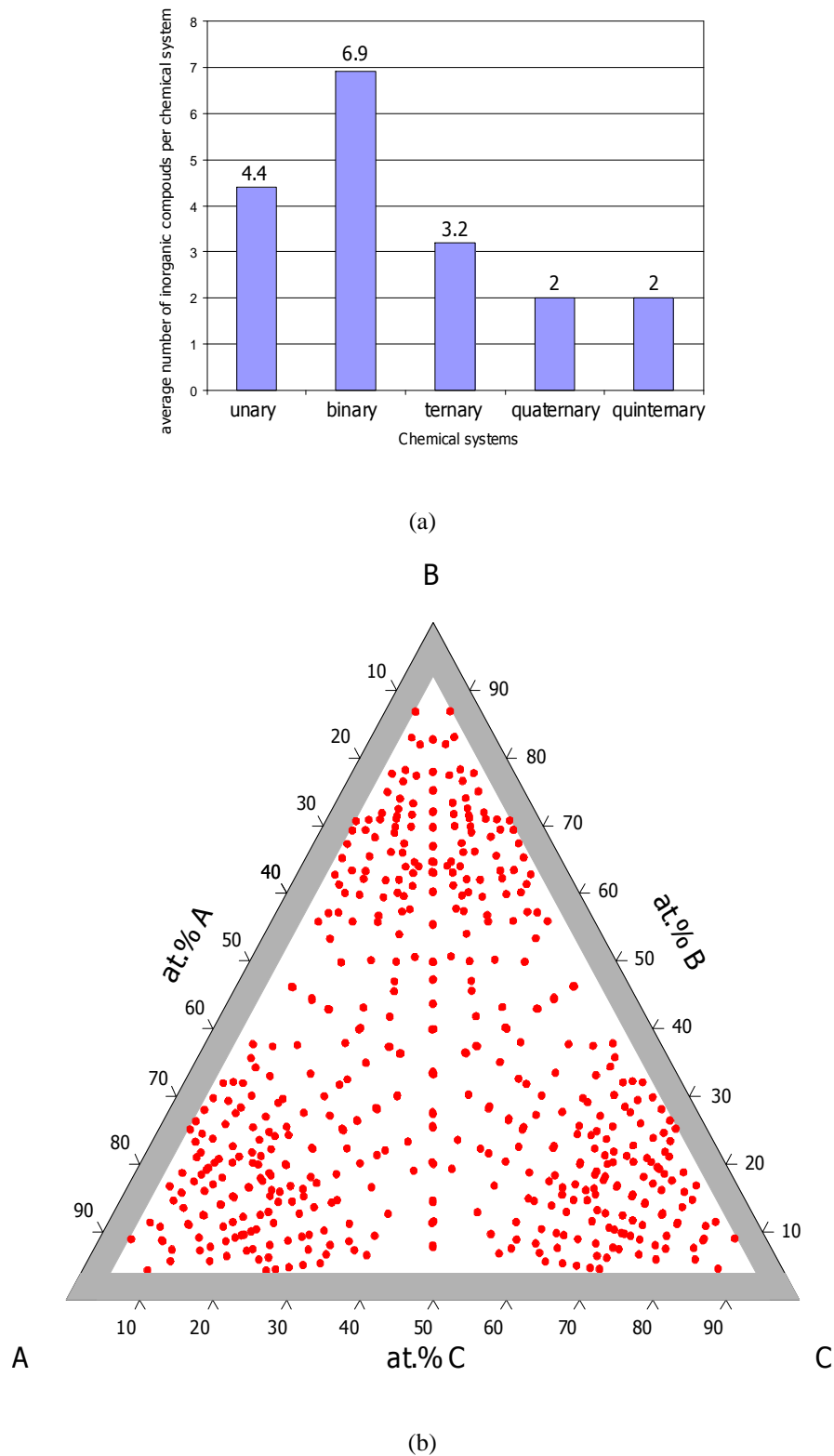


Fig. 2 (a) Average number of inorganic compounds per chemical system grouped into unary, binary, ternary, quaternary and quinternary systems. The numbers are based on data of SpringerMaterials/PAULING FILE [9]. (b) The over 250 different stoichiometric ratios of the 17'083 daltonide ternary phases shown in a ternary concentration triangle, focusing on the 1'000 most populous prototypes and their representatives in PCD-2013/14 [6].

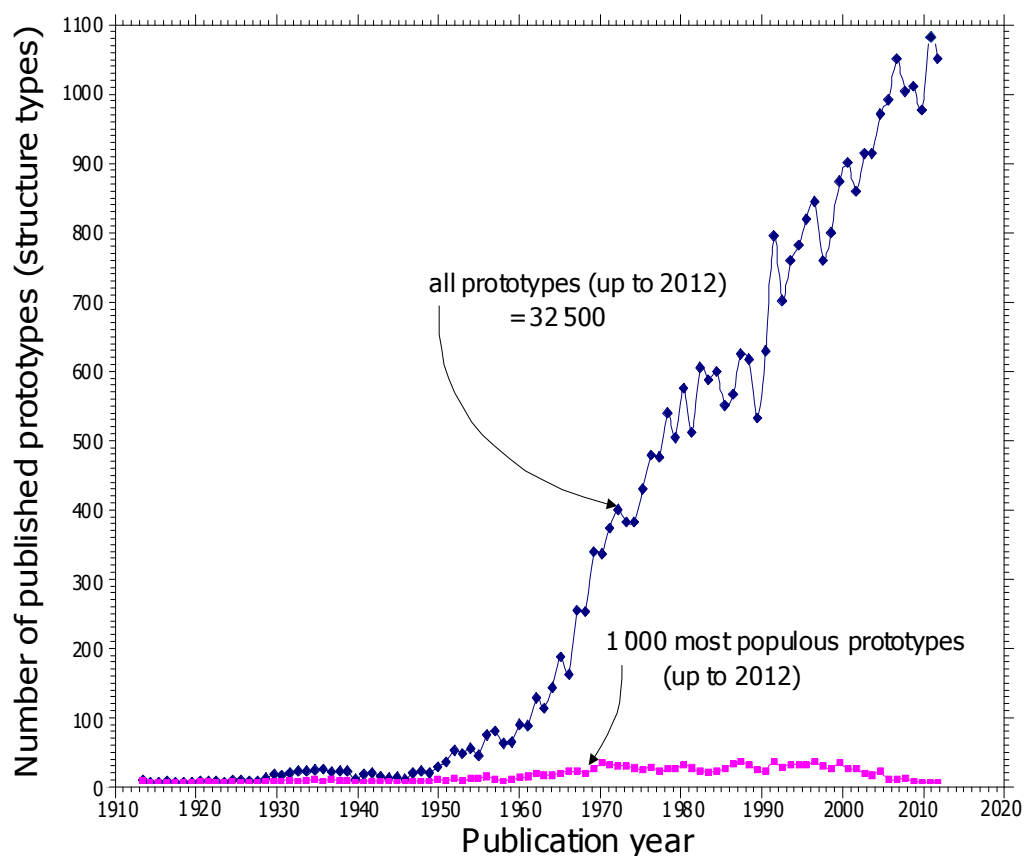


Fig. 3 Number of new prototypes as a function of the publication year. The blue curve considers all prototypes contained in PCD-2013/14 [6], the pink curve considers the 1'000 most populous prototypes in PCD-2013/14 [6].

over 32'000 prototypes for inorganic compounds, and a continued increase would lead to an infinitely large variety of 3-dimensional ways to arrange the atoms within an inorganic compound. Figs. 1-3 support, based on the data of [9], the overall conclusion of the first cornerstone: “Nature provides us with about 100 chemical elements as well as their combinations. A direct consequence of this fact is that there exist an infinite number of chemical element combinations. Furthermore, Nature has worked out a huge number of three-dimensional ways of ordering the chemical elements (atoms) within inorganic compounds. It may be added that the magnetic moments of the chemical elements may be ordered in an even higher number of four-dimensional ways.”

This leads, in practice, to a hopeless situation when trying to develop overall-valid efficient experimentation and calculation exploration strategies.

3.2. Second cornerstone: Laws that define compound formation

The second cornerstone of Nature sets strict restrictions for the formation of inorganic compounds. The enthalpy of formation has to be negative,

otherwise the chemical element combination will not lead to compound. In addition, at constant pressure, Gibbs' phase rule $P = C - F + 1$ defines the relation between the number of phases (P , here potential compounds), the number of components (C , here chemical elements) and the degree of freedom (F) of intensive properties such as temperature and composition. It is indirectly possible, with the help of materials overview – elemental-property parameters maps and statistical plots to formulate four principles. Thanks to these four principles it is possible to derive restraints, which can exclude non-former systems (chemical systems where no compounds form), and some pre-conditions to be fulfilled for compound formation.

1) Compound-formation map principle

Our work from 2008 [13] showed that the periodic number (PN) is by far the most efficient elemental-property parameter in separating formers (chemical systems where at least one distinct compound exists) from non-formers. PN represents a different enumeration of the chemical elements within the Periodic System (see Fig. 4). In contrast to the atomic number (AN), PN emphasizes the number of valence

	GN 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
QN 1	1 H 1																	2 He 112
2	3 Li 2	4 Be 8											5 B 82	6 C 88	7 N 94	8 O 100	9 F 106	10 Ne 113
3	11 Na 3	12 Mg 9											13 Al 83	14 Si 89	15 P 95	16 S 101	17 Cl 107	18 Ar 114
4	19 K 4	20 Ca 10	21 Sc 14	22 Ti 46	23 V 50	24 Cr 54	25 Mn 58	26 Fe 62	27 Co 66	28 Ni 70	29 Cu 74	30 Zn 78	31 Ga 84	32 Ge 90	33 As 96	34 Se 102	35 Br 108	36 Kr 115
5	37 Rb 5	38 Sr 11	39 Y 15	40 Zr 47	41 Nb 51	42 Mo 55	43 Tc 59	44 Ru 63	45 Rh 67	46 Pd 71	47 Ag 75	48 Cd 79	49 In 85	50 Sn 91	51 Sb 97	52 Te 103	53 I 109	54 Xe 116
6	55 Cs 6	56 Ba 12		72 Hf 48	73 Ta 52	74 W 56	75 Re 60	76 Os 64	77 Ir 68	78 Pt 72	79 Au 76	80 Hg 80	81 Tl 86	82 Pb 92	83 Bi 98	84 Po 104	85 At 110	86 Rn 117
7	87 Fr 7	88 Ra 13		104 Rf 49	105 Db 53	106 Sg 57	107 Bh 61	108 Hs 65	109 Mt 69	110 Ds 73	111 Rg 77	112 Uub 81	113 Uut 87	114 Uuq 93	115 Uup 99	116 Uuh 105	117 Uus 111	118 Uuo 118
6	57 La 16	58 Ce 18	59 Pr 20	60 Nd 22	61 Pm 24	62 Sm 26	63 Eu 28	64 Gd 30	65 Tb 32	66 Dy 34	67 Ho 36	68 Er 38	69 Tm 40	70 Yb 42	71 Lu 44			
7	89 Ac 17	90 Th 19	91 Pa 21	92 U 23	93 Np 25	94 Pu 27	95 Am 29	96 Cm 31	97 Bk 33	98 Cf 35	99 Es 37	100 Fm 39	101 Md 41	102 No 43	103 Lr 45			

Fig. 4 The four fundamental integer elemental-property parameters: atomic number (AN), periodic number (PN), group number (GN) and quantum number (QN) using Medeleev's periodic system.

electrons, *i.e.* the periodicity. **Fig. 5** shows an idealized two-dimensional compound-formation map for binary $A-B$ systems, using as axes the periodic number of the chemical element A (PN_A) versus the periodic number of the chemical element B (PN_B). PN as elemental-property parameter works well for binary, ternary, and quaternary systems. However, it was shown in the same publication [13] that the atomic size SZ_a , like the atomic reactivity RE_a , is a direct function of PN and AN strongly reflects the periodicity (valence electrons) of the periodic system. Therefore, SZ_a and RE_a can also be used as elemental-property parameters. Compared with the earlier work of Miedema for binaries [20], and the work in [21], the maps using PN separate the domains more clearly, with an overall accuracy of more than 98%. The advantage of the elemental property parameter PN with respect to parameters such as radius or reactivity is that it is an integer number.

The compound-formation map principle was formulated in 1994 [17] as follows: *The size, electrochemical, valence electron, and cohesion-*

energy factors are the factors governing compound formation. The compound-formation maps predict compound formation in binary [20] and in ternary [21] systems. This principle has been fully verified, and in addition, during the last 20 years, has been significantly improved in its simplicity and accuracy in separating formers from non-formers. Overall one can predict that about 30% of all chemical element combinations will form no inorganic compound (non-formers), this being true for binary, ternary, and quaternary systems.

2) Number of chemical elements – AET correlation principle

Fig. 6, a three-dimensional plot, shows the number of point sets with a particular atomic environment (AET) in percent of the total number of point sets for different groups of phases (unary, binary, ternary, quaternary, quinternary, sixnary), limited to the 27 most populous atomic environments and sorted by increasing coordination number (CN). In total close to 120'000 point sets were considered (all from refined

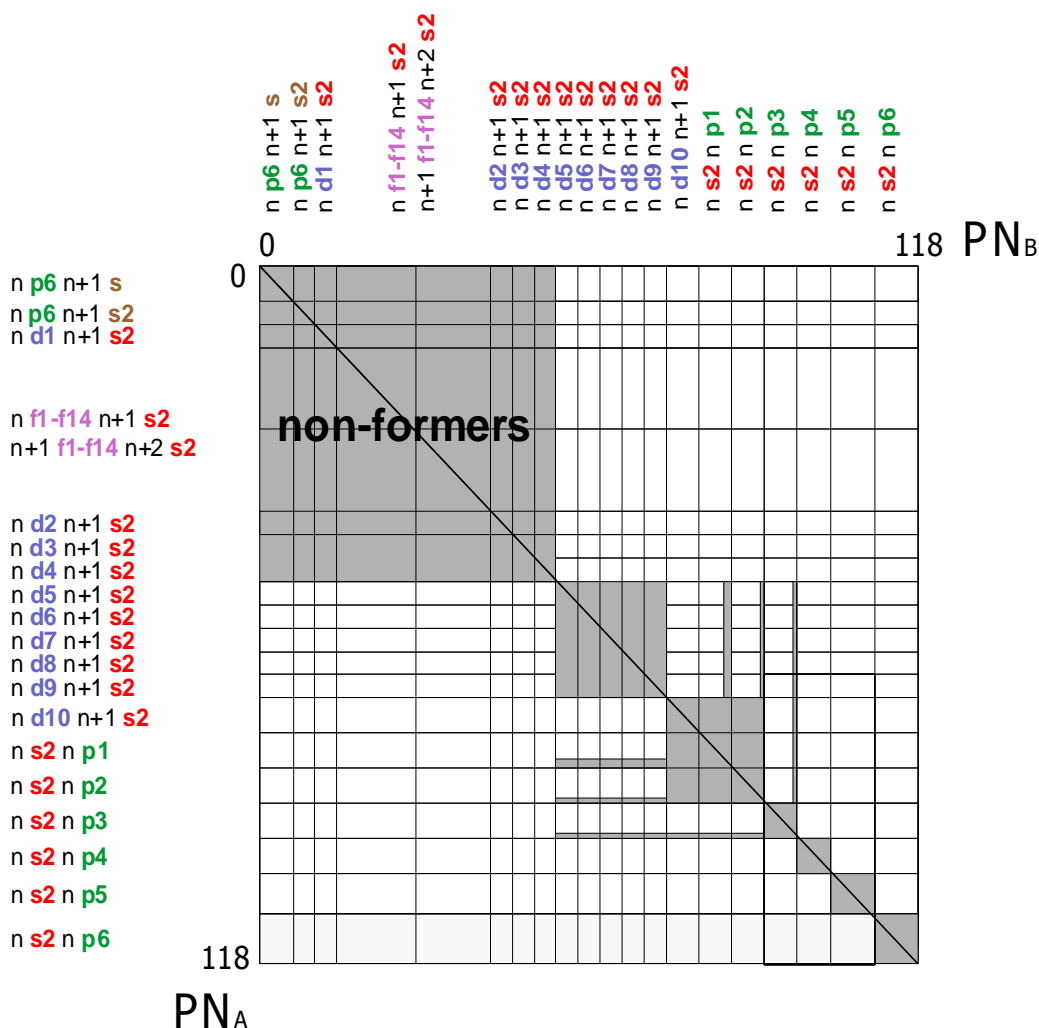


Fig. 5 Two-dimensional compound-formation map for all binary A - B systems using as axes: periodic number of the chemical element A (PN_A) versus the periodic number of the chemical element B (PN_B), highlighting the correlation to the valence electron configuration of the constituent chemical elements. The non-former areas are indicated with gray color.

structures having no mixed sites) with the following distribution: 238 (one chemical element), 6'856 (two), 17'083 (three), 4'414 (four), 600 (five), and 165 (six) (see also Table 4). It can be seen that for the chemical elements AETs with CN = 12 or higher are strongly preferred. Within binary and ternary compounds the full diversity of AETs is realized, with CN = 1 to CN = 22. In other words geometrically very different prototypes can be achieved with two or three different chemical elements within inorganic phases. Surprisingly this diversity is reduced when focusing on inorganic phases containing 4-6 chemical elements, and not extended as expected. The plot shows that for quaternary and higher-order inorganic phases AETs with CN = 6 or lower are highly preferred. The reason for this artifact can be seen from Fig. 7, where the frequency of the number of point sets (grouped in quaternary, quinary, sixnary compounds) are

plotted versus the nine most populous atomic environment types (AETs) (sorted by increasing coordination numbers (CN)) and versus the valence electron groups: s^1 , s^2 , p^1 , p^2 , p^3 , p^4 , p^5 , d^{1-10} , f^{1-14} -elements (i.e. columns 1, 2, 13, 14, 15, 16, 17, 3-12 of the Periodic System, lanthanides and actinides). In quaternary and higher-order inorganic compounds p - and/or s -elements (especially p^3 - p^5 elements) occupy the majority of the point-sets. d^{1-10} , as well as f^{1-14} elements show an overall low frequency, with the exception of point sets having as AET the tetrahedron (4-a) or the octahedron (6-a) for d^{1-10} elements as central atoms.

The summary of this newly discovered principle: The maximal atomic environment (AET) diversity is reached within binary and ternary inorganic compounds. Unaries (chemical elements) prefer AETs with high coordination number (CN = 12 or higher).

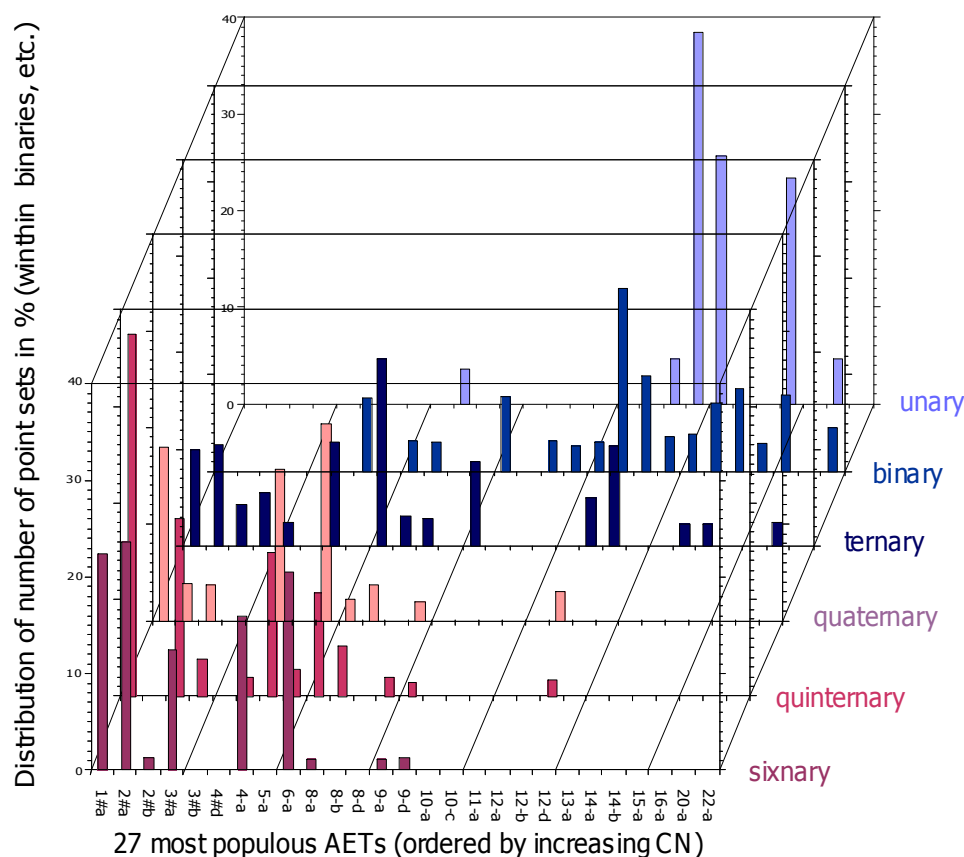


Fig. 6 Frequency of the number of point sets in percent (grouped in unary, binary, ternary, quaternary, quinternary, and sixnary phases) versus the 27 most populous atomic environment types (AETs), sorted by increasing coordination number (CN). The AET codes are explained in Table 3.

Quaternary and higher-order compounds strongly prefer AETs with low coordination number (CN = 6 or lower) with the central atoms being mainly s^1 , s^2 , p^1 , p^2 , p^3 , p^4 and p^5 elements. d^1 - d^{10} and/or f^1 - f^{14} elements. Multinary compounds predominately have as AET tetrahedra (CN = 4) or/and octahedra (CN = 6).

3) Active concentration range principle

This principle was formulated in 1994 [17] as follows: *Active composition range means that at least 5 at.% of minority elements is needed to form a compound.* This is valid for binary, ternary and quaternary inorganic systems. As a consequence, in binary systems 10%, in ternary systems 15% and in quaternary systems 38.6% of the available concentration range (area, volume) is 'inactive' for the formation of daltonide compounds (see gray area in Figs. 8 and 9). This principle has been fully confirmed.

4) Stoichiometric ratio condition principle

By now 72% of all binary systems have been investigated and we can state that 95% of all binary daltonide phases crystallize in one of the following 10 stoichiometric ratios: AB_x where $x = 6, 5, 4, 3, 2, 1.67, 1.5, 1.33, 1.25, \text{ and } 1$ (see Fig. 8). For ternaries, 20

years ago it was only possible to list 7 most often occurring stoichiometry ratios. Meanwhile 99% of the 17'083 daltonide ternary phases respect the following ternary stoichiometric ratio condition (see Fig. 9): AB_xC_y , where x takes one of the 10 above listed values for the most frequent binary stoichiometric ratios and y is equal to x or an integer divided by 1, 2, 3, or 4, and larger than x . In combination with the active composition range principle, this leads to 1 stoichiometric ratio for ABC ($x = y = 1$) compound, 9 possible stoichiometric ratios for AB_xC_x ($x = y$) and 819 possible stoichiometric ratios for AB_xC_y ($x < y$).

For the stoichiometry 1:1:1 there is only one possibility within a given ternary system $A-B-C$, but for 1: x : y , $x = y$ there are three possible compounds: AB_xC_x , A_xBC_x , and A_xB_xC , and for 1: x : y , $x \neq y$ six possibilities. This leads in total to 4'636 stoichiometric ratios assuming $A \neq B \neq C$, of which 490 have so far been found experimentally. It is worth mentioning that stoichiometric ratios following certain additional conditions are highly preferred: AB_xC_y , x equal one of the 10 above listed most frequent values for binary stoichiometric ratios, y equal x or an integer larger than x . With these conditions we end up with 769, compared with the 490 so far experimentally found stoichiometric ratios.

Table 3 50 most frequently occurring Atomic Environment Types in PCD-2013/14 [6].

Nr	Counts	AET-CODE	Name
1	239814	1#a	single atom
2	188107	2#a	non-collinear
3	143166	6-a	octahedron
4	137309	4-a	tetrahedron
5	83440	3#a	non-coplanar triangle
6	32717	12-b	cuboctahedron
7	24263	9-a	tricapped trigonal prism
8	21801	2#b	collinear
9	20129	12-a	icosahedron
10	14059	14-b	rhombic dodecahedron
11	13684	8-a	square prism (cube)
12	13548	3#b	coplanar triangle
13	12935	8-b	square antiprism
14	12763	5-a	square pyramid
15	11978	5-c	trigonal bipyramid
16	8970	7-g	monocapped trigonal prism
17	8270	14-a	14-vertex Frank-Kasper
18	7739	10-a	fourcapped trigonal prism
19	7648	6-b	trigonal prism
20	7507	16-a	16-vertex Frank-Kasper
21	7253	4#c	coplanar square
22	6099	7-h	pentagonal bipyramid
23	5895	13-a	pseudo Frank-Kasper (13)
24	5865	20-a	pseudo Frank-Kasper (20)
25	5262	12-d	anticuboctahedron
26	5136	4#d	non-coplanar square
27	4388	8-d	distorted square anti-prism, type a
28	3953	4#b	tetrahedron, central atom outside
29	3515	11-a	pentacapped trigonal prism
30	3481	6-d	pentagonal pyramid
31	3472	15-a	15-vertex Frank-Kasper
32	3286	17-d	7-capped pentagonal prism
33	3225	10-b	bicapped square prism
34	3021	8-g	double anti-trigonal prism
35	2731	18-a	eight equatorial capped pentagonal prism
36	2567	11-b	pseudo Frank-Kasper (11)
37	2469	10-c	bicapped square antiprism
38	2428	8-i	side-bicapped trigonal prism
39	2306	8-c	hexagonal bipyramid
40	2086	22-a	polarity, eight-equatorial capped hexagonal prism
41	2033	10-e	distorted equatorial four-capped trigonal prism
42	1937	5#d	square pyramid, central atom outside of base plane
43	1768	8-j	distorted square anti-prism, type b
44	1735	12-f	hexagonal prism
45	1147	14-d	bicapped hexagonal prism
46	1022	7-a	monocapped octahedron
47	945	18-d	sixcapped hexagonal prism
48	911	20-h	twelve pentagonal faced polyhedron
49	910	6-h	distorted trigonal prism
50	776	10-j	polarity bicapped square prism

Table 4 Number of chemical elements per prototype (rows) *versus* number of chemical elements of its representatives (columns) considering the 1'000 most populous prototypes and their representatives in PCD-2013/14 [6].

Number of chemical elements per prototype	Number of chemical elements per representative							no. distinct phases
	1	2	3	4	5	6	>6	
1	238	(1218)	(493)	(88)	(23)	(10)	(1)	2 071
2	0	6 856	(9538)	(1 494)	(234)	(51)	(30)	18 203
3	0	0	17 083	(8 634)	(2 481)	(365)	(107)	28 670
4	0	0	0	4 414	(2 134)	(503)	(185)	7 236
5	0	0	0	0	600	(252)	(100)	952
6	0	0	0	0	0	165	(221)	386

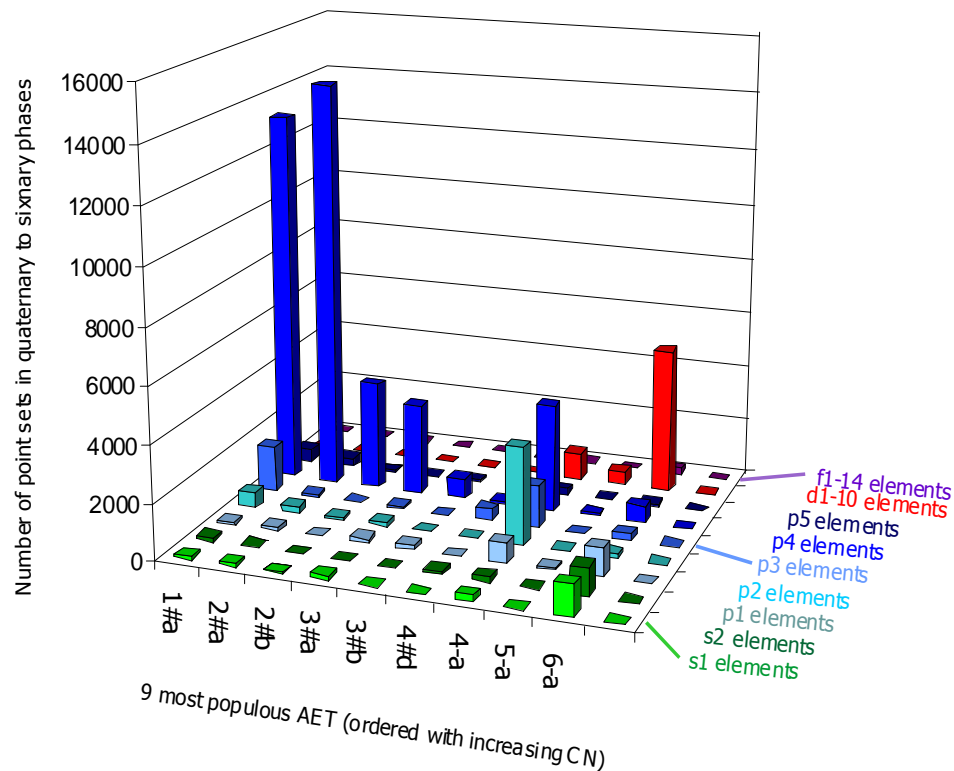


Fig. 7 Number of point sets (grouped in quaternary, quinternary, sixnary phases) *versus* the 9 most populous atomic environment types (AETs), sorted by increasing coordination number (CN), *versus* element group: $s^1, s^2, p^1, p^2, p^3, p^4, p^5, d^{1-10}, f^{1-14}$ -elements. AET: 1#a, single atom; 2#a, non-collinear; 2#b, collinear; 3#a, non-coplanar triangle; 3#b, coplanar triangle; 4#d, non-coplanar square; 4-a, tetrahedron; 5-a, square pyramid; 6-a, octahedron.

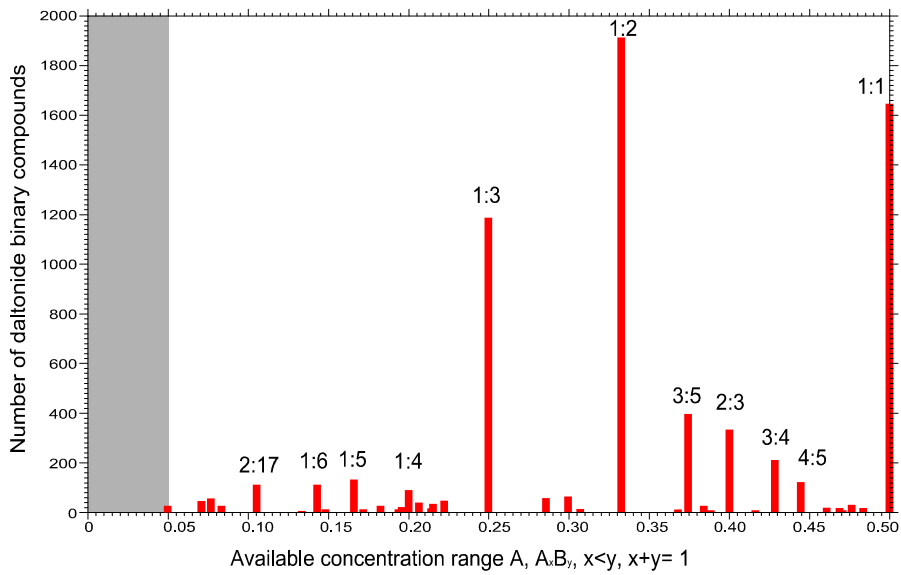


Fig. 8 Number of binary daltonide inorganic phases A_xB_y ($x < y$, $x+y = 1$) versus x . Numbers indicate frequent $m:n$ ratios (A_mB_n where m and n are integers); the gray color shows the composition range where no binary phases occur.

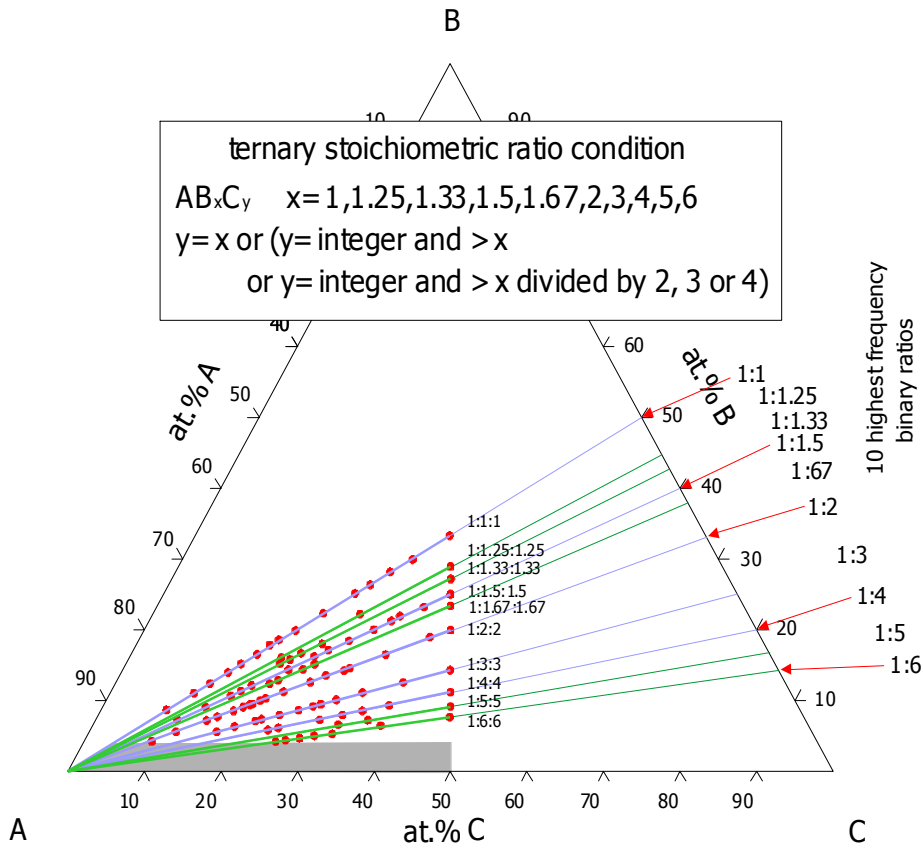


Fig. 9 Occurrence of daltonide inorganic phases in the available concentration range for ternary systems ($A_xB_yC_z$, $x < y < z$, $x+y+z = 1$); the gray area shows where no ternary phases occur.

The stoichiometric ratio condition principle was formulated in 1994 [17] as follows: *The vast majority of daltonide compounds have each point set occupied by just one kind of chemical element. This leads for binaries to the following highly preferred stoichiometric ratios: 1:1, 4:5, 3:4, 2:3, 3:5, 1:2, 1:3, 1:4, 1:5, and 1:6.* This has been fully verified in this work. For ternary intermetallic compounds, the stoichiometric ratios 1:1:1, 1:1:2, 1:1:3, 1:1:4, 1:2:2, 1:2:3, and 1:2:4 predominate. In this work the above given ternary stoichiometric ratio condition was discovered.

The four principles demonstrate with the help of Figs. 6-9, based on data of *Pearson's Crystal Data* [6] and *ASM Alloy Phase Diagram Database* [7], that the overall conclusion of the second cornerstone is correct: *"Nature provides us with a very rigid systematic framework of restraints, such as laws that define compound formation."*

Most efficient in reducing the number of potential chemical element combination is the fact that we can focus on unary to quaternary systems, since with them the full diversity of AETs (as well as prototypes) is achieved. Quinary and higher-order systems will not lead to inorganic compounds with greater diversity, and therefore not to the discovery of additional principally different geometrical atom arrangements. In other words, all principal prototypes are realized with compounds having four or less chemical elements.

The number of potential chemical element combinations (chemical systems) for unaries to quaternaries is 4'087'975. About 30% of these systems are non-formers, which leaves about 2'900'000 combinations. Taking the average number of inorganic phases for unaries as 5, for binaries as 7, for ternaries as 4 and quaternaries as 2, leads to less than 6'000'000 inorganic phases, preferring one of 10 stoichiometric ratios for binaries, or obeying the stoichiometric ratio condition for ternaries.

Taking into consideration the fact that there exist only four major groups of chemical elements (*s*-, *p*-, *d*- and *f*-elements) supports the above said. From the chemical point of view the quaternaries should represent the maximum of potential diversity that can be achieved. The basic prototypes and their basic representatives among quinary and higher-order inorganic compounds cannot be built up with chemical elements from five (or more) distinct different chemical element groups, since there exist only four groups, but one or more chemical elements will have to belong to the same group of elements.

3.3. Third cornerstone: Laws that define ordering of chemical elements within a structure type

The third cornerstone of Nature sets strict restrictions on the ordering of the chemical elements within a particular inorganic compound. When chemical elements combine to form solids, their crystal

structures are beautifully rich, yet systematic patterns underlie this process.

The most striking manifestation of this fact is the existence of so-called prototypes of crystalline inorganic compounds, which can be understood as geometrical templates for large groups of inorganic compounds, e.g. the NaCl,cF8,225 prototype has presently 1299 different basic representatives. In other words, different inorganic compounds representing the same prototype are geometrically very similar to each other. There exist, in principle, two main approaches to the classification of crystal structures, one considers as first criterion the overall symmetry (e.g. the Wyckoff sequence [11]), the other one the atomic environments, also called coordination polyhedra, of each site (leads to the atomic environment type AET classification) [22], or of part of the elements [23]. The first classification requires that the published crystallographic data are fully standardized [12], which is consequently done in the PAULING FILE [3-5] and the products derived from it, e.g. Pearson's Crystal Data [6].

The following four principles demonstrate, with the help of four statistical plots and one fundamental crystallographic consideration, the correctness of the third cornerstone.

5) Simplicity principle

Fig. 10 shows that the large majority of all prototypes of inorganic phases have less than 40 atoms per unit cell, with a maximum of types having around 8 atoms per cell. It is also demonstrated that the majority of all prototypes (and therefore also their representatives) have 3 or fewer atomic environment types (AETs), with a maximum around 2-3 AETs per prototype. The number of point sets per prototype is for the majority lower than 6, with a maximum around 2-3 point sets per prototype (see Fig. 11).

This principle was formulated in 1994 [17] as follows: *The vast majority of the intermetallic compounds have less than 24 atoms per unit cell.* By the inclusion of other inorganic compounds 24 has become 40, nevertheless the average stays at around 10 atoms per unit cell. *In addition the vast majority of all crystal structures have three or fewer atomic environment types (AETs) within the crystal structures (single-, two-, and three-environment types).* This is still supported. An analog observation can also be made focusing on the number of point sets instead of the number of different AETs.

6) Symmetry principle

Fig. 12 gives the number of inorganic phases distributed by space group number considering all entries of PCD-2013/14 [6].

The symmetry principle was formulated in 1994 [17] as follows: *The vast majority of all intermetallic compounds and alloys crystallize in one of the following 11 space groups: 12, 62, 63, 139, 166, 191, 194, 216, 221, 225, and 227.* By extending to other

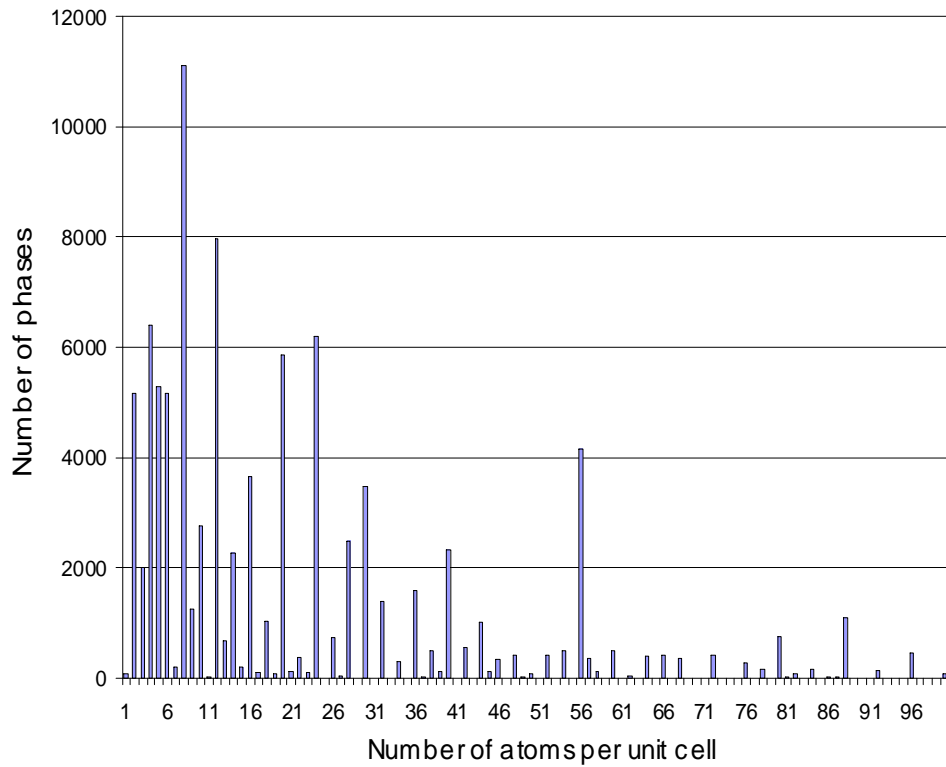


Fig. 10 Number of inorganic phases *versus* number of atoms per unit cell considering the 1'000 most populous prototypes and their representatives in PCD-2013/14 [6].

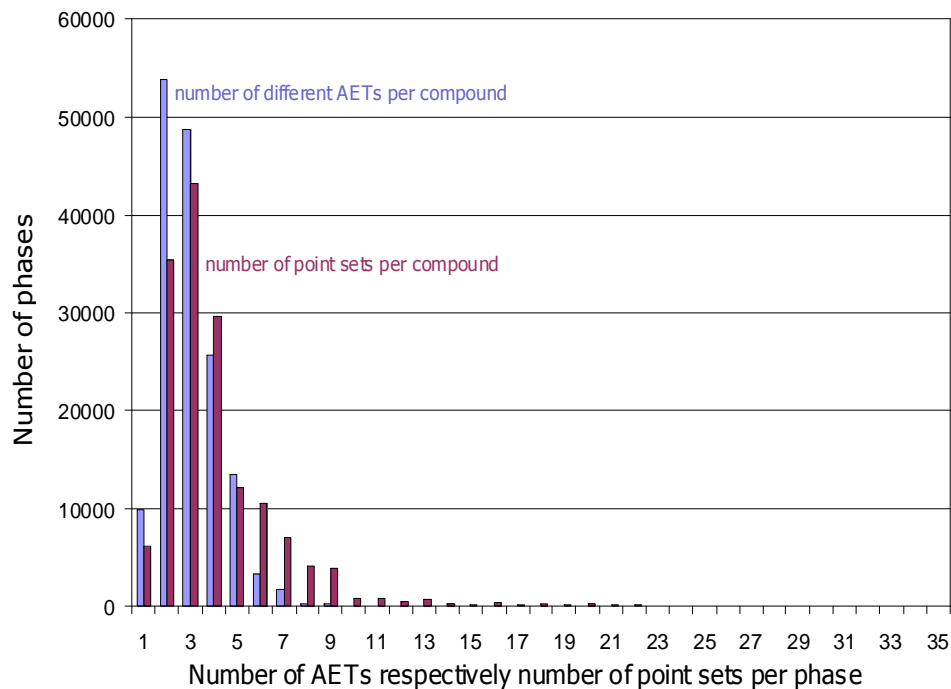


Fig. 11 Number of inorganic phases *versus* number of different AETs in the structure, respectively number of point sets in the structure, considering the 1'000 most populous prototypes and their representatives in PCD-2013/14 [6].

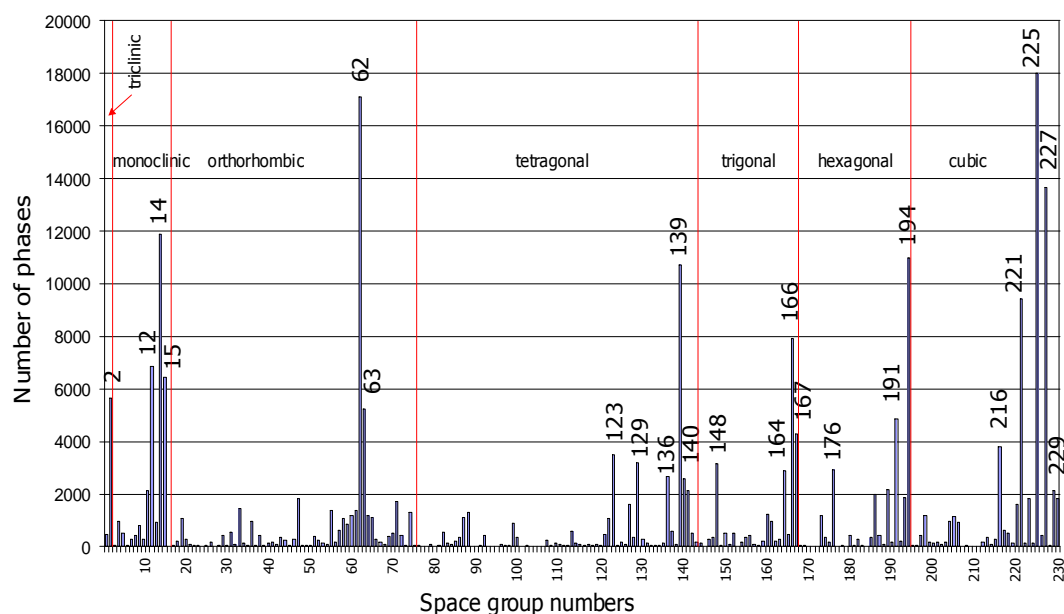


Fig. 12 Number of inorganic phases distributed by space group number considering all entries in PCD-203/14 [6].

inorganics the following space group numbers have to be added: 2, 14, 15, 123, 129, 136, 140, 148, 164, 167, 176, 229. Within each crystal system a few space groups (the ones with higher symmetry) are preferred. 10% of the space groups cover 67% of the inorganic compounds (160'000 of 237'000 compounds).

7) Atomic-environment principle

Fig. 13 shows a frequency plot for the 18 most populous AETs considering the 1'000 most populous prototypes and their representatives of the PCD-2013/14 [6]. **Table 3** lists the 50 most populous AETs among all the refined compounds of PCD-2013/14 [6].

The atomic-environment principle was formulated in 1994 [17] as follows: *The vast majority of all atoms (point sets) in intermetallic compounds have as atomic environment one of 14 polyhedra: tetrahedron, octahedron, cube, tri-capped trigonal prism, four-capped trigonal prism, icosahedron, cubooctahedron, bi-capped pentagonal pyramid, anti-cubooctahedron, pseudo Frank-Kasper (CN13), 14-vertex Frank-Kasper, rhombic dodecahedron, 15-vertex Frank-Kasper, and 16-vertex Frank-Kasper polyhedron.* The statement made 20 years ago that 14 AETs are highly preferred is still correct. After having added other classes of inorganic compounds to the intermetallics and alloys, the frequency order has changed (see also **Table 3**), and the following AETs with low coordination numbers: single atom (CN = 1), collinear (CN = 2), non-linear (CN = 2), collinear triangle (CN = 3), non-coplanar triangle (CN = 3) and square

anti-prism (CN = 8) are now among the 14 most populous AETs. In addition it can be stated that 18 out of 100 possible AETs are highly preferred, and were found for 90% of the point sets considered here (1'002'340 of 1'117'109 point sets). Consequently it appears that Nature strongly prefers certain AETs, most of them being highly symmetrical (except the AETs with CN = 1 and CN = 2).

8) Chemical element ordering principle

Table 4 gives the number of chemical elements per prototype (rows) versus the number of chemical elements of its representatives (columns), considering the 1'000 most common prototypes and their representatives in PCD-2013/14 [6]. Here we recall what we understand under basic prototype a prototype that has no mixed site occupancy.

Looking at the numbers of **Table 4** it is possible to conclude that:

- Compounds with $n-1$ chemical elements cannot be representatives of a basic prototype containing n chemical elements. This is a particular feature of the prototype classification used here, where ordered isopointal structures are distinguished (CaCu₅,hP5,191 and PrNi₂Al₃,hP5,191 or CeCo₃B₂,hP5,191).
- Compounds with n chemical elements that are representatives of a basic prototype containing n chemical elements have no mixed sites (and are therefore called basic representatives).
- Compounds with $n+m$ (where m is an integer > 0) chemical elements that are representatives of a prototype containing n chemical elements have mixed

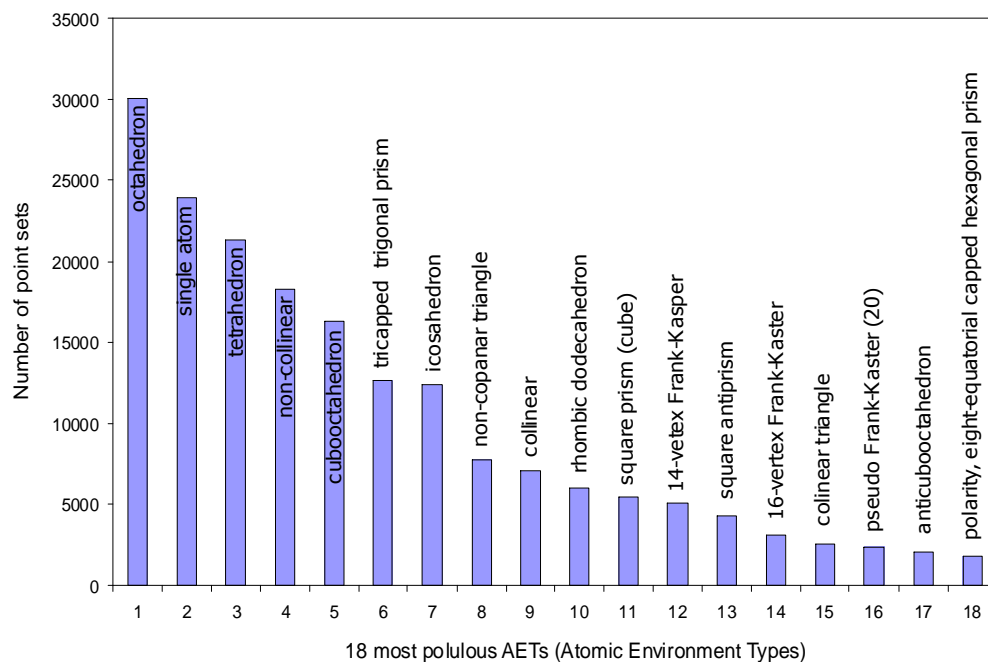


Fig. 13 Total number of point sets for the 18 most frequent AETs considering the 1'000 most populous prototypes and their representatives in PCD-release 2013/14 [6].

sites and generally belong to a solid solution of the basic prototype or one of the basic representatives with n chemical elements.

The 1'000 most populous prototypes and their representatives cover 157'308 entries in PCD-release 2013/14 [6], but represent only 57'518 distinct phases (chemical system + crystal structure). Among the 57'518 distinct phases, only 29'356 have no mixed sites. This means phases (prototypes or representatives) where the number of chemical elements n is the same as for the prototype. The remaining 28'162 phases are representatives of basic prototypes, but have mixed occupation sites, and are in general solid solutions of basic prototypes or their basic representatives. Most of them are therefore not distinct new inorganic compounds. For example, replacement of a few at.% of the chemical element A in an ABC compound by a closely related chemical element A' will lead to a quaternary representative $(A,A')BC$.

The four principles confirm, with the help of Figs. 10-13 and Table 4, the content of the third cornerstone: "When chemical elements combine to form solids, their crystal structures are beautifully rich, yet very systematic patterns underlie this process." The systematic patterns lead to restraints, nicely reflected in the below listed four principles:

- overall simplicity preferred,
- high overall symmetry preferred,
- high local symmetry reflected through the preference of regular AETs,
- high ordering tendency

The combination of the above given experimental facts reduces the number of potential prototypes to be considered in first priority to a few hundred, just 1-2% of the experimentally known prototypes. The frequency plot in Fig. 14 shows the number of representatives for the 100 most populous prototypes of PCD-2013/14 [6]. The majority of the 32'000 experimentally known prototypes (> 25'000) are unique in the sense that they have no representatives. One of the reasons for the high number of unique prototypes is the significant number of refinements with split sites or other sites very low occupancy, each distinct combination corresponding to a different prototype. For example, almost every refinement of a specific deuteride or hydride phase results in a new prototype, even if the phase is the same.

3.4. Fourth cornerstone: Laws that link the position of chemical elements within a structure type ↔ in the Periodic System

The fourth cornerstone of Nature states that there exist direct links between the position of the constituent chemical element in the Periodic System and its point set occupation within the structure of an inorganic compound.

By definition, inorganic prototypes and their representatives are geometrically very similar to each other. A closer inspection of known inorganic compounds reveals that, for each of the 32'000 prototypes, only a relatively small subset of potential representatives is actually known. The above-mentioned link may give us a possibility to predict

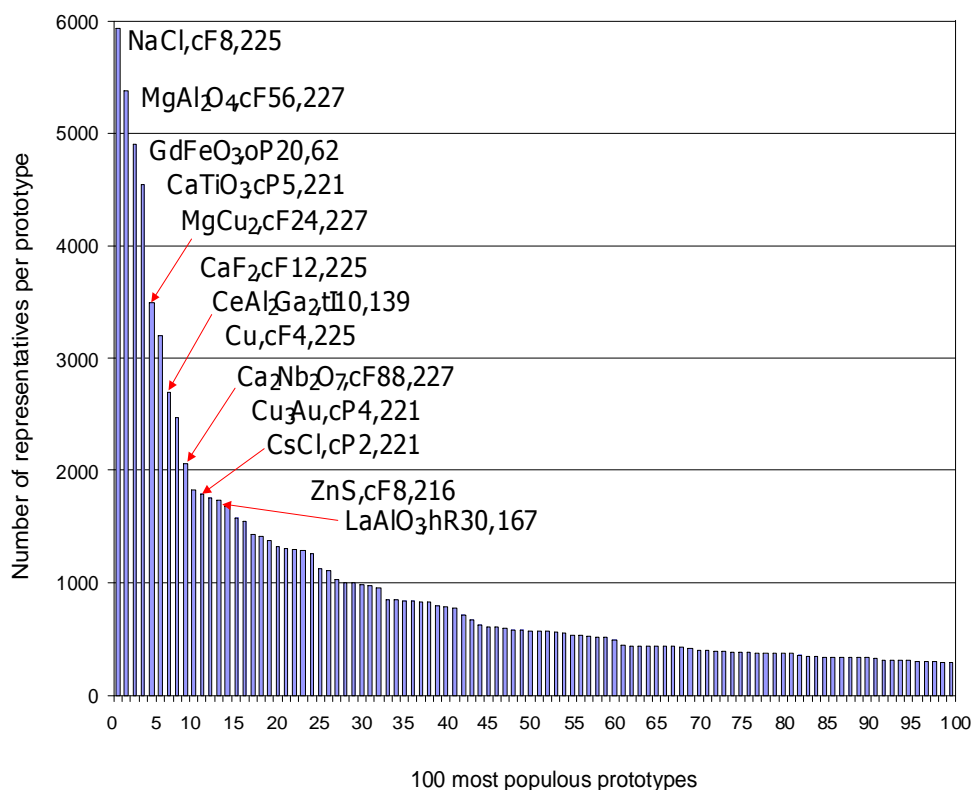


Fig. 14 Number of representatives of each of the 100 most populous prototypes in PCD-release 2013/14 [6].

which compounds will adopt a specific prototype (or a limited selection of prototypes). This is strongly supported by the following four principles:

9) *Prototype – Periodic System correlation (chemistry) principle*

This principle was formulated in 1994 [17] as follows: *The vast majority of crystal structures show a very strict regularity between the position of the chemical element in the Periodic System (s, p, d, and f elements) and its point set(s) occupation within a crystal structure.*

By looking at the 1'000 most populous prototypes we can confirm the above statement. About 50% of the prototypes, in general the ones with few representatives, reveal a simple correlation leading to a few 100 potential combinations. The other 50%, having in general many known representatives, show a broader correlation, leading to several thousand extrapolated combinations. Fig. 15 demonstrates the chemistry principle with the help of a simple point set – Periodic System correlation for the ErIr₃B₂,mS12,12 prototype, where the columns of the Periodic System to which the elements occupying the point sets in the prototype and its representatives belong, have been highlighted. This correlation leads to $6 \times 4 \times (4 \times 2) = 144$ element combinations (1:2:3), of which 23 have so far been found experimentally. Fig. 16 shows an analog correlation for a more common structure type, NaFeO₂,hR12,166.

This correlation leads to $(40 \times 103) / (1 \times 2) = 2'060$ combinations (1:1:2). 239 basic inorganic compounds are already known. With the help of these two examples, it is seen that there exists a clear link between the position of the chemical element in the structure and its position in the Periodic System.

10) *Structure stability map principle*

This principle was formulated in 1994 [17] as follows: *The size, electrochemical, valence electron, and atomic-number factors are the factors governing crystal structures of intermetallic compounds.* Structure maps separate intermetallic compounds into distinct prototype domains. There exists a whole range of different structure maps for binary and ternary intermetallic compounds. The principle has been further verified and significantly improved since the introduction of the periodic number (PN) in 2008 [13]. The by far simplest and most efficient structure stability maps use PN (and/or expressions of PN) as elemental-property parameter. The best combination was found to be a two-dimensional PN_{\max} versus mean PN_{\min} / PN_{\max} map (for a definition of mean PN_{\min} / PN_{\max} see [13]), which has the following five advantages:

- i) Inclusion of non-formers, which nicely separate from the formers, is possible.
- ii) Applicable to binary, ternary and quaternary systems.

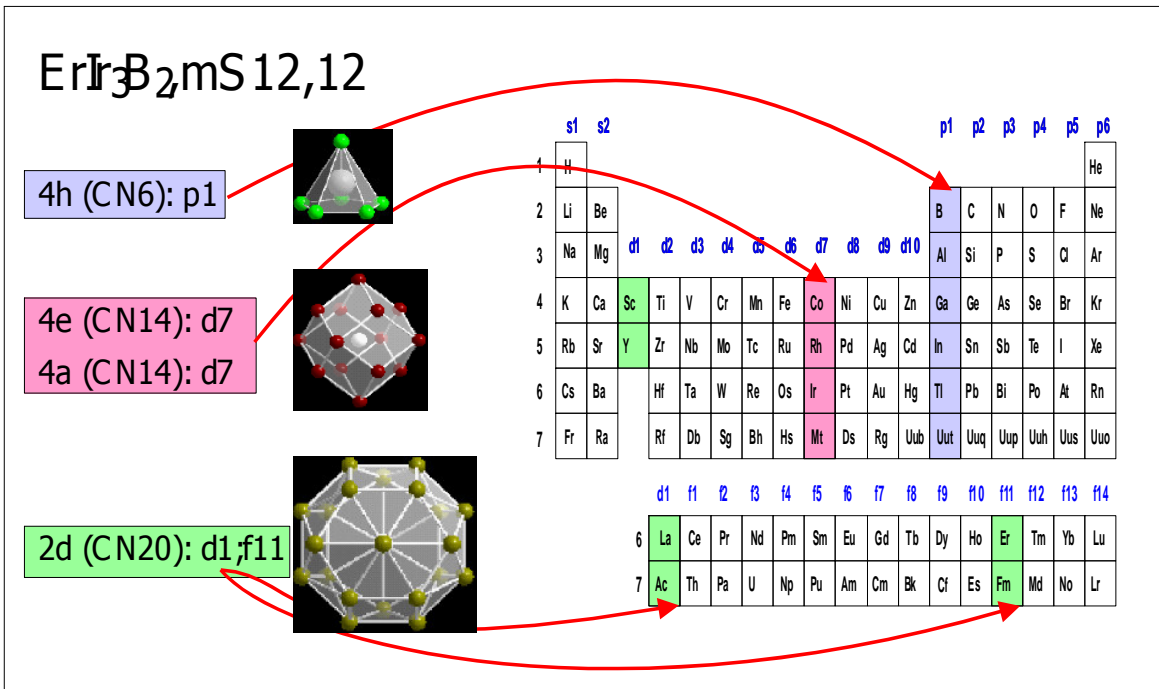


Fig. 15 Chemistry principle shown with the help of a ‘simple’ point set – *spdf*-element correlation showing the ErIr₃B₂mS_{12,12} prototype.

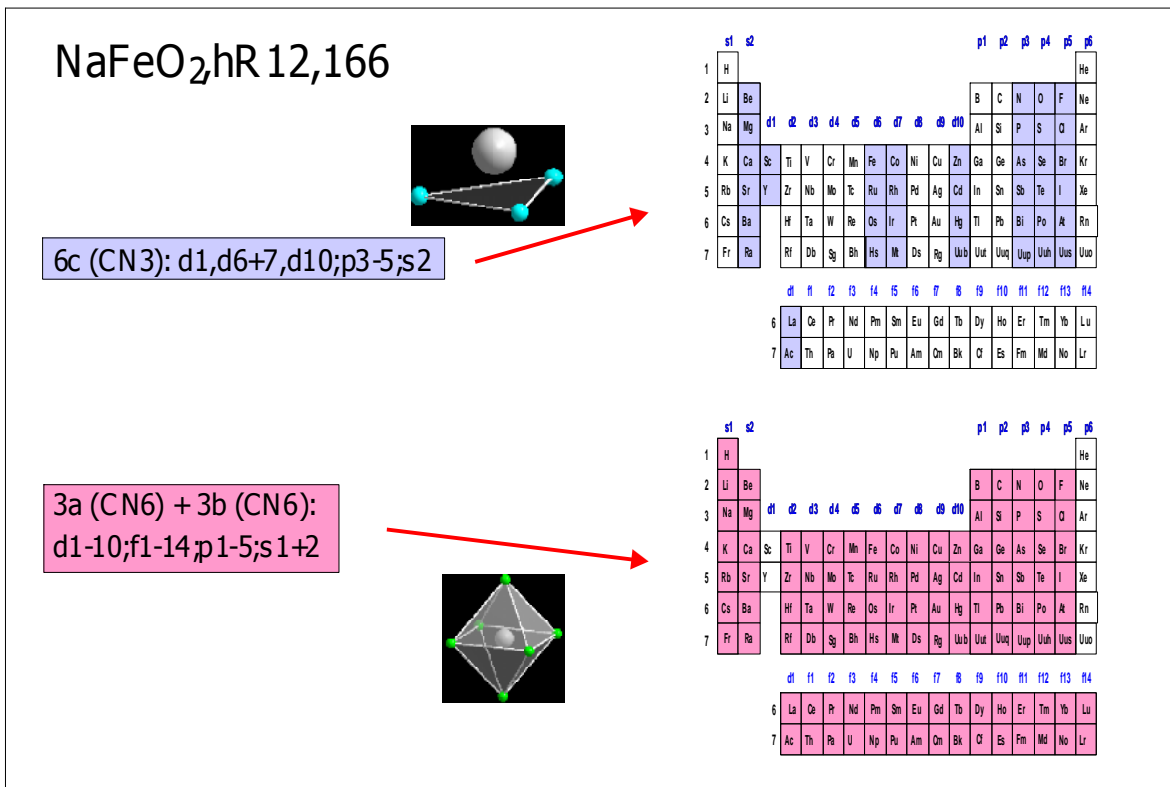


Fig. 16 Chemistry principle shown with the help of a ‘complex’ point set – *spdf*-element correlation showing the NaFeO₂hR12,166 prototype.

iii) Iso-stoichiometric prototypes and their representatives can be separated into distinct prototype stability domains in the same map, *e.g.* binary 1:3 map, ternary 1:1:4 map.

iv) Two-dimensional maps are easy to read and provide obvious prediction ability.

v) PN is an integer number, known for all chemical elements, and no inaccuracy is coming from the choice of the actual values of the elemental-property parameter.

This kind of map proved to be very efficient in revealing patterns for different materials problems,

e.g. former *versus* non-former systems; iso-stoichiometric structure stability maps; or complete solid solubility between binary compounds having the same prototype. It can also separate different atomic environments AETs into distinct stability AET domains in a generalized composition-independent map including binary and multinary compounds [15].

Fig. 17 shows a structure stability map for basic ABC_2 prototypes and their basic representatives, focusing on the following three prototypes: $NaFeO_2$, hR12,166; $CuFeS_2$, tI16,122; Cu_2MnAl , cF16,225. The majority of the

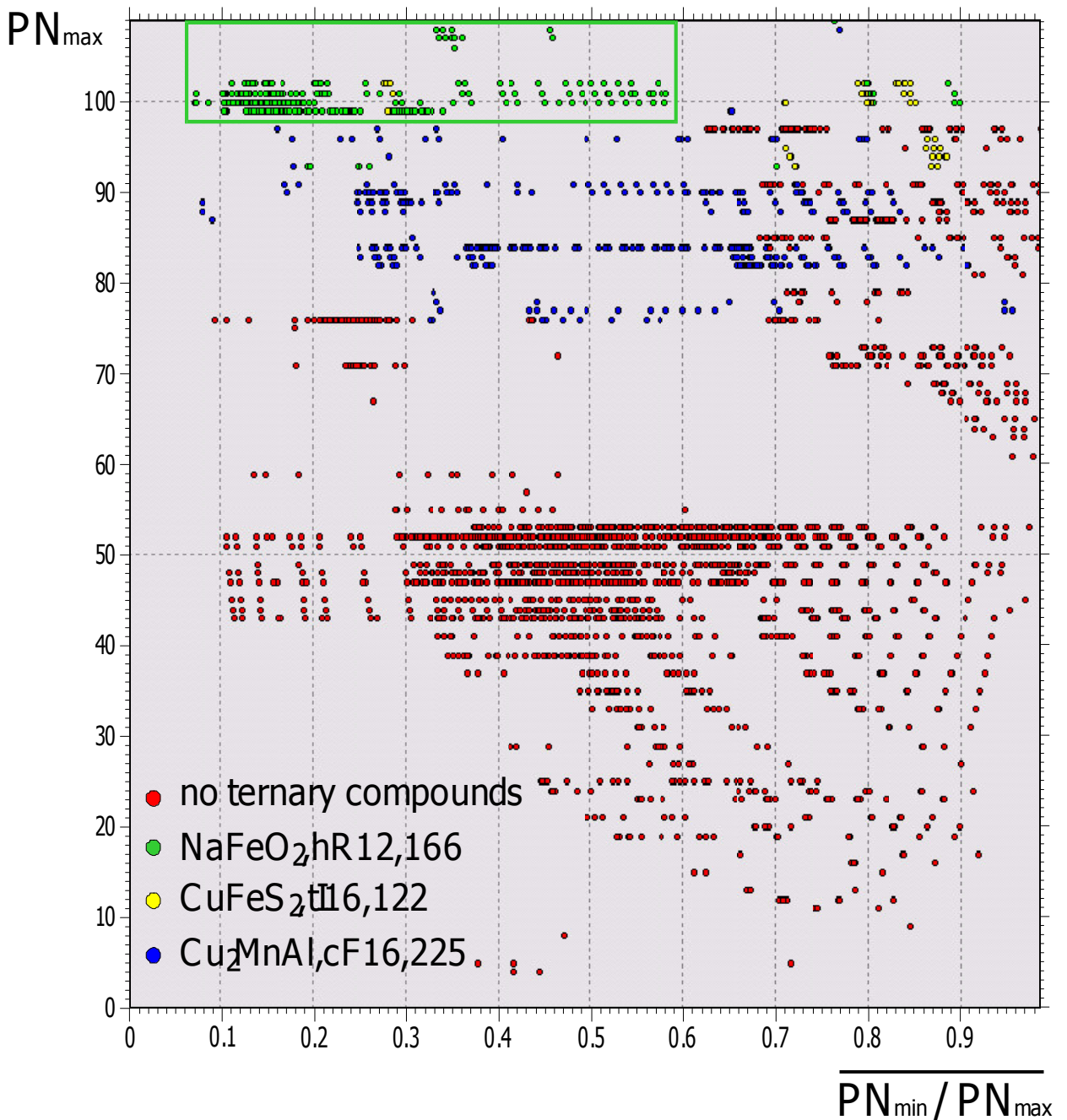


Fig. 17 Structure stability map for ABC_2 prototypes focusing on the following three prototypes: $NaFeO_2$, hR12,166; $CuFeS_2$, tI16,122; Cu_2MnAl , cF16,225 including all ternary representatives in PCD- 2013/14 [6].

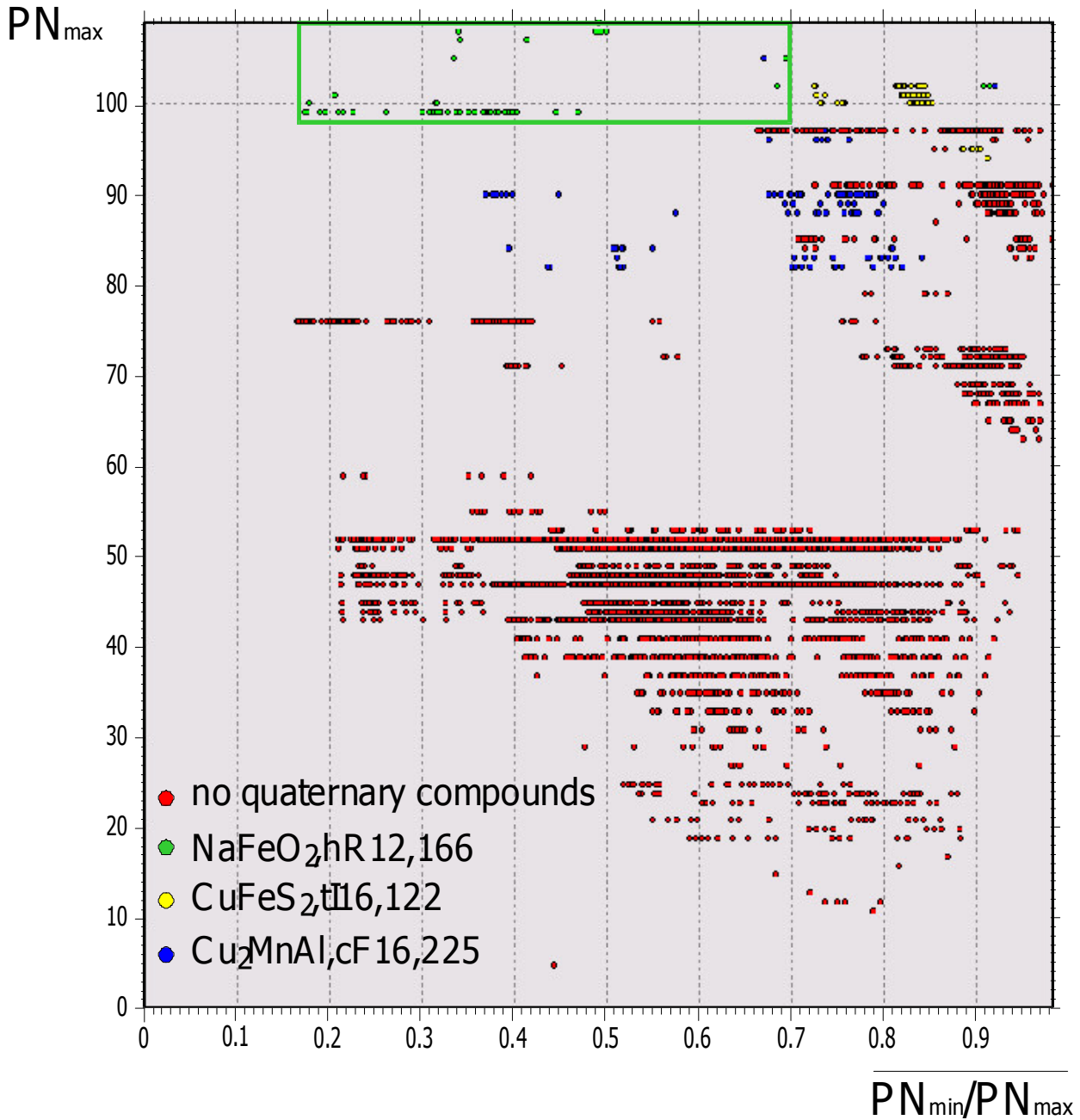


Fig. 18 Structure stability map for ABC_2 prototypes focusing on the following three prototypes: NaFeO₂,hR12,166; CuFeS₂,tI16,122; Cu₂MnAl,cF16,225 including all quaternary representatives in PCD- 2013/14 [6].

experimentally known 239 basic ternary representatives of the NaFeO₂,hR12,166 prototype are nicely located in the domain outlined by a green rectangle. Fig. 18 demonstrates that the same kind of structure stability map can also be used for quaternary ‘non-basic’ representatives (with mixed site occupancies) of the above-specified basic ABC_2 prototypes. From the two examples it is seen that there exist clearly defined domains for the representatives of the different prototypes within such iso-stoichiometric

structure stability maps. Structure stability maps having as axis expressions of the elemental-property parameter PN provide a direct link to the Periodic System.

11) Generalized AET stability map principle

This principle was formulated by us in 2008 [15]. Here we give just a short summary: The atomic environment types AETs (coordination polyhedra) observed in binary and multinary inorganic

compounds were analyzed based on a comprehensive set of literature data. A generalized AET stability map, using as coordinates PN_{\max} versus PN_{\min}/PN_{\max} , successfully sub-divided the [central atom–coordinating atoms] combinations where different atomic environment types occur, into distinct ‘AET class stability domains’. It was found that chemical elements with $PN > 54$ control the atomic environment types, independently of whether they act as central or as coordinating atoms. The same stability map also showed a clear separation between possible and impossible [central atom–coordinating atoms] combinations. Such stability maps make it possible, for a chemical element A supposed to act as central atom, to predict the AET formed by any coordinating chemical element B (C , D), independently of its concentration and the number of chemical elements involved (see Fig. 19).

The same two-dimensional PN_{\max} versus (mean) PN_{\min}/PN_{\max} coordinates separate, on the one hand different atomic environments AETs into distinct AET stability domains in a generalized composition-independent structure map including binary and multinary compounds, and on the other hand iso-stoichiometric prototypes and their representatives into distinct prototype stability domains. This means that these maps can separate element combinations into distinct stability domains, either using the atomic environment type classification (taking each atom as the center) or the prototype classification (which uses as first criterion the Wyckoff sequence). In other words in the first map we focus on the atomic environment of each point set within a prototype and its representatives, in the second map we focus on the crystal structure as an entity. The above said can be seen by comparing Fig. 19 with Fig. 20. Fig. 20 is a

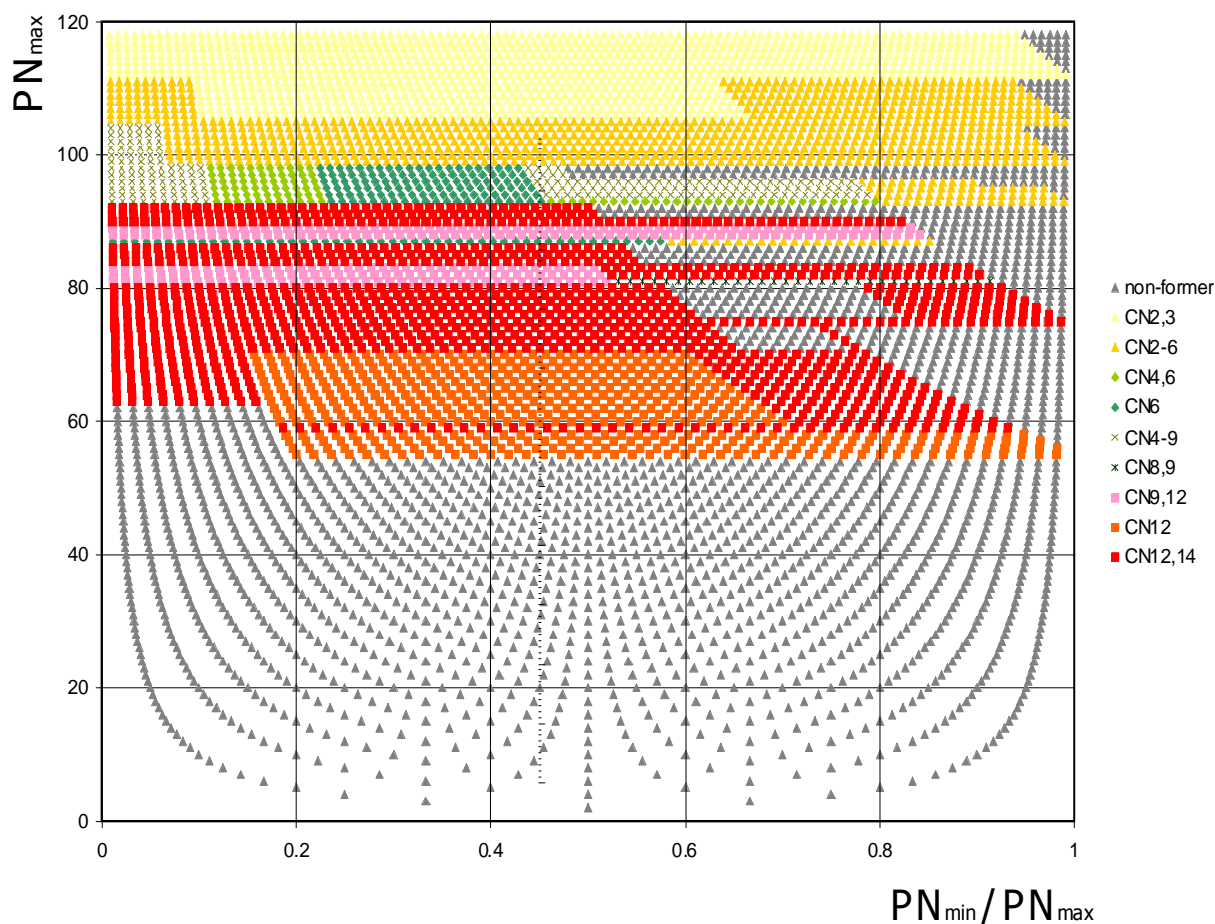


Fig. 19 Generalized AET stability map PN_{\max} versus PN_{\min}/PN_{\max} , which is independent of the stoichiometry and the number of chemical elements within an inorganic compound (based on 65'000 inorganic compounds having a [central atom–coordinating atoms]–AET occurrence higher than 3). The map is given for the central atom being PN_{\max} .

structure stability map for basic quaternary ABC_2D_4 prototypes focusing on the following three prototypes: $H_2K[PO_4]$, tI24,122; $NdBi_2ClO_4$, tP8,123; Cu_2CdSiS_4 , oP16,31 including all quaternary basic representatives. Furthermore, by comparing Figs. 19 and 20 it is seen that basic quaternary prototypes (and representatives) only occur if PN_{max} is 90 or larger. This means that at least one of the four constituents has to be one of the following chemical elements: Ge, Sn, Pb, N, P, As, Sb, Bi, O, S, Se, Te, Po, F, Cl, Br, I,

At (p^2-p^5 elements). This also supports the conclusion that the diversity of AETs is very much reduced for quaternaries, as compared to binaries and ternaries.

12) Complete solid solution stability map principle

This principle was formulated in 1994 [17] as follows: *The size, electrochemical, and valence electron factors control solid solubility.* Solid-solubility maps were found to separate regions of limited and extended solid solubility for a given chemical element

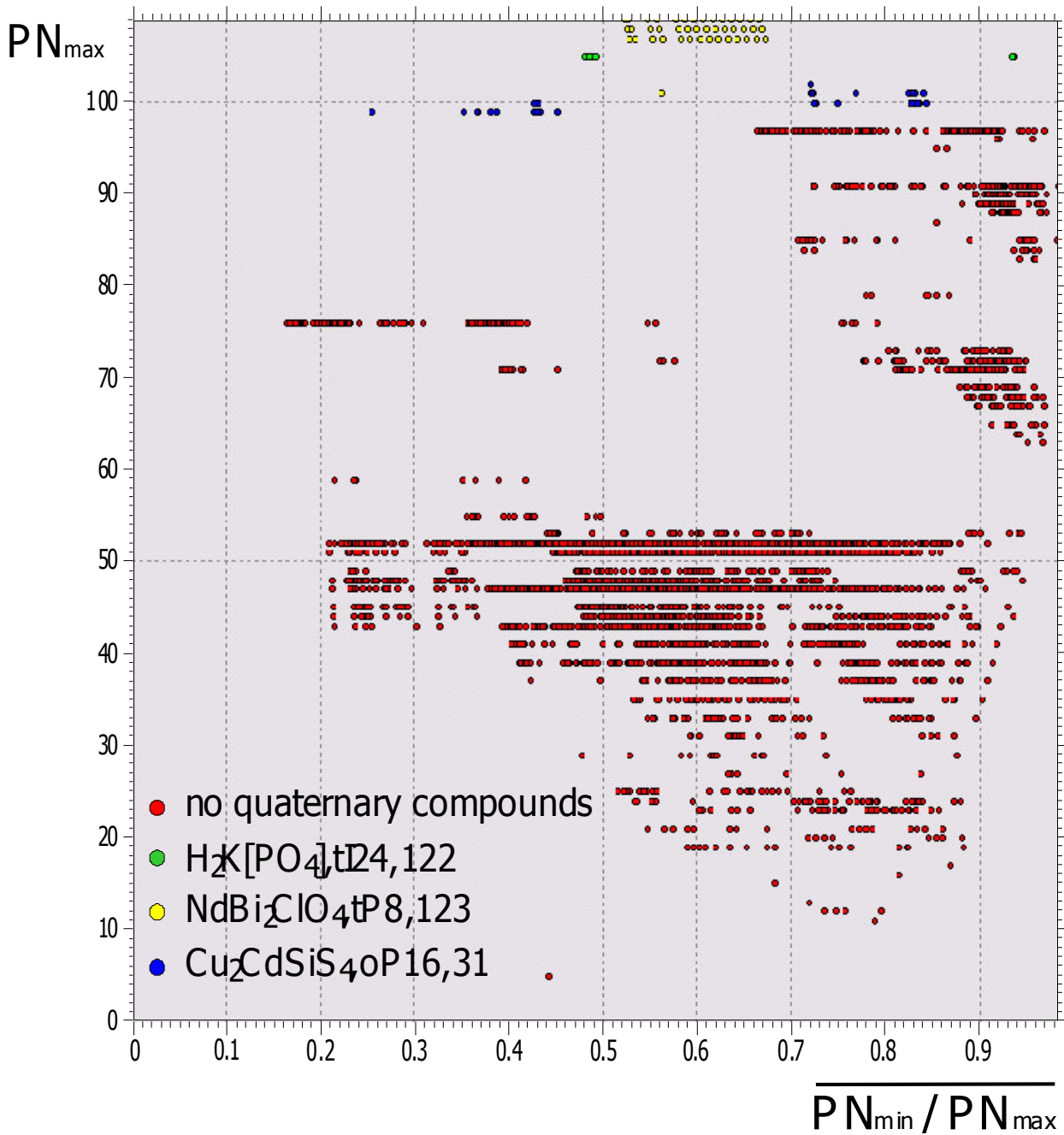


Fig. 20 Structure stability map for quaternary ABC_2D_4 prototypes focusing on the following three prototypes: $H_2K[PO_4]$, tI24,122; $NdBi_2ClO_4$, tP8,123; Cu_2CdSiS_4 , oP16,31 including all quaternary representatives in PCD-2013/14 [6].

solvent. In this work we have extended the study of solid solution behavior from chemical elements to binary compounds as solvents. Since the amount of solid solution data on binaries as solvents is very limited, we focused on a very extreme situation. We searched for ternary systems where the same prototype occurs in two of the binary boundary systems, and either a complete solid solution or limited solid solutions are formed. This requires having access to either appropriate phase diagram data or information about cell parameter(s) *versus* concentration. We selected the prototype MgCu_2 , cF24, 227 which has 238 well established binary representatives. For a 1:2 compound the complete solid solution can be either of the type $A(B,C)_2$ or $(B,C)A_2$. From the 239 binary compounds one can generate 701 $A(B,C)_2$ and 1'782

$(B,C)A_2$. Considering experimental data showing complete solid solutions, we found 61 $A(B,C)_2$ and 128 $(B,C)A_2$ cases, using data from [6,7].

The prototype MgCu_2 , cF24, 227 has two point sets, occupying Wyckoff positions $8b$ and $16c$ in the standardized description. The $16c$ site has as AET an icosahedron with $\text{CN} = 12$, and the $8b$ site a 16-vertex Frank-Kasper polyhedron with $\text{CN} = 16$. Since it is unlikely that the same chemical element will occupy sites with so different AETs, the following six potential solid solutions can be excluded: $A_2C \leftrightarrow AB_2$, $A_2C \leftrightarrow BC_2$, $AB_2 \leftrightarrow BC_2$, $A_2B \leftrightarrow AC_2$, $AC_2 \leftrightarrow B_2C$, $B_2C \leftrightarrow A_2B$ (see Fig. 21). The PN_{max} *versus* $\text{PN}_{\text{min}} / \text{PN}_{\text{max}}$ map in Fig. 22 focuses on the $A(B,C)_2$ solid solution case. Here we found experimental data for 42 systems where

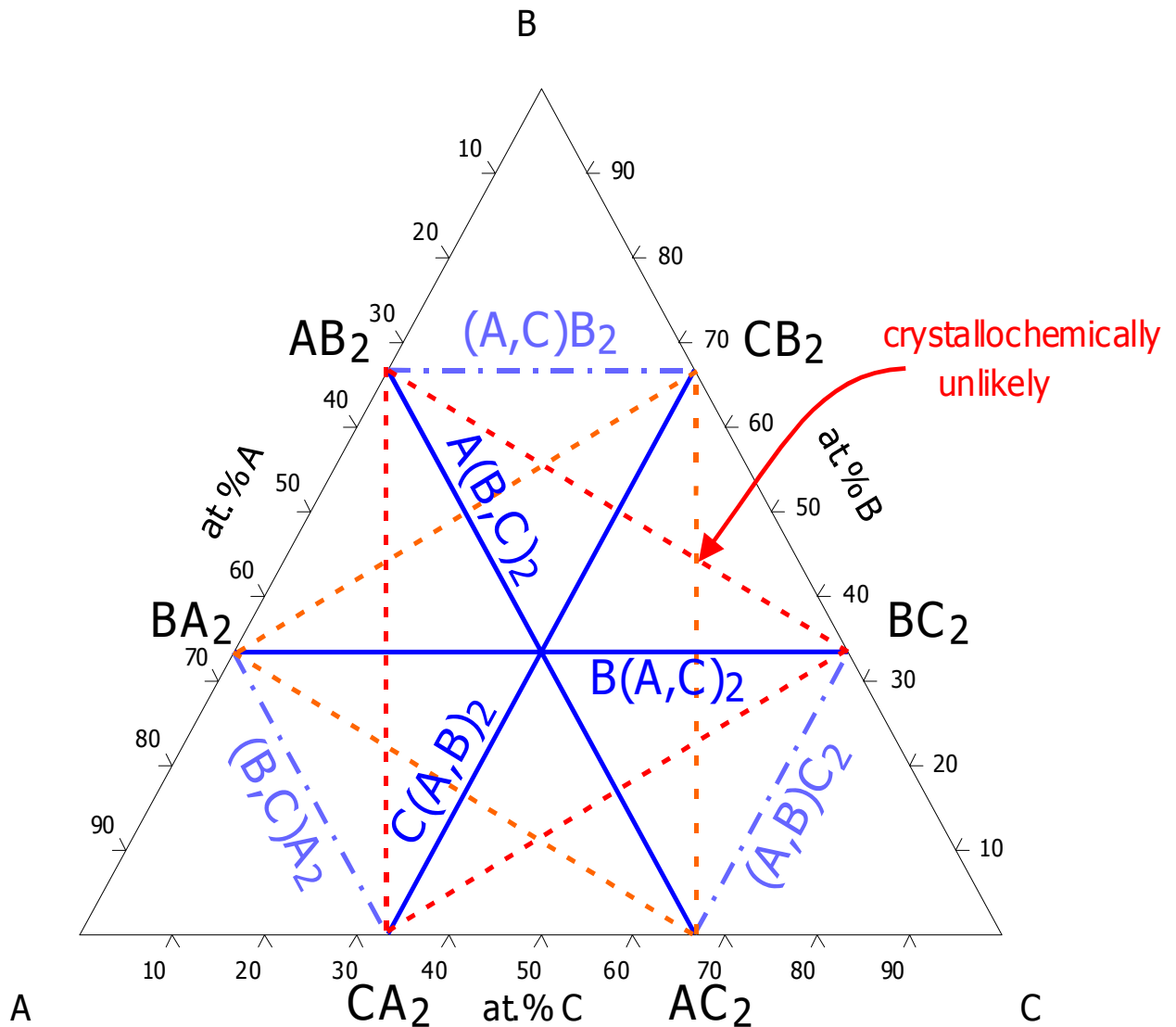


Fig. 21 Possible complete solid solutions between two binaries with the same prototype, shown on the prototype MgCu_2 , cF24, 227.

a complete solid solution does not exist, in addition to the 61 systems where complete solid solution occurs. The stability domain in the PN_{\max} versus mean PN_{\min} / PN_{\max} map for the $MgCu_2$, cF24, 227 prototype and its representatives is relatively large. Nevertheless the majority of the 61 experimentally confirmed complete solid solutions are located in a relatively small sub-domain (surrounded by a blue line). With the help of these stability domains we can extrapolate: of the 701 possible $A(B,C)_2$ combinations, 61 experimental + 267 predicted element combinations

are of the type 'complete solid solution', while 42 experimental + 331 predicted combinations are of the type 'limited solid solution'.

These four principles demonstrate, with the help of Figs. 15-22, the correctness of the fourth cornerstone of Nature: "There exist direct links between the position of the constituent chemical elements in the Periodic System and its point set(s) occupation within the structure of an inorganic compound." The existence of such links is reflected in the above outlined four principles:

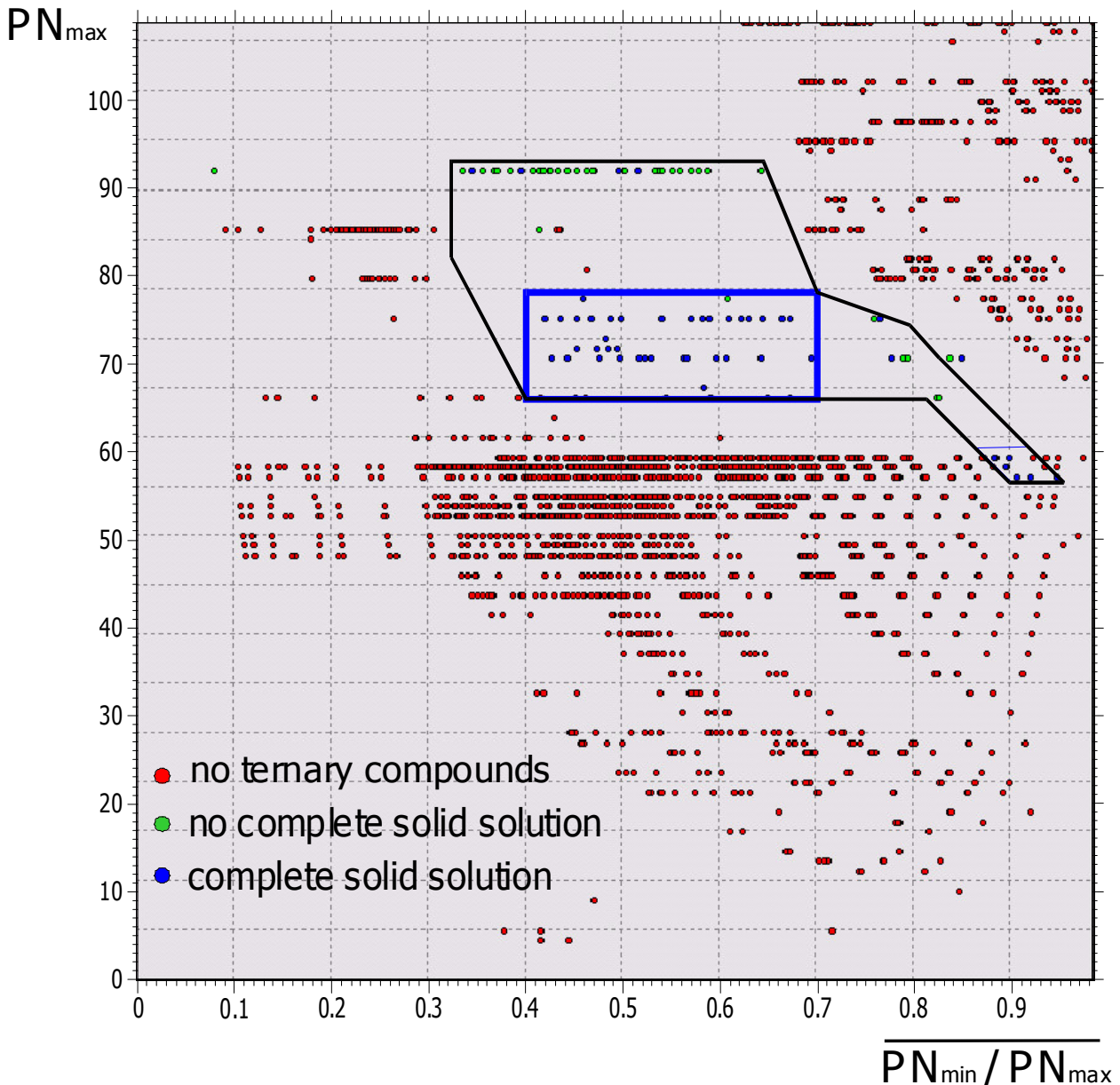


Fig. 22 PN_{\max} versus PN_{\min} / PN_{\max} map for complete solid solutions of the type $A(B,C)_2$ with the prototype $MgCu_2$, cF24, 227.

- Link between the position of the chemical element in the Periodic System and its point set position in the structure (chemistry principle),
- Existence of stability domains for different prototypes and their representatives (structure stability map principle),
- Existence of stability domains for different atomic environment types (generalized AET stability map principle),
- Existence of solid solution stability domains (solid solution stability map principle).

The above given four principles reduce the number of potential inorganic compounds to be considered from several thousands to a few hundreds for a particular prototype. In addition, it efficiently helps to select from the many potential competing prototypes the most probable ones.

4. A realistic way to build up a trustworthy calculated part of World Materials

Based on the 12 principles enumerated above (or the four cornerstones of Nature), we propose a realistic way to build up a trustworthy calculated part of World Materials. There are potentially two major risks to make it *untrustworthy*. Firstly by calculating potential inorganic compounds belonging to non-former systems, secondly by calculating potential inorganic compounds with improbable prototypes. One should keep in mind that it is still not feasible to calculate all possible 32'000 prototypes, as starting point, for each potential inorganic compound.

We propose the following calculation strategy, which must be dynamically linked to a reference database, such as the PAULING FILE [3-5]:

- Focus on unaries to quaternaries.
- Exclude from all calculations chemical systems clearly belonging to non-formers.
- Focus on known inorganic phases belonging to the most populous prototypes: A) fully refined, B) assigned.
- For each of the most populous prototypes derive point-set \Leftrightarrow periodic system (*s*-, *p*-, *d*-, *f*-element) correlations (based on known inorganic phases) and select, with the help of structure stability maps (based on known prototypes and atomic environment types), 5-10 competing prototypes: C).
- Evaluate for each inorganic phase selected under A)-C) confidence levels by comparing experimental data with own calculated data, as well as the consistency with materials-overview – elemental-property parameters maps.
- Calculate a broad range of physical properties for already known inorganic phases, as well as for predicted phases with a high confidence level of correctness, respectively probability of existence. Evaluate the confidence level for inorganic phases where data have been published.

- Store the calculated data in a fully relational database consistent with the reference database, and create dynamic links so that, in case changes will be done, derived data will be recalculated automatically.

Here we have taken advantage of several statistical plots and materials-overview – elemental-property parameters maps. In addition, the existence of correlations between the positions of the chemical elements in the Periodic System and their positions in the crystal structure, as well as the existence of a fully standardized prototype classification, has been used. Last but not least, the correlation between the number of chemical elements of a compound and the different atomic environment types has proven to be a powerful tool. This all together makes it possible, for each of the most populous prototypes, to populate the atomic sites by chemical elements that are chemically meaningful, and compute structural data and intrinsic physical properties for potentially stable inorganic phases. The resulting set of calculated data will of course include data for the experimentally known inorganic phases contained in the reference PAULING FILE [3-5], which thus serves as validation of the computational approach.

The reliability of the calculated data of World Materials depends on four major factors:

- 1) Continuous verification of the calculated data by comparison with the reference (experimentally determined data), especially for the structural data, which constitute the starting values for all calculations. This interplay generates a 'structure reliability factor' for calculated structures of inorganic compounds.
- 2) When the structure has been confirmed, its intrinsic physical properties will be calculated, and again continuously compared with the experimentally known reference data (if available). This interplay generates a 'property reliability factor' for the calculated property data. The same principle should be applied to phase diagram data.
- 3) Each calculated structure should have an active link to the structure reference data set that was used as starting values. This with the purpose of being able to compare with the experimental values, as well as maintaining the consistency (any change of the experimental reference data will automatically trigger a new calculation procedure).
- 4) Thorough quality controls of the experimentally obtained data, and data calculated by different methods are required.

5. Conclusions

It is known that for *ab initio* calculations no input other than the laws of quantum mechanics and the atomic numbers (AN) of the involved chemical elements is required. We must then trust that, through the necessary approximations in the calculations applying the laws of quantum mechanics, no relevant

information goes lost. In principle most of the experimentally determined data contained in the PAULING FILE can be confirmed by calculations, leading to a quantum mechanical materials design approach.

The cornerstones and principles presented in this work are summarized in Table 5. The first cornerstone is a direct consequence of the existence of the about 100 chemical elements given by Nature and their combinations. In practice

there exist an infinite number of chemical element combinations (and therefore inorganic compounds). This conclusion is independent of quantum mechanics.

The second to fourth cornerstones of Nature can under no way be derived from quantum mechanics, even each of the about 100 chemical elements possesses all required information, given by Nature, and 'knows what to do' when submitted to external conditions.

Table 5 Summary of cornerstones and principles proposed in this work.

Selection of chemical elements to be combined, compositions, ...:	Leading to infinitely many potential inorganic compounds:
First cornerstone of Nature i) Number of chemical element combinations: ii) Number of compounds per chemical system: iii) Number of prototypes:	infinite 0-30 > 32'000
Principles (governing factors):	Restrains:
Second cornerstone of Nature 1) Non-formers: 2) No. chemical elements per compound – AET correlation: 3) Inactive composition range: 4) Most probable stoichiometry ratios:	about 30% 4'087'975 systems 10% (2 elements); 15% (3); 38.6% (4) 10 (2 elements); 769 (3)
Third cornerstone of Nature 5) Simplicity (90% probability): 6) Symmetry (90% probability): 7) Atomic environment type (90% probability): 8) Ordering tendency:	< 40 atoms/unit cell < 10 point sets < 5 AET 23 of 230 space groups 18 of 100 AETs (<i>n</i> -1) elements impossible
Fourth cornerstone of Nature 9) Prototype – periodic system correlation: 10) Structure stability maps: 11) Generalized AET stability map: 12) Solid solution stability map:	simple to broad stability domains stability domains stability domain

The cornerstones can only be discovered by the examination of a large amount of critically evaluated experimentally determined data. Since we do not have the ability to cross-link all the consequences of the interactions of the atoms under all possible conditions (*e.g.* nature of the chemical elements to be combined, stoichiometric ratios, temperature, pressure, *etc.*), we have no other choice than to focus on different consequences of the above mentioned causes. The second cornerstone focuses on the fundamental ability to form an inorganic compound. The third cornerstone focuses on the ordering of the atoms within an inorganic compound. And, finally, the fourth cornerstone reveals the link between the position of a chemical element in the Periodic System and its position(s) in the crystal structure of a particular inorganic compound. These four cornerstones can be considered as different ways to look at the same cause. They are supported by the twelve principles outlined in **Table 5**. These ultimately lead to restraints, which are a requirement for the development of a practicable and trustworthy materials design approach.

The power of critically evaluated data is considerable, knowing that already 20 years ago, having access to only 20% of the now available data, it was possible to derive the core content for nine principles correctly. Apparently the amount of data with information on about 30'000 phases was already sufficient to derive reliable materials knowledge. The in this work verified, and where necessary, extended, nine principles, as well as the newly discovered three principles are based on a much more robust database system with information on over 140'000 inorganic phases. We thus conclude that the here-formulated twelve principles, and from them derived restraints, must be considered trustworthy.

Acknowledgements

We would like to thank Dr. Erich Wimmer and Dr. Paul Saxe of Materials Design Inc. for interesting and stimulating discussions in context with the World Materials calculated part.

References

- [1] J. Gray, *The Fourth Paradigm, Data-Intensive Scientific Discovery*, In: T. Hey, S. Tansley, K. Tolle (Eds.), Microsoft Corp., Redmond, Washington, USA, 2009, pp. xvii-xxxii.
- [2] J.P. Holdren, *Materials Genome Initiative for Global Competitiveness* 1-18, Executive Office of the President of the United States, National Science and Technology Council, 2011.
- [3] <http://www.paulingfile.com/>
- [4] P. Villars, K. Cenzual, J.L.C. Daams, F. Hulliger, T.B. Massalski, H. Okamoto, K. Osaki, A. Prince (Eds.) *PAULING FILE Binaries Edition*, on CD-ROM, ASM International, Materials Park, Ohio, USA, 2002.
- [5] http://crystdb.nims.go.jp/index_en.html
- [6] P. Villars, K. Cenzual (Eds.), *Pearson's Crystal Data: Crystal Structure Database for Inorganic Compounds*, on CD-ROM, Release 2013/14, ASM International, Materials Park, Ohio, USA, 2013.
- [7] P. Villars, H. Okamoto, K. Cenzual (Eds.), *ASM Alloy Phase Diagram Database*; <http://www.asminternational.org/AsmEnterprise/APD>, ASM International, Materials Park, Ohio, U.S.A., 2013/14.
- [8] *PDF-4⁺*, on CD-ROM, Release 2013/14, International Centre for Diffraction Data (ICDD), Newtown Square, Pennsylvania, U.S.A.
- [9] P. Villars, K. Cenzual, H. Okamoto, F. Hulliger (Eds.), *Springer Materials, The Landolt-Börnstein Database: Inorganic Solid Phases*, 2013.
- [10] K. Cenzual, M. Berndt, K. Brandenburg, V. Luong, E. Flack, P. Villars, *ESDD software package*, copyright Japan Science and Technology Corporation, 2000; updates by O. Shcherban, SCC "Structure Properties", Lviv, Ukraine.
- [11] E. Parthé, L. Gelato, B. Chabot, M. Penzo, K. Cenzual, R. Gladyshevskii, *Gmelin Handbook of Inorganic and Organometallic Chemistry, 8th Ed., TYPIC – Standardized Data and Crystal Chemical Characterization of Inorganic Structure Types*, 4 vols., Springer, Heidelberg, 1993-1994.
- [12] L.M. Gelato, E. Parthé, *J. Appl. Crystallogr.* 20 (1987) 139-143.
- [13] P. Villars, J. Daams, Y. Shikata, K. Rajan, S. Iwata, *Chem. Met. Alloys* 1 (2008) 1-23.
- [14] P. Villars, K. Cenzual, J. Daams, Y. Chen, S. Iwata, *J. Alloys Compd.* 317-318 (2004) 167-175.
- [15] P. Villars, J. Daams, Y. Shikata, Y. Chen, S. Iwata, *Chem. Met. Alloys* 1 (2008) 210-226.
- [16] P. Villars, Y. Chen, S. Iwata, *Sci. Council Jpn. J.*, 2014, accepted.
- [17] P. Villars, In: J.H. Westbrook, R.L. Fleischer (Eds.), *Intermetallic Compounds, Principles and Practice*, Wiley, New York, 1994, Vol. 1, pp. 227-275.
- [18] P. Villars, L.D. Calvert, *Pearson's Handbook of Crystallographic Data for Intermetallic Phases, Second Edition*, 4 vols., ASM International, Materials Park, Ohio, USA, 1991.
- [19] ICSD, FIZ/NIST, Fachinformations-Zentrum Karlsruhe, Germany, 2013.

- [20] F.R. de Boer, W.C.M. Mattens, A.R. Miedema, A.K. Niessen, In F.R. de Boer, D.G. Pettifor (Eds.), *Cohesion in Metals, Transition Metal Alloys (Cohesion and Structure)*, North-Holland, Amsterdam, Netherlands, Vol. 1, 120 p.
- [21] P. Villars, *J. Less-Common Met.* 119 (1986) 175-188.
- [22] J.L.C. Daams, In: J.H. Westbrook, R.L. Fleischer (Eds.), *Intermetallic Compounds, Principles and Practice*, Wiley, New York, 1994, Vol. 1, pp. 363-383.
- [23] P.I. Kripyakevich, *Structure Types of Intermetallic Compounds*, Nauka, Moscow, 1977, 288 p. (in Russian).