A systematic approach to general and structure-type formulas for minerals and other inorganic phases

DEANE K. SMITH,[1] ANDREW C. ROBERTS,[2] PETER BAYLISS,[3] AND FRIEDRICH LIEBAU[4]

[1]Department of Geosciences, Pennsylvania State University, University Park, Pennsylvania 16802, U.S.A.
[2]Geological Survey of Canada, 762-601 Booth Street, Ottawa, Ontario K1A 0E8, Canada
[3]Mineralogy, Australian Museum, 6 College Street, Sydney, N.S.W. 2000, Australia
[4]Mineralogisch-Petrologisches Institut und Museum, der Universität Kiel, Olshausenstrasse 40-60, 24098 Kiel, Germany

ABSTRACT

During the preparation of the Mineral Group List Index of the Mineral Powder Diffraction File, it was necessary to develop a consistent scheme for the representation of the structural formulas. Because the fourteen letters B, C, F, H, I, K, N, O, P, S, U, V, W, and Y of the Latin (Roman) alphabet represent chemical elements (element symbols), this system uses twelve letters to represent cations, anions, and molecules occupying various sites of the structure (site symbols). The letters available as structure-site symbols represent cations with decreasing coordination numbers (CN) as follows: $\mathbf{D}$, CN $\geq$ 9; $\mathbf{E}$, CN = 8 or 7; $\mathbf{G}$, CN = 6; $\mathbf{J}$, CN = 5; $\mathbf{Q}$, CN = 4 planar or 2 linear; $\mathbf{T}$, CN = 4 tetrahedral; and $\mathbf{R}$, CN = 3 planar. These definitions leave sufficient remaining letters for other structure-site symbols such as: $\mathbf{A}$ for all cations without regard to the coordination number; $\mathbf{L}$ for lone-electron-pair cations; $\mathbf{M}$ for neutral molecular units; $\mathbf{X}$ for monatomic anions; and $\mathbf{Z}$ for polyatomic anions. Structure sites of the same coordination type, yet distinct enough for ordered occupancy of resident cations, may be differentiated by primes (′,″,‴, etc.). Variable coordination numbers on a structure site and variable site occupancies are indicated by two symbols or subscripts with the intervening symbol $\leftrightarrow$. For example, the general structure-type formula of amphibole may be written as $\mathbf{A}_{7\leftrightarrow8}[\mathbf{T}_4\mathbf{X}_{11}]_2\mathbf{X}'_2$. A specific structure-type formula also allows for a mixture of symbols to represent chemical elements and structure sites such as $\mathbf{A}_{0\leftrightarrow1}(\mathbf{E}'\leftrightarrow\mathbf{G})_2\mathbf{G}'_3\mathbf{G}''_2[Si_4O_{11}]_2(OH)_2$ whereas an example of a chemical formula is $Ca_2Mg_5[Si_4O_{11}]_2(OH)_2$ for the amphibole tremolite.

INTRODUCTION

During the preparation of the Mineral Group List Index for the Mineral Powder Diffraction File (Bayliss et al. 1993) produced by the International Center for Diffraction Data, the need arose for a consistent technique to represent structural formulas of the various mineral groups. An examination of the mineralogical and crystal-chemical literature reveals many different methods used to represent structural formula when discussing minerals and groups of minerals having the same or similar structure types.[1] This lack of consistency leads to considerable confusion. Additional confusion results from use of the same letter in different ways, e.g., B in ABO_3 to represent either a boron atom or an unspecified cation that occupies a structure site different from an equally unspecified cation A.

The term mineral group has been used in many places (Strunz 1970; Bayliss et al. 1993; Fleischer and Mandarino 1995) without rigorous definition. For the Mineral Powder Diffraction File, a mineral group is defined as a collection of three or more mineral species with X-ray diffraction patterns that show marked similarities because the structures have considerable similarities, in particular regarding crystallographic sites, stoichiometric ratios of their sites, and coordination of the elements occupying these sites. As a consequence of such similarities, a general structure-type formula can be written for the mineral group, which contains information on these common structural properties, e.g., homologous (polysomatic) series such as humite, homeotypes such as perovskite, polytypoids such as amphibole, polytypes such as kaolinite, and isotypic series such as forsterite-fayalite. A mineral subgroup may be further defined as containing three or more minerals within a group that have the same structure type as described by Lima-di-Faria et al. (1990).

The extent of inconsistencies in structural formulas is illustrated further by an examination of a few references

---

[1] According to Lima-de-Faria et al. (1990), several structures constitute a structure type if: (1) They have the same space-group type [as defined in *International Tables for Crystallography* (1992)] or belong to a pair of enantiomorphic space-group types; (2) the atomic positions, occupied either fully or partially at random, are the same in both structures; (3) for all Wyckoff positions, both the crystallographic point configurations and their general geometrical interrelationships are similar; and (4) the corresponding atoms and corresponding bonds have similar characteristics.

in structural chemistry and mineralogy. Palache et al. (1944, 1951) used AX, AX$_2$, and A$_2$X$_3$ for simple sulfides, oxides, and halides regardless of the size of the atoms involved, the coordination number, and the valence on the atoms and structural formulas such as A$_m$B$_n$(XO$_4$)$_p$Z$_q$ for more complex situations. Wells (1984) and Lima-de-Faria (1994) used structural formulas such as ABCX$_3$ and A$_2$BX$_4$ for compounds based on close packings, also without regard to the size, coordination number, and valence. Wyckoff (1965) used structural formulas such as R$_x$(M$_n$X$_p$)$_y$. Strunz (1970) usually emphasized the main structural unit(s) in structural formulas but did use RO$_3$ and RO$_4$ where both atomic groups may occur in the same structure (as in the borates). Bloss (1971) used structural formulas similar to Palache et al. (1944, 1951), i.e., AX and A$_2$X$_3$, with no distinction as to the site coordination, and ..Z$_n$O$_m$ for the silicate groups. Muller and Roy (1974) followed Wells (1984), but made more distinction by using A for tetrahedral sites and B for octahedral sites.

Among mineralogical texts, Berry et al. (1983), Klein and Hurlbut (1985), and Blackburn and Dennen (1988) all used X$_m$Y$_n$(Z$_p$O$_q$)W$_r$ as a silicate structural formula. Smith (1972) and Zoltai and Stout (1984) used M for octahedral sites and T for tetrahedral sites. Deer et al. (1963) employed XYZ$_2$O$_6$ and WX$_2$Y$_5$Z$_8$(OH)$_2$ for pyroxene and amphibole, respectively; whereas Putnis (1992) used ABSi$_2$O$_6$ and A$_{0-1}$B$_2$C$_5$T$_8$O$_{22}$(OH,F)$_2$, respectively, to make comparisons of pyroxene and amphibole easier and used similar structural formulas for other silicate groups. Fleischer and Mandarino (1995) used ABZ$_2$O$_6$ for pyroxene and A$_{0-1}$B$_2$Y$_5$Z$_8$O$_{22}$(OH)$_2$ for amphibole, whereas AB(CO$_3$)$_2$ is used for the dolomite group.

These few examples show that there is no systematic usage of symbols for structural formulas. X and Z are used for both cation and anion structure sites. The same structural site is described by different letters (A, B, M, and X for octahedral sites; A, R, T, and Z for the tetrahedral sites) by different authors. Some letters have double meanings: B, C, W, and Y are used for structure sites and as element symbols for boron, carbon, tungsten, and yttrium. In addition to generating confusion, these examples offer no suggestion for a universally accepted approach for writing structural formulas. Unfortunately, these examples are only a representation of the techniques used in technical articles and other references. It is time to remedy this situation with a systematic approach to structural formulas.

This paper introduces an approach for a standard use of the Latin letters in a manner that does not lead to ambiguity between structure-site symbols and element symbols, yet retains a logical sequence with respect to the coordination of a site, which is correlated with the size of the cations occupying the structure site, and the role of the site within structural subunits. These structure-type formulas are also compatible with the report accepted by the IUCr Commission on Nomenclature, Subcommittee on the Formulas of Inorganic Structure Types (Lima-de-Faria et al. 1990).

A SYSTEMATIC APPROACH

Structural formulas in chemistry and mineralogy contain two types of symbols having different quality: (1) symbols to represent atoms and structure sites and (2) symbols to indicate structure details. Here the term structure site is defined as a set of symmetry-equivalent points. These sites differ from Wyckoff positions used in the *International Tables for Crystallography* (1992) by the fact that the coordinates, *x, y,* and *z,* of a structure site must be within a range that its points have the same definite coordination number. Within the first category of symbols are element symbols and structure-site symbols.

Symbols for atoms and structure sites

Element symbols. In scientific writing, chemical formulas are constructed of the letters of the Latin (Roman) alphabet regardless of the language used for the main text. Consequently, it is desirable to retain this alphabet in any represention of structural formulas. Fourteen Latin letters individually represent elements: B, C, F, H, I, K, N, O, P, S, U, V, W, and Y. These symbols are henceforth called element symbols. They should not be considered available as structure-site symbols.

Structure-site symbols. The letters **A**, **D**, **E**, **G**, **J**, **L**, **M**, **Q**, **R**, **T**, **X**, and **Z** are available for representation of sites in the structure and the species occupying them. To distinguish further the structure-site symbols, the letters are represented best using a different font (bold)[2] but it is not necessary because the letters are unique. The letters (structure-site symbols) are utilized in a systematic fashion in which each letter represents specific sites in the structure characterized by coordination and, indirectly, by the size of the occupying ions. The type of structure-site symbol will then suggest the elements that will usually substitute on that site. The various structure-site symbols are defined below.

To minimize deviation from some of the older usage indicated in the introduction and to provide a systematic usage, the early letters of the alphabet are assigned to sites in the structure occupied by the larger cations. The later letters are assigned for those sites occupied by the smaller cations and also to those occupied by anions. This set of definitions retains familiar associations such as **T** for tetrahedral sites. **X** and **Z** are used for anions. To help the memory, the following symbols are chosen logically: **A** (all cations), **L** (lone-electron-pair cations), **M** [molecular units (neutral)], and **T** (tetrahedrally coordinated cations). There are sufficient letters to cover all the structural configurations in both mineral and other inorganic structures. These structure-site symbols are based on coordination of other specific character of a site and are related to the elements as illustrated in Table 1.

The structure-site symbol **A** is reserved for general for-

---

[2] The use of bold face in the *American Mineralogist* is usually restricted to designate vectors. Its use here is with special permission to illustrate the procedure recommended in this article for the structure-site symbols so that they are easily recognized.

TABLE 1. Element distribution by valence and coordination

Symbol	CN (ion. rad.)	Block*	Valence 1	Valence 2	Valence 3
A				All	
D	≥9 (≥1.0 Å)†	s	Cs,Rb,NH₄,K,Na	(Ra),Ba,Sr,Ca	
		p			
		d		Cd	Y
		f		Nd,Sm,Eu	REE
E	8, 7 (≥0.8 Å)	s	(Cs),(Rb),NH₄,Na	(Ra),Ba,Sr,Ca,Mg	
		p			Tl,In
		d		Cd,Mn	Y,Sc,Fe
		f		Nd,Sm,Eu,Dy	REE
G	6 (≥1.0–0.6 Å)	s	(K),Na,Li	Ca,Mg,(Ba),(Sr)	
		p			Ga,Al
		d	Ag	(V),Cr,Mn,Fe,Co,Ni, Cu,Zn,Cd,(Pd)	Y,Sc,Ti,V,Cr,Fe,(Nb),(Ta), (Mo),(Co),(Ni),(Pd),(Pt)
		f			REE
J	5 (0.7–0.5 Å)	s	(Na),Li	Mg	
		p			(Ga),Al
		d		Ni,Cu,Zn	
		f			
T	4 tetrahedral (<0.6 Å)	s		Mg,Be	
		p			Ga,Al,B
		d	Ag,Cu	Fe,Co,Ni,Cu,Cd,(Hg)	Fe
		f			
R	3 planal (≤0.2 Å)	s		(Be)	
		p			B
		d	Cu,Ag	Cu,Ag	
		f			
Q	4 planar, 2 linear (1.0–0.2 Å)	s	H		
		p			
		d	Au,Ag,Cu,Hg,(Pd)	(Fe),(Ni),Cu,Pd,Ag,Pt,Hg	
		f			
L	Asymmetric (lone-electron pair)		Tl	Pb,Sn	Bi,Sb,As
X	Monoatomic anion		I,Br,Cl,F,OH	O,S,Se,Te	As,(Sb),P,N
Z	Polyatomic anion		(NO₂),NO₃,(ClO₃),(ClO₄)	CO₃,S₂TeO₃,SeO₃,SO₃ TeO₄,SeO₄,SO₄	(AsO₃),(PO₃) AsO₄,PO₄
M	Molecule		Noble gases, N₂, H₂O, CO₂, CH₄, amines, etc.		

Notes: Elements that are only rarely observed in minerals are indicated in parentheses. CN = coordination number.
*s, p, d, f = s-block elements, p-block elements, etc.
† Radii are illustrated for oxocompounds.

mulas where it represents all the cations, or all the cations not in the structural unit, or for a site which is so variable that no simple combination of other symbols may be used. [In this paper, the use of structural unit follows the definition given in Lima-de-Faria et al. (1990). A structural unit within a structure is an assemblage of atoms that are more tightly bonded together than others. Structural units can be finite groups or assemblies that are infinite in one, two, or three dimensions.] Some examples for the use of symbol **A** include **A(TX₄)** for **D(TX₄)**, **E(TX₄)**, **G(TX₄)**, and T(T′X₄), represented by barite [Ba(SO₄)], zircon [Zr(SiO₄)], zincosite [Zn(SO₄)], and berlinite [Al(PO₄)], respectively, and **A**₇₋₈ [**T**₈**X**₂₂]**X′**₂ for amphibole instead of **A**₀₋₁ (**E↔G**)₂**G′**₃**G″**₂[**T**₈**X**₂₂]**X′**₂. The latter formula is an example of the use of **A** for a highly variable site where the CN may range from 6 to 12 depending on the atom on the site.

Other structure-site symbols for cations are based on coordination numbers produced in close-packed structures, e.g., 12, 8, 6, 4, and 3, and their derivatives; the ranges were chosen following the rules described below. The illustrated radii are for structures of oxocompounds. According to the radius ratio rule for ionic structures, corresponding values for thiocompounds (sulfides) and chlorides, for example, would be larger by a factor of approximately $r(S^{2-})/r(O^{2-}) \approx r(Cl)/r(O^{2-}) \approx 1.3$.

Symbol **D** is used for cation sites with CN ≥ 9. These structure sites contain large cations with effective ionic radii (Shannon 1976; Whittaker and Mantus 1970) greater than 1.0 Å and valences 1+ to 4+. Examples for the use of **D** are the alum group [NaAl(SO₄)₂·12H₂O] where **D** represents the site occupied by K^{1+}, Na^{1+}, H_3O^{1+}, and NH_4^{1+} and the apatite group where **D** and **D′** represent the structure site occupied by Ca^{2+} or tri and quadrivalent rare-earth elements and actinides, e.g., Sr₃(CeNa)(PO₄)₃(OH). Symbol **E** is used for sites in the structure with CN = 8

TABLE 1—*Extended*

	Valence	
4	5	6
All		
Th,U,Ce		
Pb,Sn Hf,Zr Th,U,Ce	(Ta),(Nb) U	(Mo) U
(Pb),Sn,Ge,(Si) Cr,Ti,Mn,(Hf),(V),(Ta), (Nb),(Cr),(Mo),(W),(Os),(Pd)	Bi,Sb,As Ta,(V),(Nb),(Mo) U	(Te) W,Mo,(Cr) U
Ti	(V)	(Mo)
(Sn),Ge,Si,(C) Cr	As,P V,Cr,(Nb) U	(Te),Se,S W,Mo,Cr,(Mn)
C	N	
		U
Te,Se,S	(I)	
(Si),C SiO_4	(B)	

or 7. These sites usually contain uni- to quadrivalent ions with effective radii between 1.0 and 0.8 Å for oxocompounds. Examples include the structure site occupied by Ca^{2+} in amphibole and pyroxene and the cation site in monazite-(Ce) $[Ce(PO_4)]$ and zircon $[Zr(SiO_4)]$.

Symbol **G** is used for sites in the structure with CN = 6, including octahedral and triangular prismatic. Examples are Al in kaolinite $\{Al_2[Si_2O_5](OH)_4\}$ and Ni in nickeline (NiAs), respectively. For kaolinite, the effective ionic radii are in the range 1.0–0.6 Å. Symbol **J** represents sites occupied by cations with CN = 5, e.g., Al in andalusite $\{Al_2[SiO_4]O\}$ and Fe in magnetoplumbite $(PbFe_{12}O_{19})$. Their effective ionic radii are usually in the range 0.7–0.5 Å.

Symbol **L** is reserved for cations with lone-electron pairs (LEP) such as As^{3+}, Sb^{3+}, Bi^{3+}, Tl^{1+}, Sn^{2+}, Pb^{2+}, S^{4+}, Se^{4+}, and Te^{4+}. They behave, in most cases, quite differently from all other cations in that their LEP are stereoactive. A high degree of stereoactivity results in strongly asymmetric, often pyramidal, one-sided coordinations complemented by several long distances. Sulfides that contain such **L** sites include berthierite $(FeSb_2S_4)$ and bournonite $(PbCaSbS_3)$. If a **D**, **G**, or **T** cation is replaced by an **L** cation but retains the essential geometry of the site, it is written (**D/L**), (**G/L**), or (**T/L**). In the large LEP cations with low valence such as Tl^{1+} and Pb^{2+}, the stereoactive influence of the LEP is so weak that it is camouflaged during crystal-structure determination by a high atomic-displacement (temperature) factor, which is the consequence of static or dynamic disorder of the cation that actually deviates slightly from the center of a nearly regular polyhedron of anions. Examples are cerussite (PbC_3), galena (PbS), and litharge (PbO), respectively. If a **D**, **E**, or **G** cation is partially substituted by a large low-valence cation with LEP, the corresponding atomic displacement factors increase considerably, more so than in the case of substitution by other normal **D**, **E**, or **G** cations. If the substituting cation is smaller or has higher valence, the stereoactive influence of the LEP is even stronger, which may result in symmetry reduction and eventually phase transformation. In most cases, such replacement is accompanied by changes in electrical and optical properties.

Symbol **M** is used to represent sites in the structure that contain neutral molecular units, either inorganic, including H_2O, or organic, e.g., in the zeolite group. Symbol **Q** is used for structure sites with fourfold square planar or the twofold linear coordination. Mineral examples are Pt and Pd in braggite (PdS) and cooperite (PtS) and Hg in cinnabar (HgS), respectively. Symbol **R** is used for structure sites with CN = 3, occupied by cations with planar or near planar triangular coordination, e.g., **C** in calcite $[Ca(CO_3)]$ or buetschliite $[K_2Ca(CO_3)_2$ (Zemann 1981)]. Symbol **T** represents tetrahedral sites such as found in silicates, sulfates, arsenates, and phosphates. Transition element cations involving crystal-field distortions (Jahn-Teller) would still be in the appropriate structure site corresponding to the coordination number, i.e., **D**, **E**, **G**, **J**, **Q**, **R**, and **T** and not in the lone-electron pair **L** site.

Symbols **X** and **Z** are reserved for anions. Symbol **X** is used for monatomic anions that may be uni-, di-, tri-, or quadrivalent. Examples include the halide ions F_{1-}, C_{1-}, Br_{1-}, and I_{1-}; the chalcogenide ions O^{2-}, S^{2-}, and Se^{2-}; the pnictide ions As^{3-}, Sb^{3-}, and Bi^{3-}; and the carbide ion C^{4-}. The ion $(OH)^{1-}$, which is similar in size and behavior to F^{1-}, is also an **X**-type anion. **Z** is used for polyatomic anions such as $(SO_4)^{2-}$ and $(CO_3)^{2-}$ groups found in the hauyne $\{Na_3Ca[Al_3Si_3O_{12}](SO_4)\}$ and cancrinite $\{Na_3Ca[Al_3Si_3O_{12}](CO_4)\}$ groups, respectively.

Except for hydrides, the coordination number of H is nearly always 1 or 2. Where CN = 1, H is a terminal atom and, therefore, mono-bonded as in $(OH)^{1-}$ ions. Where CN = 2, it is involved in an H bridge, which is, with few exceptions, not linear. Consequently, except for NH_4^{1+} and H_3O^{1+} as **D**, H_2O as **M**, and $(OH)^{1-}$ or H^{1-} as **X**, hydrogen does not fit into any of the listed sites and should be always given as H.

In some mineral groups, the triangular CO_3 unit may

TABLE 2. Symbols useful for structural details in structure-type formulas

′ ″ ‴	For multiple distinct sites in any coordination
()	Indicates the principal radicals involving a single polyhedron or complex in formulas
[]	Structural units in formulas involving more than one polyhedral unit
↔	Indicates possible change of coordination numbers between the corresponding symbols and between subscripts to indicate a possible continuous variation between the values
,	Indicates substitution between symbols or subscripts
/	Indicates limited substitution between symbols
*	Indicates chemical substitution within a structure site
hijklmnpqrs	Integral subscript or multiplier value
tuwxyz	Non-integral subscript or multiplier value
^	Indicates inclusive integers in subscripts and multipliers
v	For the valence of the cation in subscripts
δ	Indicates small deviations from the integer in subscripts such as 1-δ

substitute for the TO_4 tetrahedral unit by disordering the three O atoms over the four tetrahedral O atom sites. The central atom is at or near the center of the triangle in this substitution and never at the tetrahedral center. Thus this situation is properly designated as $[(RX_3),(TX_4)]$.

Symbols for structural details

In addition to element and structure-site symbols represented by Latin letters, several other symbols (Table 2) are used to indicate structural details. They are used to indicate the nature of structural units, variations in structure-site occupancy, variations in subscripts and multipliers, and deviations from full occupancy.

Comma. In accord with general usage, the comma "," should be used to indicate continuous substitution, partial as well as complete, if there is no change of CN. Vacancies (symbol □) are treated here just like atoms. This usage is restricted to chemical formulas for atoms on a specific site. The comma is not appropriate for structure site formulas because different symbols indicate a change in the coordination number. Examples for proper use of the comma include forsterite-fayalite $[(Mg,Fe)_2(SiO_4)]$ and hornblende $\{(Na,□)Ca_2(Mg,Fe,Al)_5[(Si,Al)_4O_{11}]_2(OH)_2\}$. In structure-type formulas, the comma is used to mean "or" where substitution may occur. An example is in the zeolites where H_2O and **M** units may both occur.

Slash. The slash "/" between structure-site symbols indicates, without regard to possible solid solubility, that the different atom types form either isotypic members of a structure group or homeotype members of a mineral group. Examples in chemical formulas include the olivine family where $(Mg/Fe/Ca/Mn)_2(SiO_4)$ implies the isotypic members forsterite, fayalite, larnite, and tephroite and $Ca(Ti/Ge)O_3$ for isotypic perovskite types. Structure-type formulas such as $(D/E)GO_3$ indicate homeotype members of the perovskite group such as $Ca^{[12]}Ti^{[6]}O_3$ and $Mg^{[8]}Si^{[6]}O_3$. The slash is also used for isotypic structures where some of the members have non-LEP elements and other members have LEP-type elements and little, if any, substitution occurs between the two types. An example

of this usage is monazite $(G/L)(PO_4)$ where the **G/L** site could be Ca or Pb, which show little tendency to substitute for each other.

Double arrow. The double arrow "↔" is used to indicate potential continuous variation between the symbols. This use includes substitution, partial as well as complete, between atoms represented by either element symbol or structure-site symbols where there is a change in CN and the rest of the structure remains equivalent. It is also used in subscripts and multipliers to indicate that continuous variability may occur between the values listed. Its use in subscripts implies the possibility of vacancies in the structure site. If more than a minor part of a structure is affected, the structure would be too different to belong to the same structure-type or homeotype. An example is the augitic pyroxenes where $(Ca↔Mg,Fe)(Mg,Fe)[Si_2O_6]$ indicates that Ca and Mg,Fe may not have the same CN. In structure-type formulas, if the coordination number of a structure site varies continuously with temperature, pressure, or composition, then two structure-site symbols may be used written with an intervening ↔ symbol to imply a possible continuous variation in CN, e.g., $(D↔E)$ and $(E↔G)$. An example is one of the structure sites in ortho-pyroxene, which changes CN from 6 to 7 between room temperature and 1200 K as reported by Yang and Ghose (1995). The corresponding structure-type formula is, therefore, written as $(E↔G)G'[T_2X_6]$.

Primes. The prime symbols, ′,″,‴, are used in general and specific structure-type formulas where different sites may have the same structure-site symbol but distinct characteristics because of different nearest-neighbor arrangements that lead to cation ordering. Examples include amphibole $\{A_{0↔1}(E↔G)_2G'_3G''_2[T_4O_{11}]_2(OH)_2\}$ where the sixfold **G** sites have distinct differences. The first **G** site may have CN = 6, 7, or 8 as a function of composition and temperature; hence, it is written $(E↔G)$. The other **G** sites may be occupied by the same cation in some amphibole species and different cations in other amphibole species. Freieslebenite, $AgLL'S_3$, is an example where the **L** sites have different coordination numbers.

Asterisk. Note in several examples that the primes are used to imply distinct structure sites. Some formulas are written to imply the nature of the chemical substitution within a structure site, i.e., solid solution. To distinguish this chemical substitution, the * symbol may be introduced to indicate the chemical substitution. Thus, the most general formula for the forsterite-fayalite isotype series, $(Mg,Fe)_2[SiO_4]$, can be represented as $G_xG^*_{2-x}[TX_4]$ or, in more detail, as $(G_uG^*_{1-u})(G_wG^*_{1-w})'[TX_4]$ for a more ordered state. The * symbol should not be used in mineral group structure-type formulas but it is available when it is desirable to make both chemical and structural distinctions.

Table 1 shows the distribution of elements in minerals and other inorganic phases by valence, coordination number, and, indirectly, by ionic radius. In this scheme, except for **A**, **L**, **M**, **X**, and **Z**, it is the coordination of the site that determines the symbol to be used. Table 2 lists ad-

TABLE 3. Structure-type formulas for selected mineral groups

Group	Specific structure-type formula	Group	Specific structure-type formula
Amphibole	$\mathbf{A}_{0\leftrightarrow1}(\mathbf{E}\leftrightarrow\mathbf{G})_2\mathbf{G}'_3\mathbf{G}''_2[\mathbf{T}_4\mathbf{O}_{11}]_2\mathbf{X}_2$	Mica	$\mathbf{DG}_{2,3}[\mathbf{T}_4\mathbf{O}_{10}]\mathbf{X}_2$
Apatite	$(\mathbf{D/L})_3(\mathbf{D'/L'})_2(\mathbf{TO}_4)_3\mathbf{X}$	Monazite	$(\mathbf{D/L})(\mathbf{TX}_4)$
Aragonite	$(\mathbf{D/L})(\mathbf{RO}_3)$	Nickeline	$\mathbf{G}_{1\text{-}\delta}\mathbf{X}; \delta = 0\leftrightarrow0.13$
Astrophyllite	$\mathbf{D}_3\mathbf{G}_2\mathbf{G}'_2[\mathbf{T}_8\mathbf{O}_{24}]\mathbf{X}_{6,7}$	Olivine	$\mathbf{G}_2(\mathbf{TO}_4)$
Axinite	$\mathbf{E}_2\mathbf{GAl}_2[\mathbf{B}[\mathbf{Si}_2\mathbf{O}_7]_2\mathbf{O}](\mathbf{OH})$	Pentlandite	$\mathbf{GT}_8\mathbf{S}_8$
Barite	$(\mathbf{DL})(\mathbf{TX}_4)$	Perovskite	$(\mathbf{D}\leftrightarrow\mathbf{E})\mathbf{GX}_3$
Brucite	$\mathbf{GX}_2$	Pyrite	$\mathbf{GXX'}$
Cancrinite	$\mathbf{D}_{6\leftrightarrow9}[\mathbf{T}_6\mathbf{T}'_6\mathbf{O}_{24}]\mathbf{Z}_{1\wedge4}\cdot0\wedge2\mathbf{H}_2\mathbf{O}$	Pyrochlore	$\mathbf{E}_{1\wedge2}\mathbf{G}_2\mathbf{X}_6\mathbf{X'}$
Chlorite	$\mathbf{G}_{4\wedge6}[\mathbf{T}_4\mathbf{O}_{10}](\mathbf{X})_8$	Pyroxene	$(\mathbf{E}\leftrightarrow\mathbf{G})\mathbf{G'}[\mathbf{T}_2\mathbf{O}_6]$
Crichtonite	$\mathbf{DG}_{21}\mathbf{X}_{38}$	Scapolite	$\mathbf{D}_4[\mathbf{T}_4\mathbf{O}_8]_3(\mathbf{X},\mathbf{Z})_{2/v}; v = 1,2$
Epidote	$\mathbf{D}_2\mathbf{G}_3(\mathbf{SiO}_4)[\mathbf{Si}_2\mathbf{O}_7]\mathbf{X}_2$	Smectite	$\mathbf{A}_s\mathbf{G}_{2,3}[\mathbf{T}_4\mathbf{O}_{10}]\mathbf{X}_2\cdot0\wedge8\mathbf{H}_2\mathbf{O}; \delta = 0.3$
Feldspar	$\mathbf{D}_{2x/v,3}[\mathbf{T}_x\mathbf{T}'_{4-x}\mathbf{O}_8]_2; x = 1\leftrightarrow2$	Sodalite	$\mathbf{D}_4[\mathbf{T}_3\mathbf{T}'_3\mathbf{O}_{12}](\mathbf{X},\mathbf{Z})_{2/v}; v = 1,2$
Garnet	$\mathbf{E}_3\mathbf{G}_2[\mathbf{TX}_4,\mathbf{X}_4]_3$	Sphalerite	$\mathbf{T}_3(\mathbf{T'/L})\mathbf{X}_4$
Humite	$\mathbf{G}_{2n+1}(\mathbf{SiO}_4)_n\mathbf{X}_2; n = 1,4$	Spinel	$\mathbf{G}_2\mathbf{TX}_4$
Kaolinite-serpentine	$\mathbf{G}_{2,3}[\mathbf{T}_2\mathbf{O}_5](\mathbf{OH})_4\cdot0,2\mathbf{H}_2\mathbf{O}$	Tetradymite	$\mathbf{L}_n\mathbf{X}_{n\pm1}; n = 1^\wedge5$
Lazulite	$\mathbf{GG}'_2(\mathbf{PO}_4)_2(\mathbf{OH})_2$	Tetrahedrite	$\mathbf{T}_6\mathbf{T}'_6\mathbf{L}_4\mathbf{X}_{12}\mathbf{X'}$
Marcasite	$\mathbf{GXX'}$	Tourmaline	$\mathbf{DG}_3\mathbf{G}'_6(\mathbf{BO}_3)_3[\mathbf{Si}_6\mathbf{O}_{18}]\mathbf{X}_4$
Melilite	$\mathbf{E}_2[\mathbf{TT}'_2\mathbf{O}_7]$	Zeolite	$\mathbf{D}_{n/v}[\mathbf{Al}_n\mathbf{Si}_p\mathbf{O}_{2(n+p)}]\cdot x(\mathbf{H}_2\mathbf{O},\mathbf{M});$

ditional symbols used in the Mineral Group Index for structural details, including subscripts, superscripts, multipliers, and structural units, when preparing formulas. Their use allows subtle aspects of the structures to be represented, such as solid solutions, defects and vacancies, and valence, where useful. Further symbols for co-ordination, valency, symmetry, etc., are described in Lima-de-Faria et al. (1990).

Using the element symbols, the structure-site symbols, and a few rules, all structure formulas for the mineral groups may be constructed on the basis of the known crystal structures. In writing the structure formulas for mineral groups and subgroups, element symbols may be retained whenever the element is not replaced by any other element for any mineral species of the group. The structure-site symbols are used where there is substitution by additional elements and are selected to represent the nature of the site coordination.

DISCUSSION

The systematic scheme developed for the Mineral Powder Diffraction File has applications beyond representing mineral groups. There are four levels of formulas, which, in order of decreasing generality, are as follows. (1) *Formula type* as named by Muller and Roy (1974). This formula type gives only the stoichiometry of cation sites and anion sites such as $\mathbf{AA'}_2\mathbf{X}_4$. (2) *General structure-type formulas* contain only structure-site symbols. Examples in the $\mathbf{AA'}_2\mathbf{X}_4$ formula type are $[\mathbf{D}_2(\mathbf{TX}_4)]$ arcanite $[\mathbf{K}_2(\mathbf{SO}_4)]$, $(\mathbf{EL}_2\mathbf{X}_4)$ berthierite $(\mathbf{FeSb}_2\mathbf{S}_4)$, $(\mathbf{GG'}_2\mathbf{X}_4)$ wilkmanite $(\mathbf{Ni}_3\mathbf{S}_4)$, $(\mathbf{GL}_2\mathbf{X}_4)$ minimum $(\mathbf{Pb}_3\mathbf{O}_4)$, $[\mathbf{G}_2(\mathbf{TX}_4)]$ forsterite $[\mathbf{Mg}_2(\mathbf{SiO}_4)]$, $(\mathbf{GT}_2\mathbf{X}_4)$ spinel $(\mathbf{MgAl}_2\mathbf{O}_4)$, $(\mathbf{TT'}_2\mathbf{X}_4)$ phenacite $[\mathbf{Zn}(\mathbf{SiO}_4)]$, and $[\mathbf{T}(\mathbf{T'T''X}_4)]$ eucryptite $\{\mathbf{Li}[\mathbf{AlSiO}_4]\}$. (3) *Specific structure-type formulas* contain a mixture of structure-site symbols and chemical element symbols. For example, specific structure-type formulas for mineral subgroups and minerals belonging to the olivine group $\mathbf{G}_2\mathbf{TX}_4$ are $\mathbf{G}_2\mathbf{TO}_4$ for the oxidic olivines; $\mathbf{GG'TO}_4$ for triphylite $[\mathbf{LiFe}(\mathbf{PO}_4)]$ and sinhalite $[\mathbf{MgAl}(\mathbf{BO}_4)]$, $\mathbf{G}_2\mathbf{TS}_4$ for sulfidic olivines with $\mathbf{T}$ = Si,

Ge, and Sn; and $\mathbf{G}_2\mathbf{TX}_4$ for the halidic olivines with $\mathbf{T}$ = Be and Zn and $\mathbf{X}$ = F, Cl, Br, and I. (4) *Chemical formulas* contain only chemical element symbols. Examples are forsterite $[\mathbf{Mg}_2(\mathbf{SiO}_4)]$ and chrysoberyl $[\mathbf{Al}_2(\mathbf{BeO}_4)]$, in the $\mathbf{G}_2\mathbf{TO}_4$ subgroup; $[\mathbf{Mn}_2(\mathbf{SiS}_4)]$, $[\mathbf{Ca}_2(\mathbf{GeS}_4)]$, and $[\mathbf{Ca}_2(\mathbf{SnS}_4)]$ in the $\mathbf{G}_2\mathbf{TS}_4$ subgroup; and $[\mathbf{Na}_2(\mathbf{BeF}_4)]$, $[\mathbf{Li}_2(\mathbf{BeCl}_4)]$, and $[\mathbf{Li}_2(\mathbf{ZnBr}_4)]$ in the $\mathbf{G}_2\mathbf{TX}_4$ subgroup of the olivine structure type.

This scheme for representing the general or structure-type formulas of the mineral groups and subgroups has been developed specifically for the Group List Index of the Mineral Powder Diffraction File. Table 3 contains representative formulas for many of the mineral groups recognized by mineralogists and illustrates the use of the scheme described in this paper. After establishing a structure-type formula for a mineral group, a formula of a mineral species that does not fit the structure-type formula means that either the mineral is not in that group or the structure-type formula is not quite correct. As further structural information becomes available, minor changes in group assignments or structure-type formulas may be warranted.

The versatility of the scheme is illustrated by the ability to write complex structural formulas based on chemistry even when no structural data are available by using the information in Table 1. Even amorphous phases may have a formula such as $\mathbf{D}_h\mathbf{E}_i\mathbf{G}_j\mathbf{J}_k\mathbf{L}_l\mathbf{Q}_m\mathbf{R}_n\mathbf{T}_o\mathbf{X}_p\mathbf{Z}_q\cdot r\mathbf{M}$ where h, i, j, k, l, m, n, o, p, and r are different integers based on the known chemistry. The versatility of the scheme can also be applied to hypothetical structure-type families such as amphiboloids (analogous to the pyroxenoids), which could be represented as $(\mathbf{D},\mathbf{E},\mathbf{G})_{n-p}[\mathbf{T}_n\mathbf{X}_{3n+q}]$-$\mathbf{X'}_{0\wedge2}\cdot0\wedge3\mathbf{M}$.

Because this scheme successfully represents structure-type formulas, the rules adapted by the Mineral Powder Diffraction File could be adapted for all uses of inorganic structure-type formulas, including minerals described in texts and in the scientific literature. In addition, this scheme can be used to describe structure sites with mixed chemistry, e.g., amphibole group minerals.

The silicates, or any other inorganic family containing polymeric anions consisting of tetrahedral subunits such as the phosphates, may be represented in several ways with the scheme. Where there is more than one T site, the prime notation may be used. Where the tetrahedrally coordinated site follows close rules that relate the substitutions to the abundance of multiple cations in the structure, both T' and T^* may be used as illustrated by the feldspar formula. In the feldspars, there are two to eight distinct tetrahedral sites, and each of the sites may have a different composition. Thus, the most general formula for the plagioclase series is $D[T_4X_8]$. The plagioclase series could also be represented by $D_{2-w}D^*_{2x}[T_{2-x}X_8]_2$ or, in more detail, as $(D_{2-w},D^*_{2x})[(T_{2+u},T^*_{2-u})(T_{2+w},T^*_{2-w})(T_{2+y},T^*_{2-y})(T_{2+z},T^*_{2-z})'''X_8]_2$, where $x = u + w + y + z$. The value of x affects the valence of the species of elements that appear on the D sites. In some groups such as the mica group, the T substitution does not modify the site abundances, and no advantage is gained by usuig $[T_xT^*_{4-x}X_{10}]$ as compared to $[T_4X_{10}]$, although it is not incorrect to do so.

The sulfide and sulfosalt groups present different problems. Because many of the sulfide structures are derivative structures based on close-packed arrangement, several different representations from simple to complex may be used. For example, the formula for the chalcopyrite group could be expanded into a very detailed formula for the more complex subgroups; the formula $T_6T'_{2-x}L_{1+x}S_8$ could be used for the mawsonite subgroup.

ACKNOWLEDGMENTS

REFERENCES CITED

Bayliss, P., Erd, R.C., Mrose, M.E., Roberts, A.C., and Sabini, A.P. (1993) Mineral powder diffraction file (3rd Ed.) data book, Sets 1–42, 782 p., and Mineral powder diffraction file—search manual, Sets 1–42, 596 p. International Centre for Diffraction Data, Newtown Square, Pennsylvania.

Berry, L.G., Mason, B., and Dietrich, R.V. (1983) Mineralogy, concepts, descriptions and determinations, 561 p. Freeman, San Francisco, California.

Blackburn, W.H. and Dennen, W.H. (1988) Principles of mineralogy, 413 p. Brown, Dubuque, Iowa

Bloss, F.D. (1971) Crystallography and crystal chemistry, an introduction, 545 p. Holt, Reinhart, and Winston, New York.

Deer, W.A., Howie, R.A., and Zussmann, J. (1963) Rock-forming minerals, vol. 2: Chain silicates, 379 p. Longmans, London.

Fleischer, M. and Mandarino, J.A. (1995) Glossary of mineral species 1995, (7th Ed.), 288 p. The Mineralogical Record, Tuscon, Arizona.

International Tables for Crystallography (1992) Vol. A, (3rd Ed.), 854 p. KIuwer Academic Publishers, Dordrecht, The Netherlands.

Klein, C. and Hurlbut, C.S (1985) Manual of mineralogy, 651 p. Wiley, New York.

Lima-de-Faria, J. (1994) Structural mineralogy: an introduction, 346 p. Kluwer, Dordrecht, The Netherlands.

Lima-de-Faria, J., Hellner, E., Liebau, F., Makovicky, E., and Parthe, E. (1990) Nomenclature of inorganic structure types. Acta Crystallographica, A46, 1–11.

Muller, O. and Roy, R. (1974) The major ternary structural families, 487 p. Springer-Verlag, New York.

Palache, C., Berman, H., and Frondel, C. (1944) System of mineralogy, Vol. I, (7th Ed.), 834 p. Wiley, New York.

———(1951) System of mineralogy, Vol. II, (7th Ed.), 1124 p. Wiley, New York.

Putnis, A. (1992) Introduction to mineral sciences, 457 p. Cambridge University Press, Cambridge, U.K.

Shannon, R.D. (1976) Revised effective ionic radii and systematic studies of interatomic distances in halides and chalcogenides. Acta Crystallographica, A32, 751–767.

Smith, J.V. (1972) Geometrical and structural crystallography, 480 p. Wiley, New York.

Strunz, H. (1970) Mineralogische Tabellen, (5th Ed.), 621 p. Akademische, Leipzig, Germany.

Wells, A.F. (1984) Structural inorganic chemistry, (5th Ed.), 1382 p. Oxford University Press, London, U.K.

Whittaker, E.J.W. and Mantus, R. (1970) Ionic radii for use in geochemistry. Geochemica et Cosmochimica Acta, 34, 945–956.

Wyckoff, R.W.G. (1965) Crystal structures, Vol. 3: Inorganic compounds $R_x(MX_4)$, $R_x(M_nX_p)_y$, hydrates, and antimonates, 981 p. Interscience, New York.

Yang, H. and Ghose, S. (1995) A transitional structural site and anomalous Fe-Mg order-disorder in Mg-rich orthopyroxene, $(Mg_{0.75}Fe_{0.25})SiO_3$. American Mineralogist, 80, 9–20.

Zemann, J. (1981) Zur Stereochemie der Karbonate. Fortschrift der Mineralogie, 59, 95–116.

Zoltai, T. and Stout, J.H. (1984) Mineralogy, concepts and principles, 505 p. Burgess, Minneapolis, Minnesota.