CLAY MINERAL THERMOMETRY—A CRITICAL PERSPECTIVE

E. J. ESSENE AND D. R. PEACOR
Department of Geological Sciences, The University of Michigan
Ann Arbor Michigan 48109-1063

Abstract—Diagenetic clay minerals usually occur as heterogeneous assemblages of submicroscopic layers consisting of different structure types such as illite, smectite and chlorite, with variable composition within a given structure type, and with highly variable concentrations of imperfections. The dimensions of mixed-layering, the semi-coherent to coherent nature of the structures across the layering, and compositional heterogeneity occur at a scale well below that of an individual thermodynamic phase. These relations imply that most clays are not distinct minerals or phases, and that assemblages of clays in shales and mudstones are incompatible with the phase rule. Such relations are better evaluated in terms of the formation of metastable materials with each small unit having unique chemical properties, rather than as a small number of stable homogeneous phases. Consequently, treatment of most clay minerals in terms of equilibrium stability with either a thermodynamic or experimental approach is subject to error.

Chemical reactions involving most clay minerals are best understood with kinetic models. These involve a great variety of parameters such as time, fluid/rock ratio, deformation history, nature of starting materials and transformation mechanisms, as well as the variables, such as temperature, pressure and composition, that are commonly used to define equilibrium. Solubility experiments on the stabilities of clay minerals are unlikely to attain equilibrium at low temperatures. Moreover, the activity of soluble species may be controlled by surface equilibria, or by absorbed or exchangeable cations. Interpretations of available experiments on the solubility of illite vs. other mineral assemblages are in violation of Schreinemakers' rules and indicate lack of equilibrium.

Predictable sequences of clay minerals as a function of temperature are best understood through the Ostwald step rule, in which clay mineral assemblages undergo reactions in response to kinetic factors that represent reaction progress rather than an approach to equilibrium. Currently used clay mineral thermometers (illite crystallinity, smectite/illite reaction, chlorite composition) are not based on equilibrium reactions. Such systems are not accurate thermometers and therefore have questionable utility.

Key Words—Chlorite thermometry, Clay thermometry, Illite crystallinity, Illite solubility, Metastability, Ostwald step rule, Smectite solubility, Stability.

INTRODUCTION

Submission of this paper has been prompted by the authors' concerns that clay thermometers are commonly used uncritically and without appropriate ca- veats in studies of sediment diagenesis, ore deposits, sea floor metamorphism, and hot spring deposition. Recent research has demonstrated that most clay minerals are out of equilibrium with their environment, that clay mineral transformations are equivalent to Ostwaldian steps driven by the potential for minimum free energy, and that clay minerals may reach equilib- rium only in the deepest epigenetic or low grade meta- morphic zones. An implication of these relations is that clay mineral transformations are governed by kinetic factors and therefore are controlled by parameters such as duration of time, amount of stress and deformation, and ratio of fluid to rock, as well as temperature, pres- sure and composition. This assessment, however, is in contrast to that of researchers who presume that most clay mineral assemblages are well represented by an approach to equilibrium and that they provide reliable thermometric and thermodynamic systems. Evidence supporting the authors' views is marshaled in the hope that an examination will focus attention on the factors that determine clay mineral relations as a function of temperature in geological systems.

CONCEPT OF A USEFUL THERMOMETER

A thorough understanding of the concept of geo- thermometers has evolved over the last three or four decades in the field of metamorphic petrology (e.g., Essene 1982, 1989). There are now many useful geo- thermometers and geobarometers commonly used by igneous and metamorphic petrologists to determine pressures and temperatures for rocks that equilibrated at temperatures of a few hundred degrees Celsius or more. The relatively high temperature range of most metamorphic thermometers is due to the necessity of an approach to equilibrium, which is operative only at elevated temperatures. It is at low temperatures, es- pecially those corresponding to sediment diagenesis and low-grade metamorphism, where there is a paucity of accurate thermometers. It has been a goal of clay mineralogists and those studying anchizonal meta- morphic rocks to apply systems that yield accurate...
equilibration temperatures. It is therefore necessary to identify requirements for useful geological thermometers in order to create a framework within which clay thermometers can be discussed.

Requirements for a useful geothermometer include systems in chemical equilibrium, phases that are chemically and structurally homogeneous with well-characterized structures and compositions, constraint of additional variables that affect the condition of stable equilibrium, and experimental calibration at compositions, pressures and temperatures comparable to those for which it is to be applied. The condition that a system must be in a state of stable equilibrium means that the relations of equilibrium thermodynamics determine its state as a function of pressure, temperature, and composition. If involving solid solutions, the activity/composition relations must be well established. Systems initially out of equilibrium owe their existence in part to kinetic factors and specific reaction paths, which may not be reproducible solely as a function of temperature.

The requirement of homogeneity refers to all materials at a minimum scale of a few micrometers. The phases must be chemically and structurally well-defined. For example, chlorite compositions have been used as a geothermometer (Cathelineau and Nieva 1985), although Shau et al. (1990) and Jiang et al. (1994b) demonstrated that they may have a significant mixed-layer component and are unreliable for an accurate thermometer. Nonetheless, chlorite thermometry has been applied to a wide range of diagenetic, hydrothermal and metamorphic systems. The requirement of constraint of variables means that the effects of pressure and solid solutions are constrained, such that the number of phases (including minerals) and their compositions and structures are determined only by temperature, i.e., conditions correspond to those of univariant equilibrium.

Experimental calibration of thermometers is necessitated because only through experimentation can equilibrium be established by reversing reactions. The repeated synthesis of a given phase, the repeated occurrence of minerals, or the apparent approach to steady state conditions is a necessary but not sufficient criterion for equilibrium. Most experimentalists are well aware how kinetic factors inhibit the ability of silicate systems to reach equilibrium states at temperature below a few hundred degrees Celsius. Although synthesis may be possible, reversibility is generally unattainable at low temperatures, and extrapolations of high temperature experiments to lower temperature regimes are fraught with uncertainties unless accurate volume, entropy and mixing properties are available for each phase. It is on the basis of reversed experiments in a range equivalent to those in natural systems that most metamorphic thermometers are successfully applied (e.g., Essene 1982, 1989).

THE BASIS OF CLAY MINERAL THERMOMETRY

Current applications of clay mineral thermometry are based largely on empirical calibrations as observed for natural clay-bearing systems. Several kinds of clay thermometers are widely used, but some of the more prominent are: (1) illite and chlorite “crystallinity” based on variations observed in line broadening from powder X-ray diffraction (XRD) patterns (e.g., Frey 1987), (2) chlorite thermometry grounded on correlation of natural chlorite compositions with temperature (Cathelineau and Nieva 1985, Walshe 1986, Cathelineau 1988, de Caritat et al. 1993), and (3) the reaction for which smectite is a reactant and illite/smectite (I/S) or illite a product (e.g., Hower et al. 1976, Hoffman and Hower 1979). Reaction rates of clay mineral transformations have also been used to constrain time-temperature relations during sediment diagenesis (Huang et al. 1993).

The development of clay mineral thermometers is based on observations that show that clay minerals predictably and repeatedly undergo the same sequence of transformations with increasing grade of diagenesis and metamorphism. It is often assumed that equilibrium relations are implied by the repeatability of clay mineral assemblages in space and time, and as a function of increasing temperature (e.g., Velde 1985, 1992a, 1992b, Velde and Medhioub 1988, Loucks 1992). An example of reproducible reactions is found in the sequence volcanic glass → palagonite → smectite → I/S → illite → muscovite, with the proportion of illite in I/S increasing with diagenetic grade, and with crystallinity of illite and muscovite changing with increasing grade. It is often assumed that repeated associations demonstrate that clay mineral systems are in a state of stable equilibrium at each step of such continuous sequences (e.g., Velde 1985, 1992a, 1992b). However, it is not generally recognized that although repetition is a necessary condition, it is not proof of equilibrium. There are many geological examples of systems in which reproducible changes occur under non-equilibrium conditions, e.g., formation of magnesian calcite, calcian dolomite or aragonite from marine water, growth of opal from fluid supersaturated in silica, maturation of hydrocarbons during diagenesis, and formation of glass during rapid quench of a siliceous melt. Few would argue that such processes involve systems at equilibrium, and yet they are as reproducible as are sequences with clay minerals.

A better explanation for clay mineral reactions involves the Ostwald step rule, which states that a system in a state of metastable equilibrium will proceed via steps toward a state which has minimum free energy, i.e., the system will move toward a state of stable equilibrium (Ostwald 1897, van Santen 1983, May et al. 1986). The evolution of clay minerals in diagenetic and
low grade metamorphic sequences may be interpreted in terms of Ostwaldian steps, whereas equilibrium relations have yet to be demonstrated for most clay minerals during diagenesis.

DO CLAY MINERAL THERMOMETERS FULFILL REQUIREMENTS FOR GEOTHERMOMETERS?

Four requirements for a useful thermometer—characterization of phases, equilibrium conditions, constraints on variables, and experimental calibration—are each discussed in this section in the context of general relations of clay minerals.

Characterization of phases

Transmission electron microscope (TEM) studies of a wide variety of diagenetic clays reveal that they are comprised of complex submicroscopic intergrowths of minerals, including clay minerals such as smectite, chlorite and/or illite (e.g., Peacor 1992, and references therein). These studies have shown that individual packets of layers have thicknesses on the order of tens or a few hundred Ångstroms. The clays comprise complex submicroscopic intergrowths of individual packets of authigenic chlorite, berthierine, I/S, and illite, etc., which may coexist with detrital grains of phyllosilicates in a complex array. Moreover, packets of authigenic phyllosilicates that form during early diagenesis commonly contain mixed layers of other phyllosilicates, and even the smallest packets are commonly not structurally homogeneous. For example, chlorite is commonly interlayered with berthierine, and illite with smectite or chlorite (Peacor 1992). Mixed layering may occur either as locally ordered or disordered layer sequences, but many sequences are unique, even in the same sample. In addition, compositions of a given layer type are generally variable from packet to packet and within a given packet. Any bulk chemical analyses of these minerals, including electron microprobe measurements, must of necessity represent averages over the intergrowths. The effective area of electron microprobe analysis is on the order of 1–2 μm², one to two orders of magnitude greater than average packet sizes for authigenic illite (Merriman et al. 1990). Even analyses of the most carefully prepared separates do not correspond only to a single clay mineral of a single mode; TEM observations show that even 0.1 μm-sized illite separates of Gulf Coast illite contain small quantities of rutile, apatite and other clay minerals (Ohr et al. 1994). Measurement of K/Ar isotope data on separates indicates that pure authigenic material cannot be entirely separated from detrital clays (Pevear personal communication). Analysis may appear to represent relatively homogeneous materials because they are averages over the same diverse suites of materials. Even analytical electron microscope (AEM) analyses, for which the area of analysis approaches 100 Å, may be subject to contamination (Jiang et al. 1994b).

With few exceptions, analytical data for clay minerals are invariably subject to multiphase contamination and thus to inaccuracy. Such data are still useful, but they are less accurate than analyses of most minerals from higher grade rocks, for which large homogeneous crystals are generally available. Bulk analytical data on clay minerals must therefore be used cautiously. For example, analytical data for the samples leading to definition of a chlorite geothermometer and some of those for which applications of the thermometer have been applied may have been perturbed by contamination (Shau et al. 1990, Jiang et al. 1994b).

Identification of most clay minerals is obtained by powder X-ray diffraction (XRD). Although such data provide a wealth of information, they suffer from two problems. First, lower limits of detection may be considerable when more than one clay mineral is present, and small amounts of mixed-layered material may remain undetected (Shau et al. 1990). Although XRD provides characterization over large volumes as compared with TEM, only average results are obtained, and heterogeneities, whose presence is not permitted by the phase rule and which reflect lack of equilibrium, may go undetected. Identification of intergrowths at high resolution is thus a unique advantage of TEM/AEM, which has revealed many kinds of heterogeneities in clays subjected to diagenesis or low-grade metamorphism.

Assessment of equilibrium

Application of clay thermometry implicitly assumes that relevant diagenetic clays are single phase materials that have approached equilibrium compositions (Loucks 1992, Velde 1992a, 1992b). This presumption was questioned by Lippmann (1981, 1982), who concluded that weathering and diagenesis would produce a mixture of chlorite and muscovite instead of clay minerals if equilibrium were attained. Based on general observations of clay mineral compositions, Lippmann (1981) noted that the solid solution ranges of diagenetic clay minerals expand with decreasing temperature, the reverse of behavior expected for solid solutions as a function of temperature. Such relations are comprehensible only if the expanded composition ranges correspond to minerals in a state of metastable equilibrium and cannot represent states of stable equilibrium. Jiang et al. (1990a) described pyrophyllite-muscovite that formed in shales and slates, implying that all illite is metastable (Lippmann 1981). However, these interpretations were questioned by Aja and Rosenberg (1992) and by Loucks (1992). Aja and Rosenberg (1992) argued that illites are restricted in their composition, although their composition plot shows a large range of solid solution for illite; they did not address the extensive solid solution observed for smectite. We
maintain that clay minerals with extensive solid solutions, such as smectite and illite, are in high entropy states inconsistent with equilibration at low temperatures. Loucks (1992) argued that bulk analytical data indicate the hitherto neglected importance of additional H$_3$O$^+$ and H$_2$O components in illite. This interpretation was questioned by Jiang et al. (1994a), who found no evidence of such components with AEM analyses on Loucks' best example of illites with these components. We support Lippmann's (1981, 1982) conclusions that illite and smectite are metastable regardless of the incomplete nature of the evidence that was presented in support of his views.

Observations made during TEM studies of pelitic rocks commonly reveal the presence of heterogeneous detrital minerals that persist with diagenetic minerals, whereas the strict adherence to equilibrium requires phase homogeneity (e.g., Li et al. 1995). Such associations violate the phase rule and persist in a non-equilibrium state. Likewise, the ubiquitous presence of local variations in composition, stacking disorder and other features observed by TEM are inconsistent with requirements of homogeneity. The heterogeneous nature of diagenetic clay mineral systems is therefore unambiguously inconsistent with the phase rule and therefore with an attainment of stable equilibrium. In addition, the general prograde course of clay diagenesis through low grade metamorphism is characterized by increasing homogeneity (Peacor 1992). At temperatures of $\geq$ca. 300°C, mineral assemblages take on the characteristics of those at equilibrium, with a small number of homogeneous phases as consistent with requirements of the Gibbs phase rule.

Variables affecting clay mineral reactions

Even for reactions occurring under conditions of stable equilibrium, temperature can be accurately determined only if it is the dominant variable affecting the system. If the composition of a given mineral is sensitive to bulk rock composition, pressure, or activities of exotic species in solution, temperature can only be determined where these other variables are constrained and their combined effect on resultant values of temperature is assessed. Many of the systems used in thermometry of diagenetic rocks have been strongly criticized by Essene (1982, 1989) on the basis that they do not represent equilibrium and that their states are functions of variables other than temperature. He noted that illite crystallinity, conodont color indices and coal reflectivity all depend on irreversible reactions involving metastable phases that are controlled by factors that affect the rate of reaction. Tectonic stress may activate dissolution of reactant clays and neocrystallization of clay minerals (e.g., Lee et al. 1985). Difference in fluid/rock ratio, presence of organic material, or variation in methane pressure may affect clay mineral transitions (Itaya 1981, Wintsch et al. 1981, Yau et al. 1987, 1988, Whitney 1990, Eberl 1993, Small 1993, Li et al. 1995).

Transformation of metastable to stable assemblages is an irreversible process that depends on time as well as temperature. Metastable phases and assemblages tend to have disordered structures and high entropy configurations of atoms, and they may form more readily in systems initially far from equilibrium (Chernov and Lewis 1967). Irreversible reactions are sensitive to fluid/rock ratio and deformation history, as any parameter that accelerates dissolution will tend to expedite transformation to the more stable assemblage. If smectite is less stable than illite in diagenetic sediments, the transformation of smectite to illite will depend on many variables in addition to temperature, and the degree of completion of this reaction is best described in terms of reaction progress rather than by an approach to reversible equilibrium. In contrast, an equilibrium reaction will proceed in either direction depending on displacement from the equilibrium locus of the reaction. Clay mineral reactions are generally controlled by kinetic factors, and therefore are not measures of an approach to equilibrium.

Interpretation of experimental data

Conditions for experimental determination of equilibria were discussed by Fyfe (1960), but his concerns about appropriate procedures have been neglected by many subsequent workers, particularly those interested in low temperature equilibria. Direct experimental reversals of equilibria are obtained by documenting growth and/or dissolution of solid reactant(s) and product(s). Syntheses of stable run products by reaction of less stable starting materials such as oxide mixtures, gels or metastable silicates have no bearing on the relative stabilities of run products. Unfortunately, most experiments on the stability of chlorite solid solutions (Fawcett and Yoder 1966, Fleming and Fawcett 1976, McOnie et al. 1975, Cho and Fawcett 1986) and phengite solid solutions (Velde 1965, 1969, Massonne and Schreyer 1987) are based on unreversed experiments carried out at high temperatures and have limited value for equilibrium relations. Solubility experiments on the stability of clay minerals, which may be obtained in the temperature range of diagenetic systems, have additional difficulties related to problems in establishing actual reactions that control solubility products. Exchangeable cations will equilibrate with solutions far more rapidly than essential constituents of the clay sheets. Experiments that only monitor such cations may apparently be reversed yet do not document stability of the clay framework. This situation is analogous to cation exchange experiments on feldspars, where highly unstable phases are readily synthesized. As a result of surface energy effects, dissolution of very fine grained materials may comprise excessively soluble fractions that do not correspond to equilibrium solu-
bilities of a coarse homogeneous phase. Similar effects may occur on surface layers in a highly comminuted clay reactant. Once fine grained components are completely dissolved, the time frame for additional dissolution will increase dramatically, and an experimentalist may mistakenly conclude that steady state equilibrium has been achieved. Unless reaction direction can also be approached from the direction of oversaturation causing growth of the same solids, all such experiments must be viewed with skepticism. At temperatures below a few hundred degrees Celsius, reaction rates near equilibrium conditions are so slow that reversals are unlikely to be obtained. It is only with solutions that are substantially out of equilibrium with coexisting solids that measurable reaction can occur, and production of metastable solids by irreversible reactions should be the rule with most weathering reactions, low temperature experiments, and many diagenetic processes.

Even in systems with metastable phases whose survival is dependent on reaction rates, laboratory calibration may not be directly applicable to rock systems unless the transformation mechanism is the same in nature as in the laboratory. Suzuki et al (1993) proposed simplifications in kinetic models of vitrinite reflectance but cautioned that they "... represent gross simplifications of what must be an enormously complex natural system, composed of thousands of parallel reactions producing many products not even considered in these models". The same caveats apply to clay mineral systems.

EXPERIMENTAL STUDIES OF CLAY STABILITY

Synthesis of illite (Velde 1969), phengite (Velde 1965, Massonne and Schreyer 1987) and chlorite (Fawcett and Yoder 1966, McOnie et al 1975, Fleming and Fawcett 1976) solid solutions from metastable reactants document that they are more stable than starting materials but do not constrain equilibrium stability relations among possible run products (Fyfe 1960). Solubility measurements of illite (Kittrick 1984, Sass et al 1987, Aja et al 1991a, 1991b), smectite (Aja et al 1991a, 1991b) and chlorite (Kittrick 1982, Thornton et al 1986) have uncertain significance for either metastable or stable phase equilibria as they are fine-grained and likely to be submicroscopically heterogeneous. The meaning of solubility measurements on fine grained materials is difficult to ascertain as the reactivity of fine particulates is elevated compared to equivalent coarse phases as a result of the effect of surface energy (May et al 1986, Walther and Wood 1986). Composition and structure may vary in a regular way as a function of size, so that dissolution may occur preferentially for grains that do not represent the average composition and structure. The concept of a phase cannot extend down to scales of hundreds of nanometers where surface energy effects provide a significant contribution to the total free energy of each particle. At best, solubility measurements on clay minerals may provide only upper limits on equilibrium solubilities.

Garrels (1984) used water analyses of solutions in contact with clay minerals compiled by Aagaard and Helgeson (1983) to infer that illite/smectite intergrowths act as a mixture of two discrete phases rather than solid solutions. However, the line in pH/pK vs. aSiO₂ space may be controlled by the availability of exchangeable or absorbed cations rather than fixed by equilibration with a heterogeneous multiphase reaction. Sass et al (1987), Aja (1991) and Aja et al (1991a, 1991b) obtained experimental bounds on the stability of illite and other phyllosilicates at 25–250°C and low unconstrained pressure, which should lie on the liquid/vapor boundary for the aqueous solution. But here, run products were not characterized and were inferred only from the solution chemistry assuming that equilibrium had been attained in the experiments. The use of AEM and STEM studies on starting materials and run products is requisite to an understanding of reactions that take place in solubility experiments.

The phase boundaries shown by Aja (1989) and Aja et al (1991a) for the stability field of illite in pK/pH vs. T space (Figure 1a) violate Schreinemakers' 180° rule, which requires that new assemblages must constrain but not expand the stabilities of other assemblages (Zen 1966). Application of the 180° rule to the data of Aja (1989) and Aja et al (1991a) indicates that muscovite-microcline-quartz-H₂O and muscovite-kalolinite-quartz-H₂O are more stable than illite-quartz-H₂O at low temperatures (Figure 1a and b). A similar relation exists for the data presented by Aja (1991), although he did not construct graphs of pK/pH vs. T. If at equilibrium, these diagrams require that the stability field of illite is located at high temperatures and the muscovite stability field at low temperatures (Figure 1b)! This improbable inference indicates that illite stability is still more constrained than that displayed by Aja et al (1991a) and that equilibrium was not attained in their experiments.

Figure 1. a. Diagram of $\log(aK^+/aH^+)$ vs. $T$ in equilibrium with quartz from Aja et al (1991a). Points represent their experimental determinations. Dashed lines represent their calculated stability fields for kaolinite, muscovite and microcline. The kaolinite-illite boundary is metastable with respect to the calculated kaolinite-muscovite boundary if both are at equilibrium. Figure b. Diagram of $\log(aH^+/aK^+)$ vs. $T$ in equilibrium with quartz. Points represent the same experimental determinations as in Figure 1a. Application of Schreinemakers' 180° rule (Zen, 1966) to the data suggests that illite is the high temperature phase and muscovite the low temperature phase. This unreasonable result is regarded as strong evidence that some or all the experiments failed to attain equilibrium.

et al (1994). If illite or smectite have a field of stability or metastability, it will be further constrained relative to stable diaspore. Aja et al (1991a, 1991b) concluded that they had reversed their experiments by obtaining measurements from initially high as well as low $K^+/H^+$ and over- vs. undersaturated $\text{SiO}_2$, but these cannot be accepted as reversals of the heterogeneous equilibria. The measured pH/pK and 2pH/pMg may be controlled by exchangeable or absorbed cations, dissolution of fine crystallite fractions, or reactions with surface layers only. Whether or not it is metastable, illite of known composition, grain size, and structure must be documented to grow as well as dissolve to establish true experimental reversals. Until the effects of variable grain size, structure and composition are independently evaluated one should not rely on solubility experiments on clays as approaching equilibrium conditions (May et al 1986).

**THERMODYNAMIC PROPERTIES OF CLAYS**

Other than for a very few phases of simple composition such as kaolinite (e.g., Anovitz et al 1991), thermodynamic data on most clay minerals have been obtained from solubility measurements or from summation estimates. It is unclear whether solubility measurements can be shown to have attained equilibrium, as clay minerals are by definition comprised of sub-micrometer crystallites with complex physical and chemical characters (May et al 1986). Summation techniques have also been used to estimate the thermodynamic properties of complex phases from those of simple end-members (e.g., Tardy and Garrels 1974, Nriagu 1975, Mattigod and Sposito 1978, Stoessell 1979, 1981, Tardy and Fritz 1981, Merino and Ransom 1982, Aagaard and Helgeson 1983, Sposito 1986, Chernak and Rimstidt 1989, Tardy and Duplay 1992, Ransom and Helgeson 1993, 1994). Summation approaches yield approximately correct results for solids because of the plague of small free energy differences (Fyfe et al 1958), but it is those same small differences that make all the difference in terms of reaction direction and whether a given phase is stable or metastable. Such techniques generally do not correct for cation disorder (e.g., Si/Al, □/K) that may occur on cation sites of clay minerals. Strict additivity approaches imply that complex compounds have the same Gibbs free energies as simple components, which cannot exactly be correct. Any corrections to the summation technique involve model-dependent terms that arbitrarily stabilize or destabilize solid solutions. Whether illite is calculated to be more or less stable than pyrophyllite-muscovite is dependent on the activity model chosen for illite, and does not provide an independent test (Jiang et al 1990a). Summation approaches for Gibbs free energy estimates of clay minerals cannot be expected to have precisions adequate for the basis of a clay thermometer.

**IMPLICATIONS FOR METASTABILITY OF CLAY MINERALS**

The relations described above clearly show that neither laboratory experimentation nor thermodynamic calculation can demonstrate that specific clay minerals form or exist in states of stable equilibrium in rocks subject to diagenesis or low grade metamorphism. It is therefore necessary to look to natural assemblages for evidence of stable equilibrium or the lack thereof.
The relations described above provide abundant evidence for the lack of stable equilibrium in most natural clay systems. Lippmann's insightful observations (1981, 1982) that clay solid solutions decrease with increasing temperature is a general reflection of metastability. Each clay packet has unique chemical properties that are dependent on shape, size, surface area and structure. Moreover, the definition of a phase with fixed thermodynamic properties per mole for a given composition, temperature and pressure is inappropriate for application of submicroscopic materials. The ultimate optical, microprobe and SEM resolution at a scale of 0.1–1.0 μm is comparable to the minimum scale for which a crystalline phase and therefore a mineral should be defined. Certainly any material that shows line broadening on powder XRD patterns is suspect as a phase and in terms of being adequately defined with stable phase equilibria. We conclude that general observations of most clay minerals are inconsistent with stable equilibrium conditions.

APPLICATIONS TO SPECIFIC CLAY THERMOMETERS

Each of the commonly applied clay thermometers—chlorite and illite crystallinity, chlorite thermometry, and illite/smectite reaction thermometry—is addressed individually below.

Chlorite and illite crystallinity

Observations of line broadening in powder X-ray diffraction (XRD) patterns of illite and chlorite separates are generally described as "crystallinity" measurements (Kübler 1967, Kisch 1983, Frey 1987). Illite crystallinity has been used widely to evaluate variations in temperatures of sediment diagenesis or low grade metamorphism (e.g., Kisch 1983, Roberts and Merriman 1985, Robinson and Bevins 1986). Variation in illite crystallinity is generally ascribed to the general state of "perfection" of illite crystallites, which is a function of compositional heterogeneity, crystal size distribution, defect density (strain), and variation in structure of mixed layers. Correlation of TEM measures of crystal size distributions with XRD data reveal that variation in illite crystallinity is primarily related to crystallite thickness (Merriman et al 1990, Jiäng et al 1994b). Peacor (1992) emphasized that the contributions of the four factors decrease concomitantly as a function of increasing grade of metamorphism. Increase in grain size has been ascribed to Ostwald ripening, wherein larger crystals grow at the expense of smaller crystals of the same composition with increase in grade (e.g., Eberl et al 1990). This process is driven by the relatively large surface areas and therefore higher free energies of small crystals in relation to larger crystals of the same species, and also by other kinds of heterogeneities, including high defect densities. Small grain sizes, on the order of 100 Å thickness for illite, and high defect densities of initially formed crystals are an expression in part of low temperature during formation. The process of coarsening with increasing temperature is, by definition, one concerning metastable states, and crystal size is therefore kinetically controlled at low grades. Although increase in grain size is affected by increase in temperature, other parameters such as variable time of heating, stress rate and magnitude, and fluid/rock ratio are also important factors. Moreover, increase in grain size of most clay minerals is accompanied by structural and/or chemical change and therefore is not adequately represented as Ostwald ripening. Grain size of clay minerals is therefore not a reliable thermometer. Average grain size of both metamorphic and diagenetic rocks represents an interplay between grain growth, pressure solution and recrystallization, which is controlled by variable deformation and access of fluids. The use of illite crystallinity as a simple geothermometer is clearly flawed.

Chlorite thermometry

Cathelineau and Nieva (1985) proposed an empirically calibrated thermometer based on electron microprobe analyses of Al in chlorites from geothermal systems. They did not identify a specific reaction that would fix the substitution of Al in chlorite. Walshe (1986) considered chlorite solid solutions in terms of six components and wrote various reactions in equilibrium with quartz and water involving the components. His thermodynamic mixing model involves an ideal site-by-site approximation, although non-isovalent exchanges in multicomponent solid solutions are unlikely to be ideal (Essene 1989). Cathelineau (1988) obtained additional analytical data on geothermal chlorites and revised the chlorite thermometer. Shau et al (1990) found that most "chlorite" occurring in diagenetic and low grade metamorphic rocks contains interlayers of corrensite that produce anomalously high AlVIV/AlIV, Ca, Na and/or K. They noted that many of the chlorite analyses of Cathelineau (1988) have similar chemical characteristics and suggested that his chlorite thermometry records variable amounts of corrensite packets rather than representing a monotonic solid solution of a single phase responding to increasing temperature. Because the proportion of corrensite will decrease with increasing grade as a reflection of increasing homogeneity, apparent change in "chlorite" composition with grade is expected, but it is not an equilibrium function and therefore not an accurate thermometer. The amount of corrensite interlayered with chlorite may depend on the rate of nucleation and crystal growth, therefore indirectly on the degree of supersaturation in solution as well as temperature.

Aagaard and Jahren (1992) related an increase in Al content of authigenic chlorite in sandstones to a Tschermak's type exchange reaction with coexisting illite under increasing temperature. The data are EDS-
TEM analyses from Jahren and Aagaard (1992) that show a broad scatter in the Al contents of illite and chlorite; the chlorite analyses also have a substantial excess of Al$^{IV}$ over Al$^{III}$, indicative of corrensite interlayers or of substantial dioctahedral substitution. Jahren and Aagaard (1992) proposed that the illite formed from stable kaolinite and K-feldspar, but this involves an irreversible reaction with an unstable initial assemblage (Garrels 1984, Aja 1991, Jiang et al 1994a).

Literature on chlorite thermometry was reviewed by de Caritat et al (1993). They compared various approaches to chlorite thermometry of geothermal systems and concluded that no single chlorite thermometer was uniformly successful. However, six of the nine chlorite analyses listed by de Caritat et al (1993) show a large excess of Al$^{IV}$ over Al$^{III}$ and an apparent deficiency in the octahedral site, prima facie evidence for the presence of submicroscopic corrensite interlayers (Shau et al 1990). Jiang et al (1994b) compared analyses of ordinary authigenic trioctahedral chlorites as found in the literature with analyses as determined by AEM, verifying that multiphase contamination is the rule, and noting that inequalities in Al$^{IV}$ and Al$^{III}$, apparent octahedral vacancies, and small Ca, Na, K contents of so-called “chlorite” are caused largely but perhaps not entirely by contamination. We question the applicability of thermometers to bulk specimens whose chemical properties are determined by a physical mixture of two structure types of different chemistry. For true chlorites, heterogeneous reactions may be considered for calcareous, mafic and clastic rocks, representing the change in the Al content with two different chlorite components, here clinochlore, Mg$_3$Al$_2$Si$_3$O$_{10}$[OH]$_8$, and “serpentine,” Mg$_3$Si$_4$O$_{10}$[OH]$_8$. The Al content of chlorite is related to the two chlorite components and is buffered by the following assemblages and reactions:

calcite-epidote-fluid:
$$5 \text{Mg}_3\text{Si}_4\text{O}_{10}[\text{OH}]_8 + 4 \text{Ca}_2\text{Al}_2\text{Si}_3\text{O}_{12}[\text{OH}] + 16 \text{H}^+ = 6 \text{Mg}_5\text{Al}_2\text{Si}_3\text{O}_{10}[\text{OH}]_8 + 14 \text{SiO}_2 + 6 \text{H}_2\text{O} + 8 \text{Ca}^{2+} \quad (4)$$
muscovite/illite-microcline-quartz-fluid:
$$6 \text{KAl}_3\text{Si}_3\text{O}_{10}[\text{OH}]_2 + 5 \text{Mg}_3\text{Si}_4\text{O}_{10}[\text{OH}]_8 = 6 \text{Mg}_5\text{Al}_2\text{Si}_3\text{O}_{10}[\text{OH}]_8 + 6 \text{KA}\text{Si}_3\text{O}_8 + 2 \text{SiO}_2 + 2 \text{H}_2\text{O} \quad (5)$$
and muscovite/illite-quartz-fluid:
$$4 \text{KAl}_3\text{Si}_3\text{O}_{10}[\text{OH}]_2 + 5 \text{Mg}_3\text{Si}_4\text{O}_{10}[\text{OH}]_8 + 4 \text{H}^+ = 6 \text{Mg}_5\text{Al}_2\text{Si}_3\text{O}_{10}[\text{OH}]_8 + 14 \text{SiO}_2 + 2 \text{H}_2\text{O} + 4 \text{K}^+ \quad (6)$$

Reactions (1)–(6) provide constraints on the Al content of chlorite solid solutions as a function of each specific assemblage. Contrary to the implications of chlorite thermometry as it is usually applied, different specific assemblages will drive the Al content of chlorite in different directions with increasing grade. Similar reactions may be written for clinochlore and amesite components, for iron chlorite analogues, with kaolinite or pyrophyllite, with laumontite or prehnite instead of clinzoisite, or with pumpellyite instead of actinolite. These reactions have a great advantage over many of those previously proposed for chlorite thermometry as they are independent of the activity of a small amount and an uncertain dioctahedral aluminum or ferric iron component in chlorite solid solution. Liou et al (1987) and Frey et al (1991) considered reactions similar to (1) and (3) but balanced equilibria only with a fixed chlorite composition and therefore did not explicitly address chlorite thermometry. Unfortunately, all reactions currently proposed for chlorite thermometry suffer from a similar failing, the lack of reliable activity/composition relations for chlorite solid solutions experimentally calibrated in the temperature range of applications. Until such data become available, application of any reactions involving chlorite remain speculative.

Smectite to illite reaction

The groundwork for the smectite/illite geothermometer was established largely by research on diagenesis of Gulf Coast mudstones by Powers (1959, 1967), Burst (1969), and Perry and Hower (1970). Hower et al (1976) detailed progressive change of the sequence from smectite-rich R0 I/S to R1 I/S to illite-rich I/S as a function of depth and temperature based on the method of Reynolds and Hower (1970). The smectite to illite transformation has been utilized as a geothermometer in a variety of ways, including maximum temperature of reaction (e.g., Bish and Aronson 1993, Pollastro 1990, 1993), temperature directly measured by reac-
tion rate (e.g., Huang et al. 1993) and temperature where the specific dioctahedral clay species is viewed as reflecting a stable state as a function of temperature (e.g., the sequence R1 I/S through muscovite as hypothesized by Hower and Altaner 1983).

The latter two relations imply that smectite is metastable relative to illite-rich I/S or illite at the temperature at which reaction takes place, but they do not require that illite is metastable. They are also consistent with metastability of smectite even at the low temperatures corresponding to its formation, but where rates of reaction to produce illite are exceedingly slow. Although seldom stated, it is generally assumed that reaction occurs only when a temperature is reached at which illite is the stable phase relative to smectite, i.e., when the stability field of smectite is exceeded. There is an implied sense that although reaction progress is controlled by kinetic factors, the reaction proceeds only when the stability field of smectite is exceeded and it becomes metastable relative to stable illite (Sass et al. 1987, Aja et al. 1991b). The reactions for which smectite is a reactant and illite a product are generally studied with the assumption that at least illite is a thermodynamically stable phase. However, diagenetic sediments containing smectite, I/S, and/or illite display compelling evidence of a lack of chemical equilibrium in violation of the phase rule. Equilibrium has yet to be demonstrated for such reactions, and the presence of smectite, I/S, or illite is not a measure of absolute temperature. The formation of illite at very low temperatures normally regarded as corresponding to the stability field of smectite (e.g., Ahn et al. 1988, Turner and Fishman 1991), is definitive evidence for the metastability of smectite. Recognition of smectite, I/S, and illite metastability leads to some usefulness in utilizing clays as general indicators of maximum temperature ranges or as measures of rate-determined reactions (Nadeau and Reynolds 1981).

OSTWALD STEP RULE

If clay mineral reactions that involve metastable phases either as reactants or products or both are irreversible, and are functions of several variables affecting rates of reaction, why do they progress through the same sequence of products? The answer lies in an understanding of the Ostwald step rule. Morse and Casey (1988) reviewed aspects of the Ostwald step rule in application to several types of sedimentary systems, including clay mineral systems. They noted that clay diagenesis displays the general pattern of the Ostwald step rule, that the smectite-to-illite reaction may be an example, and that the heterogeneity of smectite at or near the unit cell level implies metastability. The Ostwald step rule predicts that a system that is not in a state of stable equilibrium will undergo reaction(s) in which the subsequent system state is one of lower free energy. A system in a state of metastable equilibrium will proceed via reaction steps toward the state of minimum free energy, each intermediate state being of lower free energy than the preceding one. Morse and Casey (1988) showed that a given metastable system may change via several reactions, but if the rate of reaction leading to the stable system is slow, and that leading to an intermediate metastable system is rapid, the latter will prevail at the expense of the former. Where a reactant and product have similar crystal structures, reaction rates tend to be rapid. Many TEM observations document the direct replacement by structurally related phyllosilicate reactants and products (Veblen 1992).

The sequence of reactions from smectite to I/S to illite to muscovite is one in which there is a shift in composition and crystal structure toward those of muscovite, and in which any intermediate material is more like that immediately preceding than the stable end-member muscovite. The transformation sequence proceeds inexorably to a point of maximum order where $\text{Al/Si} = 1/3$. If intermediate states (e.g., illite) did not exist for which reaction rates are rapid, muscovite would be a common mineral in sediments undergoing diagenesis or smectite would persist to greater depths (Lippmann 1981). The rate of irreversible reactions involving metastable materials determines the amount of specific reaction products. The Ostwald step rule predicts that a complex, heterogeneous clay mineral system will progress through intermediate steps toward a state of chemical equilibrium in which a small number of homogeneous phases, as is observed, is consistent with the phase rule (Peacor 1992). A corollary of these relations is that clay mineral reactions proceed via steps having large activation energies and therefore requiring input of energy by tectonic stress or pressure solution. The episodic nature of clay mineral reactions is thus consistent with the step rule (Peacor 1992).

WHAT CAN CLAYS CONTRIBUTE TO GEOTHERMOMETRY?

Even though clay mineral systems can not be used as accurate thermometers, we do not imply that they are not useful in providing information about metamorphic grade. Indeed, there is a wealth of data that demonstrate that much useful information can be derived about approximate temperatures from clay mineral assemblages. Such estimates can be obtained in three cases: (1) recognizing that rate of reaction of metastable phases is the critical factor, all pertinent variables except temperature may be constrained (Huang et al. 1993); (2) approximate maximum temperatures may be obtained through the observation of a mineral known to form only early in prograde sequences; and (3) "reaction progress" can be measured. The correlation that is generally observed in regionally metamorphosed rocks between metamorphic grade and a wide range of so-called thermometers, each of which
is subject to the problems discussed above, represents an application of the above principles. If all variables that affect reaction rates have equal effects, on average, and if pressure-temperature-time (P-T-t) paths of low-grade rocks are similar, all rocks will show the same average changes as a function of grade. Thermometers such as illite and chlorite crystallinity, CCI, and reactions in metabasites have been compared with one another (e.g., Kisch 1981, 1983, Frey 1987). The correlations are generally reasonable, and no one should doubt their value for rough estimates of temperature. Given the above analysis, however, the estimates are at best approximations. Attempts to quantify temperature for systems having different thermal or tectonic histories, bulk compositions or textures cannot give rise to accurate determinations of temperature of formation.

Many authors have indirectly evaluated the rate at which smectite is converted to illite (Waples 1980, Hower and Altaner 1983, Bethke and Altaner 1986, Pytte and Reynolds 1989, Huang et al 1993, Pollastro 1993, Price and McDowell 1993). Price and McDowell (1993) advanced a T-t model for clay transformations observed in 1.1 Ga old pelites of the Nonesuch Formation, Michigan. Pollastro (1993) inferred that different T-t relations apply to sediments of Miocene through Mississippian age and those < 3 Ma in age. Huang et al (1993) experimentally determined rates of reaction, developed a kinetic model, and successfully applied that model to several basins of relatively recent age. These studies employ the rate-dependent nature of the reaction and the significance of temperature and concentration of K+ or other cations. Huang et al (1993) experimentally investigated rates for I/S transformations with K+ concentrations of approximately 200 ppm. They acknowledged that not all relevant variables were considered, as in the case of Mg concentration, the state of the smectite used as a starting material in their experiments, and pH in relation to the presence of organic acids generated during diagenesis. Huang et al (1993) emphasized the complexity of variables that may affect the rate of the I/S reaction and cautioned that constraint of all such variables is necessary for both synthetic and analogous natural systems before even approximate T-t relations can be established. However, the complexity emphasizes the approximate nature of such results, even though they may lead to apparently useful applications.

Limitations on maximum temperatures are obtained assuming that some clay minerals cannot exist at temperatures greater than some maximum values (e.g., Bish and Aronson 1993). Except where hydrothermal alteration has overprinted higher grade white micas (Nieto et al 1994), smectite is generally not observed in rocks subjected to temperatures as high as 200°C, and its occurrence can be taken to represent a very approximate maximum temperature. Such relations are gross approximations, given the metastability of such phases and the plethora of relevant variables affecting the clay reactions, and are really measures of reaction progress.

If temperature cannot be accurately determined for any given system, what can be determined? A corollary of the Ostwald step rule is that a diagenetic system with a given bulk composition will take a well defined path by which smectite transforms to I/S and illite, and which through increase in grain size gives rise to changes in illite crystallinity through reaction steps that proceed in a prograde (dehydration) direction. The concept of reaction progress is a measure of the degree to which a given reaction has advanced. Smectite may react to form illite over a wide range of temperatures. Once illite has formed through such an irreversible reaction, it serves as a reactant in subsequent steps. Observation of the amounts of I/S, illite, or illite of a given size distribution, therefore, marks the extent of progress along a series of reactions. However, the actual reaction between smectite and illite remains poorly defined. Smectite (as a first approximation represented by a pyrophyllite-like formula) will react in the presence of K-feldspar to form illite (represented by a muscovite formula):

$$\text{Al}_2\text{Si}_4\text{O}_{10}\text{[OH]}_2 \cdot n\text{H}_2\text{O} + \text{KAlSi}_{3}\text{O}_4$$

$$= \text{KAl}_3\text{Si}_{3}\text{O}_{10}\text{[OH]}_2 + \text{SiO}_2 + n\text{H}_2\text{O} \quad (7)$$

This reaction allows reaction progress to be monitored by the increase of quartz. In many shales, however, there is inadequate K-feldspar available and quartz produced for this to be the controlling reaction. In the absence of K-feldspar, introduction of K+ ions permits ionic reactions to proceed without transport of Al3+, e.g. by:

$$2\text{K}^+ + 3\text{Al}_2\text{Si}_3\text{O}_{10}\text{[OH]}_2 \cdot n\text{H}_2\text{O}$$

$$= 2\text{KAl}_3\text{Si}_3\text{O}_{10}\text{[OH]}_2 + 6\text{SiO}_2 \text{(aq, xl)}$$

$$+ 2\text{H}^+ + 3n\text{H}_2\text{O}, \quad (8)$$

which is controlled by pK/pH as well as T (Boles and Franks 1979). Progress of this reaction (and that of similar reactions with more complicated smectite and illite formulae) may be controlled by access of external fluids to supply K ions by dissolution of K-feldspar in adjacent sandstones and H ions by reaction with organic acids. Elsewhere, kaolinite is initially present in the sediment and reacts irreversibly with K-feldspar and/or K+ to form illite. The silica may be removed in solution, precipitate as quartz, or react with detrital plagioclase to form more illite. The I/S reaction (8) involves considerable volume reduction beyond that involved with dehydration. Reaction progress will depend on the exact compositions of smectite and illite requires evaluation of rate-controlling reaction in each location by careful evaluation of modes and compo-
sitions (e.g., Ferry 1986). It is difficult to obtain accurate modes on fine-grained sediments without careful and extensive TEM studies, but some estimates have been made with powder XRD measurements (e.g., Hower et al. 1976, Awwiller 1993, Pollastro 1994). Although reaction progress generally correlates with increasing diagenetic grade, in pelitic rocks it is a measure of relative progress toward states of stable equilibrium.

CONCLUSION

This review paper was motivated by the concern that clay minerals may be used inappropriately as geo-thermometers and specifically in violation of properly constrained relations as exemplified by the Gibbs phase rule. The general relations discussed above do not obviate the use of clay minerals in determining temperatures under sub-greenschist facies conditions, but they do establish constraints under which geothermometers should rigorously be defined. It is our hope that this paper will focus attention on the pitfalls of clay mineral thermometry and perhaps serve as a focus of further discussion about these fundamental issues.

ACKNOWLEDGMENTS

The authors acknowledge the contributions of J. H. Ahn, M. Buatier, R. L. Freed, W.-T. Jiang, J. H. Lee, G. Li, R. Merriman, F. Nieto, B. Roberts, Y.-H. Shau, and L.-C. Yau to their studies of clay minerals. The authors are also grateful for discussions on the stability of phases with L. M. Anovitz, R. L. Freed, J. R. O’Neil and D. Perkins. Reviews of J. M. Budai, R. H. Pollastro, B. H. Wilkinson and another reader materially improved the quality of the manuscript. The research was supported by NSF grants EAR 91-17772 and 91-04565.

REFERENCES


Kisch, H. J. 1981. Coal rank and illite crystallinity associated with the zeolite facies of Southland and the pumpe-


Wintsch, R. P., A. F. O’Connell, B. L. Ransom, and M. J.

