ELSEVIER

Phase engineering and supercompatibility of shape memory alloys

Hanlin Gu [1,†], Lars Bumke [2,†], Christoph Chluba [2,†], Eckhard Quandt [2,*], Richard D. James [1,*]

[1] Aerospace Engineering and Mechanics, University of Minnesota, Minneapolis, USA
[2] Institute for Materials Science, Faculty of Engineering, University of Kiel, Kiel, Germany

In recent years examples of unprecedented functional and structural fatigue resistance and lowered hysteresis in shape memory alloys have been achieved by combining conditions of supercompatibility between phases with suitable grain size and a favorable array of fine precipitates. We collect, review and compare these examples to elucidate the relative roles of these factors, especially in the case of the more demanding stress-induced phase transformations, and we pose key open questions. The control of these factors lends itself to systematic alloy development. Taken together, these results point to significant opportunities to discover improved shape memory alloys as well has new reversible transforming multiferroics.

Introduction

Shape memory alloys (SMAs), also known as ferroelastics by analogy to ferromagnetic and ferroelectric materials, show two distinct properties that are attractive for many applications. First, the shape memory effect is the basis for many solid-state actuators [1–3], which relies on a reversible, thermally induced, high energy density phase transformation between austenite and martensite. Second, they exhibit superelasticity used for example in self-expanding medical implants or in elastocaloric cooling, which is based on a reversible, stress-induced austenite–martensite phase transformation [4]. These first-order phase transformations with large eigenstrains in both cases result in large work output and large enthalpy changes. Their reversibility is aided by the compatibility of the two phases. As a consequence, local stresses at the interfaces between the two phases do not exceed the elastic limit, thus the formation of dislocations is reduced.

An essential factor for the implementation of shape memory alloys in new devices is their fatigue characteristics, especially for high-cycle applications. In general, fatigue concerns two aspects: functional fatigue, which describes the cycle-dependent changes of their functional properties and structural fatigue, which refers to the integrity of the material. Commonly, both types of fatigue are closely interconnected. Often, during superelastic cycling, a residual fraction of nontransforming martensite increases with the cycle number. This degradation of the amount of transformed material reduces the overall transformation strain and enthalpy change, and also affects the lifetime of the material, as the residual martensite leads to a roughening of the material that acts as initiation sites for cracks. Previous investigations have revealed that achieving high cycle lifetime is much more demanding in stress-induced, as compared to temperature-induced, transformation, especially under tensile stress [5].

Functional fatigue and reversibility are especially relevant for the medical device industry. But even the most widely used alloy, near binary NiTi, has significant migration of transformation temperature, typically approximately 20 K after a few hundred cycles [6–9]. This migration is affected by Ni content [10], heat treatment [5,11] and thermomechanical processing history [12,13]. Complete failure after a few thousand cycles under stress-induced full transformation is typical [14]. Despite these shortcomings, near equiatomic NiTi is abundantly exploited in medical devices [15] and is currently the focus of nearly half of

* Corresponding authors.
E-mail addresses: Quandt, E. (eq@tf.uni-kiel.de), James, R.D. (james@aem.umn.edu).
† These authors contributed equally to this study.

https://doi.org/10.1016/j.mattod.2017.10.002

all research on metal alloys [16]. The successful applications of NiTi in medicine are only currently possible because they involve fully (or mostly) transforming the material only once, or just a few times. Recently, new ideas that relate to this review, such as the use of the R-phase to improve the compatibility between phases, are being exploited.

Another important motivation that drives the search for reversible transforming materials is their role as switchable multiferroics. Multiferroics are materials that combine at least two of the three "ferroic" properties ferroelectricity, ferromagnetism, and ferroelasticity [17]. Despite an intense search over the past decade, there has been limited success in finding intrinsic multiferroic compounds that exhibit strong ferroelectricity and/or ferromagnetism above room temperature, and that are switchable. However, a design route based on *multiferroic response driven by reversible phase transformation* has been quite productive. The terminology *phase engineering* refers to this tuning of composition to achieve desirable, often diverse, properties of the phases. The main point is that in materials with strong first-order phase transformations the abruptly changing crystal structure at transformation can give rise to an abruptly changing electronic structure. In this case elasticity, electricity, magnetism and temperature are coupled and can be simultaneously controlled along with transport and optical properties. Such a broad combination creates a host of possibilities for novel actuators [1], microelectronic and optical devices [18], information storage media [19], caloric cooling [20], and energy storage media [21].

Strong first-order phase transformations in crystals also have a latent heat. This allows for the coupling of heat energy to multiferrroic and transport properties, a possibility that is not realized in a significant way in intrinsic multiferroics. If latent heat is present together with an abruptly changing ferroic property then, according to the appropriate versions of the Clausius–Clapeyron equation, there is a first-order effect of electric, magnetic or stress fields on transformation temperature. Such materials have applications to refrigeration (i.e., magnetocaloric, electrocaloric, elastocaloric) [22] and energy conversion devices. Thus, a variety of interesting elastic, electromagnetic and transport properties can be switched on and off by phase transformation. In many potential applications, especially in the caloric ones, we would need to do this repeatedly.

In this review we present the state-of-the-art understanding of the reversibility of phase transformations. Great strides have been made in this direction in the past three years, and these ideas challenge conventional wisdom. But, while conditions of compatibility between phases, precipitates and grain size all clearly play critical roles, the precise influence and interaction of these factors remains a puzzle. We hope to bring some order to this puzzle via this review.

In this review, compositions are nominal and are reported as subscripts of elements taking the values 1–100 as atomic percentage. In the main bulk arc melted specimens reviewed here, this was initial composition; mass loss was checked, and was consistent with composition variation of less than 1%. All TiNiCu-based examples in Section 3 and Fig. 5 in this review are free-standing films fabricated using dc-magnetron sputtering as described in [23]. The composition of the as deposited films was determined using EDX and compared to a binary NiTi standard. The thickness of the films is typically 20–30 μm. The thickness is considerably thicker than the average grain size, therefore the described properties can be transferred to bulk materials if the same microstructure as in the sputtered films is present.

Theory of reversibility of phase transformations

In this section, we review conditions of compatibility between phases vis-à-vis the reversibility and hysteresis of phase transformations. These conditions are theorized to influence the stress in transition layers and the heights of energy barriers that relate to hysteresis, but the conditions themselves are purely geometric and depend only on the crystal structures and lattice parameters of the two phases. While conditions such as volume conservation are considered important, we will focus on two conditions for supercompatibility, $\lambda_2 = 1$ and the *cofactor conditions*, that have a particularly dramatic influence on hysteresis and reversibility.

Geometric conditions on lattice parameters, such as $\lambda_2 = 1$ and the cofactor conditions, are conditions of compatibility between the phases. That is, they allow the growing phases to fit together without slip. Phases can fit together without slip while being separated by stressed transition layers, which is the normal case, for example, of the crystallographic theory of martensite. However, $\lambda_2 = 1$ and the cofactor conditions permit the phases to fit together (in many ways) *without stressed transition layers*. We use the terminology *supercompatibility* for such cases.

Both $\lambda_2 = 1$ and the cofactor conditions are expressed in terms of the *transformation stretch matrix* and the two groups that represent the point group symmetries of the two phases. To define the transformation stretch matrix, we first note that in the most reversible martensitic phase transformations, the point group symmetries of the two phases have a group–subgroup relation, with the high temperature phase having the larger group. Thus, there is a primitive lattice describing the periodicity of the martensite crystal defined by vectors $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, $\boldsymbol{b}_3$ and a *sublattice* of the austenite with periodicity $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ having about the same unit-cell volumes. The transformation proceeds by a linear transformation $\boldsymbol{F}$ of the austenite sublattice to the martensite primitive lattice, i.e., $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$. By switching the sign of one of the vectors, if necessary, one can assume that $\boldsymbol{F}$ has positive determinant. Then, $\boldsymbol{F}$ has a polar decomposition $\boldsymbol{F} = \boldsymbol{R}\boldsymbol{U}$, where $\boldsymbol{R}$ is a 3×3 rotation matrix and $\boldsymbol{U}$ is positive-definite and symmetric. $\boldsymbol{U}$ is the *transformation stretch matrix*. The atoms within the sublattice unit cell of austenite shuffle to their final positions in the primitive unit cell of martensite.

There may be several sublattices of austenite that have about the same unit cell volume as the primitive lattice of martensite. Also, there are infinitely many choices of $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ that describe the same sublattice. It is observed that, given $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, $\boldsymbol{b}_3$, the material very often chooses $\boldsymbol{a}_1$, $\boldsymbol{a}_2$, $\boldsymbol{a}_3$ which gives the smallest strain $\|\boldsymbol{U} - \boldsymbol{I}\|$ measured in a suitable norm. The intuition behind this smallest strain criterion originates from Bain [24] and Lomer [25]. Suitable algorithms for finding $\boldsymbol{U}$ based on this principle are given in [26,27]. Also, there is a close relation between this procedure and the widely used *Cauchy-Born Rule* [28,29] for relating atomic-level deformations to macroscopic deformations. In a

common interpretation of this rule, $\boldsymbol{F} = \nabla \boldsymbol{y}$ represents the macroscopic deformation gradient. $\boldsymbol{F}$ primarily influences macroscopic stress, and the mechanical work is a suitable stress acting on $\boldsymbol{F}$.

One could also, in principle, take sublattices of the martensite too, but accepted transformation mechanisms for most martensitic transformations follow the pattern given here: a sublattice of austenite mapped to the primitive lattice of martensite. The austenite is often a relatively simple partially ordered cubic phase and, even for complex lattices with many different atoms in the unit cell, the primitive periodicity $\boldsymbol{b}_1$, $\boldsymbol{b}_2$, $\boldsymbol{b}_3$, of martensite is measured directly by X-ray methods. Thus, these procedures for finding the transformation stretch matrix are relatively easy to put into practice.

Since $\boldsymbol{U}$ is positive-definite and symmetric, it has three real-positive eigenvalues $0 < \lambda_1 \leqslant \lambda_2 \leqslant \lambda_3$. Among these three eigenvalues, λ_2 is singled out as having particular significance, because of the following theorem [30](Prop.4): *a necessary and sufficient condition that a continuous deformation* $\boldsymbol{y}(\boldsymbol{x})$ *defined on a domain* Ω *has a gradient that takes values* $\nabla \boldsymbol{y} = \boldsymbol{F} = \boldsymbol{RU}$ *(martensite) on a region* $\mathscr{R}$ *and* $\nabla \boldsymbol{y} = \boldsymbol{I}$ *(austenite) on the complementary region* $\Omega \setminus \mathscr{R}$ *for some rotation matrix* $\boldsymbol{R}$ *is that* $\lambda_2 = 1$. (To relate this statement to Prop. 4 of [30], note that $\boldsymbol{y}(\boldsymbol{x})$ of this form is continuous if and only if $\boldsymbol{F} = \boldsymbol{RU} = \boldsymbol{I} + \boldsymbol{b} \otimes \boldsymbol{m}$ for some vectors $\boldsymbol{b}$, $\boldsymbol{m}$. Take $\boldsymbol{F}^T\boldsymbol{F}$ to eliminate the rotation matrix $\boldsymbol{R}$ and note that the eigenvalues of $\boldsymbol{C} = \boldsymbol{F}^T\boldsymbol{F}$ are the squares of the eigenvalues of the positive-definite, symmetric matrix $\boldsymbol{U}$ so, in particular, $\boldsymbol{U}$ has middle eigenvalue equal to 1 if and only if $\boldsymbol{C} = \boldsymbol{F}^T\boldsymbol{F}$ has middle eigenvalue equal to 1.) If $\lambda_2 = 1$, there are two compatible interfaces between austenite and martensite (Fig. 2). Explicit formulas for the two interface normals can be found in [30].

The conditions $\lambda_2 = 1$ and the cofactor conditions are achieved by compositional tuning. That is, the lattice parameters of both phases change when the overall composition is altered. If accurate lattice parameter measurements and symmetries are made at a few compositions, a map of the relation of λ_2 or the cofactor conditions to composition can be developed [31]. As noted below, there is extreme compositional sensitivity of thermal hysteresis and measures of reversibility to these conditions. Thus, for example, when seeking the composition at which $\lambda_2 = 1$, it is best to make a composition in which λ_2 is slightly greater than 1 (accounting for experimental error), and another composition at which λ_2 is slightly less than 1, and then to interpolate the two compositions.

The compatibility condition: $\lambda_2 = 1$

The condition $\lambda_2 = 1$ primarily affects hysteresis [8,31–33], and also improves reversibility under thermal cycling [8]. A theory of the influence of $\lambda_2 = 1$ on hysteresis is given in [33]. That a purely geometric condition $\lambda_2 = 1$ has such a strong and sensitive influence on a property like hysteresis is surprising.

A fundamental observation of thermal hysteresis in martensitic materials is rate independence. The hysteresis loop can grow as the rate increases, due to effects originating from the release of latent heat and the finite rate of heat transfer [34,12], but observations suggest that, as the loop is traversed more and more slowly, there is a limiting loop. This behavior eliminates explanations for hysteresis based on various forms of viscosity, viscoelasticity or thermal activation.

The details can be found in [33] but the essence of the theory is the following. There is evidence from both theory [35] and experiment [36] that rather well-developed nuclei of martensite exist above the nominal austenite finish temperature A_f. These can be associated with defects, triple junctions, precipitates or compositional fluctuations. Similarly, below the martensite finish temperature M_f, the presence of small islands of retained austenite is well accepted in diverse alloys. For definiteness we consider only the first case of cooling from high temperature.

The theory presented in [33] is based on a concept of metastability. It hypothesizes that a twinned platelet grows when the temperature is lowered below that at which the two bulk phases have the same free energy. At small undercooling, a spontaneous thickening of the platelet leads to an increase of energy, due to the subtle interplay of bulk and interfacial energy at the twinned austenite/martensite interfaces bounding the platelet. At sufficiently large sizes the bulk energy dominates and growth leads to a decrease in the energy with size. The situation is similar to nucleation of a metastable phase in a fluid, but the accounting for elastic and interfacial energies is different. At a certain undercooling, a sufficient number of nuclei have a size beyond the barrier to cause large-scale transformation. The theory is quantitative and, when evaluated for the cubic to orthorhombic transformations in TiNiX alloys, gives a graph of hysteresis vs. λ_2 very similar to that shown in Fig. 1. The drop shown in Fig. 1 is indeed very sharp; previously, many nearby alloys were synthesized but most alloy series jumped across the singularity.

Why does λ_2 so strongly affect hysteresis? This concerns the delicate interplay of bulk and interfacial energy at the austenite/martensite interface. Of course, when $\lambda_2 = 1$ the existence of a perfect unstressed interface between phases implies that the elastic energy in the stressed transition layer is eliminated. What is less obvious, but is supported by this model, is that the bulk energy in the transition layer grows extremely rapidly as λ_2 departs from 1, particularly in the case $\lambda_2 > 1$.

Cofactor conditions

The *cofactor conditions* [10,37] consist of three conditions (i) the condition $\lambda_2 = 1$ discussed above, and (ii) a second condition that depends on the twin system chosen, and (iii) an inequality that is usually satisfied for Type I and Type II twin systems. These conditions should be taken together. (We explain below the meaning of satisfying (ii) only.) Since the cofactor conditions include the condition $\lambda_2 = 1$, they imply the existence of perfect, unstressed interfaces between the austenite phase and any given variant of martensite, as described above.

Physically, the cofactor conditions are quite analogous to conditions in structural mechanics for flexibility. A familiar example is a truss with pinned joints: if certain special geometric conditions are satisfied by the lengths of the bars and angles between bars, then an otherwise rigid truss can become flexible, i.e., have zero energy modes of motion possible. Under the cofactor conditions, crystal with both austenite and martensite present exhibits numerous zero energy (or near zero energy) modes of motion that are not be possible with generic lattice parameters.

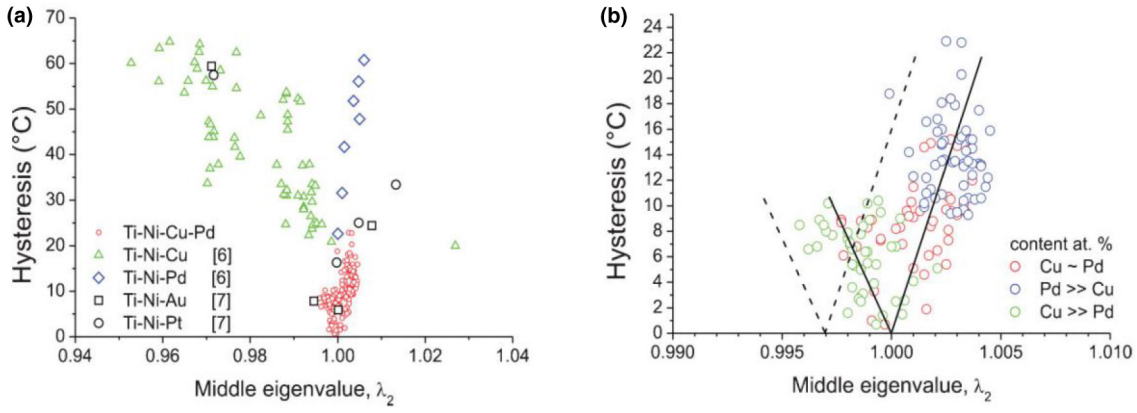


FIGURE 1

(a) Hysteresis vs. λ_2 measured on alloys in the Ti–Ni–X–Y system, (b) Close-up of graph in (a) near $\lambda_2 = 1$, taken from [8].

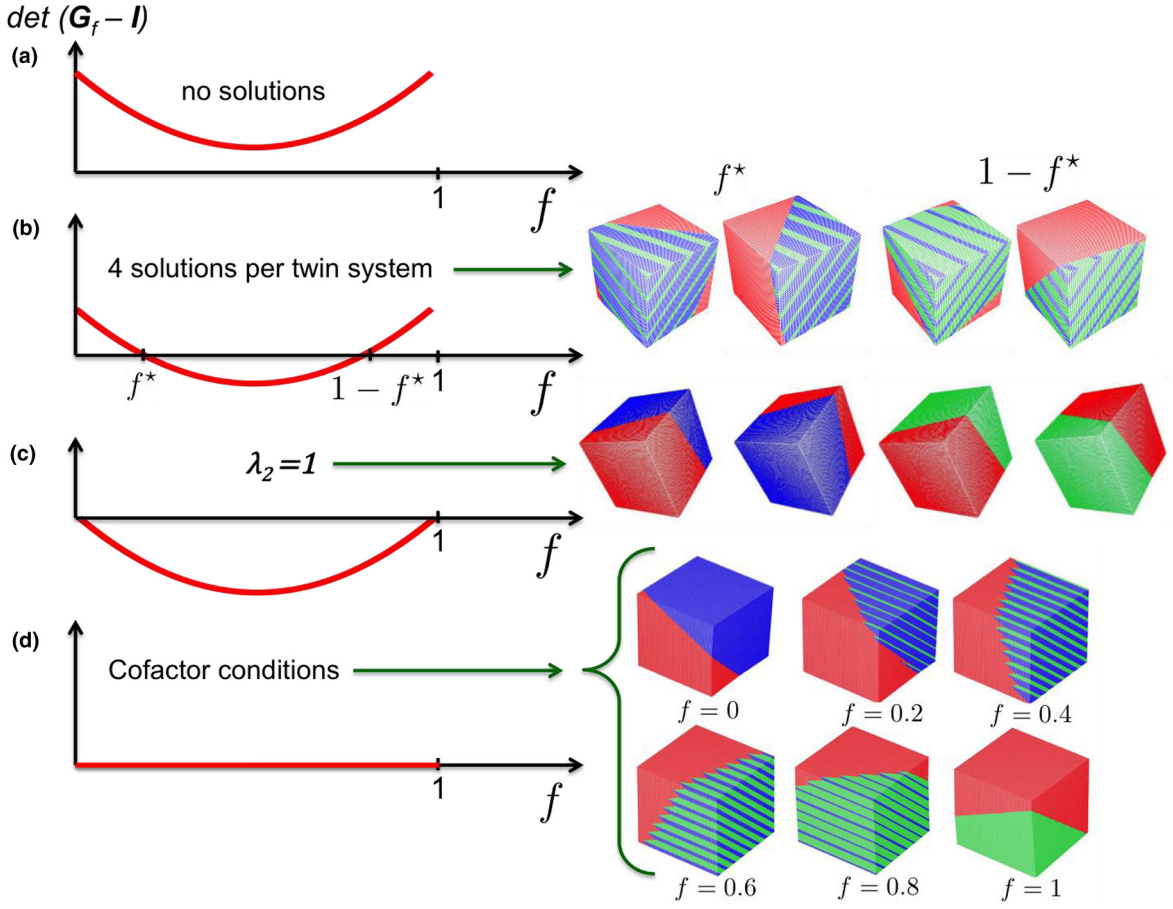


FIGURE 2

Meaning of the cofactor conditions in the context of the crystallographic theory of martensite: (a) the case of no solutions of the crystallographic theory for the given twin system, (b) the generic case of 4 solutions per twin system satisfied by many reversible martensites, (c) the case $\lambda_2 = 1$, (d) the case in which the cofactor conditions are satisfied (The accompanying illustration is a case of Type I twins). Color code: blue and green are variants of martensite, and red is austenite.

The derivation of the cofactor conditions begins with the crystallographic theory of martensite [38,39]. That theory governs the austenite/twinned martensite interface ("habit plane"), that is ubiquitous in martensitic phase transformations. It delivers necessary and sufficient conditions under which the bulk energy in the elastic transition layer between phases can be made vanishingly small by refining the twins. Given the twin system, that theory has as unknowns the volume fraction f of twins, a rigid body rotation $\boldsymbol{R}$ of the austenite, a shear vector $\boldsymbol{b}$ and a unit normal to the habit plane $\boldsymbol{m}$ [30]. $\boldsymbol{R}$, $\boldsymbol{b}$, and $\boldsymbol{m}$ can be eliminated

from the theory using the polar decomposition theorem and the theorem mentioned above concerning λ_2 (see [10,30]). Then the theory reduces to the single scalar equation $\det(\boldsymbol{G}_f - \boldsymbol{I}) = 0$ for the volume fraction f. (See Theorem 2 of [10] for the definition of $\boldsymbol{G}_f$.) It turns out [30] that in all cases $\det(\boldsymbol{G}_f - \boldsymbol{I}) = 0$ is a quadratic equation for $0 \leqslant f \leqslant 1$ that is symmetric about $f = 1/2$. The situation is illustrated in Fig. 2.

As seen in Fig. 2, there are various possibilities for this quadratic function (in red). As in Fig. 2a it can have no real roots. In most nonreversible martensites, such as many (but not all) martensitic steels, Fig. 2a holds for all twin systems. Or, as seen in Fig. 2b, it can have exactly two roots f^* and $1 - f^*$: this is the classic case satisfied by many classical NiTi-based or Cu-based shape memory alloys [29]. Also, a mild inequality (see below) must be verified for these roots to give solutions. If it holds, there are in fact two interfaces corresponding to f^* and two to $1 - f^*$ – briefly, four solutions per twin system – and these are illustrated in Fig. 2 (top, right). Of course, these two roots f^* and $1 - f^*$can occur at 0 and 1, as shown in Fig. 2c, and illustrated to the right. This is exactly the case if $\lambda_2 = 1$. Again, if a certain inequality holds, this is precisely the situation discussed in Section 2.1 above, and the four possible interfaces are seen to the right of Fig. 2c. Finally, the *cofactor conditions* are the conditions that *the quadratic function is identically zero*, Fig. 2d. Assuming again a certain inequality that is usually satisfied, this means that there are solutions of the crystallographic theory, i.e., low-energy interfaces, for any volume fraction $0 \leqslant f \leqslant 1$ of the twins. The situation is illustrated to the right of Fig. 2d.

Due to symmetry, satisfaction of the cofactor conditions for one twin system implies its satisfaction for other crystallographically equivalent twin systems. Unexpectedly, there are a host of other implications of the cofactor conditions [10]. These include solutions of the crystallographic theory for Types I and II twin systems with no elastic transition layer (i.e., perfect fitting), nucleation mechanisms with zero elastic energy, and complex "riverine" zero energy microstructures [9].

To understand the restrictions imposed by the cofactor conditions in a simple case, we refer to Fig. 3. This plot assumes $\lambda_2 = 1$. The locus of points in the λ_1, λ_3 plane at which the cofactor conditions are satisfied are the red (Type I twins) and blue (Type II twins) curves. (In cubic-to-orthorhombic systems, with only three relevant lattice parameters, a/a_0, b/a_0, c/a_0, this kind of 2D plot is possible.) Several known alloys that satisfy accurately $\lambda_2 = 1$, but obviously not the full cofactor conditions, are represented by colored dots. They fall closely on a straight line in this plot, a fact which is not understood. The intersection of this line with the blue and red curves is very close to points representing the TiNiCuCo alloys discussed below! In fact, the numbering of points 2-3-4-5 corresponds to the points 2-3-4-5 on the stress–strain curve shown in Fig. 5.

Fig. 3 suggests that it would be difficult to find new alloys with large transformation strains satisfying the cofactor conditions. For example, it is hopeless to think that, by normal alloy development procedures, one could move the point corresponding to $Ti_{50}Ni_{40.75}Pd_{9.25}$ to the blue curve while keeping $\lambda_2 = 1$. However, in cubic-to-monoclinic systems there is tremendous

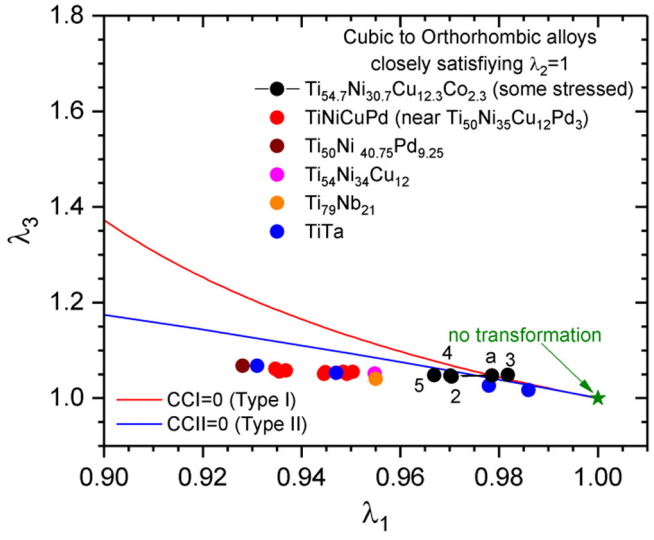


FIGURE 3

The cofactor conditions in cubic-to-orthorhombic systems (see text) for the material systems $Ti_{54.7}Ni_{30.7}Cu_{12.3}Co_{2.3}$ ($\lambda_2 = 1.00083$) [40], TiNiCuPd ($\lambda_2 = 1.00088$) [8], $Ti_{50}Ni_{40.75}Pd_{9.25}$ ($\lambda_2 = 1.0002$) [10], $Ti_{54}Ni_{34}Cu_{12}$ ($\lambda_2 = 0.9905$) [40], $Ti_{79}Nb_{21}$ ($\lambda_2 = 1$) and $Ti_{79}Ta_{21}$ ($\lambda_2 = 1$) [41]. The point $\lambda_1 = \lambda_3 = 1$ corresponds to no transformation, i.e., a continuous change of lattice parameters.

sensitivity of the cofactor conditions to the monoclinic angle, and this suggests various possible starting points.

At present there is no quantitative theory that relates the satisfaction of the cofactor conditions to reversibility, but there are strong correlations summarized in this review. Evidently, these correlations are due to the many strains and many interfaces possible in low (or zero)-energy microstructures involving both austenite and martensite, under the cofactor conditions.

There are various equivalent formulations of the cofactor conditions [10]. In this review we report them in the following way. With reference to Fig. 2d, we recall that the quadratic function $q(f) = \det(\boldsymbol{G}_f - \boldsymbol{I})$, symmetric about 1/2, vanishes identically if and only if $q(0) = 0$ and $q'(0) = 0$. In terms of the twin system $\boldsymbol{a}$, $\boldsymbol{n}$ (Types I or II) and transformation stretch matrix $\boldsymbol{U}$, the cofactor conditions are (see [10] for details):

$$q(0) = 0 \leftrightarrow \lambda_2 = 1,$$

$$q'(0) = 0 \leftrightarrow \boldsymbol{a} \cdot \boldsymbol{U}cof(\boldsymbol{U}^2 - \boldsymbol{I})\ \boldsymbol{n} = 0,\ CCI\ (Type\ I\ twins)\ or\ CCII\ (Type\ II\ twins),$$

$$tr\,\boldsymbol{U}^2 - det\,\boldsymbol{U}^2 - \frac{\boldsymbol{a}^2\boldsymbol{n}^2}{4} - 2 \geqslant 0.$$

The latter is the inequality referred to above. It is satisfied for all Types I and II twin systems discussed in this paper: thus it is not mentioned further below. Consistent with this notation $CCI = \boldsymbol{a} \cdot \boldsymbol{U}cof(\boldsymbol{U}^2 - \boldsymbol{I})\boldsymbol{n}$ evaluated for Type I twins and $CCII = \boldsymbol{a} \cdot \boldsymbol{U}cof(\boldsymbol{U}^2 - \boldsymbol{I})\boldsymbol{n}$ evaluated for Type II twins.

Simpler equivalent expressions for the cofactor conditions than the ones listed above are given in [10] in the case of twin systems that arise from twofold axes of the austenite ("Mallard's Law"). These expressions also have an appealing interpretation in terms of unstretched directions based on $\boldsymbol{U}$ or its inverse. In

all cases in this paper, we have presented the values of CCI and CCII for the twin systems that give the smallest values of CCI and CCII, and we have used up-to-date accepted values of the lattice parameters.

Stress-induced vs. temperature-induced phase transformation

Intuitively, conditions of supercompatibility should be particularly beneficial under stress, because additional low-energy microstructures are available to achieve the boundary conditions corresponding to a stressed sample.

However, the conditions $\lambda_2 = 1$ and the cofactor conditions are calculated from lattice parameters measured under zero stress. Partly this can be understood by the fact that, particularly in alloys undergoing big first-order phase transformations, elastic moduli do not soften excessively, and elastic strains due to the ambient stress are much smaller than transformation strains. The latter can be appreciated from data on TiNiCuCo presented Fig. 5, where the X-ray peaks characteristic of the orthorhombic phase broaden and shift only slightly under stresses of 400 MPa, and the data are most closely indexed as orthorhombic. As seen by the black line 2-3-4-5 in Fig. 3, corresponding to the points 2-3-4-5 on the stress–strain curve of Fig. 5, stress can make conditions of compatibility more closely satisfied, as long as this method of indexing is adopted.

To begin to understand why the cofactor conditions are important to reversibility in stress-induced transformation, consider a random polycrystal under a uniaxial stress in the direction of a unit vector $\boldsymbol{e}$, and ignore the effect of stress on lattice parameters. One can fix $\boldsymbol{e}$ and consider all possible grain orientations, or, equivalently, one can fix a grain orientation and consider uniaxial stress in all possible directions $\boldsymbol{e}$. We do the latter. Then we consider a particular alloy, in this case the TiNiCuCo and we use its lattice parameters, indexed as cubic and orthorhombic and measured under stress (see Fig. 3, with data taken from Fig. 5) on the plateau, that satisfy the cofactor conditions. Thus, we obtain many possible compatible microstructures from many twin systems, as for example shown to the right in Fig. 2d. We considered all twin systems, and the exact rigid rotations needed for compatibility. But also, we can "turn off" the cofactor conditions (Fig. 4a) by only considering the classical solutions of the crystallographic theory in this alloy. That is, in calculating the maximum resolved shear stress on the habit plane, or equivalently maximum work, we do not allow the new microstructures that are possible under the cofactor conditions to compete for the maximum. Use of the same lattice parameters in both cases allows for a meaningful comparison. In both cases we use the Schmid Law to maximize the strain ε in the direction $\boldsymbol{e}$. Using this maximizing strain ε and the associated selected variant, we then plot $1 + 10\varepsilon$ in a polar plot (multiplying ε by a factor 10 for clarity). That is, for each tensile direction $\boldsymbol{e}$, $|\boldsymbol{e}| = 1$, we plot a point at $(1 + 10\varepsilon)\boldsymbol{e}$ and we color that point according to the selected variant. If a microstructure selected by the Schmid Law is only possible under the cofactor conditions, then we color the point black in Fig. 4b.

The result is shown in Fig. 4. The lattice parameters were chosen from point "a" in Fig. 3. Fig. 4a shows the situation with cofactor conditions turned off. The color coding is by the variant selected by the Schmid Law. In Fig. 4b the black region is produced by a compatible laminate with volume fraction strictly between 0 and 1, that is possible under the cofactor conditions. It is known broadly in polycrystal martensites [42] that the measured plateau strain is determined by the worst oriented grain. While the black region is quite small, it is located *exactly at orientations* $\boldsymbol{e}$ *corresponding to the worst oriented grains*. These graphs do not highlight the fact that at many other orientations $\boldsymbol{e}$ the cofactor conditions provide a great many possible microstructures to achieve strains in these directions, even if they are not maximizing.

Fatigue determining factors in stress-induced phase transformations

Influence of compatibility

TiNiCuCo alloys

The pronounced dependence of crystallographic compatibility on the thermal hysteresis is discussed in Section 2.1. However, its influence on functional and structural fatigue especially on the stress-induced austenite–martensite transformation is only just emerging. For this investigation different free-standing sputtered films of highly compatible TiNiCuCo compositions ($Ti_{54.7}Ni_{30.7}Cu_{12.3}Co_{2.3}$ and $Ti_{50.8}Ni_{34.2}Cu_{12.5}Co_{2.5}$) with a trans-

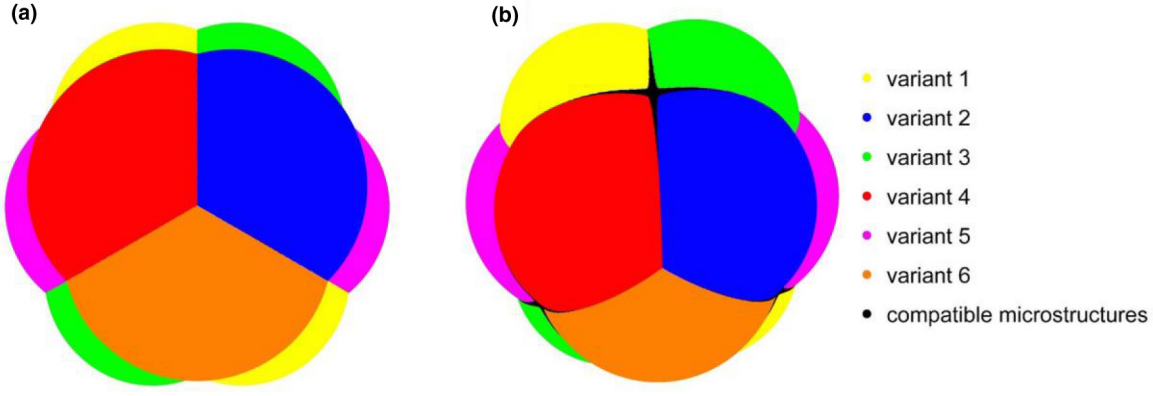
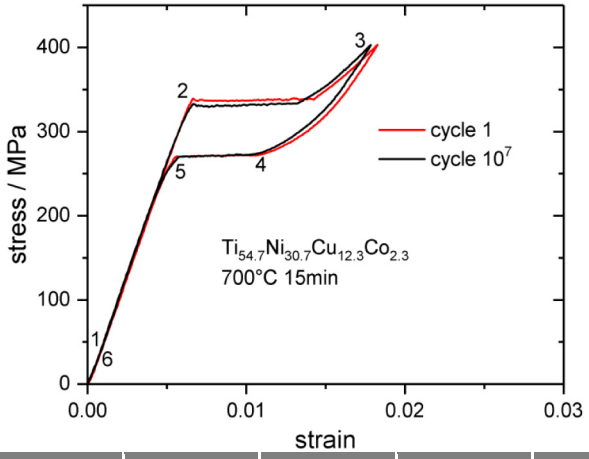


FIGURE 4

Polar plot of maximum strains selected by the Schmid Law, (a) with classical solutions of the crystallographic theory of martensite only, and (b) with solutions possible under the cofactor conditions.



	a_o (Å) of B2 at 22 °C	a (Å) of B19 at 22 °C	b (Å) of B19 at 22 °C	c (Å) of B19 at 22 °C	λ_2	CCI (10^{-5})	CCII (10^{-5})
1	3.03727						
2	3.05187	4.30821	2.9476	4.48891	1.00299	8.32571	0.769014
3	3.04113	4.31244	2.98205	4.5029	1.00398	5.02326	3.21548
4	3.03649	4.2989	2.94661	4.49807	1.00083	4.26161	3.43131
5	3.04034	4.30043	2.93665	4.50005	1.00118	6.60788	6.31816
6	3.03722						

FIGURE 5

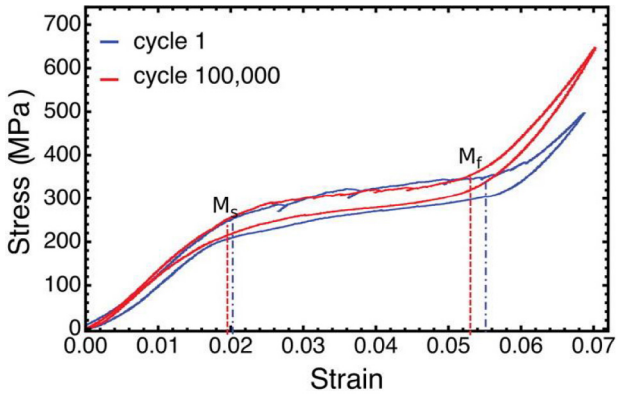
High cycle fatigue data of $Ti_{54.7}Ni_{30.7}Cu_{12.3}Co_{2.3}$ films showing no functional fatigue within 10 million full tensile cycles. The numbering corresponds to the position where the crystallographic data were recorded in in-situ tensile experiments to determine the crystallographic compatibility under tensile load. The table documents the almost perfect satisfaction of the cofactor conditions.

formation matrix middle eigenvalue of $\lambda_2 = 1 \pm 0.01$ were investigated in terms of their temperature- and stress-dependent structural properties and their fatigue behavior under tensile cyclic loading [5,40]. The thermal transformation of these alloys was found to be very stable in any case; the superelastic degradation behavior was found to depend strongly on different parameters: grain size (see chapter 3.2) and the kind and nature of precipitates (see chapter 3.3). Among these samples the composition $Ti_{54.7}Ni_{30.7}Cu_{12.3}Co_{2.3}$ showed no measurable fatigue within 10 million tensile full cycles. In stress-dependent XRD investigations it was revealed that this sample fulfills the crystallographic compatibility conditions almost perfectly (Fig. 5) and slightly better than the other samples of this investigation, namely the $Ti_{50.8}Ni_{34.2}Cu_{12.5}Co_{2.5}$ alloy (Fig. 7). As the $Ti_{50.8}Ni_{34.2}Cu_{12.5}Co_{2.5}$ sample annealed at 600 °C (Fig. 7b) has a very similar microstructure but showed a slightly higher functional fatigue, it can be concluded that very close to ideal compatibility between austenite and martensite has an influence on the fatigue behavior.

ZnAuCu alloys

The alloy $Zn_{45}Au_{30}Cu_{25}$ was found [9] by systematically tuning composition, beginning from the Heusler alloy Zn_2AuCu, which was identified as a good starting point for satisfaction of the cofactor conditions based on literature values of its lattice parameters. $Zn_{45}Au_{30}Cu_{25}$ undergoes a cubic-to-monoclinic ($L2_1$ to M18R) transformation and exhibits values of the cofactor conditions in the table of Fig. 6. Thus, the cofactor conditions are satisfied closely for both Types I and II twins. This alloy shows as little as 0.2 K hysteresis, and up to 2 K hysteresis under cyclic conditions [9], consistent with its value of λ_2. Under thermal cycling to $2^{14} = 16,384$ cycles it exhibits a steady transformation temperature and only a small loss of latent heat. While these conditions are less demanding than for other alloys discussed in this review, what is particularly interesting is that the nearby alloy $Zn_{45}Au_{27}Cu_{28}$ exhibits significant degradation under the same conditions [9]. This highlights the great sensitivity to changes of composition that is a hallmark of conditions of supercompatibility.

$Zn_{45}Au_{30}Cu_{25}$ could be considered similar to the CuZn family of shape memory alloys that have received a lot of attention, but ultimately suffer from numerous problems relating to aging and functional fatigue. However, a study of cyclic compression of 1–2-μm diameter pillars reveals unexpected resistance to functional fatigue [43]. As shown in Fig. 6, this alloy shows near perfect reversibility under conditions of 500–650 MPa and 6–7% strain each cycle for 100,000 cycles. Notice the small stress hysteresis and slight hardening of the martensite that occurs during the test. A comparison of the behavior of 1- and 2-μm diameter pillars indicates only a minor size effect. Movies of the pillar during compression [43] reveal a near homogeneous deformation, despite the large strains. Evidently, due to supercompatibility, the interfaces cost so little energy that the pillar can be full of microstructure, even at micron scale. That is, small traces of many parallel interfaces whose orientation is consistent with the perfect austenite/martensite interface can be seen on the



	a_o (Å) of B2	a (Å) of B19'	b (Å) of B19'	c (Å) of B19'	θ (°) of B19'	λ_2	CCI (10^{-5})	CCII(10^{-5})
$Au_{30}Cu_{25}Zn_{45}$	6.1606	4.4580	5.7684	40.6980	86.80	1.00056	3.78354	3.01936

FIGURE 6

Cyclic compression of a 2-μm diameter pillar of $Zn_{45}Au_{30}Cu_{25}$. The values of λ_2, CCI, $CCII$ are given in the table below [43]. Reprinted (adapted) with permission from (X. Ni, J.R. Greer, K. Bhattacharya, R.D. James, X. Chen, Exceptional Resilience of Small-Scale $Au_{30}Cu_{25}Zn_{45}$ under Cyclic Stress-Induced Phase Transformation, Nano Letters. 16 (2016) 7621–7625.). Copyright (2016) American Chemical Society.

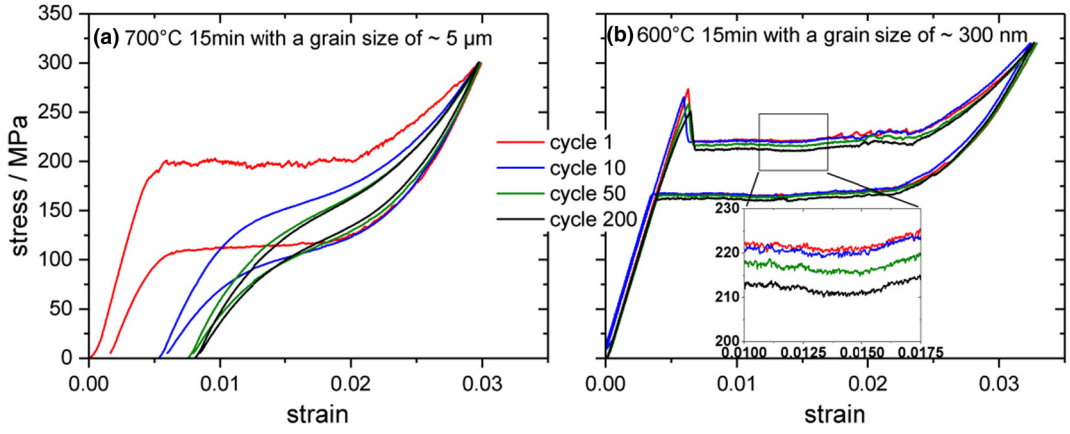


FIGURE 7

Amorphous $Ti_{50.8}Ni_{34.2}Cu_{12.5}Co_{2.5}$ thin film crystallized at (a) 700 °C for 15 min leading to a grain size of 5 μm, and (b) 600 °C for 15 min leading to a grain size of 300 nm.

specimen. However, the deformation itself viewed by SEM during a test suggests a smooth, homogeneous deformation.

Influence of grain size

In common metals the grain size has an important influence on structural parameters like tensile strength, ductility and fatigue properties associated to failure by crack nucleation and growth [13]. Shape memory alloys (SMAs) exhibit a first-order phase transformation, and are especially influenced by grain size, as discussed in the following section. Special attention will be drawn to transformation temperatures, transformation strain, transformation mechanism and hysteresis. Besides structural fatigue, other transformation properties exhibit functional degradation that correlates with grain size.

In general, it is an experimental challenge to investigate only the effect of grain size, since sample preparation variations that change the grain size also change other microstructural features like precipitates, parent phase composition or defect density. These microstructural properties are also strongly influenced by the different preparation processes like casting, casting with an additional thermo mechanical treatment or thin film deposition. This makes a comparison difficult but nevertheless some general conclusions can be drawn.

Influence on transformation properties

It is commonly accepted that the grain boundaries favor the nucleation of martensite [44]. By itself, this results in an increase of M_s temperature with decreasing grain size due to higher grain boundary volume fraction. On the other hand, grain size reduction results in an increase of the energy barrier for the martensite transformation caused by stored elastic energy and size effects that limit the self-accommodation of martensite variants in a sin-

gle grain [45]. As a consequence, the parent austenite phase is stabilized, which leads to the opposite effect, i.e., a decrease of the M_s temperature [46]. Simulations and experiments for NiTi have shown that ~50-nm grain size is the limit for the martensite transformation [46–50,55]. By doping, this limit can be shifted to lower grain size as has been shown for TiNiCu [51].

The transformation strain of the superelastic effect also depends on the grain size. An increasing volume fraction of grain boundaries accompanied by grain size reduction leads to a smaller transformation strain, since grain boundaries do not participate in the transformation [52]. Additionally, as explained before, in small grains the stress-induced transformation can be suppressed by limited accommodation possibilities of martensite within the grain.

The limited martensite accommodation can also alter the transformation mechanism in SMAs [53]. The typical B2-B19′ or B2-R-B19′ transformation sequence in NiTi can be suppressed in favor of a single B2-R transformation which allows an easier self-accommodation in a small grain due to smaller transformation strain [54].

Finally, the transformation stress hysteresis is also dependent on the grain size. By approaching a critical grain size of 50 nm in binary NiTi, the stress-induced hysteresis drops tremendously from 200 MPa at grain sizes larger than 50 nm to 20 MPa at a grain size of ~20 nm. This is explained by the change of the transformation mechanism at extremely small grain sizes and is accompanied by a significant loss in effect size [55].

Influence on structural fatigue

In SMAs two types of fatigue can be separated. First *structural fatigue* is similar to fatigue of conventional metals and is characterized by nucleation and growth of cracks that result in ultimate failure [13,50,56]. Crack nucleation mechanisms are based on defect generation like dislocation generation [57] or grain boundary gliding [58] which are both strongly correlated to the grain size [59]. Like conventional metals a small grain size in an SMA can impede defect generation which results in an increase of the ultimate tensile strength [58]. For high strength alloys the crack nucleation probability is reduced, but on the other hand the crack growth rate is increased which leads to a fast failure of the device once a crack occurs [56]. If the material is in the austenite phase, stress at the crack tip could in principal be reduced by a stress induced austenite–martensite transformation. But some investigations have shown that this transformation cannot decrease the crack growth velocity since the growth in martensite is enhanced [60].

The second type of fatigue in SMAs is the functional degradation of the transformation which is characterized in reduced critical stresses for the transformation and a broader stress distribution [52]. This effect is caused by defect generation that accumulates preferentially at the grain boundaries [61,62]. Due to the surrounding stress fields, these defects enhance the stress-induced martensitic transformation or even stabilize the martensite phase at the grain boundary above the austenite finish temperature [61]. As a consequence, the transformation strain of the superelastic transformation is reduced [52]. To overcome this problem grain size reduction has been reported to be an effective way to inhibit defect generation without obstructing the martensitic transformation [13,47,63]. Nevertheless, a small grain size is not always a sufficient criterion to obtain high functional stability in every case [64].

Comparison of TiNiCuCo alloys with different grain sizes

To investigate the influence of grain size, and to separate its influence on fatigue from precipitate and compatibility effects, we present recent work on TiNiCuCo films that have been deposited in the amorphous state by sputter deposition. This fabrication process yields also a low inclusion content [65]. The TiNiCu thin film system has been explored already by several groups [66] and a precipitate free composition has been chosen. Cobalt is added to shift the transformation temperature below room temperature. Sequential annealing of the amorphous phase enables adjustment of the grain size by varying the annealing temperature. Annealing at 600 °C results in a grain size of 300 nm and a 700 °C anneal gives a grain size of larger than 5 μm. Both heat treatments yield precipitate-free samples with similar crystallographic compatibility of the stress-induced austenite–martensite transformation, as verified by in situ synchrotron XRD [5]. Significant changes can be observed in the functional stability and yield strength, as shown in Fig. 7. While the small-grained sample exhibits high superelastic stability, the coarse-grained sample shows severe degradation even after the first transformation cycle combined with a low yield strength. In agreement with previous investigations, ex situ TEM analysis revealed the formation of dislocation networks as the reason for the changes of the transformation [5]. This different microstructure has a severe influence on the lifetime of the samples. While the 5-μm GS sample fractured after a few thousand cycles, the 300-nm GS sample survived approximately 10^5 cycles.

Influence of precipitates

Precipitates in TiNi-based shape memory alloys

Besides crystallographic compatibility and a small grain size, an important factor for the design of SMAs with increased functional stability is the presence of precipitates. They can significantly influence shape memory properties like yield strength, transformation temperature, transformation strain, transformation mechanism and hysteresis. It is well-known from the development of slightly Ni rich TiNi that special focus should be placed on the misfit of the lattice parameters between the precipitate and martensite/austenite phase, which in the ideal case should be as small as possible to preserve coherency. In that case, coherent growth can occur between the precipitate and the matrix. The presence of the coherent precipitate not only increases the resistance to dislocation formation and slip, but also can promote the nucleation of a reversible phase transformation, thus reducing the need of growth [31]. In addition, coherent precipitates do not hinder the formation of martensite variants, whereas incoherent precipitates tend to constrain the formation of martensite [66–68].

A particularly desirable situation, in an alloy satisfying the compatibility condition $\lambda_2 = 1$, is to have flat coherent precipitates whose interface plane is close to one of the two perfect interfaces between austenite and martensite [2,28]. In that case transformation can occur with high volume fraction of martensite, without the need of special stressed microstructures that

accommodate the precipitates, and still the strengthening effects of the precipitates mentioned above remain.

It is well known that classical binary NiTi is far away from fulfilling the compatibility conditions discussed above [69], and, presumably as a result, shows rather poor lifetime. A method to increase the resistance to functional fatigue in slightly Ni-rich compositions is the use of coherent Ti_3Ni_4 precipitates, which reduce the dislocation activity compared to incoherent precipitates and precipitate free samples [70]. It is often observed that with the formation of these precipitates the transformation path changes from a single-step transformation to a multistep transformation [71]. Transformation temperature changes as a consequence of the interplay of stress fields, and the change of matrix composition due to precipitation [72]. A similar trend is observed in NiTiHf/Zr SMA. In case of Ni-rich TiNiHf/Zr alloys, which undergo the cubic to monoclinic transformation, it was found that coherent H-phase precipitates (F*ddd* symmetry) [73,74] strengthen the matrix and do not interact with martensite twin formation [66]. This results in a relatively low thermal hysteresis of 28 °C and good functional stability for $Ni_{50.3}Ti_{29.7}Hf_{20}$ SMA despite having $\lambda_2 = 0.9365$ which is far away from the ideal case $\lambda_2 = 1$ [67,75]. Similar results are reported for the NiTiZr system [67,76].

Of special interest for this review is also the TiNiCu system, since it has the possibility to obtain a small grain size in the sub μm range in combination with coherent precipitates and a good crystallographic compatibility over a wide compositional variation. Ti_2Cu precipitates show an epitaxial coherent relation to the martensite–austenite interface promoting the nucleation and growth [40,77]. As a result functional fatigue is negligible at 10 million cycles [40] or reduced in other cases [66]. Besides Ti_2Cu, tetragonal precipitates like TiCu showed a coherency relation [78]. The stress field accounted to the coherent interface of precipitates or Guinier Preston zones can also be used to stabilize the austenite phase and thus reduce the transformation temperature [66,79].

Ti-rich vs. near-equiatomic TiNiCu

In this section we compare two TiNiCu alloys with approximately 12 at% Cu: an unusual Ti-rich composition with two types of Ti-rich precipitates, Ti_2Cu and Ti_2Ni, and a near-equiatomic alloy. The latter is used in a common approach to maximize the amount of phase-transforming material with a minimum volume fraction of precipitates. Thus, this comparison allows the investigation of the influence of precipitates on the functional and the structural fatigue in these materials. It was found that the presence of coherent Ti_2Cu precipitates improves the functional stability and increases the lifetime of these SMAs tremendously: only the Ti-rich sample showed a run-out at 10 million full superelastic cycles without any degradation in the effect size (Fig. 8), while the near-equiatomic sample, which exhibited no Ti_2Cu precipitates, showed a lifetime until fracture between several 1000 and 100,000 cycles depending on the grain size (Fig. 7). The coherent growth of the phase pairs B2/Ti_2Cu and B19/Ti_2Cu is considered to be primarily responsible for the almost complete reversibility of the stress–strain curves. In contrast, the Ti_2Ni precipitates, which are found in both samples, have a much larger misfit especially to the B2 phase, which

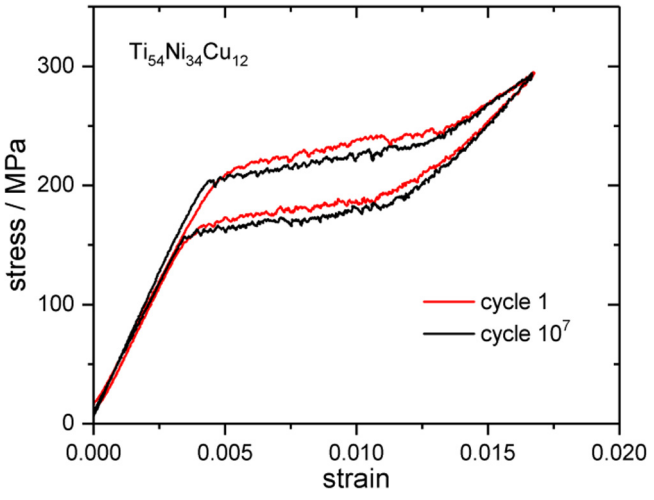


FIGURE 8

Ultralow fatigue of the Ti-rich TiNiCu demonstrated by next-to-identical 1st, 200th, and 10^7th superelastic transformation cycles. The testing temperature was 70 °C with an austenite-finish temperature of 65 °C.

prevents coherent growth to both phases for these precipitates. In the case of Ti_2Cu precipitates at each temperature and stress, the transforming phases are guided in a specific way that is directed by the internal stresses at the interfaces between the transforming matrix and the Ti_2Cu precipitates. Thus, these precipitates are considered to act like sentinels, assuring that the B2 ↔ B19 transformation follows the same low energy path during each cycle [40].

TiNiCu with different sizes of the precipitates

The TiNiCu-based alloys in the previous chapters covered the TiNiCu microstructure nearly without precipitates ($Ti_{50.8}Ni_{34.2}Cu_{12.5}Co_{2.5}$), or with a combination of Ti_2Cu and Ti_2Ni precipitates ($Ti_{54}Ni_{34}Cu_{12}$). We pointed out that precipitates of the type Ti_2Cu promote the reversible transformation whereas Ti_2Ni precipitates do not influence or improve the fatigue behavior [40,66,79]. In the following example we discuss solely the influence of the Ti_2Cu precipitates in the TiNiCu system.

For this investigation $Ti_{53.7}Ni_{24.6}Cu_{21.7}$ thin films were chosen, since this composition should show only the presence of Ti_2Cu precipitates [66]. To obtain different microstructures, amorphous 20-μm-thick films were heat treated by rapid thermal annealing at 550 °C and 700 °C, respectively. To link the compatibility and the influence of precipitates on the functional fatigue behavior, the samples underwent 200 full superelastic and 30 thermal cycles. A strong correlation of annealing temperature with the microstructure and functional fatigue was revealed. Both samples possess a comparable crystallographic compatibility with $\lambda_2 = 0.99$ (In this case $\lambda_2 = \frac{d(1\,1\,0)_{b2}}{d(0\,1\,0)_{b19}}$ [5]). TEM bright field images of the sample annealed at 550 °C reveal a fine microstructure with an average grain size of 500 nm surrounded by small and homogeneously distributed, approximately 50 nm large Ti_2Cu precipitates located at the grain boundaries (Fig. 9b).

Increasing the annealing temperature to 700 °C does not change the grain size of the matrix phase. However, the Ti_2Cu precipitates are now irregularly distributed and enlarged to

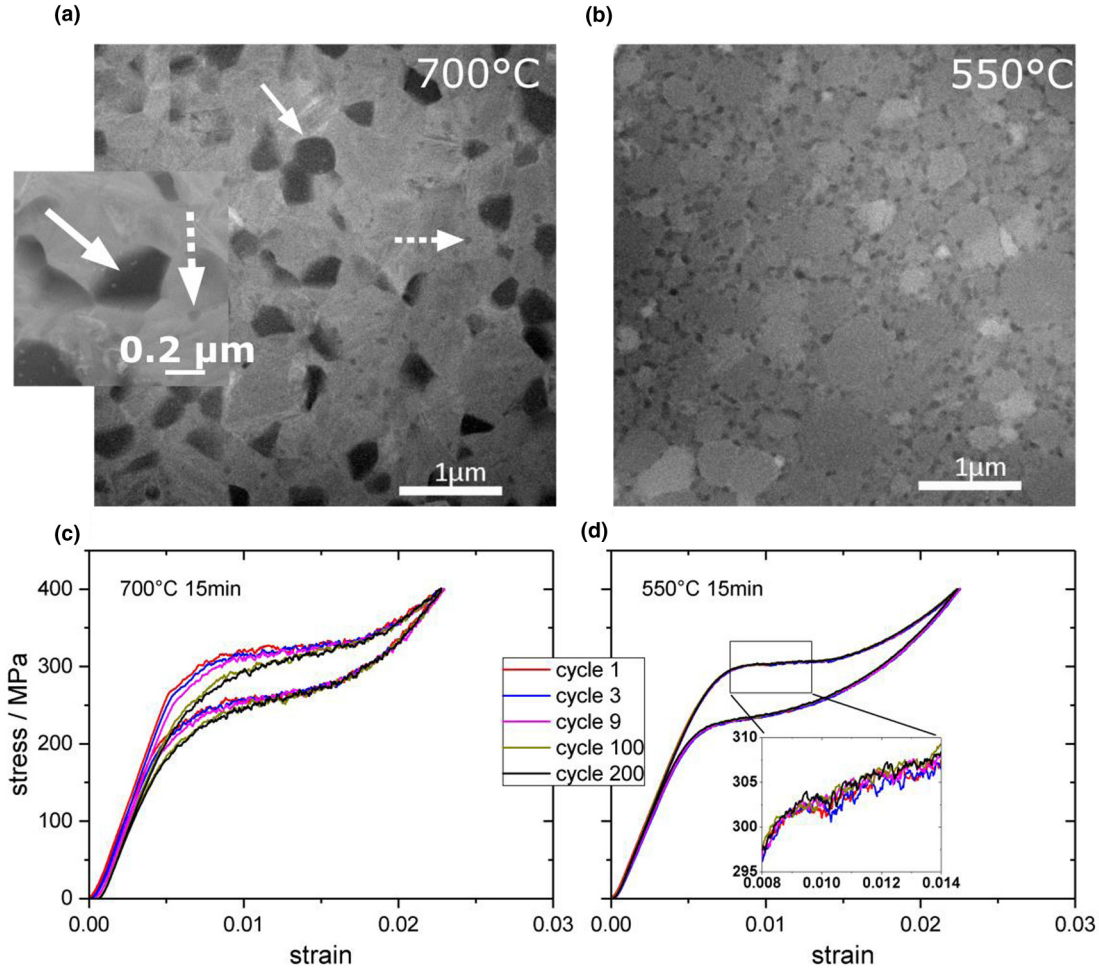


FIGURE 9

The figure correlates the microstructure of (a) and (b) with the functional fatigue of the stress-induced martensitic transformation (c, d) of $Ti_{53.7}Ni_{24.6}Cu_{21.7}$ films. (a) The TEM image shows the microstructure of the sample annealed at 700 °C. The inset shows a magnified image of the B19 phase and the precipitates. The B19 grains contain round Ti_2Ni precipitates (dashed arrow) and are surrounded by large Ti_2Cu precipitates (solid arrow). (b) The TEM image shows the microstructure of the sample annealed at 550 °C. The B2 phase (area of bright contrast) is surrounded by small Ti_2Cu precipitates (dark contrast). Figure (c) and (d) show the functional fatigue for 200 superelastic cycles for the sample annealed at 700 °C and 550 °C, respectively.

approximately 300 nm. In addition, spherical Ti_2Ni precipitates were found to grow within the grains (Fig. 9a). It was found, that the Ti_2Cu precipitates for the sample annealed at 550 °C showed a coherency relation to the matrix. A detailed crystallographic description of the coherency relation and its influence on the martensitic transformation can be found in [40,77].

In contrast no coherence relation was found for the Ti_2Cu/Ti_2Ni with the B2/B19 interface for the sample annealed at 700 °C, caused by the large size of the Ti_2Cu precipitates and the large crystallographic misfit of the Ti_2Ni and B2 phase. The loss in coherency has a tremendous effect on the functional fatigue for the stress-induced transformation. The sample annealed at 700 °C evolves small irreversible strains within the first 200 cycles, whereas the sample annealed at 550 °C shows no functional fatigue (Fig. 9c, d). In conclusion TiNiCu films annealed at 550 °C show no functional fatigue since they combine good crystallographic compatibility, small grain size and coherent Ti_2Cu precipitates. As a result, these features suppress the formation of dislocations and stabilized martensite and ease the reversible formation of martensite and austenite. In contrast, samples with incoherent precipitates cannot suppress the formation of dislocations and therefore tend to stabilize martensite.

Conclusions and outlook

In this review, we have laid out what we believe are the most important factors governing the reversibility of phase transformations in shape memory alloys. We have identified crystallographic compatibility, grain size and precipitation as important factors. We show that, if compositions are tuned so that all three of these factors are fulfilled, alloys with unprecedented resistance to functional fatigue and low hysteresis are possible. However, our results show that one or more of these factors can be compromised a little and still remarkable functional fatigue is achieved. For example, alloys satisfying the strongest conditions of compatibility (the cofactor conditions) to high accuracy have exceptional resistance to functional fatigue, but so do alloys that satisfy these compatibility conditions only approximately, but that have small (but not too small) grain size and a favorable array of fine coherent precipitates. One critical factor must be

respected: there is great sensitivity of composition or heat treatment to several of the factors mentioned here.

These results are developed and applied here in the context of shape memory alloys. Could they also apply in general to solid–solid phase transformations, for example to the multiferroic materials mentioned in the introduction, or nonmetallic materials such as oxides or molecular crystals? Since many of the conditions given here are primarily geometric or elastically dominated (conditions of compatibility, grain size, the coherency of precipitates), it is expected that they should be applicable to broad classes of phase transformations. One main factor requires further study: in ferroelectric or ferromagnetic phase transformations there are ferroelectric/ferromagnetic domains whose "compatibility" in terms of, say, pole avoidance, could be important. However, we have tuned the composition of ferromagnetic NiMnSn and NiCoMnSn Heusler type alloys to satisfy $\lambda_2 \approx 1$ with a sharp drop in hysteresis [80]. Possibly, this is a reflection of the fact that, (a) by symmetry, poles at twin boundaries are often automatically absent, and (b) elastic energies are typically dominant over polar or magnetic energies in hard materials.

There are some particular questions raised by our work. We do not understand the relative roles of $\lambda_2 = 1$ vs. the full cofactor conditions in determining hysteresis and reversibility. A comparative study of similar alloys with the same processing, one satisfying to high accuracy $\lambda_2 = 1$ but far from satisfying $CCI = 0$ or $CCII = 0$, and another satisfying the full cofactor conditions, would be illuminating. Similar comparative studies involving also grain size and precipitates would be helpful.

While the effect of λ_2 on thermal hysteresis is quite clear (Fig. 1), the factors governing stress hysteresis are unclear, as can be seen from a comparison of the various stress–strain curves shown in this review. All of them have rather small hysteresis, but there is still a wide variation. By comparison, the $\beta_1 \rightarrow \beta_1'$ transformation in the classic Cu–Al–Ni shape memory alloy has a very large transformation strain on the order of 10%, a stress hysteresis of about 5 MPa, but poor resistance to fatigue [81]. Would it be possible to have such small stress hysteresis in an alloy with high resistance to functional fatigue?

In very careful tuning of lattice parameters to make $\lambda_2 = 1$ by compositional increments of ¼ at% Pd, substituted for Ni, the thermal hysteresis of $Ti_{50}Ni_{40.75}Pd_{9.25}$ has been reduced to 2–3 °C [10]. A similar result has been obtained in the NiMnSn system. It would be interesting to know what is the ultimate lowest thermal hysteresis possible by compositional changes, and what physical phenomenon governs the remaining hysteresis.

In the near term it would be useful to have good starting points for compositional tuning. Two systems have been identified in [10] as promising. One is the VO_2 system, which has an unusual crystallographic change such that satisfaction of $\lambda_2 = 1$ automatically implies satisfaction of the second of the cofactor conditions $\boldsymbol{a} \cdot \boldsymbol{U}cof(\boldsymbol{U}^2 - \boldsymbol{I})\boldsymbol{n} = 0$ for compound twins. Additionally, these conditions are satisfied simultaneously for many of the (compound) twin systems. Thus, this is quite an interesting case, also in view of the fascinating optical and electronic properties of VO_2. Evidently, though this awaits confirmation, this tuning has been carried out in CrNb doped VO_2 by Miyazaki et al. [18]. It would be extremely interesting to examine the microstructures of this alloy. Another promising system is the CuAlMn system (see [10]).

Acknowledgment

The work from the University of Minnesota was supported by AFOSR (FA9550-15-1-0207), NSF (DMREF-1629026), ONR (N00014-14-1-0714) and the MURI program (FA9550-12-1-0458, FA9550-16-1-0566). The work at the University of Kiel was supported by the Deutsche Forschungsgemeinschaft (DFG) via the Priority Program 1599 and the Reinhart Koselleck Project QU 146/23-1. The authors acknowledge fruitful discussions with Jake Steiner and Manfred Wuttig (University of Maryland), with Torben Dankwort and Lorenz Kienle (University of Kiel) and technical support by Rodrigo Lima de Miranda (Acquandas GmbH Kiel). HG and RDJ are pleased to acknowledge the support of Medtronic Corp, via a gift to the University of Minnesota.

Appendix A. Supplementary data

Supplementary data associated with this article can be found, in the online version, at https://doi.org/10.1016/j.mattod.2017.10.002.

References

[1] J.D. Busch, A.D. Johnson, Shape-Memory Alloy Micro-Actuator, United States Patent, Patent Number 5061914 (1991).
[2] W.L. Benard, H. Kahn, A.H. Heuer, M.A. Huff, A titanium-nickel shape-memory alloy actuated micropump, in: Proceedings of International Solid State Sensors and Actuators Conference (Transducers '97), vol. 1 (1997, pp. 3–6, https://doi.org/10.1109/SENSOR.1997.613659.
[3] H. Kahn, M.A. Huff, A.H. Heuer, J. Micromech. Microeng. 8 (1998) 213–221, https://doi.org/10.1088/0960-1317/8/3/007.
[4] K. Otsuka, X. Ren, Prog. Mater Sci. 50 (2005) 511–678, https://doi.org/10.1016/j.pmatsci.2004.10.001.
[5] C. Chluba et al., Trans. R. Soc. A 374 (2016), https://doi.org/10.1098/rsta.2015.0311.
[6] S. Miyazaki, Y. Igo, K. Otsuka, Acta Metall. 34 (1986) 2045–2051, https://doi.org/10.1016/0001-6160(86)90263-4.
[7] S. Miyazaki, K. Otsuka, ISIJ Int. 29 (1989) 353–377, https://doi.org/10.2355/isijinternational.29.353.
[8] R. Zarnetta et al., Adv. Funct. Mater. 20 (2010) 1917–1923, https://doi.org/10.1002/adfm.200902336.
[9] Y. Song et al., Nature 502 (2013) 85–88, https://doi.org/10.1038/nature12532.
[10] X. Chen et al., J. Mech. Phys. Solids 61 (2013) 2566–2587, https://doi.org/10.1016/j.jmps.2013.08.004.
[11] D.A. Miller, D.C. Lagoudas, Mater. Sci. Eng., A 308 (2001) 161–175, https://doi.org/10.1016/S0921-5093(00)01982-1.
[12] J.A. Shaw, S. Kyriakides, J. Mech. Phys. Solids 43 (1995) 1243–1281, https://doi.org/10.1016/0022-5096(95)00024-D.
[13] E. Hornbogen, A. Heckmann, Materialwiss. Werkstofftech. 34 (2003) 464–468, https://doi.org/10.1002/mawe.200390094.
[14] C. Maletta et al., Smart Mater. Struct. 21 (2012) 112001, https://doi.org/10.1088/0964-1726/21/11/112001.
[15] T. Duerig, A. Pelton, D. Stöckel, Mater. Sci. Eng., A 273–275 (1999) 149–160, https://doi.org/10.1016/S0921-5093(99)00294-4.
[16] Patent iNSIGHT Pro Gridlogics Technologies Pvt. Ltd., Shape Memory Material Technology Insight Report. (2015). http://www.patentinsightpro.com/techreports/0315/Tech Insight Report – Shape Memory Materials.pdf (accessed August 3, 2017).
[17] W. Eerenstein, N.D. Mathur, J.F. Scott, Nature 442759 (2006) 759–765, https://doi.org/10.1038/nature05023.
[18] K. Miyazaki et al., AIP Adv. 6 (2016), https://doi.org/10.1063/1.4949757.
[19] M. Wuttig, N. Yamada, Nat. Mater. 6 (2007) 824–832, https://doi.org/10.1038/nmat2009.

[20] L. Mañosa, A. Planes, M. Acet, J. Mater. Chem. A 1 (2013) 4925–4936, https://doi.org/10.1039/C3TA01289A.
[21] Z. Hu et al., Appl. Phys. Lett. 104 (2014) 263902, https://doi.org/10.1063/1.4887066.
[22] S. Fähler et al., Adv. Eng. Mater. 14 (2012) 10–19, https://doi.org/10.1002/adem.201100178.
[23] R.L. De Miranda, C. Zamponi, E. Quandt, Adv. Eng. Mater. 15 (2013) 66–69, https://doi.org/10.1002/adem.201200197.
[24] E.C. Bain, N.Y. Dunkirk, Trans. Am. Inst. Min. Metall. Eng. 1 (1924) 871–876.
[25] W.M. Lomer, The β → α transformation in uranium-1. 4 at.% chromium alloy, Institute of Metals, Monograph and Report Series. (1955) 243.
[26] X. Chen et al., J. Mech. Phys. Solids 93 (2016) 34–43, https://doi.org/10.1016/j.jmps.2016.02.009.
[27] K. Koumatos, A. Muehlemann, Proc. R. Soc. London, Ser. A 472 (2016) 1–25, https://doi.org/10.1098/rspa.2015.0865.
[28] J.L. Ericksen, Math. Mech. Solids 13 (2008) 199–220, https://doi.org/10.1177/1081286507086898.
[29] K. Bhattacharya, Microstructure of Martensite: Why It Forms and How It Gives Rise to the Shape-Memory Effect, 2nd ed., Oxford University Press, Oxford, 2003.
[30] J.M. Ball, R.D. James, Arch. Ration. Mech. Anal. 100 (1987) 13–52, https://doi.org/10.1007/BF00281246.
[31] A.N. Bucsek et al., Shape Mem. Superelastic. 2 (2016) 62–79, https://doi.org/10.1007/s40830-016-0052-5.
[32] J. Cui et al., Nat. Mater. 5 (2006) 286–290, https://doi.org/10.1038/nmat1593.
[33] Z. Zhang, R.D. James, S. Müller, Acta Mater. 57 (2009) 4332–4352, https://doi.org/10.1016/j.actamat.2009.05.034.
[34] P. Leo, T.W. Shield, P. Bruno, Acta Metall. Mater. 41 (1993) 2477–2485.
[35] L. Zhang, L.Q. Chen, Q. Du, Phys. Rev. Lett. 98 (2007) 1–4, https://doi.org/10.1103/PhysRevLett. 98.265703.
[36] J.C. Fisher, J.H. Hollomon, D. Turnbull, J. Appl. Phys. 19 (1948) 775–784, https://doi.org/10.1063/1.1698202.
[37] R.D. James, Z. Zhang, Springer Ser. Mater. Sci. 79 (2005) 159–175, https://doi.org/10.1007/3-540-31631-0_9.
[38] M. Wechsler, D. Lieberman, T. Read, Trans. Am. Inst. Min. Metall. Eng. 197 (1953) 1503–1515.
[39] J.S. Bowles, J.K. Mackenzie, Acta Mater. 2 (1954) 129–137, https://doi.org/10.1016/0001-6160(54)90102-9.
[40] C. Chluba et al., Science 348 (2015) 1004–1007, https://doi.org/10.1126/science.1261164.
[41] K.A. Bywater, J.W. Christian, Phil. Mag. 25 (1972) 1249–1273, https://doi.org/10.1080/14786437208223852.
[42] K. Bhattacharya, R.V. Kohn, Acta Mater. 44 (1996) 529–542, https://doi.org/10.1016/1359-6454(95)00198-0.
[43] X. Ni et al., Nano Lett. 16 (2016) 7621–7625, https://doi.org/10.1021/acs.nanolett.6b03555.
[44] F.J. Gil, J.M. Manero, J.A. Planell, J. Mater. Sci. 30 (1995) 2526–2530, https://doi.org/10.1007/BF00362129.
[45] S. Won et al., Scripta Mater. 59 (2008) 1186–1189, https://doi.org/10.1016/j.scriptamat.2008.08.005.
[46] T. Waitz, H.P. Karnthaler, Acta Mater. 52 (2004) 5461–5469, https://doi.org/10.1016/j.actamat.2004.08.003.
[47] H. Yin et al., Int. J. Fatigue 88 (2016) 166–177, https://doi.org/10.1016/j.ijfatigue.2016.03.023.
[48] T. Waitz et al., J. Mech. Phys. Solids 55 (2007) 419–444, https://doi.org/10.1016/j.jmps.2006.06.006.
[49] M. Kabla et al., Acta Mater. 70 (2014) 79–91, https://doi.org/10.1016/j.actamat.2014.02.009.
[50] A. Ahadi, Q. Sun, Scripta Mater. 113 (2016) 171–175, https://doi.org/10.1016/j.scriptamat.2015.10.036.
[51] J.E. Schaffer, J. Mater. Eng. Perform. 18 (2009) 582–587, https://doi.org/10.1007/s11665-009-9369-y.
[52] P. Sedmák et al., Acta Mater. 94 (2015) 257–270, https://doi.org/10.1016/j.actamat.2015.04.039.
[53] T. Waitz et al., Mater. Sci. Eng., A 481–482 (2008) 479–483, https://doi.org/10.1016/j.msea.2007.03.122.
[54] X. Shi et al., J. Mater. Sci. 49 (2014) 4643–4647, https://doi.org/10.1007/s10853-014-8167-6.
[55] Q.P. Sun, Y.J. He, Int. J. Solids Struct. 45 (2008) 3868–3896, https://doi.org/10.1016/j.ijsolstr.2007.12.008.
[56] A. Lasalmonie, J. Strudel, J. Mater. Sci. 21 (1986) 1837–1852, https://doi.org/10.1007/BF00547918.
[57] C. Da Wu et al., Appl. Surf. Sci. 292 (2014) 500–505, https://doi.org/10.1016/j.apsusc.2013.11.168.
[58] A.V. Sergueeva et al., Mater. Sci. Eng., A 339 (2003) 159–165, https://doi.org/10.1016/S0921-5093(02)00122-3.
[59] N.B. Morgan, C.M. Friend, J. Phys. IV 11 (2001), https://doi.org/10.1051/jp4:2001855. Pr8-325–Pr8-332.
[60] G. Eggeler et al., Mater. Sci. Eng., A 378 (2004) 24–33, https://doi.org/10.1016/j.msea.2003.10.327.
[61] S. Miyazaki et al., Metall. Trans. A 17A (1986) 115–120, https://doi.org/10.1007/BF02644447.
[62] L.C. Brinson, I. Schmidt, R. Lammering, J. Mech. Phys. Solids 52 (2004) 1549–1571, https://doi.org/10.1016/j.jmps.2004.01.001.
[63] B. Kockar et al., Acta Mater. 56 (2008) 3630–3646, https://doi.org/10.1016/j.actamat.2008.04.001.
[64] K. Gall et al., Mater. Sci. Eng., A 486 (2008) 389–403, https://doi.org/10.1016/j.msea.2007.11.033.
[65] M. Wohlschlögel, R. Lima de Miranda, A. Schüßler, E. Quandt, J. Biomed. Mater. Res. Part B Appl. Biomater. 104 (2016) 1176–1181, https://doi.org/10.1002/jbm.b.33449.
[66] A. Ishida, M. Sato, Z. Gao, Acta Mater. 69 (2014) 292–300, https://doi.org/10.1016/j.actamat.2014.02.006.
[67] A. Evirgen et al., Acta Mater. 121 (2016) 374–383, https://doi.org/10.1016/j.actamat.2016.08.065.
[68] J.X. Zhang, M. Sato, A. Ishida, Acta Mater. 51 (2003) 3121–3130, https://doi.org/10.1016/S1359-6454(03)00124-1.
[69] J. Frenzel et al., Acta Mater. 58 (2010) 3444–3458, https://doi.org/10.1016/j.actamat.2010.02.019.
[70] K. Gall, H. Maier, Acta Mater. 50 (2002) 4643–4657, https://doi.org/10.1016/S1359-6454(02)00315-4.
[71] F. Jiang et al., Acta Mater. 57 (2009) 4773–4781, https://doi.org/10.1016/j.actamat.2009.06.059.
[72] K. Gall et al., J. Eng. Mater. Technol. 121 (1999) 28, https://doi.org/10.1115/1.2815995.
[73] F. Yang et al., Acta Mater. 61 (2013) 3335–3346, https://doi.org/10.1016/j.actamat.2013.02.023.
[74] R. Santamarta et al., Acta Mater. 61 (2013) 6191–6206, https://doi.org/10.1016/j.actamat.2013.06.057.
[75] A.P. Stebner et al., Acta Mater. 76 (2014) 40–53, https://doi.org/10.1016/j.actamat.2014.04.071.
[76] A. Evirgen et al., Mater. Sci. Eng., A 655 (2016) 193–203, https://doi.org/10.1016/j.msea.2015.12.076.
[77] T. Dankwort et al., J. Appl. Crystallogr. 49 (2016) 1009–1015, https://doi.org/10.1107/S160057671600710X.
[78] A. Ishida, M. Sato, K. Ogawa, Phil. Mag. 88 (2008) 2427–2438, https://doi.org/10.1080/14786430802350694.
[79] A. Ishida, M. Sato, Phil. Mag. 87 (2007) 5523–5538, https://doi.org/10.1080/14786430802350694.
[80] V. Srivastava, X. Chen, R.D. James, Appl. Phys. Lett. 97 (2010) 11–13, https://doi.org/10.1063/1.3456562.
[81] K. Otsuka, H. Sakamoto, K. Shimizu, Acta Metall. 27 (1979) 585–601, https://doi.org/10.1016/0001-6160(79)90011-7.