



HAL
open science

From bond strength to charge distribution: modern developments of the empirical analysis of crystal structures

Massimo Nespolo

► **To cite this version:**

Massimo Nespolo. From bond strength to charge distribution: modern developments of the empirical analysis of crystal structures. *Physical Sciences Reviews*, 2023, 8 (7), pp.983-1006. 10.1515/psr-2019-0074 . hal-03938666

HAL Id: hal-03938666

<https://hal.univ-lorraine.fr/hal-03938666>

Submitted on 13 Jan 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial - NoDerivatives 4.0 International License

1 From Bond Strength to Charge Distribution: modern developments of 2 the empirical analysis of crystal structures

3 Massimo Nespolo¹

4 ¹Université de Lorraine, CNRS, CRM2, Nancy, France. massimo.nespolo@univ-lorraine.fr

5 Abstract

6 A brief review of the empirical analysis of chemical-bond topology in crystal structures is presented, from
7 the pioneering work of Pauling to the most recent development represented by the Charge Distribution
8 analysis. The strengths and limitations of the various methods are briefly discussed.

9 1. Introduction

10 A real crystal differs from its idealized model for the presence of static and dynamic defects. Among the
11 formers, we can cite vacancies, interstitials, impurities, precipitates, dislocations and the surface itself.
12 Among the dynamic defects we have individual and collective atomic vibrations. A classical X-ray
13 diffraction experiment gives a space- and time-average structure which justifies adopting an essentially
14 idealized model of the crystal structure, where the deviations from the ideal model are accounted for in terms
15 of disordered occupation of atomic sites and a finite volume assigned to each atomic position, most often in
16 the shape of an ellipsoid that represents the atomic displacement during the diffraction experiment.

17 The stability of a crystalline edifice is governed by its thermodynamics. The calculation of the free energy
18 of a crystal is a complex task (Jacucci and Quirke, 1984; Vasileiadis, 2013), not suitable for a routine
19 screening of large samples. A rougher, empirical estimation is in most cases sufficient to rationalize known
20 crystal structures and to predict what type of environment an atom may have in a crystal. This type of
21 considerations constitutes the rationale behind approaches like the crystal field theory (Burns, 1993) or the
22 Madelung Part of the Lattice Energy (Hoppe, 1970a; 1995). The former is well-known among the readers,
23 whereas the latter has found more limited applications, possibly because of the problems related to the
24 convergence of lattice sums in Madelung's constants (Borwein *et al.*, 1985).

25 Pauling (1929) introduced the well-known empirical rules that today bring his name to estimate the
26 stability of ionic crystals. These rules are based on the description of crystal structures as compact packing of
27 hard spheres bringing a charge that corresponds to the valence of the atoms they represent. The interatomic
28 distance corresponding to a chemical bond is interpreted as the sum of the ionic radii, and the type of
29 coordination, including the coordination number, is governed by the ratio of these radii. The strength of the
30 bond is therefore estimated as the ratio between the charge of the cation centring the coordination
31 polyhedron and its coordination number, *i.e.* the number of anions occupying the corners of that polyhedron.
32 Under the very drastic assumption of a uniform coordination polyhedron (*i.e.* a polyhedron in which all the

33 cation-anion bond distances have the same length), Pauling's bond strength possesses a significant prediction
34 power, expressed by the second rule, which states that the sum of the bond strengths about a given anion is
35 equal to the formal charge of that anion:

$$36 \qquad q(A_j) = \sum_i s_i, \quad s_i = q(C_i)/N_i \qquad (1)$$

37 where q is the charge, A_j and C_i are the j -th anion and i -th cation respectively, N_i is the coordination number
38 of C_i , s_i is the bond strength and the summation is performed over the A_j - C_i bonds. Within the conditions
39 used to define Pauling's bond strength, Eq. (1) can be used to predict the number of cations bonded to a
40 given anion but also as a very simple rationale behind empirical observations like Loewenstein's rule
41 (Loewenstein, 1954), for which today we dispose of more sophisticated interpretations (Larin, 2013).

42 Whereas the sphere packing model still plays an important role in the description and interpretation of
43 non-molecular structures, at least as the aristotype from which a large number of real structures can be
44 obtained as substitution-distortion hettotypes (Umayahara and Nespolo, 2018), the assumption of ionic
45 bonding is quite restrictive on the type of structures that can be studied. For example, simple structures
46 corresponding to important aristotypes, like cuprite, Cu_2O , do not fit Pauling's scheme, because the sum of
47 the ionic radii of Cu^+ and O^{2-} (0.96 and 1.40 Å respectively: Pauling radii) is much larger than the actual
48 bond distance (about 1.85 Å), and the ratio, 0.69, suggests a tetragonal antiprism coordination for copper,
49 which instead occurs as linear O-Cu-O coordination. When non-ionic bonds are involved, availability of
50 orbitals, rather than the ratio of cation and anion radii, governs the coordination (Brown, 1988; Chiari, 1988).
51 In this context, the terms "cation" and "anion" have to be interpreted as a handy term for "electropositive
52 atom" and "electronegative atoms" respectively.

53 Even more severe is the assumption of uniform coordination polyhedra. Pauling's bond strength does not
54 include any dependence on the bond length but assigns a unique bond strength to a cation bonded to a given
55 number of anions of the same chemical species ("homoligand polyhedra": Mohri, 2000). This assumption
56 becomes less and less realistic the more a coordination polyhedron undergoes deformation. Small polyhedra
57 centred on hard ions¹ like Si-centred tetrahedra in aluminosilicates present a small range of scattering of
58 bond lengths which justifies a description in terms of Pauling's bond strength as a first approximation. In
59 contrast, soft cations like the larger alkali metals usually occur in the cavities of a structure and their
60 interaction with the surrounding anions is spread over a large interval of bond lengths; no net separation
61 appears between a first and a second coordination sphere, so that the coordination number is hardly identified
62 (Chiari, 1990; Balić Žunić & Makovicky, 1996; Makovicky & Balić Žunić, 1998). In this type of
63 coordination environment, the bond strength decreases almost asymptotically: not only assigning a unique
64 bond strength gives a completely unrealistic picture; one also needs some criterion to assign the coordination
65 number.

66 Finally, the possibility of heteroligand polyhedra (Mohri, 2000), i.e. coordination polyhedra in which the
67 atoms at the centre of the polyhedra are bonded to chemically different atoms at the corners, is not even

1 By the terms "hard" and "soft" ions we make reference to the well-known HSAB theory (Pearson, 1968).

68 considered in Pauling's approach, which explains the lack of duality in his definition of bond strength. In
69 fact, whereas in ionic solids a cation is often, although not always, bonded to a single chemical species of
70 anion, that anion is usually bonded to different types of cations. Now, heteroligand polyhedra around cations
71 do also occur: a cation can for example be bonded at the same time to oxygen, halogens, sulphur etc. This
72 type of polyhedra, non-uniform by definition, cannot be analysed in terms of Pauling's bond strength.

73 Different generalisations of Pauling's approach have been introduced, which all have in common the
74 consideration of the dependence of the bond strength from the bond distance. After a very brief critical
75 account of some of them, we concentrate on the Charge Distribution approach, which is the most recent
76 development of Pauling's original ideas.

77 **2. Bond strength as a function of bond lengths**

78 The dependence of the bond strength from the bond length can be modelled by a function relating the
79 former to the latter and a series of empirical parameters:

$$80 \quad s_i = f(d_i, \mathbf{p}) \quad (2)$$

81 where s_i and d_i are the bond strength and the bond length of the i -th interatomic distance between two given
82 chemical elements, and \mathbf{p} is a vector containing a certain number (usually up to three) of empirical
83 parameters. Baur (1970, 1971) showed a linear dependence of the bond length from the bond strength
84 received by *each* anion (cf. also Ferraris and Catti, 1973). Several functions f have been proposed by various
85 authors, where the bond strength decreases in a more or less exponential way with the increase of the bond
86 length. These functions are often called "R-s curves" (Donnay and Allman, 1970; Pyatenko, 1973; Brown
87 and Shannon, 1973; Brown and Wu, 1976; Allman, 1975; Trömel, 1983, 1984, 1986; Brown and Altermatt,
88 1985; Urusov, 1995), where R is the bond length (we use d instead of R for consistency of notation with
89 other approaches). The bond strength computed from Eq. (2) is called *bond valence*, shortened to BV
90 hereafter (Donnay and Allman, 1970); the coordination number is obtained as the number of anions after
91 which s_i is close to zero (Brown, 1978) but it has been set *a priori* in the group of structures used for
92 determining the empirical parameter. By replacing Pauling's definition of bond strength with a curve of type
93 (2), Eq. (1) becomes much more general and the restrictions to uniform or homoligand coordination
94 polyhedra are removed. Furthermore, bonds between the same type of atom become in principle treatable,
95 provided that a suitable and reliable set of parameters \mathbf{p} is available for the pairs of atoms involved in the
96 chemical bond, under the experimental conditions in which the structure was solved and refined. These
97 parameters are obtained by a fit on a large set of reliably refined structures and are frequently redefined and
98 re-refined (Bosi, 2014a,b; Gagné & Hawthorne, 2015) so that the same crystal structure gives different
99 results when analysed with the same type of curve but with a different set of empirical parameters.

100 Brown (1977, 1978) developed the theory of bond valence as a method of predicting bond lengths in
101 inorganic crystals. In case of very distorted distributions, however, the results are poor; for the case of V^{5+}
102 Gopal (1972) reported that "V-O bond lengths can be predicted with some accuracy for systems where the

103 range of bond lengths are not too great ($\sim 10\%$). Several extensions have been proposed with different
104 goals, such as: 1) to treat bonds which occur with a wide range of lengths (Brown, 1987); 2) to introduce the
105 effect of the lone-pairs on the geometry of the coordination polyhedra (Wang & Liebau, 1996a,b); 3) to take
106 into account zero and negative oxidation states (Naskar *et al.*, 1997); 4) to correlate bond strength and
107 electron density distribution (Gibbs *et al.*, 1998); 5) to treat compounds with heteroligand polyhedra (Mohri,
108 2000). All these extensions are based on a more or less complex power-law relationship between the bond
109 length and the bond valence and do not have any internal criterion to evaluate whether the method itself is
110 applicable or not to a given structure. Therefore, when a significant deviation from the expected bond
111 valences occurs, the investigator should evaluate independently, and *before* drawing any conclusion from the
112 bond valence analysis, whether the quality of the refinement of the structure model is reliable, and whether
113 the type of structure (type and pattern of bonding) is suitable for such analysis. A ‘blind’ application without
114 critical inspection may otherwise result into a discussion in terms of bond valences without solid ground.

115 Extensions in a different direction were proposed in terms of graph-theoretical approaches. Boisen *et al.*
116 (1988) considered all the Lewis structures that can be associated with a given compound, and defined the
117 *resonance bond number* (RBN) as the weighted average of the bond order over the Lewis structures. This
118 method ignores the periodicity of the crystalline structures and omits any non-nearest-neighbour interaction:
119 the calculation requires 10^7 or 10^8 Lewis graphs per atom and the resonance bond numbers obtained in this
120 way are close or identical to the bond valences, but the method has the advantage of being applicable to
121 amorphous as well as to crystalline solids (Rutherford, 1991; 1998).

122 Hoppe (1970b; 1979) criticized the common definitions and applications of the concepts on which bond
123 valence theory is based, namely:

- 124 1) *the coordination number*, which can be hardly identified in those cases in which no net separation
125 appears between a first and a second coordination sphere;
- 126 2) *the ionic radius*, which is far from being uniquely defined, because the interatomic distances cannot
127 always be described as sum of ionic radii, the variation depending partly on geometrical constraints
128 on the crystal structure and especially on the overall connectivity of the atoms in the structure
129 (Rutherford, 1991).

130 To generalize the concepts of coordination number and ionic radius, Hoppe (1979) introduced the
131 definitions of FIR (Fictive Ionic Radii), MEFIR (MEan Fictive Ionic Radii) and ECoN (Effective
132 Coordination Number), and Hoppe *et al.* (1989) applied ECoN to the calculation of the distribution of
133 charges in crystalline structures. This “charge” actually has a strict correspondence with the BV, as defined
134 by Brown (1978). Hoppe *et al.* (1989) showed that the charge distribution method systematically gives better
135 results than BV when applied to oxide compounds. Despite a large and increasing number of publications
136 where this approach has shown its power, it still remains underestimated with respect to bond valence,
137 possibly because of its more demanding computation requirements. In the following sections we present a
138 systematic analysis of the approach originally developed by Hoppe and a critical comparison with the bond

139 valence methods. The step-by-step developments are available in our series of articles on the subject
140 (Nespolo *et al.*, 1999, 2001; Eon and Nespolo, 2015; Nespolo, 2016).

141 3. FIR, MEFIR and ECoN

142 Atoms in a non-molecular crystal structure can be classified in “polyhedron-centring atoms” (**PC-atoms**)
143 and “vertex (corner) atoms” (**V-atoms**). The assignment of cations (“electropositive atoms”) or anions
144 (“electronegative atoms”) to each of the two categories can be done interchangeably. The majority of non-
145 molecular structures are described as corner, edge or, more rarely, face sharing of polyhedra centred on
146 cations, but anion-centred polyhedra represent an alternative, and sometimes preferable description in a
147 number of cases (Krivovichev, 2009; Krivovichev *et al.*, 2013).

148 In the following sections, PC-atoms and V-atoms are indicated as PC(*ij*) and V(*rs*) respectively, where *i*
149 and *r* identify the atomic site (i.e. the chemical element in absence of isomorphic substitution; otherwise, a
150 statistical mix of chemical elements), *j* and *s* the crystallographic type. The latter indices, *j* and *s*, are
151 necessary when atoms of the same chemical element are not all symmetrically equivalent but two or more
152 occur in the asymmetric unit. The formal charges (which become weighted average charges in case of
153 isomorphic substitutions) are indicated as *q*(*ij*) and *q*(*rs*) and depend on *j* or *s* through the site occupation
154 factor (*sof*): for *sof* = 1, all the PC-atoms with the same *i*, as well as all the V-atoms with the same *r*, have the
155 same *q*. Similarly, the multiplicity of the Wyckoff position is indicated as *h*(*ij*) and *h*(*rs*). A bond between the
156 atoms PC(*ij*) and V(*rs*) is indicated as *ij*→*rs*, the corresponding distance as *d*(*ij*→*rs*). The *L*-th bond length
157 between PC(*ij*) and V(*rs*) is then indicated as *d*(*ij*→*rs*)_{*L*}. For computational purposes, the bond lengths are
158 ordered in increasing length with respect to *i*, *j* and *r*, where *d*(*ij*→*rs*)₁ is the shortest one and *s* a sort of
159 dummy index at this stage.

160 Pauling's (1929) statement that in ionic compounds the distance between two bonded ions with charges of
161 opposite sign should be nearly identical to the sum of the ionic radii R(*i*) and R(*r*) was generalised by Hoppe
162 (1979), with the introduction of the Fictive Ionic Radii FIR(*ij*→*rs*)_{*L*} concept, defined as:

$$163 \text{FIR}(ij \rightarrow rs)_L = d(ij \rightarrow rs)_L \cdot \frac{R(i)}{R(i) + R(r)}. \quad (3)$$

164 In case of uniform coordination polyhedra, all the *d*(*ij*→*rs*)_{*L*} are identical and the same holds for
165 FIR(*ij*→*rs*)_{*L*}; if the crystal is ionic, *d*(*ij*→*rs*) = R(*i*)+R(*r*) so that FIR(*ij*→*rs*) should correspond to R(*i*).
166 Deviations from the ionic radius may be read as a measure of the departure from the pure ionic character of
167 the bond. For non-uniform polyhedra, instead, the concept of ionic radius is less well defined; accordingly, a
168 different Fictive Ionic Radius is assigned to each bond, which is then used to compute a weighted average,
169 called the MEan Fictive Ionic Radius (MEFIR). Concretely, an ion at the centre of a uniform polyhedron can
170 be described as a sphere whose dimension, FIR, is directly related to the ionic radius, R(*i*). For non-uniform
171 polyhedra, instead, the ionic radius shows a dependence on the direction, which leads to a polyhedral shape
172 for the atomic basins (Pendás *et al.*, 1998). The computation of MEFIR corresponds to assigning to the ion a

173 spherical basin whose radius is a weighted average of the ionic radius for each direction. MEFIR is
 174 calculated through a convergent iterative process:

$$\begin{aligned}
 & \sum_s \sum_L \text{FIR}(ij \rightarrow rs)_L \cdot \exp \left\{ 1 - \left[\frac{\text{FIR}(ij \rightarrow rs)_L}{{}^{n-1}\text{MEFIR}(ij \rightarrow r)} \right]^6 \right\} \\
 175 \quad {}^n\text{MEFIR}(ij \rightarrow r) = & \frac{\sum_s \sum_L \text{FIR}(ij \rightarrow rs)_L \cdot \exp \left\{ 1 - \left[\frac{\text{FIR}(ij \rightarrow rs)_L}{{}^{n-1}\text{MEFIR}(ij \rightarrow r)} \right]^6 \right\}}{\sum_s \sum_L \exp \left\{ 1 - \left[\frac{\text{FIR}(ij \rightarrow rs)_L}{{}^{n-1}\text{MEFIR}(ij \rightarrow r)} \right]^6 \right\}} \quad (4)
 \end{aligned}$$

176 where n is the number of iterations and where the starting value, ${}^0\text{MEFIR}(ij \rightarrow r)$, is simply $\text{FIR}(ij \rightarrow rs)_1$. The
 177 exponential term is responsible for the asymptotic decrease of the contribution of each FIR with the the
 178 increase of the bond distance. MEFIR gives a quantitative meaning to the *rule of mutual influence* (Beck,
 179 2014) of different cations on their seeming size and polarity, according to which the size of an ion depends
 180 on the chemical relations with its neighbours. A global ${}^n\text{MEFIR}(ij)$ for each PC-atom can also be defined,
 181 because ${}^n\text{MEFIR}(ij \rightarrow r)$ does not change significantly as a function of r :

$$\begin{aligned}
 & \sum_r \sum_s \sum_L \text{FIR}(ij \rightarrow rs)_L \cdot \exp \left\{ 1 - \left[\frac{\text{FIR}(ij \rightarrow rs)_L}{{}^{n-1}\text{MEFIR}(ij)} \right]^6 \right\} \\
 182 \quad {}^n\text{MEFIR}(ij) = & \frac{\sum_r \sum_s \sum_L \text{FIR}(ij \rightarrow rs)_L \cdot \exp \left\{ 1 - \left[\frac{\text{FIR}(ij \rightarrow rs)_L}{{}^{n-1}\text{MEFIR}(ij)} \right]^6 \right\}}{\sum_r \sum_s \sum_L \exp \left\{ 1 - \left[\frac{\text{FIR}(ij \rightarrow rs)_L}{{}^{n-1}\text{MEFIR}(ij)} \right]^6 \right\}} \quad (5) \\
 {}^0\text{MEFIR}(ij) = & \text{FIR}(ij \rightarrow rs)_1
 \end{aligned}$$

183 The concepts of FIR and MEFIR have been introduced for ionic bonds but they do apply as well to polar
 184 bonds. According to Pauling's (1960) classification, chemical bonds can be divided into ionic, polar and non-
 185 polar depending on the electronegativity difference of the bonded atoms, this difference being respectively
 186 higher than 1.7, between 0.4 and 1.7, and lower than 0.4. For polar bonds, one could simply use covalent
 187 radii instead of ionic radii in Eq. (3): $\text{FIR}(ij \rightarrow rs)_1$ would then correspond to the covalent radius for bond
 188 order 1, or be smaller than it for higher bond orders.

189 FIR and MEFIR clearly aim at treating atoms as spherical although they are not necessarily spherical. The
 190 purpose is to keep the description of a crystal structure as simple as possible, with nevertheless some
 191 consideration of the effect of the environment in which each atom is embedded. In the same line, the
 192 classical definition of the coordination number of an atom as “the number of other atoms directly linked to
 193 that specified atom” (IUPAC, 1997) becomes less and less satisfactory with the distribution of bond distances
 194 over a large interval, from uniform polyhedra to cavities which can hardly be recognized as polyhedra, like
 195 in the typical example of alkaline metals showing a very irregular coordination. These considerations led
 196 Hoppe (1970) to call the coordination number an “inorganic chameleon” and generalise it to a function of the
 197 weighted average of bond distances which he called the Effective Coordination Number, ECoN (Hoppe,
 198 1979). ECoN is computed as a normalised sum of bond strengths defined as function of the bond lengths to
 199 which a weight is associated that decreases with the increase of the bond length. The approach is therefore
 200 somehow similar to that behind Eq. (2), with however a fundamental difference: instead of using empirical
 201 parameters which depend on the pair PC-V, like in the Bond Valence approach, the function is based on the

202 ratio of experimental bond distances in each polyhedron; the bond strength computed in this way is called
 203 *bond weight*, shortened to BW hereafter.

204 The first step in the computation of the bond weight is the calculation of a weighted mean distance
 205 ${}^n d(ij \rightarrow r)$ through a convergent iterative process similar to that used to obtain ${}^n \text{MEFIR}$:

$$206 \quad {}^n d(ij \rightarrow r) = \frac{\sum_s \sum_L d(ij \rightarrow rs)_L \cdot \exp \left\{ 1 - \left[\frac{d(ij \rightarrow rs)_L}{{}^{n-1} d(ij \rightarrow r)} \right]^6 \right\}}{\sum_s \sum_L \exp \left\{ 1 - \left[\frac{d(ij \rightarrow rs)_L}{{}^{n-1} d(ij \rightarrow r)} \right]^6 \right\}} \quad (6)$$

207 with ${}^0 d(ij \rightarrow r) = d(ij \rightarrow rs)_1$. The weighted average $d(ij \rightarrow r)$ is actually simply MEFIR scaled by the ratio $[R(i)$
 208 $+R(r)]/R(i)$; in fact, solving Eq. (3) for $d(ij \rightarrow rs)_L$, we rewrite Eq. (6) at step $n = 1$ as follows:

$$\begin{aligned} {}^1 d(ij \rightarrow r) &= \frac{\sum_s \sum_L \text{FIR}(ij \rightarrow rs)_L \cdot \frac{R(i) + R(r)}{R(i)} \cdot \exp \left\{ 1 - \left[\frac{\text{FIR}(ij \rightarrow rs)_L \cdot \frac{R(i) + R(r)}{R(i)}}{\text{FIR}(ij \rightarrow rs)_1 \cdot \frac{R(i) + R(r)}{R(i)}} \right]^6 \right\}}{\sum_s \sum_L \exp \left\{ 1 - \left[\frac{\text{FIR}(ij \rightarrow rs)_L \cdot \frac{R(i) + R(r)}{R(i)}}{\text{FIR}(ij \rightarrow rs)_1 \cdot \frac{R(i) + R(r)}{R(i)}} \right]^6 \right\}} = \\ &= \frac{\sum_s \sum_L \text{FIR}(ij \rightarrow rs)_L \cdot \frac{R(i) + R(r)}{R(i)} \cdot \exp \left\{ 1 - \left[\frac{\text{FIR}(ij \rightarrow rs)_L}{\text{FIR}(ij \rightarrow rs)_1} \right]^6 \right\}}{\sum_s \sum_L \exp \left\{ 1 - \left[\frac{\text{FIR}(ij \rightarrow rs)_L}{\text{FIR}(ij \rightarrow rs)_1} \right]^6 \right\}} = \\ &= \frac{R(i) + R(r)}{R(i)} \frac{\sum_s \sum_L \text{FIR}(ij \rightarrow rs)_L \cdot \exp \left\{ 1 - \left[\frac{\text{FIR}(ij \rightarrow rs)_L}{\text{FIR}(ij \rightarrow rs)_1} \right]^6 \right\}}{\sum_s \sum_L \exp \left\{ 1 - \left[\frac{\text{FIR}(ij \rightarrow rs)_L}{\text{FIR}(ij \rightarrow rs)_1} \right]^6 \right\}} = \frac{R(i) + R(r)}{R(i)} \cdot {}^1 \text{MEFIR}(ij \rightarrow r) \end{aligned}$$

209 and by iteration we obtain ${}^n d(ij \rightarrow r) = {}^n \text{MEFIR}(ij \rightarrow r) \cdot [R(i) + R(r)]/R(i)$.

210 The shortest PC-V distance in each polyhedron is used as *normalising parameter* at the zero-th stage of
 211 the iteration, to be replaced by the weighted average until convergence is reached. This exponential term,
 212 which acts as a weight applied to each bond length, is precisely the bond weight ${}^n w(ij \rightarrow rs)$:

$$213 \quad {}^n w(ij \rightarrow rs)_L = \exp \left\{ 1 - \left[\frac{d(ij \rightarrow rs)_L}{{}^n d(ij \rightarrow r)} \right]^6 \right\} = \exp \left\{ 1 - \left[\frac{\text{FIR}(ij \rightarrow rs)_L}{{}^n \text{MEFIR}(ij \rightarrow r)} \right]^6 \right\} \quad (7)$$

214 The sum over the bond weights gives ${}^n \text{ECoN}(ij \rightarrow r)$:

$$215 \quad {}^n \text{ECoN}(ij \rightarrow r) = \sum_s \sum_L {}^n w(ij \rightarrow rs)_L \quad (8)$$

216 Because ${}^n w(ij \rightarrow rs)$ is equivalently defined in terms of FIR/ ${}^n \text{MEFIR}$ and of $d^n d$ (Eq. 7), ECoN too is

217 equivalently defined in terms of MEFIR (as in Hoppe, 1979) and of $d(ij \rightarrow r)$ (as in Hoppe *et al.*, 1989). Eq.
 218 (6) can therefore be rewritten as:

$$219 \quad {}^n d(ij \rightarrow r) = \frac{\sum_s \sum_L d(ij \rightarrow rs)_L \cdot {}^{n-1} w(ij \rightarrow rs)_L}{\sum_s \sum_L {}^{n-1} w(ij \rightarrow rs)_L} \quad (6')$$

220 MEFIR is the only variable that depends, through FIR, on the ionic radii. BW, ECoN, and the computed
 221 ‘charges’, defined and illustrated in the next section, which are a function of ECoN, do not depend on the
 222 ionic radii. For homoligand polyhedra ${}^n \text{ECoN}(ij \rightarrow r)$ is actually ${}^n \text{ECoN}(ij)$, because r takes only a single
 223 value for each coordination polyhedron. Instead, for heteroligand polyhedra, ${}^n \text{ECoN}(ij)$ defined simply as the
 224 sum of ${}^n \text{ECoN}(ij \rightarrow r)$ results from r coordination shells corresponding to each type of anion.

225 ECoN, as a generalisation of the classical coordination number, is a *real* number that has to become equal
 226 to the integer coordination number for uniform polyhedra. This is the condition imposed to obtain the
 227 *contraction parameter* 6 in the definition of the bond weights. It was in fact obtained by finding the highest
 228 value giving an ECoN equal to the number of first neighbours in the structure of simple metals, where a clear
 229 separation between a first and a second coordination sphere exists (Hoppe, 1979). One exception was later
 230 found, for which the contraction parameter had to be modified. In the case of hydrogen bonds, the ratio of
 231 two short distances (donor-H and H-acceptor) with a high relative gap makes the weight ${}^n w(ij \rightarrow rs)$ for the
 232 second bond (hydrogen-acceptor bond) negligible so that its presence is not taken into account in the
 233 iterative calculation. A revised contraction parameter of 1.6 was then introduced (Nespolo *et al.*, 2001).

234 Although Eq. (7) can be equally expressed in terms of ${}^n d(ij \rightarrow r)$ or of ${}^n \text{MEFIR}(ij \rightarrow r)$, the results
 235 significantly differ in the case of heteroligand polyhedra. If r takes only one value (homoligand polyhedra),
 236 the two expressions give exactly the same result. If instead r takes multiple values, Eq. (7) has to be
 237 computed separately for each r when using ${}^n d(ij \rightarrow r)$ so that each heteroligand polyhedron is divided into
 238 homoligand subpolyhedra, treated independently, while this is not the case when it is computed in terms of
 239 ${}^n \text{MEFIR}(ij \rightarrow r)$. In fact, by using the bond lengths one automatically selects a different normalising parameter
 240 (minimal distance) for each type of V-atom: it would make no sense to normalise bond lengths of a V-atom
 241 with respect to the shortest bond length of a *different* V-atom having in general a different size. It would also
 242 make no sense to define a unique weighted mean distance ${}^n d(ij)$ by making r in ${}^n d(ij \rightarrow r)$ variable over all the
 243 V-atoms coordinated by the same PC-atom because the corresponding bond distances can be significantly
 244 different so that global average would over-estimate bonds with smaller V-atoms and under-estimate bonds
 245 with larger V-atoms. Accordingly, ${}^n \text{ECoN}(ij \rightarrow r)$, Eq. (8), is defined unambiguously for each homoligand
 246 subpolyhedron, and the resulting ${}^n \text{ECoN}(ij \rightarrow r)$ are then summed up to obtain ${}^n \text{ECoN}(ij)$:

$$247 \quad {}^n \text{ECoN}(ij) = \sum_r {}^n \text{ECoN}(ij \rightarrow r) = \sum_r \sum_s \sum_L {}^n w(ij \rightarrow rs)_L \quad (9)$$

248 ignoring thus the effect of the size of the different V-atoms at this stage.

249 **3.1. Examples**

250 Table 1 shows the example of a regular octahedron with six PC-V distances of 2.000 Å, undergoing three
251 gradual deformations of bond distances, which however do not affect the arithmetic average, kept constant at
252 2.000 Å. With increasing distortion, the weighted average distance $\bar{d}(ij \rightarrow r)$ (Eq. 6') decreases with respect
253 to the arithmetic average. This is a normal effect of the exponential decay of the bond weights. When two
254 distances d_1 and d_2 depart from the value d_{uni} corresponding to a uniform polyhedron by the same amount but
255 in the opposite sense ($d_1 = d_{\text{uni}} - \Delta$; $d_2 = d_{\text{uni}} + \Delta$), the increase of the bond weight of d_1 is larger than the
256 decrease of the bond weight of d_2 . This leads to a shortening of the weighted average distance and a
257 reduction of ECoN, which is no longer equal to the classical value of the coordination number. Accordingly,
258 the smaller ECoN becomes, the more distorted is the polyhedron. In this example, ECoN decreases slowly
259 for the two first distortions, but for the third distortion is only 4.73, and the distribution of the bond weights
260 shows that the coordination is closer 4+2 than to 6 (Figure 1).

261 Table 2 shows the example of KNaTiO_3 , which contains three types of cation (Ti, K, Na; one
262 crystallographic type each) and one type of anion (O; two crystallographic types: O1 and O2). The index i
263 runs from 1 to 3, and for each of them j takes the only value 1 ($ij = 11$: Ti; $ij = 21$: K; $ij = 31$: Na). The index
264 r takes the only value 1 and the index s runs from 1 to 2 ($rs = 11$: O1; $rs = 12$: O2). Na. where K is in contact
265 with twelve oxygen atoms (six pairs of distances), Ti and Na with five. The K-O distances increase in a more
266 or less continuous way, until a jump is observed after the eighth distance. FIR has a parallel behaviour and
267 the bond weight decreases. The contribution of the last four oxygen atoms is almost negligible (lower than
268 0.01) and the value of ECoN is practically 8. Only the first eight oxygen atoms contribute significantly to the
269 coordination of K. The opposite occurs for Na, where the contribution of each to the five bonds to ECoN is
270 comparable, leading to ECoN practically 5. An intermediate situation occurs for Ti, which makes five bonds
271 with oxygen atoms but ECoN is only 4.56.

272 Table 3 shows the results for α' - Ba_2TiO_4 . This orthorhombic polymorph of Ba_2TiO_4 crystallizes in $P2_1nb$,
273 which is an unconventional setting of $Pna2_1$ (space-group type No. 33). It is a Bieberbach-type of space
274 group (Nespolo et al. 2018), which does not contain any special Wyckoff position; the multiplicity of the
275 general Wyckoff position is 4. For this compound $Z = 12$, so that the unit cell contains 24 barium atoms (six
276 sites), 12 titanium atoms (three sites) and 48 oxygen atoms (twelve sites). The six polyhedra centred on
277 barium have eight to ten bond contacts whereas the three titanium atoms are all in tetrahedral coordination.
278 ECoN from the three titanium atoms is 3.99, showing that the distortion is minimal. On the other hand, for
279 barium ECoN spans a large interval, from 5.54 to 7.69. The minimal value corresponds to Ba1, with eighth
280 Ba-O contacts, of which seven are bonding contacts, whereas the last two contribute for less than 2% to
281 ECoN ($0.069 + 0.035 = 0.104$, i.e. 1.9 % of 5.54) and can therefore be considered negligible. The
282 discrepancy between ECoN and the number of oxygen atoms bonded to Ba (6) is a measure of the degree of
283 distortion of the polyhedron. Ba2 has nine bond contacts, none of which can be considered negligible. The
284 bond weights span however a very large interval, which explains why the nine contacts correspond to ECoN

285 of only 6.16. Ba3 presents a similar situation, the tenth Ba-O distance having a bond weight negligible. The
 286 last three barium polyhedra are less deformed: with eight Ba-O bonds each, their ECoN goes from 7.43 to
 287 7.69.

288 4. Charge Distribution Analysis of homoligand polyhedra

289 ECoN($ij \rightarrow r$) computed through Eq. (8) can be distributed among all the bonds around the PC-atom
 290 defining a homoligand (sub)polyhedron, obtaining in this way the fractional contribution by each V-atom to
 291 ECoN itself.

$$292 \quad \Delta\text{ECoN}(ij \rightarrow rs) = \frac{\sum_L {}^n w(ij \rightarrow rs)_L}{\text{ECoN}(ij \rightarrow r)}. \quad (10)$$

293 If one sums up $\Delta\text{ECoN}(ij \rightarrow rs)$, one gets obviously 1 (cfr. Eq. 8):

$$\sum_s \Delta\text{ECoN}(ij \rightarrow rs) = \frac{\sum_s \sum_L {}^n w(ij \rightarrow rs)_L}{\text{ECoN}(ij \rightarrow r)} = \frac{\text{ECoN}(ij \rightarrow r)}{\text{ECoN}(ij \rightarrow r)} = 1$$

294 Each bond $ij \rightarrow rs$ connects a PC-atom with formal charge $q(ij)$ and a V-atom with formal charge $q(rs)$. We
 295 can read this network of bonds as a graph along which the formal charge $q(ij)$ is transferred (“distributed”)
 296 from the PC-atom to the V-atoms bonded to it, exactly like Pauling’s scheme in Eq. (1). Instead of using a
 297 single bond strength calculated as q/N , which does not take into account the variation of bond strength with
 298 bond lengths, we use $\Delta\text{ECoN}(ij \rightarrow rs)$ as a factor to distribute the formal charge of the PC-atom. Because the
 299 result is entirely dependent on the geometry of the coordination polyhedron, the distribution scheme is
 300 usually read the other way round, saying that the formal charge $q(ij)$ is used as a *scaling parameter* applied
 301 to ΔECoN , providing the contribution by each PC-V bond to the distributed ‘charge’:

$$302 \quad \Delta q(ij \rightarrow rs) = \Delta\text{ECoN}(ij \rightarrow rs) q(ij). \quad (11)$$

303 Finally, by summing the $\Delta q(ij \rightarrow rs)$ on the PC-atoms around a V-atom, one should obtain the expected
 304 ‘charge’ of the V-atom itself:

$$305 \quad Q(rs) = - \sum_i \sum_j \Delta q(ij \rightarrow rs) \frac{h(ij)}{h(rs)}. \quad (12)$$

306 where q is the input charge (formal oxidation number) and Q is the charge computed as a function of the
 307 distribution of ECoN. The ratio of the multiplicities of the respective Wyckoff positions, $h(ij)/h(rs)$, is
 308 introduced to avoid counting multiple contributions when a bond connects atoms on Wyckoff positions with
 309 different multiplicities. Eq. (12) is the extension of Pauling’s second rule (Eq. 1) to general coordination
 310 polyhedra.

311 A similar distribution is repeated the other way round and summed up on the V-atoms about a PC-atom.
 312 This time, however, instead of using $q(rs)$ as scaling parameter, in a perfectly symmetrical way, the ratio
 313 $q(rs)/Q(rs)$ for the $V(rs)$ bonded to $PC(ij)$ is used:

$$Q(ij) = \left[\sum_r \sum_s \Delta \text{ECoN}(ij \rightarrow rs) \frac{q(rs)}{Q(rs)} \right] q(ij) = \sum_r \sum_s \Delta q(ij \rightarrow rs) \frac{q(rs)}{Q(rs)}. \quad (13)$$

315 If $q(rs)/Q(rs) = 1$ for each r and s , *i.e.* if the distributed charge and formal charge are identical, then the
 316 bracket in Eq. (13) goes to 1 (it becomes the sum of the fractions of ECoN about each PC-atom) and thus
 317 $Q(ij) = q(ij)$. This shows that in a structure correctly solved and perfectly valence-balanced the distribution of
 318 ECoN scaled by the formal charges gives back those charges. On the other hand, when a structure has
 319 developed some structural strains, its V-atoms may not be perfectly balanced; nevertheless, the distribution
 320 of this unbalance, measured by the ratio $q(rs)/Q(rs)$, should give back the expected formal charges on the
 321 PC-atoms, otherwise the whole set of coordination polyhedra would be unbalanced and the structure would
 322 be unstable. In other words, CHARDI possesses an internal criterion to evaluate the quality of its analysis:
 323 the ratio $q(ij)/Q(ij)$ for the PC-atoms. When this ratio is reasonably close to 1 for all the PC atoms, then the
 324 analysis of the connectivity can be approached by the study of the $Q(rs)$, which diverge from $q(rs)$
 325 proportionally to the structural strains inside the structure. One speaks of *over-under-bonding (OUB) effect*.
 326 Now, because the core of the method is a *distribution* of ECoN scaled by the formal charges (oxidation
 327 numbers), in the trivial case of a structure containing only one type of V-atom, when there is nothing to
 328 distribute, $Q(rs)$ is identical to $q(rs)$ and the CHARDI analysis simply does not give any information.

329 Tables 4 and 5 and Figures 2 and 3 show the Charge Distribution analysis for KNaTiO_3 and $\alpha\text{-Ba}_2\text{TiO}_4$,
 330 whose ECoN analysis was presented in Tables 2 and 3. The same tables provide the Bond Valence Sums
 331 (BVS hereafter) as well. The results of the CHARDI analysis are closer to the formal charges (oxidation
 332 numbers) than the BVS. The two approaches are fundamentally different: whereas CHARDI, as its name
 333 says, *distributes* the charges and sums back those distributions, BVS *computes* the charges. With CHARDI it
 334 is therefore impossible to get charges that are globally lower (or higher) than the oxidation numbers, as it
 335 occurs instead with BVS (*i.e.* both crystallographic types of oxygen in KNaTiO_3 have a BVS lower, as
 336 absolute value, than the formal oxidation numbers, whereas with CHARDI the slight excess of one oxygen is
 337 compensated by the slight deficit of the other, once the site multiplicity is taken into account). However,
 338 because these figures are often used for structure validation, it is evident that the conclusions drawn from
 339 BVS are subject to criticism in the spotlight of the better agreement one gets from CHARDI. As a matter of
 340 fact, when the method does apply (on the basis of the internal validation criterion $q(ij)/Q(ij)$), the results are
 341 usually better than those obtained with BVS.

342 5. Charge Distribution Analysis of heteroligand polyhedra

343 Eq. (5) allows treating all the PC-V distances in a polyhedron at once; this however leads to neglecting
 344 the differences between the various V-atoms. In other words, it corresponds to replacing each V-atom with a
 345 dimensionless charged point. While this does make sense in the calculation of $\text{ECoN}(ij)$, which simply gives
 346 a measure of the number of coordinated V-atoms weighted by the respective bond distances, it would prevent
 347 any charge distribution analysis. In fact, although every V-atom retains its charge, the relation between the

348 bond distance and the strength of the bond would be lost. There is however one aspect in which the use of
349 FIR and MEFIR instead of the individual and average distances presents a clear advantage.

350 In a structure built on homoligand polyhedra, every PC(*ij*) is bonded to V(1*s*) for some *s* and the minimal
351 PC-V distance, $d(ij \rightarrow 1s)_1$, obviously corresponds to a chemical bond between the two atoms. Every PC-atom
352 distributes its charge to all the V-atoms bonded to it, in an inverse proportion to the bond length. On the other
353 hand, in a structure containing heteroligand polyhedra not all the PC(*ij*) are necessarily bonded to every
354 V(*rs*); this results in two types of problems.

355 1. $d(ij \rightarrow rs)_1$ for some *r* may correspond to a non-bonded contact. If the distances are used directly, then
356 $d(ij \rightarrow rs)_1$ is scaled to itself in the first step of the calculation of the average distance, $d(ij \rightarrow r)$, and to
357 avoid considering non-bonded contacts as chemical bonds one needs an external criterion, like the
358 sum of the radii scaled by an expansion factor to allow some tolerance. If FIR and MEFIR are used
359 instead, each polyhedron is treated as homoligand at this stage and the risk of counting non-bonded
360 contacts as chemical bonds is avoided. Concretely, in the analysis of the coordination polyhedra, a
361 first step is performed in terms of FIR and MEFIR, to discriminate between bonding and non-
362 bonding contacts; then, the calculation is repeated by using the distances to treat each coordination
363 sub-polyhedron separately.

364 2. Each PC-atom distributes a fraction of its charge to each homoligand subpolyhedron. This fraction
365 depends on the relative distances: V-atoms closer to the PC-atom receive a larger fraction than V-
366 atoms farther from it. The problem is how to define this fraction. This is straightforward for
367 structures which are built by homoligand polyhedra in one of the two possible descriptions (cation-
368 centred or anion-centred), while it requires a convergence algorithm when a structure is built by
369 heteroligand polyhedra in both descriptions.

370 The treatment slightly differs depending on whether the structure under investigation contains only
371 homoligand polyhedra in one description (either cation-centred or anion-centred) or not. Numerous examples
372 have been presented and discussed in Nespolo (2016), which we do not repeat here for lack of space.

373 5.1. Structures built by homoligand polyhedra in one description

374 If a structure contains only one chemical species of cation or of anion, the calculation is straightforward.
375 The single-type of element (cation or anion) is first assigned to V-atoms ($r = 1$) and the computation
376 described above is performed. Next, the fraction of $Q(ij)$ coming from each V-atom is calculated by a simple
377 modification of Eq. (13):

$$378 \quad \Delta Q(i \rightarrow 1s) = \sum_j \Delta q(ij \rightarrow 1s) \frac{q(1s) h(ij)}{Q(1s) h(1s)} = \frac{\sum_j \Delta \text{ECoN}(ij \rightarrow 1s) q(ij) q(1s) h(ij)}{Q(1s) h(1s)}. \quad (14)$$

379 The summation is here over *j*, *i.e.* the crystallographic type of PC-atom, instead than *r* and *s*. The result is
380 thus the fraction of the formal charge that each V-atom (identified by their crystallographic type *s*) shares
381 with the *i*-th chemical species of PC-atoms. The role of PC-atoms and V-atoms is then swapped so that Eq.

382 (14) gives in the next step the fraction that each PC-atom (V-atom in the previous step) shares with each
 383 chemical species of V-atom (PC-atom in the previous step):

$$384 \quad \Delta Q(1j \rightarrow r) = -\Delta Q(i \rightarrow 1s). \quad (15)$$

385 This corresponds to re-labelling the contribution ΔQ so that the only chemical species of V atom ($r = 1$)
 386 becomes the only chemical species of PC-atom ($i = 1$); the index on the crystallographic type is also
 387 relabelled from s to j . Finally, the index on the chemical species of PC-atoms (i) is relabelled as index on the
 388 chemical species of V-atoms (r). In the case of KNaTiO_3 discussed in section 3.1 as cation-centred
 389 homoligand cation-centred polyhedra, we had $ij = 11$: Ti, $ij = 21$: K, $ij = 31$: Na, $rs = 11$: O1 and $rs = 12$: O2.
 390 In the anion-centred description, the two oxygen-centred polyhedra become heteroligand; $ij = 11$: O1; $ij =$
 391 12 : O2; $rs = 11$: Ti; $rs = 21$: K; $rs = 31$: Na.

392 $\Delta Q(ij \rightarrow r)$ is then used instead of $q(ij)$ in Eq. (13), which is replaced by Eq. (16):

$$393 \quad Q(ij) = \sum_r \sum_s \Delta \text{ECoN}(ij \rightarrow rs) \frac{q(rs)}{Q(rs)} \Delta Q(ij \rightarrow r). \quad (16)$$

394 5.2. Structures built by heteroligand polyhedra in both descriptions

395 When a structure contains heteroligand polyhedra in both descriptions, we miss the starting point to
 396 compute Eq. (14). The problem is solved with an iterative procedure in which the starting value of $\Delta Q(ij \rightarrow r)$
 397 in Eq. (15) is assigned as a function of $\text{ECoN}(ij \rightarrow r)/\text{ECoN}(ij)$:

$$398 \quad \Delta Q(ij \rightarrow r) = q(ij) \frac{\text{ECoN}(ij \rightarrow r)}{\text{ECoN}(ij)} \quad (15')$$

399 Eq. (15') corresponds to Pauling's definition of bond strength, Eq. (1) applied to each sub-polyhedron.
 400 From the second step Eq. (14) and (15) can be used directly and the role of PC and V-atoms is swapped back
 401 and forth between cations and anions until the same values of $\Delta Q(ij \rightarrow r)$ are obtained in each description,
 402 meaning that convergence has been obtained.

403 6. Comparison of Charge Distribution and Bond Valence approaches

404 The concept of bond strength has its roots in the XIX century theory of valence and is based on the idea
 405 that the total bond valence received by each atom is equal to its atomic valence. Charge Distribution and
 406 Bond Valence are both expressions of strengths, but they differ in the way they describe a crystal structure.
 407 The expected outcomes of the two approaches also do not fully overlap, although both methods can be used
 408 for structure validation.

409 Bond Valence assigns a strength (valence) to each bond on the basis of some empirical parameters. These
 410 parameters are far from being established once for all and indeed in the literature these parameters are often
 411 re-refined or redefined. If a reliable and optimised set of parameters is available, then the Bond Valence
 412 approach is potentially able to predict the geometry of polyhedra and molecules in a crystal structure
 413 (Brown, 2014). However, the empirical parameters depend, in general, on the conditions under which the set

413 of crystal structures used to obtain these parameters have been solved and refined. As a matter of fact, bond
414 lengths do depend on the temperature, pressure, electromagnetic field or other external constraints to which
415 the structure is subjected; empirical parameters that have been optimized for ambient conditions are less
416 adapted to describe structures under conditions in which the bond lengths are significantly shrunk or
417 elongated. Furthermore, bond lengths are also affected by the spin state of the atoms: for example, the Fe-N
418 distance can change as much as 10% following a high-spin to low-spin transition (Legrand *et al.*, 2007).
419 However, the spin state does not normally enter in the definition of the Bond Valence parameters. Attempts
420 have been made to relate the Bond Valence parameters to electron density (Adams, 2014), although doubts
421 can be reasonably expressed on the ground on which simple empirical parameters can bear information
422 related to quantum mechanics. Finally, Bond Valence lacks an internal criterion to evaluate whether the
423 results of the analysis can be considered correct or not.

424 Charge Distribution fundamentally differs from Bond Valence in being essentially a geometric analysis of
425 the coordination polyhedra exploiting the observed distances only, rather than comparing them with some
426 ‘standard values’. Atoms are replaced by point charges in a Madelung-type scheme and the bond weight is a
427 purely geometric concept which quantitatively tells how strongly the presence of a V-atom at the corner of a
428 polyhedron is felt by the PC-atom at the centre of it. The bond weight is not directly related with the electron
429 density of the bond and the charge is distributed, rather than computed, among each bond. The equivalent of
430 the Bond Valence Sum strictly applies to computed charges of the PC-atoms, not to those of the V-atoms, the
431 two quantities bringing different information: a measure of the OUB effects for latter, the evaluation of the
432 applicability – including structure reliability – for the former.

433 Despite the Madelung-type approach, CHARDI has proven to be able to treat a wide range of structures
434 with a large chemical heterogeneity, although some types of compounds still remain outside its possibilities.
435 Apart from these rare exceptions, CHARDI can bring important information not always easily obtained from
436 experimental data:

- 437 • a structure validation based only on internal criteria, without the reference to empirical parameters;
- 438 • the presence of structural anomalies;
- 439 • the site-occupancy factor of incompletely occupied sites;
- 440 • the most likely oxidation state of atoms with multiple valences;
- 441 • the most likely oxidation state of an atomic site in compounds presenting isomorphic substitutions, which
442 gives a clear indication of the fractional occupancy by each type of atom: this information may be
443 difficult to obtain from conventional X-ray diffraction experiments in case of heterovalent substitutions
444 involving atoms of close atomic number or atoms with different valence states;
- 445 • the most likely positions of missing light atoms (like hydrogen), who are hidden by the presence of
446 relatively heavy atoms.

447 CHARDI does not possess the same predictive power as Bond Valence, because it cannot treat single
448 bond lengths separated from the others in the same coordination polyhedron. This difficulty could in

449 principle be overcome in the search for missing atoms by computing a CHARDI map reminding the Fourier
450 difference map routinely used in the refinement of crystal structures. Such a development is not available to
451 date and one may reasonably question whether the computational effort would be justified, considering the
452 simplified, Madelung-type description of the structure.

453 A computationally economical approach could use Bond Valence to predict the bond lengths between an
454 already located atoms and missing atoms. The likely positions could then be estimated from the expected
455 coordination geometry and the chemical good sense. This estimation can then be verified with CHARDI,
456 whose results are normally in better agreement with the expected charges than those provided by the BVS. A
457 few trial-and-error cycles can provide a reasonable estimation of the most likely atomic position.

458 **Figure captions**

459 **Figure 1.** A regular octahedron with six distances $d(ij \rightarrow rs)$ all equal to 2.000 Å (a) undergoing increasing
460 deformation of bond distances (b) to (d), while preserving the arithmetic average of 2.000 Å. The weighted
461 average distance decreases as does ECoN too. Values in the figures are bond distances in Ångströms. Figure
462 drawn with VESTA (Momma and Izumi, 2011).

463 **Figure 2.** Graphical representation of the Charge Distribution (CHARDI) and Bond Valence Sum (BVS)
464 for the structure of KTiNO_3 . ΔQ shows the departure from the ideal value of the computed charges (valences).
465 The higher departure calculated for BVS might suggest a more important over-under bonding effect, whereas
466 this actually comes from the limits of the method itself, as shown by the significantly lower departure
467 calculated for CHARDI.

468 **Figure 3.** Graphical representation of the Charge Distribution (CHARDI) and Bond Valence Sum (BVS)
469 for the structure of α' - Ba_2TiO_4 . Same interpretation as for Figure 1.

470 **References**

- 471 Adams, S. (2014). Practical Considerations in Determining Bond Valence Parameters. *Struct Bond.* **158**, 91–128.
472 Allman, R. (1975). Beziehungen zwischen Bindungslängen und Bindungsstärken in Oxidstrukturen. *Monatsh. Chem.*
473 **106**, 779-793.
474 Balić Žunić, T. & Makovicky, E. (1996). Determination of the centroid or 'the best centre' of a coordination polyhedron.
475 *Acta Crystallogr.* **B52**, 78-81.
476 Baur, W.H. (1970). Bond length variation and distorted coordination polyhedra in inorganic crystals. *Trans. Am. Cryst.*
477 *Ass.*, **6**, 129-155.
478 Baur, W.H. (1971). The prediction of bond length variations in silicon-oxygen bonds. *Am. Mineral.* **56**, 1573-1599.
479 Beck, H. P. (2014). The co-ordination number rule and the rule of hardness, powerful tools to rationalize inorganic
480 structures. *Z. Kristallogr.* **229**, 473–488.
481 Boisen, M.B., Gibbs, G.V. & Zhang, Z.G. (1988). Resonance bond numbers: A graph-theoretic study of bond length
482 variations in silicate crystals. *Phys. Chem. Miner.*, **15**, 409-415.

483 Borwein, D., Borwein, J.M. and Taylor, K.F. (1985). Convergence of lattice sums and Madelung's constant. *J. Math.*
484 *Phys.* **26**, 2999-3009.

485 Bosi, F. (2014a). Mean bond-length variation in crystal structures: a bond-valence approach. *Acta Crystallogr.* **B70**,
486 697-704.

487 Bosi, F. (2014b). Bond valence at mixed occupancy sites. I. Regular polyhedra. *Acta Crystallogr.* **B70**, 864-870.

488 Brown, I.D. (1977). Predicting bond lengths in inorganic crystals. *Acta Crystallogr.* **B33**, 1305-1310.

489 Brown, I. D. (1978). Bond valences—a simple structural model for inorganic chemistry. *Chem. Soc. Rev.* **7**, 359-376.

490 Brown, I.D. (1987). Recent developments in the bond valence model of inorganic bonding. *Phys. Chem. Mineral.*, **15**,
491 30-34.

492 Brown, I.D. (1988). What factors determine cation coordination numbers? *Acta Crystallogr.*, **B44**, 545-553.

493 Brown, I.D. (2014). Bond Valence Theory. *Struct. Bond.* **158**, 11–58.

494 Brown, I.D. & Altermatt, D. (1985). Bond-valence parameters obtained from a systematic analysis of the Inorganic
495 Crystal Structure Database. *Acta Crystallogr.*, **B41**, 244-247.

496 Brown, I.D, & Shannon, R.D. (1973). Empirical bond-strength–bond-length curves for oxides. *Acta Crystallogr.*, **A29**,
497 266-282.

498 Brown, I.D. & Wu, K.K. (1976). Empirical parameters for calculating cation–oxygen bond valences. *Acta Crystallogr.*,
499 **B32**, 1957-1959.

500 Burns, R.G. (1993). *Mineralogical Applications of Crystal Field Theory*. 2nd edition. Cambridge: Cambridge University
501 Press, xxiv+551 pp.

502 Chiari, G. (1988). Survey on the calcium oxygen coordination. *Z. Kristallogr.* **185**, 504.

503 Chiari, G. (1990). On metal–oxygen coordination. A statistical method to determine coordination number. I. Calcium.
504 *Acta Crystallogr.*, **B46**, 717-723.

505 Donnay, G. & Allman, R. (1970). How to recognize O²⁻, OH⁻, and H₂O in crystal structures determined by X-rays. *Am.*
506 *Mineral.*, **55**, 1003-1015.

507 Eon, J.-G. & Nespolo, M., (2015). Charge distribution as a tool to investigate structural details. III. Extension to
508 description in term of anion-centred polyhedra. *Acta Crystallogr.* **B71**, 34-47.

509 Ferraris, G. & Catti, M. (1973). Generalization of Baur's correlations between bond length and bond strength in
510 inorganic structures. *Acta Crystallogr.* **B29**, 2006-2009.

511 Gagné, O.C. & Hawthorne, F. C. (2015). Comprehensive derivation of bond-valence parameters for ion pairs involving
512 oxygen. *Acta Crystallogr.* **B71**, 562-578.

513 Gibbs, G.V., Hill, F.C., Boisen, M.B. & Downs, R.T. (1998). Power law relationships between bond length, bond
514 strength and electron density distributions. *Phys. Chem. Miner.*, **25**, 585-590.

515 Gopal, R. (1972). *The Crystal Chemistry And Bonding In Vanadates Of Divalent Metal Ions And The Crystal Structure*
516 *Of Whitlockite*. Ph. D. Thesis, McMaster University, x+238 p.

517 Guenter, J.R. & Jameson, G.B. (1984). Orthorhombic barium orthotitanate, α' -Ba₂TiO₄. *Acta Crystallogr.* **C40**, 207-
518 210.

519 Hoppe, R., (1970a). *Madelung constants as a new guide to the structural chemistry of solids*. In: *Advances in Fluorine*
520 *Chemistry*, Vol. 6, p. 387–438. London: Butterworths.

521 Hoppe, R. (1970b). The Coordination Number – an “Inorganic Chameleon”. *Angew. Chem. Int. Ed. Engl.* **9**, 25-34.

522 Hoppe, R. (1979). Effective coordination numbers (ECoN) and mean fictive ionic radii (MEFIR). *Z. Kristallogr.* **150**,
523 23-52.

524 Hoppe, R. (1995). On the Madelung Part of Lattice Energy. New Pathways to use it as a Tool in Solid State Chemistry
525 (Part 1). *Z. Naturforsch.* **50a**, 555-567.

526 Hoppe, R., Voigt, S., Glaum, H., Kissel, J., Müller, H.P. & Bernet, K. (1989). *J. Less-Comm. Met.* **156**, 105-122.

527 IUPAC (1997). *Compendium of Chemical Terminology, 2nd ed.*. Compiled by A. D. McNaught and A. Wilkinson.
528 Oxford: Blackwell Scientific Publications.

529 Jacucci G., Quirke N. (1984). *Free energy calculations for crystals*. In: Catlow C.R.A., Mackrodt W.C. (eds) *Computer*
530 *Simulation of Solids. Lecture Notes in Physics*, vol 166. Springer, Berlin, Heidelberg.

531 Knizek, K. (2017). KDist, version 4.31. part of the Kalvados software suite (<http://www.fzu.cz/~knizek/kalvados>).
532 Institute of Physics, ASCR, Praha, Czech Republic.

533 Krivovichev, S.V. (2009). *Structural Crystallography of Inorganic Oxysalts*. IUCr Monographs on Crystallography No.
534 22. IUCr/Oxford University Press.

535 Krivovichev, S.V., Mentré, O., Siidra, O.I., Colmont, M. & Filatov, S.K. (2013). Anion-Centered Tetrahedra in
536 Inorganic Compounds. *Chem. Rev.*, **113**, 6459-6535.

537 Larin, A.V. (2013). The Loewenstein rule: the increase in electron kinetic energy as the reason for instability of Al–O–
538 Al linkage in aluminosilicate zeolites. *Phys Chem Minerals* **40**, 771-780.

539 Legrand, V. Pillet, S., Weber, H.P., Souhassou, M., Létard, J.F., Guionneau, P. and Lecomte, C. (2007). On the precision
540 and accuracy of structural analysis of light-induced metastable states. *J. Appl. Crystallogr.* **40**, 1076-1088.

541 Loewenstein, W. (1954). The distribution of aluminum in the tetrahedra of silicates and aluminates. *Am Mineral.* **39**, 92-
542 96.

543 Makovicky, E. & Balić Žunić, T. (1998). New Measure of Distortion for Coordination Polyhedra. *Acta Crystallogr.*,
544 **B54**, 766-773.

545 Mohri, F. (2000). A new relation between bond valence and bond distance. *Acta Crystallogr.* **B56**, 626-638.

546 Momma, K. & Izumi, F. (2011). VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology
547 data. *J. Appl. Crystallogr.*, **44**, 1272-1276.

548 Naskar, J.P., Hati, S. & Datta, D. (1997). New Bond-Valence Sum Model. *Acta Crystallogr.*, **B53**, 885-894.

549 Nespolo, M. (2016). Charge distribution as a tool to investigate structural details. IV. A new route to heteroligand
550 polyhedra. *Acta Crystallogr.* **B72**, 51-66.

551 Nespolo, M., Aroyo, M.I. & Souvignier, B. (2018). Crystallographic shelves: space-group hierarchy explained. *J. Appl.*
552 *Crystallogr.* **51**, 1481-1491.

553 Nespolo, M., Ferraris, G. & Ohashi, H. (1999). Charge Distribution as a tool to investigate structural details: meaning
554 and application to pyroxenes. *Acta Crystallogr.* **B55**, 902-916.

555 Nespolo, M., Ferraris, G., Ivaldi, G. & Hoppe, R. (2001). Charge Distribution as a tool to investigate structural details.
556 II. Extension to hydrogen bonds, distorted and hetero-ligand polyhedra. *Acta Crystallogr.* **B57**, 652-664.

557 Nespolo, M. & Guillot, B. (2016). CHARDI2015: Charge Distribution analysis of non-molecular structures. *J. Appl.*
558 *Crystallogr.*, **49** 317-321

559 Pauling, L. (1929). *J. Am. Chem. Soc.* **51**, 1010-1026.

560 Pauling, L. (1960). *The Nature of the Chemical Bond*, 3rd edition. Oxford: Oxford University Press.

- 561 Pearson, R.G. (1968). Hard and soft acids and bases, HSAB, part 1: Fundamental principles. *J. Chem. Educ.* **45**, 581-
562 586.
- 563 Pendás, A.M., Costales A. and Luaña V. (1998). *J. Phys. Chem. B.* **102**, 6937-6948.
- 564 Pyatenko, Yu. A. (1973). *Sov. Phys. Crystallogr.*, **17**, 677-682.
- 565 Rutherford, J.S. (1991). Graph theory applied to bond length variations in silicate minerals. *Trans. Am. Cryst. Soc.*, **27**,
566 315-321.
- 567 Rutherford, J.S. (1998). Theoretical Prediction of Bond-Valence Networks. II. Comparison of the Graph-Matrix and
568 Resonance-Bond Approaches. *Acta Crystallogr.*, **B54**, 204-210.
- 569 Trömel, M. (1983). Empirische Beziehungen zu den Bindungslängen in Oxiden. 1. Die Nebengruppenelemente Titan
570 Bis Eisen. *Acta Crystallogr.* **B39**, 664-669.
- 571 Trömel, M. (1984). Empirische Beziehungen zu den Bindungslängen in Oxiden. 2. Leichtere Hauptgruppenelemente
572 sowie Kobalt, Nickel und Kupfer. *Acta Crystallogr.* **B40**, 338-342.
- 573 Trömel, M. (1986). Empirische Beziehungen zu den Bindungslängen in Oxiden. 3. Die offenen Koordinationen um Sn,
574 Sb, Te, I und Xe in deren niederen Oxidationsstufen. *Acta Crystallogr.* **B42**, 138-141.
- 575 Umayahara, A., Nespolo, M. (2018). Derivative structures based on the sphere packing. *Z. Kristallogr.* **233**, 179-203.
- 576 Urusov, V.S. (1995). Semi-empirical groundwork of the bond-valence model. *Acta Crystallogr.*, **B51**, 641-649.
- 577 Vasileiadis, M. (2013). Calculation of the free energy of crystalline solids. Ph.D. thesis, Imperial College, London,
578 xii+202 pp.
- 579 Wang, X & Liebau, F. (1996a). Studies on bond and atomic valences. I. correlation between bond valence and bond
580 angles in Sb^{III} chalcogen compounds: the influence of lone-electron pairs. *Acta Crystallogr.*, **B52**, 7-15.
- 581 Wang, X & Liebau, F. (1996b). Effective coordination numbers (ECoN) and mean fictive ionic radii (MEFIR). *Z.*
582 *Kristallogr.*, **211**, 437-439.
- 583 Werthmann, R. and Hoppe, R. (1985). Ein neues einfaches Titanat: KNaTiO₃. *Z. Anorg. Allg. Chem.*, **523** 54-62.

Alphabetical Index

bond strength.....	2 ff., 6 f., 10, 13, 16
Bond strength.....	3
Bond Strength.....	1
bond valence.....	3 f., 13, 16 ff.
Bond valence.....	16
Bond Valence.....	6, 11, 13 ff.
bond weight.....	7 ff., 14
charge distribution.....	4, 11
Charge distribution.....	16 f.
Charge Distribution.....	1, 3, 10 f., 13 ff., 17
coordination polyhedron.....	1 f., 8, 10, 15
ECoN.....	4 ff., 13, 15, 17 f.
Effective coordination number.....	17 f.
Effective Coordination Number.....	4, 6
Fictive Ionic Radius.....	4 f., 17 f.
FIR.....	4 f.

heteroligand.....	2 ff., 8, 11 ff., 17
homoligand.....	2 f., 8, 10, 12 f.
MEan Fictive Ionic Radius.....	4 f.
MEFIR.....	4 f.
resonance bond number.....	4
Resonance bond number.....	15

Table 1. Simulation of the progressive distortion of a regular octahedron and its effect of the weighted average distance ${}^n d(ij \rightarrow r)$ and effective coordination number ${}^n \text{ECoN}(ij \rightarrow r)$.

	Regular octahedron: Figure 1(a)			Distortion 1: Figure 1(b)			Distortion 2: Figure 1(c)			Distortion 3: Figure 1(d)		
	$d(ij \rightarrow rs)_L$	$\frac{d(ij - rs)_L}{d(ij - rs)_1}$	${}^n w(ij \rightarrow rs)_L$	$d(ij \rightarrow rs)_L$	$\frac{d(ij - rs)_L}{d(ij - rs)_1}$	${}^n w(ij \rightarrow rs)_L$	$d(ij \rightarrow rs)_L$	$\frac{d(ij - rs)_L}{d(ij - rs)_1}$	${}^n w(ij \rightarrow rs)_L$	$d(ij \rightarrow rs)_L$	$\frac{d(ij - rs)_L}{d(ij - rs)_1}$	${}^n w(ij \rightarrow rs)_L$
$d(ij \rightarrow rs)_1$	2.000	1.000	1.000	1.940	1.000	1.166	1.890	1.000	1.278	1.800	1.000	1.355
$d(ij \rightarrow rs)_2$	2.000	1.000	1.000	1.970	1.015	1.075	1.940	1.026	1.125	1.870	1.039	1.133
$d(ij \rightarrow rs)_3$	2.000	1.000	1.000	1.990	1.026	1.014	1.970	1.042	1.033	1.920	1.067	0.975
$d(ij \rightarrow rs)_4$	2.000	1.000	1.000	2.000	1.031	0.984	2.020	1.069	0.883	1.990	1.106	0.763
$d(ij \rightarrow rs)_5$	2.000	1.000	1.000	2.040	1.052	0.866	2.070	1.095	0.739	2.180	1.211	0.302
$d(ij \rightarrow rs)_6$	2.000	1.000	1.000	2.060	1.062	0.808	2.110	1.116	0.631	2.240	1.244	0.205
${}^n d(ij \rightarrow r)$	2.000			1.995			1.918			1.912		
${}^n \text{ECoN}(ij \rightarrow r)$	6.00			5.91			5.69			4.73		

Table 2. Calculation of ECoN and the parameters contributing to it for the structure of KNaTiO_3 (structural data from Werthmann and Hoppe, 1985). h is the multiplicity of the Wyckoff position.

PC	h	V	h	$d(ij \rightarrow rs)$	FIR	${}^n \text{MEFIR}$	${}^n d(ij \rightarrow r)$	${}^n w$	${}^n \text{ECoN}(ij \rightarrow r)$
Ti1	4	O1	4	1.7570	0.556	0.596	1.887	1.416	4.56
		O2	8	2×1.9098	0.604			0.927	
		O2	8	2×2.0034	0.634			0.648	
K1	4	O1	4	2×2.7924	1.662	1.760	2.959	1.341	8.04
		O1	4	2×2.9093	1.732			1.101	
		O2	8	2×3.0757	1.831			0.770	
		O2	8	2×3.1244	1.860			0.679	
		O2	8	2×3.6514	2.174			0.079	

		O2	8	2×3.7320	2.222			0.048	
Na1	4	O1	4	2.2963	1.282	1.295	2.318	1.057	4.99
		O2	8	2×2.3136	1.292			1.012	
		O2	8	2×2.3355	1.304			0.956	

Table 3. Calculation of ECoN and the parameters contributing to it for the structure of α' -Ba₂TiO₄ (structural data from Guenter and Jameson, 1984).

PC	V		$d(ij \rightarrow rs)$	FIR	ⁿ MEFIR	ⁿ d($ij \rightarrow r$)	ⁿ w	ⁿ ECoN($ij \rightarrow r$)	
Ba1	4	O3	4	2.5393	1.401	1.509	2.736	1.434	5.54
		O12	4	2.7154	1.498			1.045	
		O10	4	2.7240	1.503			1.026	
		O9	4	2.7587	1.522			0.950	
		O8	4	2.9594	1.633			0.548	
		O5	4	3.0263	1.670			0.435	
		O12	4	3.3986	1.876			0.069	
		O9	4	3.4941	1.928			0.035	
Ba2	4	O7	4	2.5718	1.419	1.564	2.837	1.561	6.16
		O8	4	2.7522	1.519			1.181	
		O5	4	2.7872	1.538			1.106	
		O12	4	3.0013	1.656			0.669	
		O2	4	3.0071	1.659			0.659	
		O10	4	3.1208	1.722			0.462	
		O1	4	3.3184	1.831			0.210	
		O8	4	3.3600	1.854			0.172	
		O6	4	3.4019	1.877			0.139	

Ba3	4	O11	4	2.5389	1.401	1.563	2.836	1.624	6.26
		O6	4	2.7956	1.543			1.085	
		O1	4	2.8319	1.563			1.008	
		O4	4	2.9213	1.612			0.822	
		O4	4	2.9579	1.632			0.750	
		O4	4	3.1532	1.740			0.411	
		O2	4	3.3056	1.824			0.221	
		O2	4	3.3578	1.853			0.173	
		O1	4	3.4173	1.886			0.127	
		O5	4	3.6230	1.999			0.035	
Ba4	4	O10	4	2.6305	1.452	1.529	2.772	1.309	7.43
		O9	4	2.7123	1.497			1.130	
		O2	4	2.7473	1.516			1.053	
		O11	4	2.7772	1.533			0.988	
		O1	4	2.7878	1.538			0.966	
		O8	4	2.8327	1.563			0.870	
		O12	4	2.9010	1.601			0.730	
		O3	4	3.1039	1.713			0.379	
Ba5	4	O2	4	2.6350	1.454	1.523	2.760	1.274	7.69
		O1	4	2.6857	1.482			1.163	
		O6	4	2.7063	1.493			1.117	
		O7	4	2.7870	1.538			0.941	
		O4	4	2.7995	1.545			0.914	
		O7	4	2.8391	1.567			0.831	

		O8	4	2.8751	1.587			0.757	
		O5	4	2.9076	1.605			0.693	
Ba6	4	O6	4	2.6684	1.473	1.530	2.774	1.230	7.69
		O9	4	2.6762	1.477			1.213	
		O5	4	2.6841	1.481			1.196	
		O3	4	2.7682	1.528			1.012	
		O12	4	2.8298	1.562			0.880	
		O4	4	2.9132	1.608			0.710	
		O10	4	2.9146	1.608			0.707	
		O11	4	3.0045	1.658			0.540	
Ti1	4	O1	4	1.7884	0.566	0.570	1.802	1.045	3.99
		O3	4	1.7895	0.566			1.042	
		O2	4	1.8152	0.574			0.956	
		O4	4	1.8173	0.575			0.949	
Ti2	4	O7	4	1.7994	0.569	0.575	1.818	1.062	3.99
		O6	4	1.8183	0.575			1.000	
		O5	4	1.8212	0.576			0.990	
		O8	4	1.8360	0.581			0.941	
Ti3	4	O11	4	1.7970	0.569	0.575	1.819	1.073	3.99
		O10	4	1.8251	0.577			0.979	
		O9	4	1.8252	0.577			0.979	
		O12	4	1.8307	0.579			0.961	

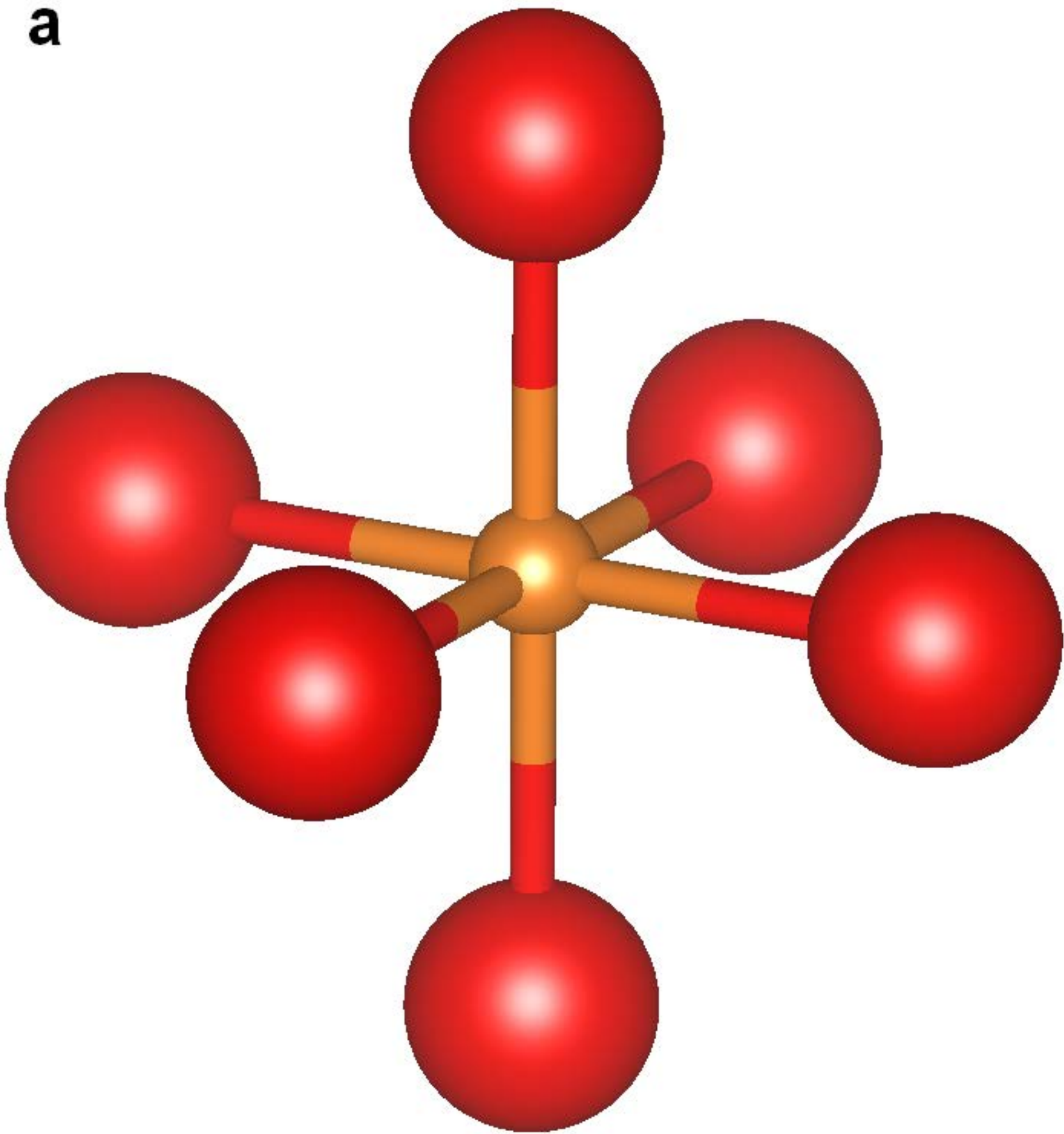
Table 4. Charge Distribution (CHARDI) and Bond Valence Sum (BVS) for the structure of KTiNO_3 . CHARDI computed with CHARDI2015 (Nespolo and Guillot, 2016); BVS computed with Kdist (Knizek, 2017).

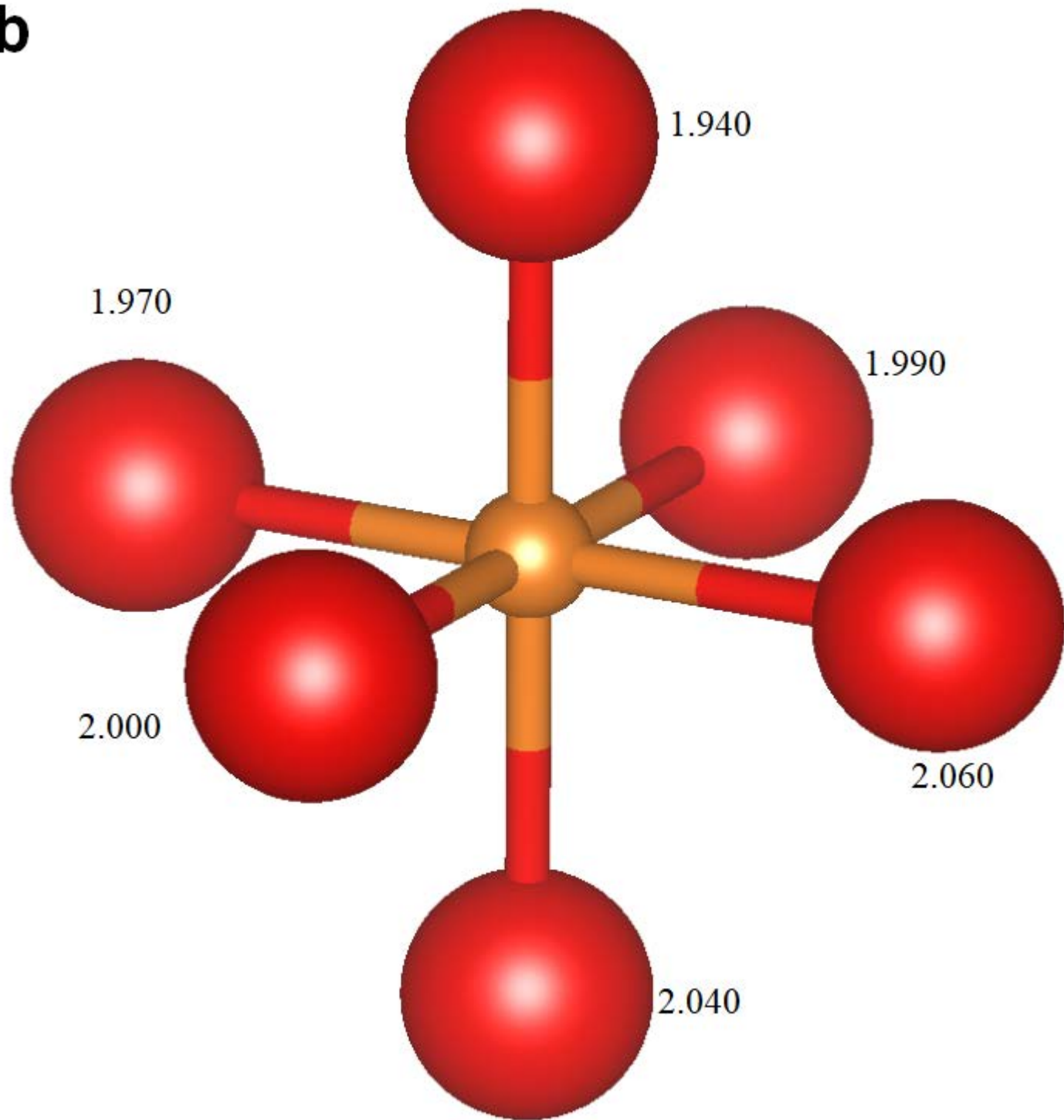
	CHARDI	BVS
Ti	4.01	3.92
K	0.99	0.83
Na	1.01	1.09
O1	-2.06	-1.95
O2	-1.97	-1.94

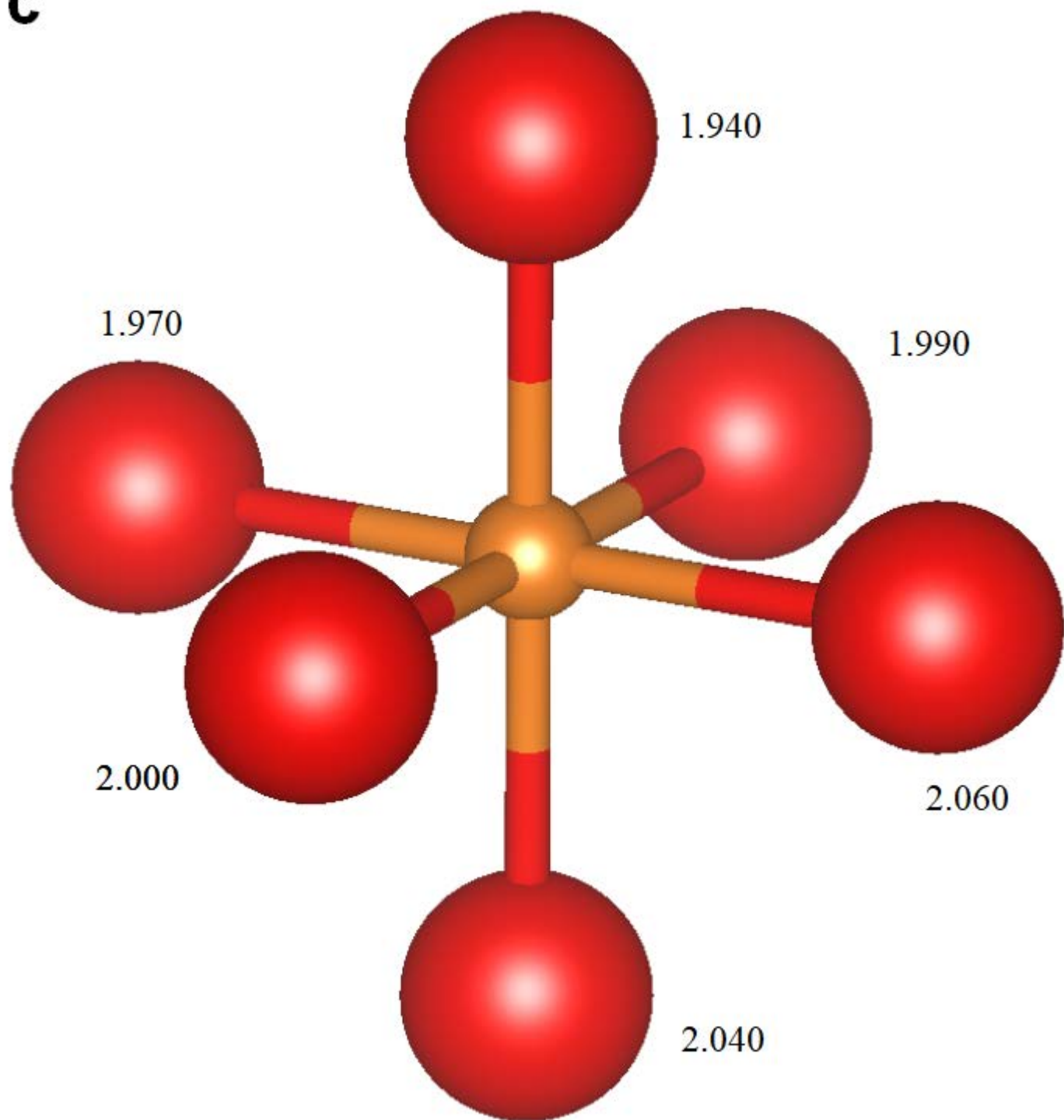
Table 5. Charge Distribution (CHARDI) and Bond Valence Sum (BVS) for the structure of α' - Ba_2TiO_4 . CHARDI computed with CHARDI2015 (Nespolo and Guillot, 2016); BVS computed with Kdist (Knizek, 2017).

	CHARDI	BVS		CHARDI	BVS
Ba1	2.01	1.78	O1	-2.04	-2.01
Ba2	1.99	1.56	O2	-1.91	-1.93
Ba3	1.99	1.59	O3	-1.93	-1.97
Ba4	2.01	2.04	O4	-2.01	-1.86
Ba5	2.00	2.17	O5	-2.02	-1.90
Ba6	2.00	2.05	O6	-2.01	-1.97
Ti1	4.05	4.14	O7	-2.03	-2.00
Ti2	3.96	3.98	O8	-2.01	-1.87
Ti3	3.99	3.97	O9	-1.96	-1.95
			O10	-2.04	-1.96
			O11	-2.00	-1.96
			O12	-2.01	-1.88

a



b

C

d

