Raphael J.F. Berger*, Pär Håkansson and Raúl Mera-Adasme*

A consistent model for the key complex in chronic beryllium disease

https://doi.org/10.1515/znb-2020-0010
Received January 24, 2020; accepted March 8, 2020

Abstract: A hypothesis on the structure of the key complex in chronic beryllium disease (CBD) is discussed with respect to the current knowledge on CBD, and with respect to the constraints implied by the coordination chemistry of beryllium and experimental data on the engaged protein complexes. The structure hypothesis is based on the \([\text{Be}_4\text{O}]^{6+}\) moiety as a coordination center, which is also found in the so called “basic beryllium carboxylates”. The structure of a small molecular model, optimized at the DFT level of theory, is used to compare the structural demands of this coordination center with a structure of the in vitro model of a beryllium immunoprotein complex determined previously by protein crystallography (Clayton & al., Cell 2014, 158, 132). \(^{9}\text{Be}\) NMR chemical shielding values, quadrupole coupling constants and asymmetry parameters (\(\eta\)) have been calculated.

Keywords: beryllium disease; protein structure \(4p4k\); toxicity of beryllium.

1 Introduction

Many beryllium-based compounds and materials exhibit unique properties [1, 2]. For instance the long sought-for beryllium analog of MOF-5 [3–5] shows a superior hydrogen absorption capacity [6] or a beryllium based rechargeable battery is estimated to offer both the highest volume- and gravimetric energy densities known so far [7].

Beryllium is not rare and it appears that the only major obstacle for a much more wide-spread use of the element are the severe health adverse effects [8] an incorporation can cause in humans. The two most common conditions are the so called “acute beryllium disease” and the “chronic beryllium disease” (CBD). Especially much progress has been achieved in the recent decades concerning the investigation of the bio-chemical mechanism of the latter [9, 10].

CBD is mainly known as an occupational lung disease, that resembles sarcoidosis and auto-immune disorders in its symptoms [11, 12] and is also linked with lung cancer [13].

Interestingly, despite the recent general progress on the inorganic and bio-inorganic chemistry of beryllium [14–16] and despite various suggested model mechanisms for metal hyper-sensitivity [17] and model structures of the key complex of CBD [12, 18] the molecular mechanism of CBD is still obscure and the molecular structure of the key complexes of beryllium with the engaged immuno-proteins remains unknown [12, 19]. Summarizing the current knowledge of the bio-chemical species and mechanisms involved in CBD is beyond the scope of this work, and we instead refer to the literature for this [10].

In the course of the investigation of the glutamyl-glutamate bispeptide as a ligand superior to monofunctional carboxylic acids in the so called “basic beryllium carboxylates” we have suggested the \([\text{Be}_4\text{O}]^{6+}\) core of such complexes as a model for the beryllium species that is engaged in CBD [19]. The aim of this work is a detailed elaboration on this model and the computational prediction of some of its properties.

2 Methods

2.1 Model description

Our model for the key complex of CBD (1) is derived from a model of Clayton et al. (2) [12], which essentially is a ternary complex with the following components

- \(\text{DP}2\): a protein from the class II of the major histocompatibility complexes (MHC-II), the HLA-DP2 (human leucocyte antigen of type DP2), more specifically its \(\beta\)-chain subunit (that is sometimes also written as HLA-DP2\(\beta1\)) in specific HLA-DPB1 alleles. The (various) HLA-DPB1 alleles that have glutamic acid (E) residues at positions \(\beta26E, \beta68E\) and \(\beta69E\).
– M2: a mimotope peptide called M2 of sequence FWILDLETIG with aspartic acid (D) at position 4 (p4D) and another glutamic acid at position 7 (p7E)

– AV22: a CD4+ T-helper cell antigen receptor protein (TCR-AV22)

– a Be2+ and a Na+ cation

The refined single crystal X-ray structure 2 (PDB-ID: 4p4k) [12] consists (apart from solvent, water and co-crystallizing agents) mainly of V-shaped and C2-symmetric dimeric units of complexes containing HLA-DP2β1, HLA-DP2α1, M2, AV22, Na+ and Be2+. Clayton et al. could confirm that the binding site (S, see Fig. 1a and b) of the metal cations agrees with what previously has been deduced from mutant and allele variant studies of both the DP2 [20–23] and bound (self)-peptides [20], and it shows a close spatial proximity and possibly chemically bonding interactions between the metal cations and the five previously mentioned carboxylate moieties β26E, β68E, β69E in the DP2 complex and p4D, and p7E in the M2 peptide and in addition with the carbonyl group of the Leu residue at position 5 (p5L). A section of the structure model 2 from two perspectives is shown in Fig. 1a and b.

The five carboxylate groups in S are arranged in two roughly perpendicular arrays, the one consisting of those from β26E, β68E, β69E of the DP2 and the other of those of p4E, and p7E of the mimotope peptide M2. A diagram that schematically illustrates this specific geometric arrangement is given in Fig. 1c which can be compared with the X-ray structure model shown in Fig. 1a.

Our new model for S [19] contains a [Be4O]6+ complex ion as it is well known from the so called “basic beryllium carboxylate” compounds [3] as the coordination center (see Fig. 1d) binding the M2 to the DP2 protein. Its structure can be described as an oxide centered tetrahedron of Be2+ cations which is coordinated by six carboxylate ligands in an η2-mode, i.e. all six edges of the Be2+ tetrahedron are each bridged by one carboxylate group, thus total electro-neutrality and at the same time the optimal tetrahedral arrangement of oxygen around Be2+ is achieved. We note that also one example of a non-homoleptic derivative of a basic beryllium carboxylate has been described in the

![Fig. 1](image-url)

**Fig. 1**: The beryllium binding site S in the key complex of CBD. (a) Binding site S in the X-ray structure model of 4p4k (2) [12], first perspective, with the three carboxylate moieties of DP2 arranged vertically on the left and the two carboxylate moieties of M2 arranged horizontally and in top view. In the center is the original Na+ + Be2+ model. (b) Binding site S in the X-ray structure model of 4p4k (2) [12], second perspective. (c) Schematic illustration of the bis-meridional binding site S in the metal free coordination pocket of the DP2 · M2 complex, roughly from the same perspective as in Fig. 1a. (d) Schematic illustration of the binding site S in our new model 1 containing the Be4O6+ moiety, roughly from the same perspective as in Fig. 1a.
literature [24], which is a product of the formal substitution $[\text{RCO}_2]^{-} \rightarrow [\text{OH}]^{-}$ of composition $\text{Be}_4\text{O}(\text{OH})_5$.

2.2 DFT structure optimizations

The structures of 3 (see Fig. 2) and $\text{Be}_4\text{O}(\text{OAc})_6$ have been optimized using Turbomole [25, 26] (version 7.2) at the DFT level of theory using the PBE functional [27, 28], Grimmes D3 dispersion correction together with the Becke-Johnson damping [29–31] and the RI approximation for the Coulomb integrals [25]. The numerical quadrature grid M4 [32] for the exchange-correlation contribution and the def2-TZVP [33] basis set was used. Prior to the chemical shielding calculations with Turbomoles mpshift module [34], single point calculations with stricter convergence criteria (denconv 0.1d-07 and scfconv 8) have been performed.

The Hessian and vibrational frequencies were calculated with Turbomoles aoforce module [35], in order to confirm the structure as a minimum on the potential energy surface.

2.3 $^9\text{Be}$ quadrupolar coupling and linewidth

In this section the quantum chemistry details of nuclear quadrupole coupling (NQC) are provided, followed by a discussion of half width of the $^9\text{Be}$ NMR signals.

As a starting point for estimating nuclear quadrupole coupling (NQC), we consider the quadrupole coupling constants ($C_Q$) and asymmetry parameters ($\eta$):

$$C_Q = \frac{e^2 Q V_{zz}}{h},$$

$$\eta = \frac{V_{xx} - V_{yy}}{V_{zz}},$$

where the absolute electron charge is denoted $e$, the nuclear quadrupole moment $Q$, the Planck constant $h$ and the eigenvalues of the electric field gradient tensor $V_{aa}$, $aa = xx, yy, zz$, respectively. We provide NQC for the models 3, the basic beryllium formiate: $\text{Be}_4\text{O}(\text{HCO}_2)_6$ (4) and 4 where one formiate is replaced by formaldehyde as a minimal model for the carbonyl group of the p7L residue of the self peptide (M2): $[\text{Be}_4\text{O}(\text{HCO}_2)_5\text{OCH}_2]^+$ (5) in Table 2.

2.3.1 NQC with quantum chemistry

The NQC is given by the ground state single configuration expectation value of the electric field gradient [36]. Possible challenges in computing NQC with quantum chemistry may come from the need to consider relativistic effects [36] and finding a level of theory that captures electron-electron correlation to sufficient accuracy. In addition, when considering relaxation theory there is the large computational cost if explicit dynamics is considered [37]. In this study we do not consider relativistic effects, since these are not expected to dominate for the low mass $^9\text{Be}$ atom. The optimized structures (see Section 2.2) of models 3–5 in vacuum enter the NQC calculations. The NQC calculations are performed at the DFT level (functional B3PW91) with self consistent field gradient norm less than $10^{-5}$ and Dalton software [38].

To explore sensitivity to basis set size the smallest and largest $C_Q$ for model 5 [see eqs. (1) and (2)] are provided in Table 1, with Pople-style and Dunning’s

![Fig. 2: DFT (PBE0 + D3/def2-TZVP) optimized structure of the minimal $[\text{Be}_4\text{O}(\text{EGGGE})(\text{DLFE})]^+$ model (3) for the coordination site S in 1, hydrogen atoms omitted for clarity and labels corresponding to the respective positions in 1, see also Fig. 1d).](image-url)
correlation-consistent basis sets [38]. It is concluded that 6-31G* is sufficient for the sought model comparison.

In Table 2 the NQC values for the three models are provided. In particular 
CQ shows an exceptionally large difference between model 4 and 5, in the range of factors 4–19, resulting in a distinct difference in NMR relaxation for these two models.

2.3.2 9Be NMR linewidth and relaxation

To discuss the scope of discriminating between proposed models 3–5 with NMR linewidth and relaxation studies we provide the transverse relaxation rate [39]:

\[
\frac{1}{T_2} \propto C_Q [9 J_0(0) + 15 J_1(\omega_0) + 6 J_2(2\omega_0)],
\]

that in addition to the square of \( C_Q \), depends on the spectral density \( f(n\omega) \) computed in the laboratory frame, where \( \omega_0 \) is the 9Be resonance frequency. Although this is an approximate form and in general has a dependence on \( \eta \) as well it is sufficient to make some key points. Note (i) we have a linewidth of the 9Be quadrupole spectra that is directly proportional to \( 1/T_2 \) as well as experimental relaxation observables that all depend on the square of \( C_Q \). Secondly (ii), extracting properties via relaxation and line-shape studies require modelling of the spectral density \( f(n\omega) \). Hence, in the case of testing the structure models 3–5 in a protein framework and in aqueous solution we require a model how the EFG undergoes dynamics due to local bonds and the structure around the Be atoms as well as how the whole protein reorient in solution to allow a computation model of the relaxation rates [39]. Fortunately, modelling of relaxation in complex materials has progressed [37, 40], involving cost effective methods for the required quantum chemistry EFG-dynamics combined with methods for important slower dynamics. Modelling along this line, combined with collection of experimental data is ongoing to provide both a test of best candidate structural model as well as providing information on the local Be dynamics and the dynamics of the protein as a whole.

3 Discussion

The initial heuristic basis for our model was the close spatial accumulation of five carboxylate donor functions in a comparably small region (S) and also the occurrence of the glutamyl-glutamate motif (EE) in 2 [19], the latter of which we had shown to be able to act as a chelating ligand providing superior stability to \( \text{Be}_4\text{O}[(\text{O}_2\text{CR})_6]_n \) type compounds as compared to such with non-chelating mono-carboxylate ligands [\text{RCO}_2]^{-} [19]. This hypothesis on the structure of 1 (see Section 2.1 and Fig. 1 for a detailed description) shall be tested against the following requirements: It shall

A. match the binding site S in terms of the structural and chemical constraints of the coordination chemistry of beryllium
B. due to the uniqueness of CBD explain the chemical uniqueness of beryllium with respect to this coordination mode
C. be in line with the previous knowledge about and lead to a consistent explanation and prediction of the specific immunological and biochemical mechanisms involved in CBD
D. explain experimentally obtained analytical data

3.1 A.

The binding site contains five singly negatively charged carboxylate groups in close spatial proximity. As was previously noted [12] this accumulation of negative charges poses a problem for a model containing a mononuclear Be-complex, since the \( \text{Be}^{2+} \)-cation has a formal electronic charge of only 2+, while the \( \text{Be}_4\text{O}^{2+} \) coordination center gives an almost perfect match.

If the carbonyl moiety of the p7L residue is taken as an additional donor into consideration (see also Fig. 1c and d and Fig. 3d in ref. [12]), topologically, both the M2– and DP2–ligands provide a meridional arrangement around the coordination site. This bis-meridional arrangement circumscribes (thus requires) an octahedral coordination center, which \( \text{Be}_4\text{O}^{2+} \) indeed is.

<table>
<thead>
<tr>
<th>Model</th>
<th>( C_Q (1) )</th>
<th>( \eta (1) )</th>
<th>( C_Q (2) )</th>
<th>( \eta (2) )</th>
<th>( C_Q (3) )</th>
<th>( \eta (3) )</th>
<th>( C_Q (4) )</th>
<th>( \eta (4) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>3</td>
<td>-0.303</td>
<td>0.905</td>
<td>0.239</td>
<td>0.992</td>
<td>1.567</td>
<td>0.210</td>
<td>0.259</td>
<td>0.563</td>
</tr>
<tr>
<td>4</td>
<td>-0.090</td>
<td>0.450</td>
<td>0.073</td>
<td>0.293</td>
<td>-0.081</td>
<td>0.625</td>
<td>-0.0876</td>
<td>0.535</td>
</tr>
<tr>
<td>5</td>
<td>0.367</td>
<td>0.764</td>
<td>0.366</td>
<td>0.772</td>
<td>1.600</td>
<td>0.197</td>
<td>0.591</td>
<td>0.503</td>
</tr>
</tbody>
</table>
The size of the coordination site in 2 can be estimated from the distance between trans-positioned carbon atoms of the six donor functions, i.e. \( C^{\beta 5-\beta 7} = 5.98 \, \text{Å} \), \( C^{\beta 4-\beta 6} = 6.58 \, \text{Å} \), and \( C^{\beta 2-\beta 6} = 9.06 \, \text{Å} \), respectively. A Comparison with the values from a DFT optimization for a simplified model 3 (vide infra) which yields 5.98 Å, 6.88 Å and 6.17 Å, respectively, shows that the steric requirements for the [Be₄O]⁶⁺ moiety are very well met by the coordination site S. The larger mismatch of the value for \( C^{\beta 2-\beta 6} \) is discussed under D.

The [Be₄O]⁶⁺ moiety is composed of four doubly charged cations but still matches the relatively small region of the coordination site S. This can be explained by the short average distance of about 1.6 Å between Be and the central oxygen atom. Thus effectively, the [Be₄O]⁶⁺ coordination center is only about the 1.5-fold size of a methane molecule (CH₄) and the [Be₄O]⁶⁺ moiety meets almost perfectly the electrostatic, topologic and spatial requirements of the coordination site S.

### 3.2 B.

To the best of our knowledge, the only other metal (M) except beryllium, that is known to form stable, isolatable compounds of composition M₄O(RCO₂)₆ is zinc.¹ The two most significant differences between Zn₄O(RCO₂)₆ and Be₄O(RCO₂)₆ type compounds, are firstly, that the former are soluble in water while the latter are insoluble in any aqueous media including strong acids and bases. In aqueous media Be₄O(RCO₂)₆ species are not accessible, they in fact are prepared usually by solid state reactions and sublimation or at least under phase-transfer conditions [19]. The reason for this markedly different behavior of Zn and Be can be explained by the inability of the small Be²⁺ cation to extend its coordination number to numbers larger than four, unlike the larger and electronically more flexible Zn²⁺. Hence a hydrolysis of the Be species is kinetically hindered, or in fact impossible. Secondly, the Zn–O distance in Zn₄O(RCO₂)₆ is with about 1.9 Å, significantly larger than the corresponding Be–O distance of about 1.6 Å. For that it seems likely that the formation of a [Zn₄O]⁶⁺ at the coordination site S is unfavorable or least kinetically strongly hindered. However, if such a complex eventually is formed it is likely to undergo hydrolysis, dissociation or be present only in low concentration in chemical equilibrium, while the formation of the [Be₄O]⁶⁺/M₂/DP₂ complex is more likely to be practically irreversible. The same arguments as for zinc apply for any other metal cation of charge 2⁺, like those of Mg, Mn, Co, Fe or Ni. In this sense the M₄O(RCO₂)₆ type compounds are both by structure and reactivity unique for beryllium among all other elements.

### 3.3 C.

The composition of a [Be₄O]⁶⁺/M₂/DP₂ complex (1) suggests that its formation under physiological conditions should be a rather slow, rarely occurring process, since four Be²⁺ cations have to accumulate in the small coordination site S and also since an oxide dianion has to be formed. This very general expectation would meet the fact that CDB has long, and partially very long latency times.

On the other hand its not implausible, that eventually such complexes will form: We suggest that the mechanism could involve a step-wise transport of one single or two Be²⁺ cations via the peptide inside the lysosome to the MHC-II. The M₂ peptide with its three donor functions is a plausible chelating ligand for Be²⁺ cations at the position between p4D and the p5L carboxyl oxygen, (or between p7E and one of the former). The Be²⁺-loaded self-peptide can then act as “vector” transporting Be²⁺ to the β2E⁻, β6E⁻, and β6E⁻ sites of the DP₂. The higher negative charge accumulation at the DP₂ (3⁻ as compared to 2⁻ in M₂) can lead to a transfer of Be²⁺ from the M₂ to the DP₂. If more Be²⁺ is available to further peptide molecules inside the lysosome this process can take place again and can lead to an accumulation of possibly up to four Be²⁺ cations at coordination site S.

However, the irreversibility (vide supra) of its formation implies severe consequences for the immunological mechanisms involved. The self peptide, that is distorted by the coordination of [Be₄O]⁶⁺ cannot be removed by CD4-T cells, which could trigger the emergence of CD4-T cells that are adapted to distorted self-peptides, explaining the auto-immuno symptomatics of CBD. Our model in addition explains the enhanced interleukin-2 production of T-cell hybridomas when plexin self-peptides are presented (together with Be), that contain three acidic residues (D, D, E) [23]. In this way the ideal number of six carboxylate ligands is yielded in the complex at site S. This trend is also reflected in the high thermal stability of the Be/plexin A4/DP2/AV22 complex [12].

### 3.4 D.

However plausible 1 is as a model, the structure factors of 4p4k are inconsistent with it [41, 42]. In particular the proposed conformation of the β6E⁻ carboxylate group is in contradiction with the steric requirements for the [Be₄O]⁶⁺

---

¹ Manganese and iron are known to form compounds that contain the M₄O(RCO₂)₆ moiety, but in addition other ligands that are coordinating the respective metal ions.
coordination center. This is also reflected in the $C^{\text{BeO}}-C^{\text{OAc}}$ distance of 9.06 Å, which is about 3 Å larger than expected for the new model (while the other two C–C distances are matching very well). It is known, that the $\beta 69$E carboxylate group can be engaged in binding the AV22 protein of the CD4 T-cell. Hence, this carboxylate group is pointing towards the AV22 site in the protein complex which is ultimately responsible for the $\text{in vivo}$ development of the chronic beryllium disease (CBD). This hypotheses is motivated by the high charge accumulation found experimentally for the binding site $S$ of a CBD model complex [12]. The hypotheses is backed by the topology and geometric features found for $S$, it is backed by the fact that this syndrome is specific for beryllium by the specific chemistry of $[\text{Be}_4\text{O}]+$-type compounds and it is backed by the sym­ptomatics of CBD. We have moreover suggested a mechanism of formation of this complex that can guide attempts of an vitro reproduction.

Due to the predictive power of the $[\text{Be}_4\text{O}]+$ hypotheses, we are inclined to conclude that the $\text{in-vitro}$ preparation [12] did not lead to the key complex of CBD, but possibly to some intermediate product. This was already considered by the authors of the study:

"While our data support the conclusion that the cation captured in this crystal is Na$^+$, our structure here does not rule out the possibility that, in vivo or tissue culture, cations other than Na$^+$ might occupy this position and may even complete the TCR ligand." [12]

A mere reorientation of the carboxylate group $\beta 69$E towards the binding site $S$ could lead to a shortening of the $C^{\text{BeO}}-C^{\text{OAc}}$ distance from 9.06 to 6 Å as it is required for the new model 1.

The presence of beryllium in the complex in principle opens another source of experimental data, namely $\text{^9Be}$ NMR spectroscopy for a chemical characterization. $\text{^9Be}$ NMR data on CBD related beryllium-protein complexes are scarce [20], nevertheless the method is an ideal complement to single crystal X-ray diffractometry, since it allows to monitor $\text{in-vitro}$ reaction progress in solution. For that we have calculated the isotropic chemical shielding values for the model complex 3 (relative to $\text{Be}_4\text{O(OAc)}_6$ in CDCl$_3$ with $\delta^{\text{Be}} = 0.9$ ppm), yielding 0.6, 0.7, 1.4 and 1.0 ppm, respectively. If these four signals can be resolved depends critically on their half-width. The $\text{^9Be}$ NMR signal half width of $\text{Be}_4\text{O(OAc)}_6$ is very small ($W_{1/2} = 4$ Hz at 20°C) [43], which is unusual for a nucleus with a non-zero nuclear quadrupole moment. The small peak width is explained by the high local symmetry of the ligand spheres around the Be nuclei, which eventually leads to a very small electric field gradient at the nucleus such that the fast quadrupole relaxation mechanism becomes prohibited.

We provide in Section 2.3 details on the calculation of the quadrupole coupling constants for models 3–5 and introduce additional calculations needed to characterize such species in a relaxation study.

### 4 Summary and outlook

We suggest that the $[\text{Be}_4\text{O}]+$ moiety is the coordination center in the protein complex which is ultimately responsible for the $\text{in vivo}$ development of the chronic beryllium disease (CBD). This hypotheses is motivated by the high charge accumulation found experimentally for the binding site $S$ of a CBD model complex [12]. The hypotheses is backed by the topology and geometric features found for $S$, it is backed by the fact that this syndrome is specific for beryllium by the specific chemistry of $[\text{Be}_4\text{O}]+$-type compounds and it is backed by the symptomatics of CBD. We have moreover suggested a mechanism of formation of this complex that can guide attempts of an vitro reproduction.

A proof for this structure hypotheses could be provided via a syntheses guided by $\text{^9Be}$ NMR spectroscopy and followed by a protein crystallographic characterization. Our ongoing quantum chemical calculations and simulations of $\text{^9Be}$ NMR parameters may provide important complementary informations. Due to the peculiar nature of this special type of coordination center, the knowledge of the structure of the CBD key complex is likely to be of decisive importance for the success of such works.

### 5 Supporting information

The electronically available supporting informations contain the DFT optimized atomic coordinates of molecules 3 (S1.1) and $\text{Be}_4\text{O(OAc)}_6$ (S1.1.1) and the atomic coordinates as used for the $\text{^9Be}$ relaxation time calculations of 3 (S2.1), 4 (S2.1), and 5 (S2.3).

**Acknowledgment:** RB acknowledges financial support from Prof. Dage Sundholm and the Magnus Ehrnrooth foundation, and the CSC – IT Center for Science, Finland, for generous provision of computational resources. PH acknowledges support from the Kvantum institute at University of Oulu and a grant for computer capacity from the Finnish Grid and Cloud Infrastructure (persistent identifier urn:nbn:fi:research-infras-2016072533).

### References
