**Predicting the crystal structure of beryl from chemical composition**

**Complete set of figures, plus supplementary figures**

Chart, scatter chart

Description automatically generated Al-SAIR (Å) *vs*. Be-SAIR (Å). Samples are plotted considering the categories of Aurisicchio *et al*. (1988) based on the c/a ratio. Square: normal beryl; circle: octahedral beryl; plus sign: tetrahedral beryl. Samples used to determine the correlation are solid black. Samples to test the model are outlines; these were not used in creating the separation line. There is a clear divergence of the octahedral and tetrahedral beryl. The line Al-SAIR (Å) = (0.45 × Be-SAIR (Å)) + 0.414 separates sample trends. Above this line, beryl is considered Octahedrally Trending (OT) and below this line beryl is considered Tetrahedrally Trending (TT).

Chart, scatter chart

Description automatically generated Si-*x* and Si-*y* fractional coordinate positions *vs*. Al-SAIR (Å) and Be-SAIR (Å). Samples used to create the model are unfilled. HMHF samples are red. OT test samples are orange; TT test samples are blue. The trendlines are solid lines. Confidence intervals are between the dashed lines. Prediction intervals are between the dotted lines. Trendlines and statistical intervals for HMHF beryl marked in red. The Si-*x* and Si-*y* positions change in opposite directions with Al-SAIR and Be-SAIR for OT and TT beryl respectively. The Si-*x* position changes at a different rate then other OT beryl.

Chart, scatter chart

Description automatically generated O1-*x* and O1-*y* fractional coordinate positions *vs*. Al-SAIR (Å) and Be-SAIR (Å). Samples used to create the model are unfilled. HMHF samples are red. OT test samples are orange; TT test samples are blue. The trendlines are solid lines. Confidence intervals are between the dashed lines. Prediction intervals are between the dotted lines. Trendlines and statistical intervals for HMHF beryl marked in red. The O1-*x* and O1-*y* positions change similarly with Al-SAIR and Be-SAIR for both OT and TT beryl, but slightly less dramatically for HMHF beryl with Al-SAIR, and O1*-y* with Be-SAIR for TT beryl.

Chart, scatter chart

Description automatically generated O2-*x*, O2-*y*, and O2-*z* fractional coordinate positions *vs*. Al-SAIR (Å) and Be-SAIR (Å). Samples used to create the model are unfilled. HMHF samples are red. OT test samples are orange; TT test samples are blue. The trendlines are solid lines. Confidence intervals are between the dashed lines. Prediction intervals are between the dotted lines. Trendlines and statistical intervals for HMHF beryl marked in red. The O2-*y* position has the most discrepancy between OT and TT beryl with respect to Al-SAIR and Be-SAIR respectively.

Chart, scatter chart

Description automatically generated Unit cell lengths *a* (Å), *c* (Å), and *c*/*a* ratio *vs*. Al-SAIR (Å) and Be-SAIR (Å). Samples used to create the model are unfilled. HMHF samples are red. OT test samples are orange; TT test samples are blue. The trendlines are solid lines. Confidence intervals are between the dashed lines. Prediction intervals are between the dotted lines. Trendlines and statistical intervals for HMHF beryl marked in red. The *a* length increases with Al-SAIR for OT beryl, and the *c* length increases with Be-SAIR for TT beryl. The *c*/*a* ratio decreases swiftly with Al-SAIR for OT beryl, and increases with Be-SAIR for TT beryl.

Scatter chart

Description automatically generated Be-O2 (Å) and Al-O2 (Å) *vs*. Al-SAIR (Å) and Be-SAIR (Å). Samples used to create the model are unfilled. HMHF samples are red. OT test samples are orange; TT test samples are blue. The trendlines are solid lines. Confidence intervals are between the dashed lines. Prediction intervals are between the dotted lines. Trendlines and statistical intervals for HMHF beryl marked in red. The Be- and Al-O2 bond lengths are highly variable as these consider the sites where cation substitutions occur. There is a strong linear correlation between the Al-O2 bond length and Al-SAIR for OT and HMHF beryl, and strong linear correlation between the Be-O2 bond length and Be-SAIR (Å) for TT beryl.

Chart, scatter chart

Description automatically generated Si-O1*a* (Å), Si-O1*b* (Å) and Si-O2 (Å) *vs*. Al-SAIR (Å) and Be-SAIR (Å). Samples used to create the model are unfilled. HMHF samples are red. OT test samples are orange; TT test samples are blue. The trendlines are solid lines. Confidence intervals are between the dashed lines. Prediction intervals are between the dotted lines. Trendlines and statistical intervals for HMHF beryl marked in red. The Si site is not substituted by alternative cations; variation is due to the changes induced by surrounding sites. There is a strong linear correlation between Si-O1*a* and Si-O1*b* bond lengths and Al-SAIR, or Be-SAIR, for OT and TT beryl respectively. There is a strong negative linear correlation between Si-O2 and Al-SAIR, or Be-SAIR, for OT and TT beryl respectively. There is minimal variation in the Si-O bond lengths with Al- or Be-SAIR for HMHF beryl.

Chart, scatter chart

Description automatically generated Na-O1 (Å) and Ow-O2 (Å) *vs*. Al-SAIR (Å) and Be-SAIR (Å). Samples used to create the model are unfilled. HMHF samples are red. OT test samples are orange; TT test samples are blue. The trendlines are solid lines. Confidence intervals are between the dashed lines. Prediction intervals are between the dotted lines. Trendlines and statistical intervals for HMHF beryl marked in red. The Na-O1 and Ow­-O2 bond lengths are proxies for the channel width. The narrow portion of the channel contracts around the 2*b* site (site occupied partially by Na) linearly with Al-SAIR or Be-SAIR, respectively for OT and TT beryl. The Na-O1 length remains effectively unchanged by Al-SAIR for HMHF beryl (which tends to contain little Na, and this is a better measure of vacant 2*b*-O1). The wider portion of the channel around the 2*a* site (site occupied partially by H2O and sometimes large alkali cations) minimally expands with Al-SAIR for OT and HMHF beryl, but contracts with Be-SAIR for TT beryl.

Chart, scatter chart

Description automatically generated BeO4, AlO6, and SiO4 polyhedral volumes (Be, Al, and Si PV) (Å3) *vs*. Al-SAIR (Å) and Be-SAIR (Å). Samples used to create the model are unfilled. HMHF samples are red. OT test samples are orange; TT test samples are blue. The trendlines are solid lines. Confidence intervals are between the dashed lines. Prediction intervals are between the dotted lines. Trendlines and statistical intervals for HMHF beryl marked in red. The BeO4 polyhedral volume increases linearly with Al-SAIR (Å) for OT and HMHF beryl, and Be-SAIR (Å) for TT beryl. The AlO6 polyhedral volume increases effectively only with Al-SAIR for both OT and HMHF beryl. The SiO4 polyhedra only increase in size minimally with Al-SAIR for OT beryl.

Chart, arrow, scatter chart

Description automatically generated O2-Be-O2*c* °, O2-Be-O2*d* °, and O2-Be-O2*e* ° *vs*. Al-SAIR (Å) and Be-SAIR (Å). Samples used to create the model are unfilled. HMHF samples are red. OT test samples are orange; TT test samples are blue. The trendlines are solid lines. Confidence intervals are between the dashed lines. Prediction intervals are between the dotted lines. Trendlines and statistical intervals for HMHF beryl marked in red. The O2-Be-O2 bond angles change more drastically with Al-SAIR for OT beryl than with Be-SAIR for TT beryl. The O2-Be-O2*d* bond angle is effectively unchanged with Al-SAIR for HMHF beryl.

Chart, scatter chart

Description automatically generated O2*a*-Al-O2*f* °, O2*a*-Al-O2*g* °, O2*a*-Al-O2*h* ° *vs*. Al-SAIR (Å) and Be-SAIR (Å). Samples used to create the model are unfilled. HMHF samples are red. OT test samples are orange; TT test samples are blue. The trendlines are solid lines. Confidence intervals are between the dashed lines. Prediction intervals are between the dotted lines. Trendlines and statistical intervals for HMHF beryl marked in red. The O2-Al-O2 bond angles change linearly with Al-SAIR for OT and HMHF beryl, with Be-SAIR for TT beryl, but with divergent trends comparing OT and TT beryl for O2*a*-Al-O2*g* and O2*a*-Al-O2*h*.

Chart, scatter chart

Description automatically generated O1*a*-Si-O1*b* °, O1*a*-Si-O2 °, O1*b*-Si-O2 °, O2-Si-O2 ° *vs*. Al-SAIR (Å) and Be-SAIR (Å). Samples used to create the model are unfilled. HMHF samples are red. OT test samples are orange; TT test samples are blue. The trendlines are solid lines. Confidence intervals are between the dashed lines. Prediction intervals are between the dotted lines. Trendlines and statistical intervals for HMHF beryl marked in red. There are linear correlations between the O-Si-O bond angles and Al-SAIR and Be-SAIR for OT and TT beryl, respectively. The correlations between O-Si-O bond angles and Al-SAIR for HMHF beryl are distinctly different than for other OT beryl.

Chart, scatter chart

Description automatically generated Be, Al, and Si bond angle variance (Be, Al, and Si BAV) (°2) and Si distortion index *vs*. Al-SAIR (Å) and Be-SAIR (Å). Samples used to create the model are unfilled. HMHF samples are red. OT test samples are orange; TT test samples are blue. The trendlines are solid lines. Confidence intervals are between the dashed lines. Prediction intervals are between the dotted lines. Trendlines and statistical intervals for HMHF beryl marked in red. The bond angle variances and Si distortion index for OT and TT beryl with respect to Al-SAIR and Be-SAIR respectively is similar, except for the Al BAV that increases for OT beryl, but decreases for TT beryl. HMHF beryl have nearly invariant Si BAV and Si distortion index across Al-SAIR.

Chart, scatter chart

Description automatically generated Unit cell volume V (Å3) and unit cell density *d* (g/cm3) *vs*. Al-SAIR (Å) and Be-SAIR (Å). Samples used to create the model are unfilled. HMHF samples are red. OT test samples are orange; TT test samples are blue. The trendlines are solid lines. Confidence intervals are between the dashed lines. Prediction intervals are between the dotted lines. Trendlines and statistical intervals for HMHF beryl marked in red. The unit cell volume and density increase with Al-SAIR or Be-SAIR for OT or TT beryl respectively, except for HMHF beryl for which the density decreases.

Chart, scatter chart

Description automatically generated H2O (*apfu*) *vs*. Na (*apfu*) and H2O (wt.%) *vs*. Na2O (wt.%). Left: H2O (*apfu*) *vs*. Na (*apfu*). Right: H2O (wt.%) *vs*. Na2O (wt.%). Trendlines are solid lines, confidence intervals are within dashed lines, and prediction intervals are within dotted lines. OT beryl trends and intervals in orange. TT beryl trends and intervals in blue. HMHF beryl do not have a discernable trend. For OT beryl: H2O (*apfu*) = 0.413 + (1.36 × Na (*apfu*)), or H2O (wt.%) = 1.375 + (0.77 × Na2O (wt.%). For TT beryl: H2O (*apfu*) = 0.441 + (1.41 × Na (*apfu*)), or H2O (wt.%) = 1.467 + (0.79 × Na2O (wt.%).