Journal of Undergraduate Chemistry Research, **2023**, 22(3), 44

THE DFT COMPUTATIONAL INVESTIGATION OF THE β-SILA-α-AMINO ACIDS AND THEIR β-PERMETHYLSILA-ANALOGS: SILICON-CONTAINING AMINO ACIDS AS A VIABLE FOUNDATION FOR SILICON-BASED LIFE

Tara S. Caso*, Robert W. Zoellner†

Department of Chemistry, California State Polytechnic University, Humboldt, Arcata, CA 95521

Abstract

The theoretical potential for the synthetic accessibility of silicon-containing α-amino acids has been addressed through density functional computational investigations at the M06-2X/6-311++G level of theory for twenty-two proteinogenic β-sila-α-amino acids and their β-permethylsila-analogs. These β-sila-α-amino acids, which could form a basis for silicon-based life, were derived from the common proteinogenic α-amino acids (excluding glycine, which has no β-carbon) and the three less-common α-amino acids *N*-formylmethionine, pyrrolysine, and selenocysteine. Density functional computational investigations at the M06-2X/6-311++G** level of theory are reported for twenty-two proteinogenic β-sila-α-amino acids and their β-permethylsila-analogs. These β-sila-α-amino acids were derived from the common proteinogenic α-amino acids (excluding glycine, which has no β-carbon) and the three less-common α-amino acids N-formylmethionine, pyrrolysine, and selenocysteine. The β-sila-α-amino acids were calculated as if in a neutral aqueous solution using the C-PCM method; only those β-sila-α-amino acids derived from the levo-α-amino acid stereoisomers were investigated. As expected, while both the uncharged and zwitterionic forms of the β-sila-α-amino acids were calculated to be stable minima, the zwitterionic forms were more stable than the uncharged forms in the simulated aqueous solutions. Some of the zwitterionic β-sila-α-amino acids exhibited intramolecular hydrogen bonds; these hydrogen-bonded molecules were calculated to be more stable than the non-hydrogen-bonded counterparts (except for two hydrogen-bonded β-sila-lysines, which were less stable than the open-chain molecules). The silicon-carbon bond lengths for the zwitterionic forms of the β-sila-α-amino acids were calculated to be within the range of the sum of the tetrahedral covalent radii of silicon and carbon, lending credence to the potential for these molecules to be synthetically accessible.**

†Corresponding author: rwz7001@humboldt.edu *Undergraduate researcher and co-author
Keywords: β-sila-α-amino acids, β-permethylsila-α-amino acids, α-silyl-α-amino acids, density functional, M06-2X
Received: June 6, 2023 Accepted: June 22, 2023

Introduction

The foundation of all organic life (on Earth, at least) is dictated by the central dogma[1] of biology: simplistically, DNA (deoxyribonucleic acid) replicates, portions of the DNA (the genes) are transcribed into RNA (ribonucleic acid), and the code within the RNA is translated into proteins using α-amino acids as the building blocks. The first of the twenty common proteinogenic amino acids important to biological life on Earth, although not recognized as such at the time, was discovered[2] as early as 1806 with the isolation of asparagine; with the isolation[3] of threonine in 1935, the last of these twenty common α-amino acids was identified. Three additional α-amino acids, *N*-formylmethionine, discovered[4] in 1964, selenocysteine, isolated[5] in 1974, and pyrrolysine, identified[6] in 2002, are also often included among the proteinogenic amino acids. More recently, the question arose as to whether the familiar carbon-based — and α-amino acid-based — system of life was the only option; perhaps life on other planets[7] could be fundamentally different.

An obvious hypothetical variation on carbon-based life might involve silicon, as silicon is in the same group — sometimes termed the crystallogens[8,9] — and directly beneath carbon in the Periodic Table of the Elements. Unfortunately, while this proximity makes silicon similar to carbon in some respects, and the two elements share some chemical characteristics (while exhibiting significant differences[8,10]), the prospects for silicon-based life may only be possible in environments completely different from those found in the solar system, perhaps even in environments[11] in which carbon-based life could not survive. However, investigations of novel chemistries such as those based on organohalides[12] may lead to hitherto unknown bases for life, including the incorporation of silicon[13] into known biological molecules.

The syntheses and properties of some sila-substituted amino acids have recently[8] been reviewed. Perhaps the earliest synthesis[14] of a β-sila-α-amino acid is that of the highly-protected β-sila-alanine derivative, ethyl [2-(*N*-*t*-butoxycarbonyl-*N*-benzylamino)-2-trimethylsilyl]acetate, in which a butoxycarbonyl group and a benzyl group protected the α-amino nitrogen, while the α-acid group was converted to the ethyl ester. This amino acid is the protected version of β-trimethylsila-alanine (**Ala-Si**^{Me}-u in Table 1), but could not be isolated and was identified in a two-component mixture through an overlapping NMR spectrum. [Note that, in some articles, the **β-sila**-α-amino acids are often referred to as **α-silyl**-α-amino acids as if the silicon-containing moiety was a *substituent* at the α-carbon atom of the amino acid rather than as an amino acid containing a silicon atom *in place of* the β-carbon atom of the amino acid.]

Interestingly, the analogs of glycine, alanine, and valine in which *all* carbon atoms have been replaced by silicon atoms[15] have been computationally investigated at the *ab initio* QCSID/aug-cc-pVTZ and MP2/aug-cc-pVTZ levels of theory. These "per-sila" α-amino acids were determined to be stable as uncharged, non-zwitterionic molecules in the gas phase. A more recent report[16] details the use of the semiempirical PM3/PM7 method to investigate the conformations of the twenty common, uncharged, non-zwitterionic α-amino acids in which, again, all carbon atoms have been replaced by silicon atoms. While the calculated stabilities

Table 1. Selected electronic and structural properties of the β-sila-α-amino acids and their permethylsila-analogs bearing neutral side-chains

molecule[a]	electronic energy (au)	dipole moment (D)	molecular orbital energies (eV) HOMO	LUMO	difference[c]	selected bond distances[b] (pm) C–O	C–O^{H}	C^{α}–N	C^{α}–C	C^{α}–Si	X–Y
Ala-Si-u	-575.095002	3.72	-8.73	0.00	8.73	120.6	134.3	146.1	149.7	191.4	
Ala-Si-z	-575.097787	13.91	-8.91	-0.08	8.83	124.6		149.7	154.9	190.8	
Ala-Si^{Me}-u	-693.058509	2.61	-8.53	-0.10	8.43	121.4	133.8	146.3	149.3	193.6	187.4$^{Si\text{-}Cy}$
Ala-Si^{Me}-z	-693.063251	14.27	-8.67	-0.13	8.54	125.0		149.9	154.0	193.0	187.3$^{Si\text{-}Cy}$
Asn-Si-u	-743.802605	6.81	-8.24	-0.02	8.22	121.1	133.4	146.6	150.0	191.8	192.6$^{Si\text{-}Cy}$ 123.2$^{Cy\text{-}O}$ 134.4$^{Cy\text{-}N}$
Asn-Si-z-N	-743.808295	12.48	-8.78	-0.09	8.69	124.6		149.5	154.9	190.0	192.3$^{Si\text{-}Cy}$ 123.2$^{Cy\text{-}O}$ 134.3$^{Cy\text{-}N}$
Asn-Si-z-H {H_2NH···OC}	-743.811746	13.65	-8.84	-0.06	8.78	124.6 124.2$^{Cy\text{-}O}$		149.5	154.8	190.7	193.0$^{Si\text{-}Cy}$ 133.5$^{Cy\text{-}N}$ 186.5$^{CyO\cdots HNH2}$
Asn-Si-z-H {OCO···HNH}	-743.811112	11.43	-8.69	-0.11	8.58	124.3 125.1$^{C\text{-}O\cdots H}$		149.7	154.7	191.1	193.1$^{Si\text{-}Cy}$ 133.9$^{Cy\text{-}N}$ 123.3$^{Cy\text{-}O}$ 191.2$^{OCO\cdots HNH}$
Asn-Si^{Me}-u	-822.449814	7.77	-8.06	-0.12	7.94	121.2	133.7	146.9	149.5	192.7	193.1$^{Si\text{-}Cy}$ 186.7$^{Si\text{-}Me}$ 123.5$^{Cy\text{-}O}$ 134.6$^{Cy\text{-}N}$
Asn-Si^{Me}-z-N	-822.454561	12.94	-8.57	--0.17	8.40	124.8		149.8	154.5	191.5	192.8$^{Si\text{-}Cy}$ 186.8$^{Si\text{-}Me}$ 123.4$^{Cy\text{-}O}$ 134.5$^{Cy\text{-}N}$
Asn-Si^{Me}-z-H {H_2NH···OC}	-822.459195	13.06	-8.61	-0.07	8.54	124.8		149.7	154.7	191.8	193.6$^{Si\text{-}Cy}$ 186.3$^{Si\text{-}Me}$ 133.9$^{Cy\text{-}N}$ 124.3$^{Cy\text{-}O}$ 185.6$^{CyO\cdots HNH2}$
Asn-Si^{Me}-z-H {OCO···HNH}	-822.457763	12.03	-8.49	-0.17	8.32	124.5 125.2$^{C\text{-}O\cdots H}$		150.0	154.4	192.3	193.5$^{Si\text{-}Cy}$ 186.6$^{Si\text{-}Me}$ 134.1$^{Cy\text{-}N}$ 123.7$^{Cy\text{-}O}$ 188.5$^{OCO\cdots HNH}$
Gln-Si-u	-783.124321	4.07	-8.70	0.22	8.92	121.2	133.5	146.3	149.9	191.6	189.2$^{Si\text{-}Cy}$ 150.5$^{Cy\text{-}Cδ}$ 134.8$^{Cδ\text{-}N}$ 123.0$^{Cδ\text{-}O}$
Gln-Si-z-H {H_2NH···OC}	-783.135647	8.97	-8.86	0.24	9.10	124.8		149.3	154.7	191.1	189.2$^{Si\text{-}Cy}$ 150.6$^{Cy\text{-}Cδ}$ 133.8$^{Cδ\text{-}N}$ 123.9$^{Cδ\text{-}O}$ 191.7$^{CδO\cdots HNH2}$
Gln-Si-z-H {OCO···HNH}	-783.131330	14.17	-9.03	0.17	9.20	124.6 125.0$^{C\text{-}O\cdots H}$		149.8	154.6	191.4	189.3$^{Si\text{-}Cy}$ 150.9$^{Cy\text{-}Cδ}$ 134.7$^{Cδ\text{-}N}$ 123.0$^{Cδ\text{-}O}$ 193.0$^{OCO\cdots HNH}$
Gln-Si^{Me}-u	-861.770284	3.90	-8.56	0.08	8.64	121.3	133.7	146.4	149.5	192.7	190.1$^{Si\text{-}Cy}$ 186.9$^{Si\text{-}Me}$ 150.1$^{Cy\text{-}Cδ}$ 134.9$^{Cδ\text{-}N}$ 123.1$^{Cδ\text{-}O}$
Gln-Si^{Me}-z-N	-861.774710	10.17	-8.74	0.02	8.76	124.8		150.1	154.7	191.8	189.7$^{Si\text{-}Cy}$ 187.1$^{Si\text{-}Me}$ 150.1$^{Cy\text{-}Cδ}$ 135.1$^{Cδ\text{-}N}$ 123.1$^{Cδ\text{-}O}$
Gln-Si^{Me}-z-H {H_2NH···OC}	-861.779900	9.74	-8.72	0.06	8.78	124.9 124.0$^{Cδ\text{-}O}$		149.6	154.5	192.0	190.1$^{Si\text{-}Cy}$ 187.1$^{Si\text{-}Me}$ 150.3$^{Cy\text{-}Cδ}$ 134.0$^{Cδ\text{-}N}$ 192.2$^{CδO\cdots HNH2}$
Gln-Si^{Me}-z-H {OCO···HNH}	-861.777820	15.16	-8.83	0.03	8.86	124.6 125.3$^{C\text{-}O\cdots H}$		149.8	153.9	192.8	190.3$^{Si\text{-}Cy}$ 187.0$^{Si\text{-}Me}$ 150.5$^{Cy\text{-}Cδ}$ 134.9$^{Cδ\text{-}N}$ 123.1$^{Cδ\text{-}O}$ 191.8$^{CO\cdots HN}$
Gly-Si — excluded; no β-carbon											
Ile-Si-u	-693.036231	2.96	-8.55	-0.05	8.50	121.3	133.7	146.3	149.4	192.8	187.4$^{Si\text{-}Me}$ 187.8$^{Si\text{-}Cy}$
Ile-Si-z	-693.040771	14.43	-8.78	-0.10	8.68	124.9		150.0	154.5	192.3	187.6$^{Si\text{-}Me}$ 187.7$^{Si\text{-}Cy}$
Ile-Si^{Me}-u	-732.358053	2.56	-8.49	-0.07	8.42	121.4	133.9	146.3	149.3	193.5	187.5$^{Si\text{-}Me}$ 188.0$^{Si\text{-}Cy}$
Ile-Si^{Me}-z	-732.362670	14.19	-8.65	-0.09	8.56	125.0		150.0	154.1	193.1	187.4$^{Si\text{-}Me}$ 188.2$^{Si\text{-}Cy}$
Leu-Si-u	-693.017162	2.07	-8.53	-0.02	8.51	121.3	133.5	146.2	149.8	192.6	188.3$^{Si\text{-}Cy}$
Leu-Si-z	-693.022749	13.37	-8.85	-0.05	8.80	124.8		150.0	154.6	191.8	188.0$^{Si\text{-}Cy}$
Leu-Si^{Me}-u	-771.660957	2.51	-8.38	-0.07	8.31	121.4	133.8	146.3	149.4	[194.1]	189.0$^{Si\text{-}Cy}$ 187.6$^{Si\text{-}Me}$
Leu-Si^{Me}-z	-771.666228	13.82	-8.70	-0.09	8.61	125.0		150.3	154.5	192.8	188.7$^{Si\text{-}Cy}$ 187.6$^{Si\text{-}Me}$
Met-Si-u	-1051.900531	3.18	-7.73	-0.03	7.70	121.2	133.5	146.3	149.9	191.8	188.4$^{Si\text{-}Cy}$ 181.5$^{Cy\text{-}S}$ 181.4$^{S\text{-}Me}$
Met-Si-z	-1051.904546	12.12	-7.72	-0.05	7.67	124.7		149.7	154.8	190.8	187.9$^{Si\text{-}Cy}$ 181.7$^{Cy\text{-}S}$ 181.5$^{S\text{-}Me}$
Met-Si^{Me}-u	-1130.546082	2.47	-7.65	-0.08	7.57	121.3	133.8	146.3	149.6	193.0	187.1$^{Si\text{-}Cy}$ 187.0$^{Si\text{-}Me}$
Met-Si^{Me}-z	-1130.550214	12.28	-7.64	-0.12	7.52	124.8		150.2	154.6	191.8	188.6$^{Si\text{-}Cy}$ 187.2$^{Si\text{-}Me}$ 181.8$^{Cy\text{-}S}$ 181.4$^{S\text{-}Me}$
Phe-Si-u	-806.120900	3.57	-8.49	-0.21	8.28	121.2	133.5	146.1	149.8	192.3	187.1$^{Si\text{-}Ph}$
Phe-Si-z	-806.125536	15.27	-8.55	-0.23	8.32	124.8		149.6	154.4	191.8	187.0$^{Si\text{-}Ph}$
Phe-Si^{Me}-u	-884.766586	2.62	-8.36	-0.11	8.25	121.3	133.8	146.2	149.5	193.7	188.2$^{Si\text{-}Ph}$ 187.0$^{Si\text{-}Me}$
Phe-Si^{Me}-z	-884.771850	14.29	-8.45	-0.13	8.32	124.9		149.8	154.2	192.8	188.5$^{Si\text{-}Ph}$ 186.8$^{Si\text{-}Me}$
Pro-Si-u	-652.506504	4.15	-8.06	0.01	8.07	120.8	134.4	146.6	149.6	[194.5]	188.2$^{Si\text{-}Cy}$ 147.3$^{N\text{-}Cδ}$ 154.5$^{Cy\text{-}Cδ}$
Pro-Si-z	-652.513887	14.99	-8.85	-0.06	8.79	124.8		150.5	154.5	191.8	188.7$^{Si\text{-}Cy}$ 150.8$^{N\text{-}Cδ}$ 152.6$^{Cy\text{-}Cδ}$
Pro-Si^{Me}-u	-731.150181	5.08	-7.87	-0.07	7.80	120.9	134.7	146.9	149.0	[195.4]	188.6$^{Si\text{-}Cy}$ 187.2$^{Si\text{-}Me}$ 147.5$^{N\text{-}Cδ}$ 154.2$^{Cy\text{-}Cδ}$
Pro-Si^{Me}-z	-731.159140	14.65	-8.72	-0.06	8.66	125.0		150.5	153.8	193.0	189.5$^{Si\text{-}Cy}$ 186.6$^{Si\text{-}Me}$ 150.8$^{N\text{-}Cδ}$ 152.5$^{Cy\text{-}Cδ}$
Ser-Si-u	-650.375767	1.80	-8.22	-0.08	8.14	121.1	133.7	147.0	149.7	191.5	166.2$^{Si\text{-}O}$
Ser-Si-z	-650.382793	13.68	-8.93	-0.13	8.80	124.8		149.5	154.1	190.2	166.5$^{Si\text{-}O}$
Ser^{OMe}-Si-u	-689.657394	4.17	-8.31	-0.01	8.30	121.1	133.5	146.9	150.0	190.6	165.9$^{Si\text{-}O}$ 142.5$^{O\text{-}Me}$
Ser^{OMe}-Si-z	-689.664572	13.84	-8.94	-0.08	8.86	124.7		149.5	154.1	190.1	166.4$^{Si\text{-}O}$ 142.7$^{O\text{-}Me}$
Ser-Si^{Me}-u	-729.023917	4.26	-8.02	-0.14	7.88	121.3	133.9	147.2	149.1	192.9	167.2$^{Si\text{-}O}$ 186.1$^{Si\text{-}Me}$
Ser-Si^{Me}-z	-729.031380	13.27	-8.80	-0.15	8.65	124.8		149.6	154.0	191.3	167.7$^{Si\text{-}O}$ 187.0$^{Si\text{-}Me}$
Ser^{OMe}-Si^{Me}-u	-768.305210	4.65	-8.09	-0.08	8.01	121.2	133.8	147.2	149.2	191.8	167.0$^{Si\text{-}O}$ 142.3$^{O\text{-}Me}$ 186.6$^{Si\text{-}Me}$
Ser^{OMe}-Si^{Me}-z	-768.313488	13.51	-8.80	-0.10	8.70	124.9		149.6	153.9	191.3	167.6$^{Si\text{-}O}$ 142.4$^{O\text{-}Me}$ 186.1$^{Si\text{-}Me}$
Thr-Si-u	-689.699285	4.36	-8.19	-0.01	8.18	121.1	133.8	147.2	149.4	191.3	166.7$^{Si\text{-}O}$ 185.8$^{Si\text{-}Me}$
Thr-Si-z	-689.707365	13.63	-8.86	-0.08	8.78	124.8		149.5	154.1	190.9	167.1$^{Si\text{-}O}$ 185.3$^{Si\text{-}Me}$
Thr^{OMe}-Si-u	-728.981075	4.16	-8.20	0.02	8.22	121.1	133.7	147.1	149.5	191.3	166.4$^{Si\text{-}O}$ 142.4$^{O\text{-}Me}$ 185.9$^{Si\text{-}Me}$
Thr^{OMe}-Si-z	-728.989368	13.91	-8.87	-0.04	8.83	124.8		149.5	154.1	190.5	166.9$^{Si\text{-}O}$ 142.5$^{O\text{-}Me}$ 185.4$^{Si\text{-}Me}$
Thr-Si^{Me}-u ≡ **Ser-Si^{Me}**-u; **Thr-Si^{Me}**-z ≡ **Ser-Si^{Me}**-z; **Thr^{OMe}-Si^{Me}**-u ≡ **Ser^{OMe}-Si^{Me}**-u; **Thr^{OMe}-Si^{Me}**-z ≡ **Ser^{OMe}-Si^{Me}**-z											
Trp-Si-u	-937.693379	5.80	-7.28	-0.07	7.21	121.3	133.6	146.0	149.6	192.3	184.4$^{Si\text{-}In}$
Trp-Si-z	-937.698400	17.79	-7.36	-0.12	7.24	124.8		149.6	154.0	192.4	184.5$^{Si\text{-}In}$
Trp-Si^{Me}-u	-1016.338213	5.25	-7.20	-0.10	7.10	121.4	133.8	146.1	149.4	193.7	185.6$^{Si\text{-}In}$ 187.3$^{Si\text{-}Me}$
Trp-Si^{Me}-z	-1016.343616	16.86	-7.29	-0.12	7.17	125.0		149.7	153.8	193.5	185.7$^{Si\text{-}In}$ 187.1$^{Si\text{-}Me}$
Val-Si-u	-653.736288	3.44	-8.05	-0.01	9.06	121.2	134.0	147.2	149.2	193.3	187.2$^{Si\text{-}Cy}$
Val-Si-z	-653.741748	13.32	-8.77	-0.02	8.79	124.9		149.9	153.4	192.3	186.8$^{Si\text{-}Cy}$
Val-Si^{Me}-u ≡ **Ala-Si^{Me}**-u; **Val-Si^{Me}**-z ≡ **Ala-Si^{Me}**-z											

[a] The three-letter symbols for the α-amino acids, plus the suffix "-Si", identify the β-sila-α-amino acid, the "-Si^{Me}" denotes a fully-methylated silicon atom, and "-u" and "-z" indicate, for the amino acid portion of the molecule, whether the α-amino and α-acid groups are uncharged or zwitterionic, respectively. For those zwitterionic molecules in which an intramolecular hydrogen-bond was observed, the "-H" indicates the molecule with, and the "-N" without, such a bond; the hydrogen bond connectivity is given in curly brackets when hydrogen bonding was calculated to be present. For the β-sila-serines and the β-sila-threonines, the superscript "OMe" indicates that the molecule has been methylated at the –OH group on the side-chain.
[b] Mean values, where appropriate. Superscripts denote the X–Y atoms in the specific bond; "Me" indicates the methyl group, "Ph" indicates the phenyl group, and "In" indicates the indole group. Bond distances in square brackets indicate distances approaching (arbitrarily defined as >194.0 pm) or exceeding 194.9 pm, the sum of the tetrahedral covalent radii for carbon and silicon.
[c] $E_{LUMO} - E_{HOMO}$.

of these per-sila-α-amino acids are intriguing, the results ignore the difficulties in the syntheses[17–21] of per-sila-benzene and similar per-sila-aromatic molecules. Such synthetic barriers, for example, diminish the possibilities regarding the potential syntheses of per-sila-phenylalanine or per-sila-tyrosine, given the aromatic moieties present in these molecules.

Computationally, the properties of α-sila-glycine, α-sila-alanine, and β-sila-alanine (**Ala-Si**-u in Table 1), as gas-phase molecules calculated at the B3LYP/6-311G** level of theory, have been reported. These sila-amino acids were investigated as part of a study[22] of intermolecular hydrogen bonding[23] in amino acids, and involved only the uncharged molecules rather than the zwitterionic systems; the parent α-amino acids are well-known to exist in aqueous solutions as the zwitterions rather than as the uncharged molecules.

In order to address the synthetic accessibility of silicon-containing α-amino acids and, thus, the potential for such sila-α-amino acids to provide a basis for silicon-based life, herein is reported a comprehensive computational investigation of the proteinogenic β-sila-α-amino acids (apart from glycine, which has no β-carbon, but including the β-sila-analogs of three less-common, proteinogenic, α-amino acids *N*-formylmethionine, pyrrolysine, and selenocysteine), in which a β-silicon atom is present in place of the β-carbon atom. The density functional M06-2X/6-311++G** level of theory was employed with the C-PCM (conductor-like polarizable continuum model) to simulate an aqueous environment for these β-sila-α-amino acids.

Computational methods

The *Spartan '18* computational package[24] was used to perform all calculations, and the hybrid density functional method M06-2X/6-311++G** was applied to all molecules. This method reportedly performs well for the investigation[25A] of zwitterionic molecules[25b] and isolated biomolecules in general. In addition, the method is recommended[26] for main-group kinetics and thermochemistry. (No attempts were made to determine whether this method was optimal for the investigation of the β-sila-α-amino acids herein, as such an investigation was beyond the scope of this investigation.) All molecules were calculated as derived from the *levo*-forms of the α-amino acids with silicon atoms replacing the β-carbon atoms. An aqueous, neutral pH solution with a dielectric constant set to 78.30 using the C-PCM (conductor-like polarizable continuum model[27]) method was employed to simulate an aqueous environment for the molecules. This method has been shown[27] to accurately calculate the structures and properties of α-amino acids in such a simulated aqueous solution. Vibrational spectra were calculated for each molecule in order to confirm that a stationary point had been located and that no imaginary frequencies existed. Multiple initial trial geometries (ITGs) were employed for each molecule. Variations in bond lengths and angles in the ITGs determined the starting points for each calculation to better ascertain whether an energy minimum for the molecule had been identified.

For the β-sila-α-amino acids derived from those proteinogenic α-amino acids bearing a neutral side-chain, the β-sila-α-amino acids were calculated as both the uncharged ("-u") molecules [R–

CH(COOH)(NH_2)] (expected to be the stable structures in the gas phase) and as the zwitterionic ("-z") molecules [R–CH(COO^-)(NH_3^+)] (expected to be the structures that exist in aqueous solution). (The methoxy-analogs of β-sila-serine and β-sila-threonine were also investigated.) However, in the cases in which the β-sila-α-amino acids were derived from the α-amino acids bearing potentially-anionic or potentially-cationic side-chains, the molecules were investigated as uncharged and zwitterionic systems both with and without a negative or positive charge, respectively, on the side-chain through the removal or addition of a hydrogen ion. Further, these molecules were calculated as the neutral "charge isomers" bearing a negative or positive charge on the side-chain and no charge on either the α-carboxylic acid group or the α-amino group, respectively. Histidine (commonly represented as **His**) represents a special case, as the imidazole side-chain could be protonated at both nitrogen atoms with a positive charge (**Hip**) or could bear only one N–H bond with a neutral side-chain. In the latter case, the nitrogen atom in the δ-position (**Hid**) or in the ε-position (**Hie**) could bear the hydrogen atom while the side-chain remains neutral. All of these molecules were investigated for the β-sila-histidines.

Results and Discussion

Selected properties and structural data for the β-sila-α-amino acids are listed in the tables: Table 1 details the β-sila-α-amino acids bearing neutral side-chains; Table 2, the β-sila-α-amino acids with potentially anionic or anionic side-chains; Table 3, the β-sila-histidines; Table 4, the β-sila-arginines; Table 5, the β-sila-α-amino acids with potentially cationic or cationic side-chains (excluding the β-sila-histidines and the β-sila-arginines); Table 6, the β-sila-*N*-formylmethionines and the β-sila-selenocysteines; and Table 7, the β-sila-pyrrolysines. The molecules in which the hydrogen atoms attached to silicon were fully replaced by methyl groups were also investigated to produce a broader picture of the types of β-sila-α-

Table 2. Selected electronic and structural properties of the β-sila-α-amino acids and their dimethylsila-analogs bearing anionic or potentially-anionic side-chains

molecule[a]	electronic energy (au)	dipole moment (D)[c]	molecular orbital energies (eV) HOMO	LUMO	difference[d]	selected bond distances[b] (pm) C–O	C–O^H	C^{α}–N	C–C^{α}	Si–C^{α}	X–Y
Asp-Si-u	-763.672966	3.47	-8.76	-0.02	8.74	121.0	133.1	146.4	150.6	190.1	$193.2^{Si\text{-}Cy}$ $120.8^{Cy\text{-}O}$ $133.9^{Cy\text{-}OH}$
Asp-Si-z-N	-763.677197	13.43	-9.01	-0.12	8.89	124.6		149.3	155.0	189.8	$192.3^{Si\text{-}Cy}$ $120.8^{Cy\text{-}O}$ $133.8^{Cy\text{-}OH}$
Asp-Si-z-H {$H_2NH\cdots OCO^H$}	-763.678706	12.92	-8.97	-0.09	8.88	124.6		149.4	154.8	190.5	$193.5^{Si\text{-}Cy}$ $121.5^{Cy\text{-}O}$ $133.5^{Cy\text{-}OH}$ $196.2^{CO\cdots HN}$
Asp^{R-}**-Si**	-763.667039	24.29	-7.82	-0.20	7.62	120.4	133.1	149.5	150.2	193.8	$193.4^{Si\text{-}Cy}$ $125.0^{Cy\text{-}O}$
Asp^{R-}**-Si**-u	-763.217708	—	-7.57	0.08	7.65	121.2	133.7	145.8	150.5	192.1	$193.5^{Si\text{-}Cy}$ $125.4^{Cy\text{-}O}$
Asp^{R-}**-Si**-z-N	-763.219149	—	-7.62	0.02	7.64	124.8		149.8	154.3	191.6	$193.3^{Si\text{-}Cy}$ $125.2^{Cy\text{-}O}$
Asp^{R-}**-Si**-z-H {$H_2NH\cdots OCO$}	-763.227623	—	-7.80	0.08	7.88	124.9		149.5	154.0	192.0	$193.8^{Si\text{-}Cy}$ $124.4^{Cy\text{-}O}$ $133.5^{Cy\text{-}OH}$ $126.8^{Cy\text{-}O\cdots H}$ $168.3^{CO\cdots HN}$
Asp-Si^{Me}-u	-842.319384	4.38	-8.63	-0.07	8.56	121.2	133.4	146.5	150.2	191.1	$193.7^{Si\text{-}Cy}$ $186.4^{Si\text{-}Me}$ $121.0^{Cy\text{-}O}$ $134.4^{Cy\text{-}OH}$
Asp-Si^{Me}-z-N	-842.324559	14.21	-8.87	-0.19	8.68	124.7		149.8	154.8	190.8	$192.7^{Si\text{-}Cy}$ $186.4^{Si\text{-}Me}$ $121.1^{Cy\text{-}O}$ $134.1^{Cy\text{-}OH}$
Asp-Si^{Me}-z-H {$H_2NH\cdots OCO^H$}	-842.326830	12.24	-8.81	-0.08	8.73	124.8		149.6	154.6	191.6	$194.0^{Si\text{-}Cy}$ $186.0^{Si\text{-}Me}$ $121.7^{Cy\text{-}O}$ $133.5^{Cy\text{-}OH}$ $193.7^{CO\cdots HN}$
Asp^{R-}**-Si**^{Me}	-842.312687	24.72	-7.60	-0.23	7.37	120.5	133.4	149.7	149.7	[195.6]	$193.0^{Si\text{-}Cy}$ $187.0^{Si\text{-}Me}$ $125.4^{Cy\text{-}O}$
Asp^{R-}**-Si**^{Me}-u	-841.859229	—	-7.35	-0.01	7.34	121.2	134.2	146.2	149.8	193.0	$193.2^{Si\text{-}Cy}$ $187.6^{Si\text{-}Me}$ $125.7^{Cy\text{-}O}$
Asp^{R-}**-Si**^{Me}-z-N	-841.862626	—	-7.40	-0.07	7.37	125.0		149.9	154.0	192.7	$193.0^{Si\text{-}Cy}$ $187.6^{Si\text{-}Me}$ $125.7^{Cy\text{-}O}$
Asp^{R-}**-Si**^{Me}-z-H {$H_2NH\cdots OCO$}	-841.872001	—	-7.60	0.05	7.65	125.1		149.6	153.7	192.6	$193.9^{Si\text{-}Cy}$ $187.2^{Si\text{-}Me}$ $124.7^{Cy\text{-}O}$ $127.3^{Cy\text{-}O\cdots H}$ $164.6^{CO\cdots HN}$
Glu-Si-u	-802.993740	4.04	-8.72	-0.05	8.67	121.1	133.3	146.3	150.1	191.4	$189.6^{Si\text{-}Cy}$ $149.0^{Cy\text{-}C\delta}$ $120.9^{C=O}$ $134.2^{C\text{-}OH}$
Glu-Si-z-N	-802.998650	13.97	-8.94	-0.15	8.79	124.6		149.6	155.0	190.3	$190.4^{Si\text{-}Cy}$ $149.0^{Cy\text{-}C\delta}$ $120.9^{C=O}$ $134.2^{C\text{-}OH}$
Glu-Si-z-H {$H_2NH\cdots OCO^H$}	-803.003981	12.90	-8.88	-0.04	8.84	124.7		149.3	154.7	190.9	$189.3^{Si\text{-}Cy}$ $149.3^{Cy\text{-}C\delta}$ $132.9^{C\text{-}OH}$ $121.7^{C\text{-}O\cdots H}$ $191.1^{CO\cdots HN}$
Glu^{R-}**-Si**	-802.986762	25.13	-8.37	-0.12	8.25	120.6	133.0	149.3	149.6	[196.2]	$186.5^{Si\text{-}Cy}$ $154.0^{Cy\text{-}C\delta}$ $125.4^{C=O}$
Glu^{R-}**-Si**-u	-802.532513	—	-8.26	0.07	8.19	120.7	134.3	146.2	149.5	192.2	$187.6^{Si\text{-}Cy}$ $153.5^{Cy\text{-}C\delta}$ $125.6^{C=O}$
Glu^{R-}**-Si**-z-N	-802.535105	—	-8.26	-0.02	8.24	124.8		149.9	154.7	192.1	$187.8^{Si\text{-}Cy}$ $153.7^{Cy\text{-}C\delta}$ $125.4^{C=O}$
Glu^{R-}**-Si**-z-H {$H_2NH\cdots OCO$}	-802.548530	—	-8.46	0.04	8.50	124.9		149.3	154.4	191.4	$188.8^{Si\text{-}Cy}$ $152.7^{Cy\text{-}C\delta}$ $124.1^{C=O}$ $127.6^{C=O\cdots H}$ $155.1^{CO\cdots HN}$
Glu-Si^{Me}-u	-881.640650	3.19	-8.59	-0.11	8.48	121.3	133.6	146.4	149.6	192.5	$190.9^{Si\text{-}Cy}$ $148.6^{Cy\text{-}C\delta}$ $186.6^{Si\text{-}Me}$ $121.1^{C=O}$ $134.3^{C\text{-}OH}$
Glu-Si^{Me}-z-N	-881.645132	13.14	-8.80	-0.15	8.65	124.8		150.0	154.8	191.3	$191.5^{Si\text{-}Cy}$ $148.6^{Cy\text{-}C\delta}$ $186.5^{Si\text{-}Me}$ $121.1^{C=O}$ $134.4^{C\text{-}OH}$
Glu-Si^{Me}-z-H {$H_2NH\cdots OCO^H$}	-881.649206	14.02	-8.74	-0.12	8.62	124.8		149.7	154.4	192.0	$190.5^{Si\text{-}Cy}$ $148.7^{Cy\text{-}C\delta}$ $186.8^{Si\text{-}Me}$ $121.9^{C=O}$ $133.2^{C\text{-}OH}$ $193.0^{CO\cdots HN}$
Glu^{R-}**-Si**^{Me}	-881.632257	25.80	-8.33	-0.19	8.14	120.6	133.0	149.3	149.7	[195.8]	$187.5^{Si\text{-}Cy}$ $153.5^{Cy\text{-}C\delta}$ $187.0^{Si\text{-}Me}$ $125.5^{C=O}$
Glu^{R-}**-Si**^{Me}-u	-881.177088	—	-8.17	-0.02	8.15	120.9	134.5	146.3	149.2	193.9	$188.3^{Si\text{-}Cy}$ $153.0^{Cy\text{-}C\delta}$ $187.5^{Si\text{-}Me}$ $125.8^{C=O}$
Glu^{R-}**-Si**^{Me}-z-N	-881.180199	—	-8.17	-0.03	8.14	125.0		150.1	154.5	193.0	$189.0^{Si\text{-}Cy}$ $153.4^{Cy\text{-}C\delta}$ $187.0^{Si\text{-}Me}$ $125.6^{C=O}$
Glu^{R-}**-Si**^{Me}-z-H {$H_2NH\cdots OCO$}	-881.192912	—	-8.32	0.01	8.33	125.0		149.6	154.1	192.7	$189.5^{Si\text{-}Cy}$ $152.3^{Cy\text{-}C\delta}$ $187.2^{Si\text{-}Me}$ $124.3^{C=O}$ $127.6^{C=O\cdots H}$ $155.4^{O\cdots H}$

[a] The three-letter symbols for the α-amino acids, plus the suffix "**-Si**", identify the β-sila-α-amino acid, the "-Si^{Me}" denotes a fully-methylated silicon atom, and "-u" and "-z" indicate, for the amino acid portion of the molecule, whether the α-amino and α-acid groups are uncharged or zwitterionic, respectively. When the side-chain bears a negative charge, a superscripted "R-" indicates this charge; the "R-" without either the "-u" or "-z" suffixes indicates that the molecule is a "charge-isomer" which bears a positive charge on the α-ammonium group, no charge on the α-carboxylic acid group, and a negative charge on the side-chain. When an intramolecular hydrogen bond can exist, the "-H" designates the hydrogen-bonded molecule and the "-N" designates the molecule without a hydrogen bond; the hydrogen bond connectivity is given in curly brackets.
[d] Dipole moments are undefined for a species bearing an overall negative charge.
[c] Mean values, where appropriate. Superscripts denote the X–Y atoms in the specific bond; "Me" indicates a methyl group. Bond distances in square brackets indicate Si–C distances approaching (arbitrarily defined as >194.0 pm) or exceeding 194.9 pm, the sum of the tetrahedral covalent radii for carbon and silicon.
[d] $E_{LUMO} - E_{HOMO}$.

Table 3. Selected electronic and structural properties of the β-sila- and β-dimethylsila-histidines

molecule[a]	electronic energy (au)	dipole moment (D)[c]	molecular orbital energies (eV) HOMO	LUMO	difference[d]	selected bond distances[b] (pm) C–O	C–O^H	C–N	C–C^{α}	Si–C^{α}	X–Y[e]
Hid-Si-u	-800.124270	5.04	-8.03	-0.07	7.96	121.2	133.4	146.3	149.8	191.8	$185.6^{Si\text{-}Cy}$ $138.3^{Cy\text{-}N\delta}$ $134.9^{N\delta\text{-}C\varepsilon}$ $131.8^{C\varepsilon\text{-}N\varepsilon}$ $137.2^{N\varepsilon\text{-}C\delta}$ $137.9^{C\delta\text{-}Cy}$
Hid-Si-z-N	-800.128359	11.32	-8.12	-0.12	8.00	124.7		149.5	154.7	191.2	$186.2^{Si\text{-}Cy}$ $138.3^{Cy\text{-}N\delta}$ $134.8^{N\delta\text{-}C\varepsilon}$ $131.8^{C\varepsilon\text{-}N\varepsilon}$ $137.1^{N\varepsilon\text{-}C\delta}$ $138.0^{C\delta\text{-}Cy}$
Hid-Si-z-H {$OCO\cdots HN^{\delta}$}	-800.133582	12.96	-7.98	-0.09	7.89	124.4 $125.1^{C\text{-}O\cdots H}$		149.6	154.8	191.1	$185.3^{Si\text{-}Cy}$ $137.9^{Cy\text{-}N\delta}$ $134.7^{N\delta\text{-}C\varepsilon}$ $132.1^{C\varepsilon\text{-}N\varepsilon}$ $137.2^{N\varepsilon\text{-}C\delta}$ $138.0^{C\delta\text{-}Cy}$ $195.2^{O\cdots H}$
Hid-Si^{Me}-u	-878.770665	4.38	-7.91	-0.10	7.81	121.3	133.7	146.4	149.4	193.2	$186.7^{Si\text{-}Cy}$ $186.8^{Si\text{-}Me}$ $138.4^{Cy\text{-}N\delta}$ $135.0^{N\delta\text{-}C\varepsilon}$ $131.7^{C\varepsilon\text{-}N\varepsilon}$ $137.4^{N\varepsilon\text{-}C\delta}$ $137.9^{C\delta\text{-}Cy}$
Hid-Si^{Me}-z-N	-878.776410	12.01	-8.04	-0.12	7.92	124.8		149.7	154.3	192.5	$187.0^{Si\text{-}Cy}$ $186.4^{Si\text{-}Me}$ $138.4^{Cy\text{-}N\delta}$ $135.0^{N\delta\text{-}C\varepsilon}$ $131.7^{C\varepsilon\text{-}N\varepsilon}$ $137.4^{N\varepsilon\text{-}C\delta}$ $138.0^{C\delta\text{-}Cy}$
Hid-Si^{Me}-z-H {$OCO\cdots HN^{\delta}$}	-878.779525	13.71	-7.86	-0.15	7.73	124.6 $125.1^{C\text{-}O\cdots H}$		150.0	154.6	192.1	$186.4^{Si\text{-}Cy}$ $187.0^{Si\text{-}Me}$ $137.9^{Cy\text{-}N\delta}$ $134.7^{N\delta\text{-}C\varepsilon}$ $132.1^{C\varepsilon\text{-}N\varepsilon}$ $137.3^{N\varepsilon\text{-}C\delta}$ $138.0^{C\delta\text{-}Cy}$ $191.3^{O\cdots H}$
Hie-Si-u	-800.123772	7.75	-8.11	-0.02	8.09	121.2	133.6	146.1	149.8	192.1	$185.4^{Si\text{-}Cy}$ $139.3^{Cy\text{-}N\delta}$ $131.2^{N\delta\text{-}C\varepsilon}$ $135.7^{C\varepsilon\text{-}N\varepsilon}$ $136.7^{N\varepsilon\text{-}C\delta}$ $137.5^{C\delta\text{-}Cy}$
Hie-Si-z-N	-800.129052	18.40	-8.16	-0.09	8.07	124.8		149.6	154.5	191.7	$185.6^{Si\text{-}Cy}$ $139.3^{Cy\text{-}N\delta}$ $131.1^{N\delta\text{-}C\varepsilon}$ $135.7^{C\varepsilon\text{-}N\varepsilon}$ $136.6^{N\varepsilon\text{-}C\delta}$ $137.6^{C\delta\text{-}Cy}$
Hie-Si-z-H {$H_2NH\cdots N^{\delta}$}	-800.132553	19.24	-8.18	-0.05	8.13	124.8		149.7	154.4	191.5	$186.3^{Si\text{-}Cy}$ $139.5^{Cy\text{-}N\delta}$ $131.2^{N\delta\text{-}C\varepsilon}$ $135.3^{C\varepsilon\text{-}N\varepsilon}$ $136.8^{N\varepsilon\text{-}C\delta}$ $137.3^{C\delta\text{-}Cy}$ $190.1^{N\delta\cdots H}$
Hie-Si^{Me}-u	-878.769258	7.35	-7.99	-0.08	7.91	121.3	133.8	146.2	149.3	193.1	$186.4^{Si\text{-}Cy}$ $186.8^{Si\text{-}Me}$ $139.4^{Cy\text{-}N\delta}$ $131.3^{N\delta\text{-}C\varepsilon}$ $135.6^{C\varepsilon\text{-}N\varepsilon}$ $136.9^{N\varepsilon\text{-}C\delta}$ $137.5^{C\delta\text{-}Cy}$
Hie-Si^{Me}-z-N	-878.771665	13.84	-7.94	-0.13	7.81	124.9		149.8	154.3	192.9	$186.3^{Si\text{-}Cy}$ $187.3^{Si\text{-}Me}$ $139.4^{Cy\text{-}N\delta}$ $131.2^{N\delta\text{-}C\varepsilon}$ $135.6^{C\varepsilon\text{-}N\varepsilon}$ $136.9^{N\varepsilon\text{-}C\delta}$ $137.5^{C\delta\text{-}Cy}$
Hie-Si^{Me}-z-H {$H_2NH\cdots N^{\delta}$}	-878.778498	18.59	-8.06	-0.07	7.99	125.0		149.9	154.1	192.4	$187.1^{Si\text{-}Cy}$ $187.0^{Si\text{-}Me}$ $139.6^{Cy\text{-}N\delta}$ $131.3^{N\delta\text{-}C\varepsilon}$ $135.3^{C\varepsilon\text{-}N\varepsilon}$ $137.0^{N\varepsilon\text{-}C\delta}$ $137.3^{C\delta\text{-}Cy}$ $189.9^{N\delta\cdots H}$
Hip^{R+}**-Si**-u	-800.573637	—	-8.43	-0.10	8.33	120.3	134.4	146.1	150.0	191.7	$188.5^{Si\text{-}Cy}$ $139.1^{Cy\text{-}N\delta}$ $132.4^{N\delta\text{-}C\varepsilon}$ $132.6^{C\varepsilon\text{-}N\varepsilon}$ $137.4^{N\varepsilon\text{-}C\delta}$ $136.6^{C\delta\text{-}Cy}$
Hip^{R+}**-Si**-z-N	-800.578555	—	-8.94	-0.29	8.65	124.6		149.4	155.0	190.1	$188.4^{Si\text{-}Cy}$ $139.0^{Cy\text{-}N\delta}$ $132.4^{N\delta\text{-}C\varepsilon}$ $132.8^{C\varepsilon\text{-}N\varepsilon}$ $137.5^{N\varepsilon\text{-}C\delta}$ $136.6^{C\delta\text{-}Cy}$
Hip^{R+}**-Si**-z-H {$OCO\cdots HN^{\delta}$}	-800.585690	—	-9.13	-0.18	8.95	123.9 $125.5^{C\text{-}O\cdots H}$		149.4	154.8	190.8	$187.3^{Si\text{-}Cy}$ $138.5^{Cy\text{-}N\delta}$ $132.1^{N\delta\text{-}C\varepsilon}$ $133.2^{C\varepsilon\text{-}N\varepsilon}$ $137.4^{N\varepsilon\text{-}C\delta}$ $136.6^{C\delta\text{-}Cy}$ $102.9^{N\delta\text{-}H}$ $181.2^{O\cdots H}$
Hip^{R+}**-Si**	-800.117458	29.79	-8.19	-0.04	8.15	125.2		146.8	154.4	188.4	$189.5^{Si\text{-}Cy}$ $138.9^{Cy\text{-}N\delta}$ $132.6^{N\delta\text{-}C\varepsilon}$ $132.7^{C\varepsilon\text{-}N\varepsilon}$ $137.6^{N\varepsilon\text{-}C\delta}$ $136.6^{C\delta\text{-}Cy}$
Hip^{R+}**-Si**^{Me}-u	-879.221578	—	-8.69	-0.19	8.50	121.2	133.6	146.2	149.9	192.2	$188.8^{Si\text{-}Cy}$ $186.2^{Si\text{-}Me}$ $139.1^{Cy\text{-}N\delta}$ $132.5^{N\delta\text{-}C\varepsilon}$ $132.8^{C\varepsilon\text{-}N\varepsilon}$ $137.6^{N\varepsilon\text{-}C\delta}$ $136.6^{C\delta\text{-}Cy}$
Hip^{R+}**-Si**^{Me}-z-N	-879.226324	—	-8.84	-0.23	8.61	125.2		149.8	155.0	191.2	$189.3^{Si\text{-}Cy}$ $185.8^{Si\text{-}Me}$ $139.1^{Cy\text{-}N\delta}$ $132.5^{N\delta\text{-}C\varepsilon}$ $132.8^{C\varepsilon\text{-}N\varepsilon}$ $137.6^{N\varepsilon\text{-}C\delta}$ $136.6^{C\delta\text{-}Cy}$
Hip^{R+}**-Si**^{Me}-z-H {$OCO\cdots HN^{\delta}$}	-879.233669	—	-8.98	-0.23	8.75	124.0 $125.7^{C\text{-}O\cdots H}$		149.8	154.3	192.2	$188.3^{Si\text{-}Cy}$ $186.4^{Si\text{-}Me}$ $138.6^{Cy\text{-}N\delta}$ $132.1^{N\delta\text{-}C\varepsilon}$ $133.2^{C\varepsilon\text{-}N\varepsilon}$ $137.5^{N\varepsilon\text{-}C\delta}$ $136.7^{C\delta\text{-}Cy}$ $103.8^{N\delta\text{-}H}$ $172.4^{O\cdots H}$
Hip^{R+}**-Si**^{Me}	-878.762608	29.75	-7.98	-0.10	7.88	125.4		147.0	154.1	189.5	$190.1^{Si\text{-}Cy}$ $186.8^{Si\text{-}Me}$ $139.1^{Cy\text{-}N\delta}$ $132.6^{N\delta\text{-}C\varepsilon}$ $132.7^{C\varepsilon\text{-}N\varepsilon}$ $137.8^{N\varepsilon\text{-}C\delta}$ $136.5^{C\delta\text{-}Cy}$

[a] The three-letter symbols, plus the suffix "**-Si**", identify the β-sila-histidines with a neutral side-chain and a hydrogen atom at N^{δ} (**Hid-Si**) or N^{ε} (**Hie-Si**), or with a cationic side-chain with hydrogen atoms at both N^{δ} and N^{ε} (**Hip**$^{R+}$**-Si**); the "-Si^{Me}" denotes the dimethylsila-group; the "-u" and "-z" indicate whether the α-amino and α-acid groups are uncharged or zwitterionic, respectively. Without "-u" or "-z", the molecule is a "charge-isomer" with no charge on the α-amino group, a negative charge on the α-carboxylate group, and a positive charge on the side-chain. The "-N" and "-H" indicate, respectively, a β-sila-histidine without and with an intramolecular hydrogen bond; the hydrogen bond connectivity is given in curly brackets.
[b] Mean values, where appropriate. Superscripts denote the X–Y atoms in the specific bond; "Me" indicates a methyl group.
[c] Dipole moments are undefined for a species bearing an overall positive charge.
[d] $E_{LUMO} - E_{HOMO}$.

amino acids that might exist; the addition of methyl groups has been reported[8] to increase the stability of these molecules. No methylation was investigated at the α-amino or α-carboxylic acid

Table 4. Selected electronic and structural properties of the β-sila- and β-dimethylsila-arginines

molecule[a]	electronic energy (au)	dipole moment (D)[c]	molecular orbital energies (eV) HOMO	LUMO	difference[d]	selected bond distances[b] (pm) C–O	C–O^{H}	C–N	C–C$^{\alpha}$	Si–C$^{\alpha}$	X–Y^{e}
Arg-Si-u	-857.887110	6.51	-8.05	-0.06	7.99	121.2	133.5	146.3	149.8	192.0	188.0$^{Si\text{-}Cγ}$ 152.6$^{Cγ\text{-}Cδ}$ 145.7$^{Cδ\text{-}Nε}$ 137.1$^{Nε\text{-}Cζ}$ 129.1$^{Cζ\text{-}NH}$ 138.7$^{Cζ\text{-}NH2}$
Arg-Si-z-N	-857.890770	14.27	-7.96	-0.10	7.86	124.7		149.9	154.9	191.1	187.9$^{Si\text{-}Cγ}$ 153.2$^{Cγ\text{-}Cδ}$ 146.5$^{Cδ\text{-}Nε}$ 138.2$^{Nε\text{-}Cζ}$ 129.0$^{Cζ\text{-}NH}$ 138.0$^{Cζ\text{-}NH2}$
Arg-Si-z-H {OCO···HNH}	-857.896338	13.84	-8.38	-0.08	8.30	124.4 125.4$^{CO\cdots H}$		149.6	153.8	191.6	188.5$^{Si\text{-}Cγ}$ 152.9$^{Cγ\text{-}Cδ}$ 148.7$^{Cδ\text{-}Nε}$ 141.0$^{Nε\text{-}Cζ}$ 128.3$^{Cζ\text{-}NH}$ 137.6$^{Cζ\text{-}NH2}$ 199.2$^{O\cdots H}$
Arg-Si-z-H {OCO···HN^{E}}	-857.893577	17.89	-7.82	-0.11	7.71	124.9 124.8$^{CO\cdots H}$		149.9	154.5	191.6	188.2$^{Si\text{-}Cγ}$ 154.1$^{Cγ\text{-}Cδ}$ 145.9$^{Cδ\text{-}Nε}$ 137.5$^{Nε\text{-}Cζ}$ 129.3$^{Cζ\text{-}NH}$ 137.9$^{Cζ\text{-}NH2}$ 210.0$^{O\cdots H}$
Arg-Si-z-H {OCO···HNC^{E} + H_2NH···N(H)C^{E}}	-857.893911	12.05	-7.97	-0.10	7.87	124.6 125.1$^{CO\cdots H}$		149.8	154.4	191.5	188.1$^{Si\text{-}Cγ}$ 153.6$^{Cγ\text{-}Cδ}$ 146.2$^{Cδ\text{-}Nε}$ 137.5$^{N\text{-}Cζ}$ 129.5$^{Cζ\text{-}NH}$ 137.8$^{Cζ\text{-}NH2}$ 222.1$^{O\cdots H}$ 230.5$^{N\cdots H}$
Arg-Si-z-H {H_2NH···N(H)C^{E}}	-857.901240	8.38	-8.24	-0.05	8.19	124.7		149.4	154.5	192.2	188.5$^{Si\text{-}Cγ}$ 153.3$^{Cγ\text{-}Cδ}$ 145.4$^{Cδ\text{-}Nε}$ 135.4$^{Nε\text{-}Cζ}$ 130.3$^{Cζ\text{-}NH}$ 138.2$^{Cζ\text{-}NH2}$ 181.4$^{H\cdots N}$
Arg-Si-z-H{H_2NH···NH_2}; **Arg-Si**-z-H{H_2NH···N^{E}}						[unable to be optimized]					
Arg^{R+}-Si	-857.899660	30.29	-7.99	-0.08	7.91	125.4		147.1	154.1	189.3	188.9$^{Si\text{-}Cγ}$ 152.1$^{Cγ\text{-}Cδ}$ 146.7$^{Cδ\text{-}Nε}$ 132.8$^{Nε\text{-}Cζ}$ 133.4$^{Cζ\text{-}NH2}$
Arg^{R+}-Si-u	-858.358377	—	-8.75	-0.15	8.60	121.2	133.5	146.3	149.9	191.7	188.3$^{Si\text{-}Cγ}$ 152.2$^{Cγ\text{-}Cδ}$ 146.6$^{Cδ\text{-}Nε}$ 132.9$^{Nε\text{-}Cζ}$ 133.4$^{Cζ\text{-}NH2}$
Arg^{R+}-Si-z-N	-858.364059	—	-8.90	-0.17	8.73	124.8		149.7	154.4	191.5	188.4$^{Si\text{-}Cγ}$ 153.1$^{Cγ\text{-}Cδ}$ 146.5$^{Cδ\text{-}Nε}$ 133.0$^{Nε\text{-}Cζ}$ 133.2$^{Cζ\text{-}NH2}$
Arg^{R+}-Si-z-H {OCO···HNH}	-858.371404	—	-9.07	-0.12	8.95	124.3 125.3$^{CO\cdots H}$		149.9	154.2	191.9	187.5$^{Si\text{-}Cγ}$ 153.0$^{Cγ\text{-}Cδ}$ 146.1$^{Cδ\text{-}Nε}$ 133.2$^{Nε\text{-}Cζ}$ 180.8$^{O\cdots H}$ 133.8$^{Cζ\text{-}NH2}$ 132.9$^{Cζ\text{-}N(H)H\cdots O}$
Arg^{R+}-Si-z-H {OCO···HN^{E}}	-858.365311	—	-9.04	-0.25	8.79	123.8 126.0$^{CO\cdots H}$		149.9	154.4	192.4	188.5$^{Si\text{-}Cγ}$ 153.9$^{Cγ\text{-}Cδ}$ 146.0$^{Cδ\text{-}Nε}$ 132.9$^{Nε\text{-}Cζ}$ 184.3$^{O\cdots H}$ 133.5$^{Cζ\text{-}NH2}$
Arg^{R+}-Si-z-H{H_2NH···NH_2}; **Arg^{R+}-Si**-z-H{H_2NH···N^{E}}						[unable to be optimized]					
Arg-Si^{Me}-u	-936.531513	5.71	-8.02	-0.12	7.90	121.3	133.8	146.4	149.4	192.8	188.6$^{Si\text{-}Cγ}$ 187.4$^{Si\text{-}Me}$ 152.5$^{Cγ\text{-}Cδ}$ 145.8$^{Cδ\text{-}Nε}$ 136.9$^{Nε\text{-}Cζ}$ 129.1$^{Cζ\text{-}NH}$ 138.8$^{Cζ\text{-}NH2}$
Arg-Si^{Me}-z-N	-936.534735	15.16	-7.93	-0.13	7.80	124.8		150.1	154.5	192.4	188.4$^{Si\text{-}Cγ}$ 187.4$^{Si\text{-}Me}$ 153.1$^{Cγ\text{-}Cδ}$ 146.7$^{Cδ\text{-}Nε}$ 138.1$^{Nε\text{-}Cζ}$ 129.0$^{Cζ\text{-}NH}$ 138.2$^{Cζ\text{-}NH2}$
Arg-Si^{Me}-z-H {OCO···HNH}	-936.540552	14.74	-8.32	-0.12	8.20	124.5 125.6$^{CO\cdots H}$		149.8	153.4	192.8	189.1$^{Si\text{-}Cγ}$ 187.4$^{Si\text{-}Me}$ 152.9$^{Cγ\text{-}Cδ}$ 148.8$^{Cδ\text{-}Nε}$ 140.8$^{Nε\text{-}Cζ}$ 196.9$^{O\cdots H}$ 128.4$^{Cζ\text{-}NH}$ 137.6$^{Cζ\text{-}NH2}$
Arg-Si^{Me}-z-H {OCO···HN^{E}}	-936.537255	18.69	-7.69	-0.14	7.55	125.1 125.6$^{CO\cdots H}$		150.1	154.3	192.9	188.9$^{Si\text{-}Cγ}$ 187.4$^{Si\text{-}Me}$ 153.9$^{Cγ\text{-}Cδ}$ 145.6$^{Cδ\text{-}Nε}$ 136.7$^{Nε\text{-}Cζ}$ 206.5$^{O\cdots H}$ 129.5$^{Cζ\text{-}NH}$ 138.4$^{Cζ\text{-}NH2}$
Arg-Si^{Me}-z-H {OCO···HNC^{E} + H_2NH···N(H)C^{E}}	-936.536712	12.86	-7.93	-0.12	7.81	124.4 125.6$^{CO\cdots H}$		150.2	154.1	192.9	188.6$^{Si\text{-}Cγ}$ 187.4$^{Si\text{-}Me}$ 153.6$^{Cγ\text{-}Cδ}$ 146.4$^{Cδ\text{-}Nε}$ 137.7$^{Nε\text{-}Cζ}$ 222.9$^{O\cdots H}$ 129.4$^{Cζ\text{-}NH}$ 137.8$^{Cζ\text{-}NH2}$ 240.7$^{N\cdots H}$
Arg-Si^{Me}-z-H {H_2NH···N(H)C^{E}}	-936.542782	9.43	-8.16	-0.10	8.06	125.0		149.8	154.3	193.0	188.9$^{Si\text{-}Cγ}$ 187.8$^{Si\text{-}Me}$ 153.2$^{Cγ\text{-}Cδ}$ 145.6$^{Cδ\text{-}Nε}$ 135.5$^{Nε\text{-}Cζ}$ 188.0$^{N\cdots H}$ 120.3$^{Cζ\text{-}NH}$ 138.3$^{Cζ\text{-}NH2}$
Arg-Si^{Me}-z-H{H_2NH···NH_2}; **Arg-Si^{Me}**-z-H{H_2NH···N^{E}}						[unable to be optimized]					
Arg^{R+}-Si^{Me}	-936.540642	32.12	-7.80	-0.13	7.67	125.6		147.5	153.5	190.1	190.2$^{Si\text{-}Cγ}$ 188.2$^{Si\text{-}Me}$ 152.6$^{Cγ\text{-}Cδ}$ 146.7$^{Cδ\text{-}Nε}$ 132.8$^{Nε\text{-}Cζ}$ 133.6$^{Cζ\text{-}NH2}$
Arg^{R+}-Si^{Me}-u	-937.003015	—	-8.60	-0.19	8.41	121.3	133.8	146.4	149.5	193.0	189.0$^{Si\text{-}Cγ}$ 187.2$^{Si\text{-}Me}$ 152.2$^{Cγ\text{-}Cδ}$ 146.7$^{Cδ\text{-}Nε}$ 132.8$^{Nε\text{-}Cζ}$ 133.4$^{Cζ\text{-}NH2}$
Arg^{R+}-Si^{Me}-z-N	-937.007939	—	-8.78	-0.23	8.55	124.8		150.0	154.6	192.6	189.4$^{Si\text{-}Cγ}$ 186.8$^{Si\text{-}Me}$ 152.1$^{Cγ\text{-}Cδ}$ 146.6$^{Cδ\text{-}Nε}$ 132.9$^{Nε\text{-}Cζ}$ 133.4$^{Cζ\text{-}NH2}$
Arg^{R+}-Si^{Me}-z-H {OCO···HNH}	-937.017232	—	-8.96	-0.18	8.78	124.7 125.3$^{CO\cdots H}$		150.0	153.8	193.1	188.1$^{Si\text{-}Cγ}$ 187.2$^{Si\text{-}Me}$ 152.9$^{Cγ\text{-}Cδ}$ 146.3$^{Cδ\text{-}Nε}$ 133.1$^{Nε\text{-}Cζ}$ 179.6$^{O\cdots H}$ 134.1$^{Cζ\text{-}NH2}$ 132.8$^{Cζ\text{-}NHH\cdots O}$
Arg^{R+}-Si^{Me}-z-H {OCO···HN^{E}}	-937.010287	—	-8.91	-0.23	8.68	123.8 126.3$^{CO\cdots H}$		150.2	154.1	193.4	189.4$^{Si\text{-}Cγ}$ 186.9$^{Si\text{-}Me}$ 153.7$^{Cγ\text{-}Cδ}$ 146.1$^{Cδ\text{-}Nε}$ 132.8$^{Nε\text{-}Cζ}$ 182.2$^{O\cdots H}$ 133.6$^{Cζ\text{-}NH2}$
Arg^{R+}-Si^{Me}-z-H{H_2NH···NH_2}; **Arg^{R+}-Si^{Me}**-z-H{H_2NH···N(H)C^{E}}						[unable to be optimized]					

[a] The three-letter symbol for arginine, plus the suffix "**-Si**", identifies the parent β-sila-arginine, the "**-Si^{Me}**" denotes a dimethylsila-group, and "-u" and "-z" indicate whether the α-amino and α-acid groups are uncharged or zwitterionic, respectively. When the side-chain bears a positive charge, a superscripted "R+" indicates this charge; without either "-u" or "-z" the molecule is a "charge-isomer" which bears no charge on the α-amino group, a negative charge on the α-carboxylic acid group, and a positive charge on the side-chain. For the molecules which may exhibit an intramolecular hydrogen-bond, an "-N" indicates the molecule without, and an "-H" with, such a bond; the connectivity of the hydrogen bond is noted in curly brackets

[b] Mean values, where appropriate. Superscripts denote the X–Y atoms in the bond; "Me" indicates a methyl group.

[c] Dipole moments are undefined for a species bearing an overall positive charge.

[d] $E_{LUMO} - E_{HOMO}$.

Table 5. Selected electronic and structural properties of the β-sila-α-amino acids and their dimethylsila-analogs bearing cationic or potentially-cationic side-chains, excluding the β-sila-histidines and the β-sila-arginines

molecule[a]	electronic energy (au)	dipole moment (D)[c]	molecular orbital energies (eV) HOMO	LUMO	difference[d]	selected bond distances[b] (pm) C–O	C–O^{H}	C–N	C–C$^{\alpha}$	Si–C$^{\alpha}$	X–Y^{e}
Cys-Si-u	-973.320650	5.12	-8.41	-0.02	8.39	120.6	133.9	146.3	149.9	191.7	216.0$^{Si\text{-}S}$
Cys-Si-z	-973.324703	12.40	-8.68	-0.14	8.54	124.6		149.1	154.5	191.2	216.2$^{Si\text{-}S}$
Cys^{R+}-Si; **Cys^{R+}-Si**-u; **Cys^{R+}-Si**-z						[unable to be optimized]					
Cys-Si^{Me}-u	-1052.968670	6.48	-8.24	-0.10	8.14	120.8	134.1	146.6	149.5	192.8	217.8$^{Si\text{-}S}$ 186.6$^{Si\text{-}Me}$
Cys-Si^{Me}-z	-1052.973451	11.37	-8.55	-0.12	8.43	124.8		149.3	154.5	192.1	218.2$^{Si\text{-}S}$ 186.0$^{Si\text{-}Me}$
Cys^{R+}-Si^{Me}; **Cys^{R+}-Si^{Me}**-u; **Cys^{R+}-Si^{Me}**-z						[unable to be optimized]					
Lys-Si-u	-748.368886	3.97	-8.47	-0.01	8.46	121.2	133.6	146.3	149.7	192.2	187.5$^{Si\text{-}Cγ}$ 153.4$^{Cγ\text{-}Cδ}$ 152.4$^{Cδ\text{-}Cε}$ 146.8$^{Cε\text{-}Nζ}$
Lys-Si-z-N	-748.373758	14.56	-8.50	-0.05	8.45	124.8		149.8	154.5	191.7	187.3$^{Si\text{-}Cγ}$ 153.4$^{Cγ\text{-}Cδ}$ 152.4$^{Cδ\text{-}Cε}$ 146.8$^{Cε\text{-}Nζ}$
Lys-Si-z-H {OCO···HN$^{\zeta}$}	-748.371393	14.60	-8.24	-0.02	8.22	125.0 124.6$^{C\text{-}O\cdots H}$		149.9	154.5	192.7	187.3$^{Si\text{-}Cγ}$ 210.8$^{O\cdots H}$ 154.7$^{Cγ\text{-}Cδ}$ 153.1$^{Cδ\text{-}Cε}$ 146.7$^{Cε\text{-}Nζ}$
Lys-Si-z-H {H_2NH···N$^{\zeta}$}	-748.382658	15.25	-8.70	-0.04	8.66	124.9		149.0	154.5	191.3	188.8$^{Si\text{-}Cγ}$ 170.5$^{Nζ\cdots H}$ 153.9$^{Cγ\text{-}Cδ}$ 152.6$^{Cδ\text{-}Cε}$ 147.7$^{Cε\text{-}Nζ}$
Lys^{R+}-Si	-748.371301	31.55	-7.95	-0.07	7.88	125.4		147.2	154.0	189.5	188.4$^{Si\text{-}Cγ}$ 153.2$^{Cγ\text{-}Cδ}$ 151.9$^{Cδ\text{-}Cε}$ 149.9$^{Cε\text{-}Nζ}$
Lys^{R+}-Si-u	-748.830763	—	-8.69	-0.12	8.57	121.2	133.6	146.3	149.8	191.8	187.8$^{Si\text{-}Cγ}$ 153.3$^{Cγ\text{-}Cδ}$ 151.9$^{Cδ\text{-}Cε}$ 149.9$^{Cε\text{-}Nζ}$
Lys^{R+}-Si-z-N	-748.835546	—	-8.88	-0.14	8.74	124.8		149.7	154.6	191.3	187.6$^{Si\text{-}Cγ}$ 153.3$^{Cγ\text{-}Cδ}$ 152.0$^{Cδ\text{-}Cε}$ 149.8$^{Cε\text{-}Nζ}$
Lys^{R+}-Si-z-H {OCO···HN$^{\zeta}$}	-748.839805	—	-9.24	-0.16	9.08	124.3 125.6$^{C\text{-}O\cdots H}$		149.7	153.9	191.8	188.4$^{Si\text{-}Cγ}$ 181.9$^{O\cdots H}$ 153.9$^{Cγ\text{-}Cδ}$ 153.1$^{Cδ\text{-}Cε}$ 149.5$^{Cε\text{-}Nζ}$
Lys-Si^{Me}-u	-827.012182	3.64	-8.33	-0.08	8.25	121.4	133.9	146.4	149.3	193.4	188.0$^{Si\text{-}Cγ}$ 187.4$^{Si\text{-}Me}$ 153.4$^{Cγ\text{-}Cδ}$ 152.4$^{Cδ\text{-}Cε}$ 146.9$^{Cε\text{-}Nζ}$
Lys-Si^{Me}-z-N	-827.017546	14.85	-8.47	-0.10	8.37	125.0		150.0	154.2	192.5	188.0$^{Si\text{-}Cγ}$ 187.5$^{Si\text{-}Me}$ 153.5$^{Cγ\text{-}Cδ}$ 152.7$^{Cδ\text{-}Cε}$ 146.9$^{Cε\text{-}Nζ}$
Lys-Si^{Me}-z-H {OCO···HN$^{\zeta}$}	-827.015304	15.15	-8.19	-0.11	8.08	125.1 124.8$^{C\text{-}O\cdots H}$		150.0	154.1	193.8	188.0$^{Si\text{-}Cγ}$ 187.6$^{Si\text{-}Me}$ 209.8$^{O\cdots H}$ 154.7$^{Cγ\text{-}Cδ}$ 153.1$^{Cδ\text{-}Cε}$ 146.8$^{Cε\text{-}Nζ}$
Lys-Si^{Me}-z-H {H_2NH···N$^{\zeta}$}	-827.023681	15.63	-8.55	-0.09	8.46	125.0		149.4	154.1	192.5	189.6$^{Si\text{-}Cγ}$ 187.7$^{Si\text{-}Me}$ 171.3$^{Nζ\cdots H}$ 153.8$^{Cγ\text{-}Cδ}$ 152.7$^{Cδ\text{-}Cε}$ 147.5$^{Cε\text{-}Nζ}$
Lys^{R+}-Si^{Me}	-827.012425	31.85	-7.80	-0.12	7.68	125.6		147.4	153.5	190.7	188.9$^{Si\text{-}Cγ}$ 188.1$^{Si\text{-}Me}$ 153.2$^{Cγ\text{-}Cδ}$ 151.9$^{Cδ\text{-}Cε}$ 150.0$^{Cε\text{-}Nζ}$
Lys^{R+}-Si^{Me}-u	-827.475236	—	-8.55	-0.18	8.37	121.3	133.8	146.4	149.4	193.0	188.7$^{Si\text{-}Cγ}$ 187.2$^{Si\text{-}Me}$ 153.3$^{Cγ\text{-}Cδ}$ 151.9$^{Cδ\text{-}Cε}$ 149.9$^{Cε\text{-}Nζ}$
Lys^{R+}-Si^{Me}-z-N	-827.479896	—	-8.74	-0.19	8.55	124.9		149.9	154.3	192.9	188.2$^{Si\text{-}Cγ}$ 187.3$^{Si\text{-}Me}$ 153.3$^{Cγ\text{-}Cδ}$ 151.9$^{Cδ\text{-}Cε}$ 149.9$^{Cε\text{-}Nζ}$
Lys^{R+}-Si^{Me}-z-H {OCO···HN$^{\zeta}$}	-827.483744	—	-9.08	-0.17	8.91	124.6 125.6$^{C\text{-}O\cdots H}$		150.0	153.7	193.1	189.0$^{Si\text{-}Cγ}$ 187.0$^{Si\text{-}Me}$ 181.5$^{O\cdots H}$ 154.0$^{Cγ\text{-}Cδ}$ 153.0$^{Cδ\text{-}Cε}$ 149.5$^{Cε\text{-}Nζ}$
Tyr-Si-u	-881.352551	5.50	-7.84	-0.06	7.78	121.1	133.8	146.9	149.4	192.7	186.4$^{Si\text{-}Ph}$ 135.8$^{Ph\text{-}O}$
Tyr-Si-z	-881.356528	15.26	-7.86	-0.04	7.82	124.8		149.7	154.5	191.9	186.1$^{Si\text{-}Ph}$ 135.8$^{Ph\text{-}O}$
Tyr^{R+}-Si	-881.280095	42.01	-7.74	-0.51	7.23	125.4		146.9	153.9	190.4	189.0$^{Si\text{-}Ph}$ 145.3$^{Ph\text{-}O}$
Tyr^{R+}-Si-u	-881.740428	—	-8.35	-0.68	7.67	120.7	134.1	146.5	149.6	192.4	188.1$^{Si\text{-}Ph}$ 145.2$^{Ph\text{-}O}$
Tyr^{R+}-Si-z	-881.743120	—	-8.85	-0.60	8.25	124.7		149.7	154.7	190.7	187.9$^{Si\text{-}Ph}$ 145.3$^{Ph\text{-}O}$
Tyr-Si^{Me}-u	-959.997950	5.54	-7.74	-0.11	7.63	121.2	134.0	147.2	149.1	193.8	187.2$^{Si\text{-}Me}$ 187.6$^{Si\text{-}Ph}$ 136.0$^{Ph\text{-}O}$
Tyr-Si^{Me}-z	-960.002252	15.61	-7.76	-0.14	7.62	125.0		149.9	154.2	193.0	187.2$^{Si\text{-}Me}$ 187.1$^{Si\text{-}Ph}$ 136.0$^{Ph\text{-}O}$
Tyr^{R+}-Si^{Me}	-959.924384	41.09	-7.63	-0.38	7.25	125.5		147.3	153.7	191.2	187.4$^{Si\text{-}Me}$ 190.1$^{Si\text{-}Ph}$ 145.4$^{Ph\text{-}O}$
Tyr^{R+}-Si^{Me}-u	-960.386676	—	-8.15	-0.54	7.61	120.9	134.1	146.8	149.1	193.4	187.0$^{Si\text{-}Me}$ 189.2$^{Si\text{-}Ph}$ 145.3$^{Ph\text{-}O}$
Tyr^{R+}-Si^{Me}-z	-960.389651	—	-8.68	-0.54	8.12	124.8		149.8	154.3	192.5	187.0$^{Si\text{-}Me}$ 188.6$^{Si\text{-}Ph}$ 145.3$^{Ph\text{-}O}$

[a] The three-letter symbols for the common α-amino acids, plus the suffix "**-Si**", identify the parent β-sila-α-amino acid, the "**-Si^{Me}**" denotes a dimethylsila-group, and "-u" and "-z" indicate whether the α-amino and α-acid groups are uncharged or zwitterionic, respectively. When the side-chain bears a positive charge, a superscripted "R+" indicates this charge; without either "-u" or "-z" the molecule is a "charge-isomer" which bears no charge on the α-amino group, a negative charge on the α-carboxylic acid group, and a positive charge on the side-chain. For those molecules which may exhibit an intramolecular hydrogen-bond, an "-N" indicates the molecule without, and an "-H" with, such a bond; the connectivity of the hydrogen bond is noted in curly brackets and no hydrogen bonds were observed for the β-sila-cysteines or the β-sila-tyrosines.

[b] Mean values, where appropriate. Superscripts denote the X–Y atoms in the bond; "Me" indicates a methyl group.

[c] Dipole moments are undefined for a species bearing an overall positive charge.

[d] $E_{LUMO} - E_{HOMO}$.

Table 6. Selected electronic and structural properties of the β-sila-*N*-formylmethionines and the β-sila-selenocysteines and their dimethylsila-analogs

molecule[a]	electronic energy (au)	dipole moment (D)[c]	molecular orbital energies (eV) HOMO	LUMO	difference[d]	selected bond distances (pm)[b] C–O	C–O^{H}	C$^{\alpha}$–N	C$^{\alpha}$–C	C$^{\alpha}$–Si	X–Y
fMet-Si-u	-1165.242676	7.24	-7.54	0.00	7.54	120.8	133.4	145.1	149.8	193.1	187.8$^{Si\text{-}Cγ}$ 182.8$^{Cγ\text{-}S}$ 181.4$^{S\text{-}Cε}$ 134.5$^{N\text{-}CH}$ 122.4$^{C=O}$
fMet^{O-}-Si	-1164.788825	—	-7.37	0.17	7.54	125.2		145.3	154.0	191.1	188.4$^{Si\text{-}Cγ}$ 182.7$^{Cγ\text{-}S}$ 181.4$^{S\text{-}Cε}$ 133.5$^{N\text{-}CH}$ 123.0$^{C=O}$
fMet-Si^{Me}-u	-1243.889150	7.64	-7.42	-0.08	7.34	120.9	133.7	145.3	149.5	[194.3]	188.3$^{Si\text{-}Cγ}$ 183.1$^{Cγ\text{-}S}$ 186.9$^{Si\text{-}Me}$ 181.5$^{S\text{-}Cε}$ 134.3$^{N\text{-}CH}$ 122.4$^{C=O}$
fMet^{O-}-Si^{Me}	-1243.432360	—	-7.24	0.04	7.28	125.3		145.6	153.9	192.2	189.1$^{Si\text{-}Cγ}$ 182.9$^{Cγ\text{-}S}$ 187.6$^{Si\text{-}Me}$ 181.5$^{S\text{-}Cε}$ 133.4$^{N\text{-}CH}$ 123.0$^{C=O}$
Sec-Si-u	-2976.684462	3.49	-8.19	-0.06	8.13	120.9	133.6	146.5	150.0	192.0	230.1$^{Si\text{-}Se}$
Sec-Si-z	-2676.690712	12.58	-8.16	-0.18	7.98	124.6		149.3	154.7	190.4	230.3$^{Si\text{-}Se}$
Sec^{R+}-Si; **Sec^{R+}-Si**-u; **Sec^{R+}-Si**-z						[unable to be optimized]					
Sec-Si^{Me}-u	-3055.332306	4.54	-8.01	-0.13	7.88	121.0	133.8	146.8	149.6	192.8	[232.1]$^{Si\text{-}Se}$ 186.6$^{Si\text{-}Me}$
Sec-Si^{Me}-z	-3055.337828	11.53	-8.03	-0.16	7.87	124.8		149.5	154.6	191.7	[232.2]$^{Si\text{-}Se}$ 186.3$^{Si\text{-}Me}$
Sec^{R+}-Si^{Me}; **Sec^{R+}-Si^{Me}**-u; **Sec^{R+}-Si^{Me}**-z						[unable to be optimized]					

[a] The symbols for these two less-common β-sila-α-amino acids are **fMet-Si**, β-sila-*N*-formylmethionine, and **Sec-Si**, β-sila-selenocysteine; the "**Si^{Me}**" denotes a fully-methylated silicon atom and "-u" and "-z" indicate whether the α-amino and α-acid groups are uncharged or zwitterionic, respectively. When the side-chain bears a positive charge, a superscripted "R+" indicates this charge; without either "-u" or "-z" the molecule is a "charge-isomer" which bears no charge on the α-amino group, a negative charge on the α-carboxylic acid group, and a positive charge on the side-chain, except in the cases of **fMet^{O-}-Si** and **fMet^{O-}-Si^{Me}** wherein the superscripted "O-" without the "-u" or "-z" suffixes indicates the presence of an α-carboxylate group with the *N*-formyl group remaining neutral. No intramolecular hydrogen bonds were observed for these molecules.

[b] Mean values, where appropriate. Superscripts denote the X–Y atoms in the specific bond; "Me" indicates a methyl group. Bond distances in square brackets arbitrarily indicate Si–C distances approaching (>194.0 pm) or exceeding 194.9 pm, the sum of the tetrahedral covalent radii for C and Si, or Si–Se distances approaching (>231.0 pm) or exceeding 232.1 pm, the sum of the tetrahedral covalent radii for Si and Se.

[c] Dipole moments are undefined for a species bearing an overall positive or negative charge.

[d] $E_{LUMO} - E_{HOMO}$.

Table 7. Selected electronic and structural properties of the β-sila-pyrrolysines and the β-dimethylsila-pyrrolysines

molecule[a]	electronic energy (au)	dipole moment (D)[c]	molecular orbital energies (eV) HOMO	LUMO	difference[d]	selected bond distances (pm)[b] C–O	C–O^{H}	C^{α}–N	C^{α}–C	C^{α}–Si	X–Y
Pyl-Si-u	-1111.174556	5.37	-8.65	-0.13	8.52	121.1 123.1$^{C_{\eta}=O}$	133.6	146.4	149.7	192.2	187.7$^{Si\text{-}C_{\gamma}}$ 153.6$^{C_{\gamma}\text{-}C_{\delta}}$ 152.4$^{C_{\delta}\text{-}C_{\varepsilon}}$ 145.6$^{C_{\varepsilon}\text{-}N}$ 133.9$^{N\text{-}C_{\eta}}$ 152.0$^{C_{\eta}\text{-}C_{\beta}}$ 147.2$^{C_{\beta}\text{-}N}$ 127.1$^{N\text{-}C_{\kappa}}$ 150.3$^{C_{\kappa}\text{-}C_{\lambda}}$ 153.9$^{C_{\lambda}\text{-}C_{\mu}}$ 155.0$^{C_{\mu}\text{-}C_{\beta}}$ 152.2$^{C_{\mu}\text{-}Me}$
Pyl-Si-z-N	-1111.176627	10.73	-8.81	-0.11	8.70	124.8 122.9$^{C_{\eta}=O}$		149.8	154.6	191.7	187.8$^{Si\text{-}C_{\gamma}}$ 153.4$^{C_{\gamma}\text{-}C_{\delta}}$ 152.3$^{C_{\delta}\text{-}C_{\varepsilon}}$ 145.7$^{C_{\varepsilon}\text{-}N}$ 134.2$^{N\text{-}C_{\eta}}$ 152.1$^{C_{\eta}\text{-}C_{\beta}}$ 147.4$^{C_{\beta}\text{-}N}$ 127.2$^{N\text{-}C_{\kappa}}$ 150.4$^{C_{\kappa}\text{-}C_{\lambda}}$ 153.7$^{C_{\lambda}\text{-}C_{\mu}}$ 154.7$^{C_{\mu}\text{-}C_{\beta}}$ 152.1$^{C_{\mu}\text{-}Me}$
Pyl-Si-z-H {$H_2NH\cdots N^{Py}$}	-1111.187719	9.50	-8.72	-0.04	8.68	124.9 122.9$^{C_{\eta}=O}$		149.6	154.6	192.2	188.0$^{Si\text{-}C_{\gamma}}$ 153.6$^{C_{\gamma}\text{-}C_{\delta}}$ 152.6$^{C_{\delta}\text{-}C_{\varepsilon}}$ 145.9$^{C_{\varepsilon}\text{-}N}$ 134.1$^{N\text{-}C_{\eta}}$ 153.1$^{C_{\eta}\text{-}C_{\beta}}$ 148.0$^{C_{\beta}\text{-}N}$ 127.2$^{N\text{-}C_{\kappa}}$ 150.0$^{C_{\kappa}\text{-}C_{\lambda}}$ 154.2$^{C_{\lambda}\text{-}C_{\mu}}$ 154.8$^{C_{\mu}\text{-}C_{\beta}}$ 152.7$^{C_{\mu}\text{-}Me}$ 184.3$^{H\cdots N}$
Pyl-Si-z-H {OCO···HN}	1111.185472	13.09	-8.64	-0.10	8.54	125.0 124.8$^{CO\cdots H}$ 123.1$^{C_{\eta}=O}$		150.1	154.3	193.1	188.2$^{Si\text{-}C_{\gamma}}$ 153.6$^{C_{\gamma}\text{-}C_{\delta}}$ 152.7$^{C_{\delta}\text{-}C_{\varepsilon}}$ 145.8$^{C_{\varepsilon}\text{-}N}$ 134.2$^{N\text{-}C_{\eta}}$ 152.2$^{C_{\eta}\text{-}C_{\beta}}$ 147.2$^{C_{\beta}\text{-}N}$ 127.1$^{N\text{-}C_{\kappa}}$ 150.3$^{C_{\kappa}\text{-}C_{\lambda}}$ 153.9$^{C_{\lambda}\text{-}C_{\mu}}$ 154.7$^{C_{\mu}\text{-}C_{\beta}}$ 152.3$^{C_{\mu}\text{-}Me}$ 199.8$^{H\cdots O}$
Pyl-Si-z-H {$H_2NH\cdots OC$}	-1111.183488	20.40	-8.75	-0.09	8.66	124.8 123.9$^{C_{\eta}=O}$		149.8	155.5	191.7	188.0$^{Si\text{-}C_{\gamma}}$ 154.0$^{C_{\gamma}\text{-}C_{\delta}}$ 153.0$^{C_{\delta}\text{-}C_{\varepsilon}}$ 145.5$^{C_{\varepsilon}\text{-}N}$ 133.3$^{N\text{-}C_{\eta}}$ 152.3$^{C_{\eta}\text{-}C_{\beta}}$ 147.0$^{C_{\beta}\text{-}N}$ 127.2$^{N\text{-}C_{\kappa}}$ 150.3$^{C_{\kappa}\text{-}C_{\lambda}}$ 151.3$^{C_{\lambda}\text{-}C_{\mu}}$ 155.4$^{C_{\mu}\text{-}C_{\beta}}$ 152.7$^{C_{\mu}\text{-}Me}$ 174.7$^{H\cdots O}$
Pyl^{R+}**-Si**	-1111.160027	56.38	-7.94	-1.08	6.86	125.4 122.6$^{C_{\eta}=O}$		147.2	153.8	189.4	188.6$^{Si\text{-}C_{\gamma}}$ 153.2$^{C_{\gamma}\text{-}C_{\delta}}$ 152.3$^{C_{\delta}\text{-}C_{\varepsilon}}$ 146.0$^{C_{\varepsilon}\text{-}N}$ 133.6$^{N\text{-}C_{\eta}}$ 152.5$^{C_{\eta}\text{-}C_{\beta}}$ 147.3$^{C_{\beta}\text{-}N}$ 127.7$^{N\text{-}C_{\kappa}}$ 148.2$^{C_{\kappa}\text{-}C_{\lambda}}$ 154.3$^{C_{\lambda}\text{-}C_{\mu}}$ 154.8$^{C_{\mu}\text{-}C_{\beta}}$ 152.0$^{C_{\mu}\text{-}Me}$
Pyl^{R+}**-Si**-u	-1111.620449	—	8.64	-1.22	7.42	121.2 122.5$^{C_{\eta}=O}$	133.6	146.5	149.7	192.2	187.4$^{Si\text{-}C_{\gamma}}$ 153.3$^{C_{\gamma}\text{-}C_{\delta}}$ 153.1$^{C_{\delta}\text{-}C_{\varepsilon}}$ 146.0$^{C_{\varepsilon}\text{-}N}$ 134.1$^{N\text{-}C_{\eta}}$ 152.3$^{C_{\eta}\text{-}C_{\beta}}$ 147.6$^{C_{\beta}\text{-}N}$ 128.0$^{N\text{-}C_{\kappa}}$ 148.1$^{C_{\kappa}\text{-}C_{\lambda}}$ 154.2$^{C_{\lambda}\text{-}C_{\mu}}$ 154.9$^{C_{\mu}\text{-}C_{\beta}}$ 151.9$^{C_{\mu}\text{-}Me}$
Pyl^{R+}**-Si**-z-N	-1111.624386	—	-8.83	-1.10	7.73	124.8 122.5$^{C_{\eta}=O}$		149.8	154.6	191.8	187.9$^{Si\text{-}C_{\gamma}}$ 153.4$^{C_{\gamma}\text{-}C_{\delta}}$ 152.3$^{C_{\delta}\text{-}C_{\varepsilon}}$ 146.0$^{C_{\varepsilon}\text{-}N}$ 133.6$^{N\text{-}C_{\eta}}$ 152.5$^{C_{\eta}\text{-}C_{\beta}}$ 147.3$^{C_{\beta}\text{-}N}$ 127.7$^{N\text{-}C_{\kappa}}$ 158.2$^{C_{\kappa}\text{-}C_{\lambda}}$ 154.3$^{C_{\lambda}\text{-}C_{\mu}}$ 154.8$^{C_{\mu}\text{-}C_{\beta}}$ 152.0$^{C_{\mu}\text{-}Me}$
Pyl^{R+}**-Si**-z-H {$OCO\cdots HN^{Py}$}	-1111.641173	—	-8.96	-0.67	8.29	123.7 126.6$^{CO\cdots H}$ 122.5$^{C_{\eta}=O}$		149.7	153.4	192.7	187.1$^{Si\text{-}C_{\gamma}}$ 153.4$^{C_{\gamma}\text{-}C_{\delta}}$ 153.2$^{C_{\delta}\text{-}C_{\varepsilon}}$ 145.9$^{C_{\varepsilon}\text{-}N}$ 134.0$^{N\text{-}C_{\eta}}$ 153.1$^{C_{\eta}\text{-}C_{\beta}}$ 147.7$^{C_{\beta}\text{-}N}$ 127.4$^{N\text{-}C_{\kappa}}$ 148.7$^{C_{\kappa}\text{-}C_{\lambda}}$ 154.3$^{C_{\lambda}\text{-}C_{\mu}}$ 154.0$^{C_{\mu}\text{-}C_{\beta}}$ 152.1$^{C_{\mu}\text{-}Me}$ 149.5$^{O\cdots H}$
Pyl^{R+}**-Si**-z-H {OCO···HN}	-1111.631782	—	-8.99	-1.08	7.91	123.9 126.0$^{CO\cdots H}$ 122.8$^{C_{\eta}=O}$		149.9	154.1	192.8	187.5$^{Si\text{-}C_{\gamma}}$ 154.0$^{C_{\gamma}\text{-}C_{\delta}}$ 152.6$^{C_{\delta}\text{-}C_{\varepsilon}}$ 145.9$^{C_{\varepsilon}\text{-}N}$ 133.2$^{N\text{-}C_{\eta}}$ 153.2$^{C_{\eta}\text{-}C_{\beta}}$ 147.8$^{C_{\beta}\text{-}N}$ 127.8$^{N\text{-}C_{\kappa}}$ 148.2$^{C_{\kappa}\text{-}C_{\lambda}}$ 154.2$^{C_{\lambda}\text{-}C_{\mu}}$ 154.2$^{C_{\mu}\text{-}C_{\beta}}$ 152.0$^{C_{\mu}\text{-}Me}$ 187.5$^{O\cdots H}$
Pyl^{R+}**-Si**-z-H {$H_2NH\cdots OC$ + $OCO\cdots HN^{Py}$}	-1111.637417	—	-9.02	-0.93	8.09	123.9 126.2$^{CO\cdots H}$ 123.4$^{C_{\eta}=O}$		149.6	153.6	193.7	188.3$^{Si\text{-}C_{\gamma}}$ 153.4$^{C_{\gamma}\text{-}C_{\delta}}$ 154.4$^{C_{\delta}\text{-}C_{\varepsilon}}$ 146.2$^{C_{\varepsilon}\text{-}N}$ 133.0$^{N\text{-}C_{\eta}}$ 153.0$^{C_{\eta}\text{-}C_{\beta}}$ 147.3$^{C_{\beta}\text{-}N}$ 127.5$^{N\text{-}C_{\kappa}}$ 148.6$^{C_{\kappa}\text{-}C_{\lambda}}$ 154.4$^{C_{\lambda}\text{-}C_{\mu}}$ 154.5$^{C_{\mu}\text{-}C_{\beta}}$ 167.2$^{O\cdots H}$ 152.0$^{C_{\mu}\text{-}Me}$ 179.3$^{NH\cdots O}$
Pyl-Si^{Me}-u	-1189.819050	4.16	-8.51	-0.16	8.35	121.4 123.2$^{C_{\eta}=O}$	133.8	146.4	149.3	193.6	188.6$^{Si\text{-}C_{\gamma}}$ 187.2$^{Si\text{-}Me}$ 153.5$^{C_{\gamma}\text{-}C_{\delta}}$ 152.5$^{C_{\delta}\text{-}C_{\varepsilon}}$ 145.7$^{C_{\varepsilon}\text{-}N}$ 133.9$^{N\text{-}C_{\eta}}$ 152.0$^{C_{\eta}\text{-}C_{\beta}}$ 147.2$^{C_{\beta}\text{-}N}$ 127.1$^{N\text{-}C_{\kappa}}$ 150.3$^{C_{\kappa}\text{-}C_{\lambda}}$ 153.9$^{C_{\lambda}\text{-}C_{\mu}}$ 155.0$^{C_{\mu}\text{-}C_{\beta}}$ 152.2$^{C_{\mu}\text{-}Me}$
Pyl-Si^{Me}-z-N	-1189.821648	10.62	-8.67	-0.12	8.55	125.0 123.0$^{C_{\eta}=O}$		150.0	154.2	192.6	188.6$^{Si\text{-}C_{\gamma}}$ 187.2$^{Si\text{-}Me}$ 153.4$^{C_{\gamma}\text{-}C_{\delta}}$ 152.3$^{C_{\delta}\text{-}C_{\varepsilon}}$ 145.7$^{C_{\varepsilon}\text{-}N}$ 134.2$^{N\text{-}C_{\eta}}$ 152.1$^{C_{\eta}\text{-}C_{\beta}}$ 147.4$^{C_{\beta}\text{-}N}$ 127.2$^{N\text{-}C_{\kappa}}$ 150.5$^{C_{\kappa}\text{-}C_{\lambda}}$ 153.7$^{C_{\lambda}\text{-}C_{\mu}}$ 154.7$^{C_{\mu}\text{-}C_{\beta}}$ 152.1$^{C_{\mu}\text{-}Me}$
Pyl-Si^{Me}-z-H {$H_2NH\cdots N^{Py}$}	-1189.827321	11.09	-8.69	-0.10	8.59	125.0 123.0$^{C_{\eta}=O}$		149.7	154.2	193.8	188.7$^{Si\text{-}C_{\gamma}}$ 188.2$^{Si\text{-}Me}$ 153.7$^{C_{\gamma}\text{-}C_{\delta}}$ 152.5$^{C_{\delta}\text{-}C_{\varepsilon}}$ 145.8$^{C_{\varepsilon}\text{-}N}$ 134.0$^{N\text{-}C_{\eta}}$ 153.0$^{C_{\eta}\text{-}C_{\beta}}$ 148.2$^{C_{\beta}\text{-}N}$ 127.2$^{N\text{-}C_{\kappa}}$ 150.1$^{C_{\kappa}\text{-}C_{\lambda}}$ 154.2$^{C_{\lambda}\text{-}C_{\mu}}$ 154.4$^{C_{\mu}\text{-}C_{\beta}}$ 152.7$^{C_{\mu}\text{-}Me}$ 182.7$^{H\cdots N}$
Pyl-Si^{Me}-z-H {OCO···HN}	-1189.823751	17.74	-8.56	-0.13	8.43	124.7 125.4$^{CO\cdots H}$ 123.1$^{C_{\eta}=O}$		149.8	153.7	193.6	189.4$^{Si\text{-}C_{\gamma}}$ 187.6$^{Si\text{-}Me}$ 154.4$^{C_{\gamma}\text{-}C_{\delta}}$ 152.8$^{C_{\delta}\text{-}C_{\varepsilon}}$ 145.5$^{C_{\varepsilon}\text{-}N}$ 134.1$^{N\text{-}C_{\eta}}$ 153.0$^{C_{\eta}\text{-}C_{\beta}}$ 148.4$^{C_{\beta}\text{-}N}$ 127.2$^{N\text{-}C_{\kappa}}$ 150.4$^{C_{\kappa}\text{-}C_{\lambda}}$ 153.7$^{C_{\lambda}\text{-}C_{\mu}}$ 153.6$^{C_{\mu}\text{-}C_{\beta}}$ 152.2$^{C_{\mu}\text{-}Me}$ 189.7$^{H\cdots O}$
Pyl-Si^{Me}-z-H {$H_2NH\cdots OC$}	-1189.826275	20.05	-8.61	-0.13	8.48	124.9 123.9$^{C_{\eta}=O}$		150.1	154.9	193.4	188.3$^{Si\text{-}C_{\gamma}}$ 187.6$^{Si\text{-}Me}$ 154.0$^{C_{\gamma}\text{-}C_{\delta}}$ 153.0$^{C_{\delta}\text{-}C_{\varepsilon}}$ 145.6$^{C_{\varepsilon}\text{-}N}$ 133.2$^{N\text{-}C_{\eta}}$ 152.3$^{C_{\eta}\text{-}C_{\beta}}$ 147.0$^{C_{\beta}\text{-}N}$ 127.2$^{N\text{-}C_{\kappa}}$ 150.2$^{C_{\kappa}\text{-}C_{\lambda}}$ 154.0$^{C_{\lambda}\text{-}C_{\mu}}$ 155.5$^{C_{\mu}\text{-}C_{\beta}}$ 152.7$^{C_{\mu}\text{-}Me}$ 175.6$^{H\cdots O}$
Pyl^{R+}**-Si**^{Me}	-1189.801553	57.03	-7.75	-1.08	6.67	125.6 122.6$^{C_{\eta}=O}$		147.5	153.5	190.6	189.2$^{Si\text{-}C_{\gamma}}$ 188.0$^{Si\text{-}Me}$ 153.2$^{C_{\gamma}\text{-}C_{\delta}}$ 152.3$^{C_{\delta}\text{-}C_{\varepsilon}}$ 146.1$^{C_{\varepsilon}\text{-}N}$ 133.5$^{N\text{-}C_{\eta}}$ 152.5$^{C_{\eta}\text{-}C_{\beta}}$ 147.3$^{C_{\beta}\text{-}N}$ 127.7$^{N\text{-}C_{\kappa}}$ 148.2$^{C_{\kappa}\text{-}C_{\lambda}}$ 154.3$^{C_{\lambda}\text{-}C_{\mu}}$ 154.8$^{C_{\mu}\text{-}C_{\beta}}$ 152.0$^{C_{\mu}\text{-}Me}$
Pyl^{R+}**-Si**^{Me}-u	-1190.264960	—	-8.49	-1.22	7.27	121.3 122.5$^{C_{\eta}=O}$	133.9	146.6	149.3	193.4	188.3$^{Si\text{-}C_{\gamma}}$ 187.3$^{Si\text{-}Me}$ 153.3$^{C_{\gamma}\text{-}C_{\delta}}$ 153.0$^{C_{\delta}\text{-}C_{\varepsilon}}$ 146.0$^{C_{\varepsilon}\text{-}N}$ 134.1$^{N\text{-}C_{\eta}}$ 152.4$^{C_{\eta}\text{-}C_{\beta}}$ 147.6$^{C_{\beta}\text{-}N}$ 127.9$^{N\text{-}C_{\kappa}}$ 148.1$^{C_{\kappa}\text{-}C_{\lambda}}$ 154.2$^{C_{\lambda}\text{-}C_{\mu}}$ 155.0$^{C_{\mu}\text{-}C_{\beta}}$ 151.8$^{C_{\mu}\text{-}Me}$
Pyl^{R+}**-Si**^{Me}-z-N	-1190.269443	—	-8.69	-1.09	7.60	125.0 122.6$^{C_{\eta}=O}$		149.9	154.2	192.9	188.4$^{Si\text{-}C_{\gamma}}$ 187.3$^{Si\text{-}Me}$ 153.4$^{C_{\gamma}\text{-}C_{\delta}}$ 152.3$^{C_{\delta}\text{-}C_{\varepsilon}}$ 146.0$^{C_{\varepsilon}\text{-}N}$ 133.6$^{N\text{-}C_{\eta}}$ 152.5$^{C_{\eta}\text{-}C_{\beta}}$ 147.3$^{C_{\beta}\text{-}N}$ 127.7$^{N\text{-}C_{\kappa}}$ 148.2$^{C_{\kappa}\text{-}C_{\lambda}}$ 154.4$^{C_{\lambda}\text{-}C_{\mu}}$ 154.6$^{C_{\mu}\text{-}C_{\beta}}$ 152.0$^{C_{\mu}\text{-}Me}$
Pyl^{R+}**-Si**^{Me}-z-H {$OCO\cdots HN^{Py}$}	-1190.285252	—	-8.87	-0.64	8.23	123.8 126.7$^{CO\cdots H}$ 122.5$^{C_{\eta}=O}$		149.9	153.2	192.7	187.6$^{Si\text{-}C_{\gamma}}$ 187.4$^{Si\text{-}Me}$ 153.3$^{C_{\gamma}\text{-}C_{\delta}}$ 153.2$^{C_{\delta}\text{-}C_{\varepsilon}}$ 146.0$^{C_{\varepsilon}\text{-}N}$ 134.0$^{N\text{-}C_{\eta}}$ 153.1$^{C_{\eta}\text{-}C_{\beta}}$ 147.7$^{C_{\beta}\text{-}N}$ 127.4$^{N\text{-}C_{\kappa}}$ 148.7$^{C_{\kappa}\text{-}C_{\lambda}}$ 154.2$^{C_{\lambda}\text{-}C_{\mu}}$ 154.0$^{C_{\mu}\text{-}C_{\beta}}$ 152.1$^{C_{\mu}\text{-}Me}$ 147.9$^{O\cdots H}$
Pyl^{R+}**-Si**^{Me}-z-H {OCO···HN}	-1190.277730	—	-8.84	-1.06	7.78	124.0 126.2$^{CO\cdots H}$ 122.8$^{C_{\eta}=O}$		149.9	153.6	193.9	188.1$^{Si\text{-}C_{\gamma}}$ 188.0$^{Si\text{-}Me}$ 153.9$^{C_{\gamma}\text{-}C_{\delta}}$ 152.4$^{C_{\delta}\text{-}C_{\varepsilon}}$ 146.1$^{C_{\varepsilon}\text{-}N}$ 133.2$^{N\text{-}C_{\eta}}$ 152.4$^{C_{\eta}\text{-}C_{\beta}}$ 147.3$^{C_{\beta}\text{-}N}$ 127.9$^{N\text{-}C_{\kappa}}$ 148.2$^{C_{\kappa}\text{-}C_{\lambda}}$ 154.3$^{C_{\lambda}\text{-}C_{\mu}}$ 154.8$^{C_{\mu}\text{-}C_{\beta}}$ 151.9$^{C_{\mu}\text{-}Me}$ 188.4$^{O\cdots H}$
Pyl^{R+}**-Si**^{Me}-z-H {$H_2NH\cdots OC$ + $OCO\cdots HN^{Py}$}	-1190.278301	—	-8.90	-0.95	7.95	124.2 125.9$^{CO\cdots H}$ 123.4$^{C_{\eta}=O}$		149.7	153.2	[194.2]	189.5$^{Si\text{-}C_{\gamma}}$ 187.9$^{Si\text{-}Me}$ 153.2$^{C_{\gamma}\text{-}C_{\delta}}$ 153.4$^{C_{\delta}\text{-}C_{\varepsilon}}$ 146.0$^{C_{\varepsilon}\text{-}N}$ 133.1$^{N\text{-}C_{\eta}}$ 152.5$^{C_{\eta}\text{-}C_{\beta}}$ 147.5$^{C_{\beta}\text{-}N}$ 127.6$^{N\text{-}C_{\kappa}}$ 148.5$^{C_{\kappa}\text{-}C_{\lambda}}$ 154.4$^{C_{\lambda}\text{-}C_{\mu}}$ 154.1$^{C_{\mu}\text{-}C_{\beta}}$ 152.0$^{C_{\mu}\text{-}Me}$ 175.7$^{O\cdots H}$ 180.1$^{NH\cdots O}$

portions of the molecules, as these positions will need to remain unsubstituted to potentially allow the polymerization of these β-sila-α-amino acids into peptides or proteins.

No methylation sites on the side-chains, save for methylation at the silicon atom, were investigated, with the exceptions of β-sila-serine (**Ser-Si**) and β-sila-threonine (**Thr-Si**). The substitution of silicon for carbon at the β-position in these β-sila-α-amino acids will result in the creation of silanols, molecules containing an Si–OH moiety. Such silanols could be reactive[28] and, thus, methylation at the β-sila-serine and β-sila-threonine hydroxy groups was also investigated with the aim of specifically broadening the potential for stability in these two β-sila-α-amino acids. However, despite the potential for reactivity, an α-amino acid containing a silane**di**ol group — 2-amino-6-dihydroxymethylsilylhexanoic acid[29] — has been synthesized and characterized. In this molecule, the silicon atom has replaced a carbon atom at the ζ-position, and the molecule was found to be both soluble and stable in water despite containing two OH groups attached to the silicon atom. Above pH ~11, the molecule was monomeric, but existed in an equilibrium between the monomer and the dimer at lower pH levels. Thus, free silanol-containing β-sila-α-amino acids may also be realistic synthetic targets.

In all cases, the zwitterionic [R–CH(COO^{-})(NH_3^{+})] forms of the molecules were calculated to be more stable in a simulated aqueous solution than were the uncharged isomers [R–CH(COOH)(NH_2)]. Further, the alternative zwitterionic "charge isomers" in which the molecules are neutral overall but carry a charge on the side-chain — either in the form [R^{-}–CH(COOH)(NH_3^{+})] or [R^{+}–CH(COO^{-})(NH_2)] — were generally (*vide infra*) found to be less stable than the zwitterionic molecules and were expected to rearrange to the more stable zwitterionic forms.

The carbon-silicon single bond lengths in the β-sila-α-amino acids were unremarkable, except for the Si–C^{α} bonds. In general, experimental Si–C bond lengths range between 184 pm and 188 pm, *i.e.*, $Cl_2Si(CH_3)_2$, 185.8 pm (electron diffraction[30]) and 185.0 pm (microwave spectra and electron diffraction[31]); $ClSi(CH_3)_3$, 184.3 pm (X-ray diffraction[32]); $Si(CH_3)_4$, 187.5 pm (electron diffraction[33]); and $Si_2(CH_3)_6$, 187.7 pm (electron diffraction[33]); the average Si–C bond distance has been reported[8] to be ~187 pm. For carbon and silicon, the tetrahedral covalent radii have been determined[34] to be 77.3 pm and 117.6 ppm, respectively, resulting in a sum of 194.9 pm for a Si–C bond. In those molecules in which the Si–C^{α} bonds exceeded the sum of the tetrahedral covalent radii of carbon and silicon, the molecules were the uncharged systems rather than the zwitterionic molecules that were calculated to exist in aqueous solution, except for two β-dimethylsila-pyrrolysines (*vide infra*). Thus, although the Si–C^{α} bonds were often calculated to be slightly longer than expected (at ~190 pm), the bonds generally remained shorter than the sum of the tetrahedral covalent radii for carbon and silicon. Such longer bonds would be expected to be weaker and more easily broken, however, and may require the

[a] The three-letter symbol for β-sila-pyrrolysine is **Pyl-Si**; the "-**Si**^{Me}" denotes a fully-methylated silicon atom and "-u" and "-z" indicate whether the α-amino and α-acid groups are uncharged or zwitterionic, respectively. When the side-chain bears a positive charge, a superscripted "R+" indicates this charge; without either "-u" or "-z" the molecule is a "charge-isomer" which bears no charge on the α-amino group, a negative charge on the α-carboxylic acid group, and a positive charge on the side-chain. For molecules in which intramolecular hydrogen bonds were observed, the "-H" indicates the molecules with, and the "-N" without, such a bond; the hydrogen bond geometries are given in curly brackets (the "N^{Py}" indicates the nitrogen atom in the five-membered pyrroline ring).

[b] Mean values, where appropriate. Superscripts denote the X–Y atoms in the specific bond; "Me" indicates a methyl group. Bond distances in square brackets arbitrarily indicate Si–C distances approaching (>194.0 pm) or exceeding 194.9 pm.

[c] Dipole moments are undefined for a species bearing an overall positive or negative charge.

[d] $E_{LUMO} - E_{HOMO}$.

addition of methyl groups on silicon to sterically inhibit reactivity at the Si–C$^{\alpha}$ bond. However, such methylation did not significantly alter the calculated bond lengths for the methylated molecules in comparison to the parent, un-methylated molecules. Further, the methylated analogs often were calculated to exhibit structures very similar to the unmethylated parent molecules, despite the methyl group steric impacts.

The Mulliken bond orders[35] for Si–C bonds can be compared to the bond lengths in these molecules to provide an estimate of the relative strengths of the bonds. For example, in the zwitterionic β-dimethylsila-alanine (**Ala-Si**^{Me}-z), the Si–C$^{\alpha}$ bond, with a length of 193.0 pm, was calculated to exhibit a Mulliken bond order of 0.83, while the Si–CH_3 bonds, with an average length of 187.3 pm, had an average bond order of 0.92. Similar values were observed throughout the β-sila-α-amino acids investigated. However, when the calculated bond lengths exceeded ~194 pm, such as with the charge-isomer of β-sila-glutamic acid (**Glu**$^{R-}$**-Si**) in which the molecule contains a side-chain with a negative charge, a positively-charged α-ammonium group and a neutral α-carboxylic acid group, the Mulliken bond order decreased sharply. In this case, the Si–C$^{\alpha}$ bond length of 196.2 pm was calculated to have a Mulliken bond order of only 0.41. Such weak bonds would be expected to be reactive — or even cleaved — in an aqueous solution, perhaps again underlining the case for methylation at the silicon center, as such methylation is known[8] to stabilize these molecules. However, in the methylated charge isomer (**Glu**$^{R-}$**-Si**^{Me}), for example, the bond distance was only marginally shorter at 195.8 pm with only a slightly-increased Mulliken bond order of 0.49. (The neutral charge-isomers of β-sila-glutamic acid were less stable than the zwitterionic forms, however, indicating that the zwitterionic forms would be the isomers present in solution.)

β-Sila-α-Amino Acids Containing Neutral Side-Chains

Only two structures needed to be investigated for most of the β-sila-α-amino acids derived from those α-amino acids bearing a neutral side-chain, and two structures for each of the permethylsila-analogs, as listed in Table 1. These twelve β-sila-α-amino acids — β-sila-alanine (**Ala-Si**), β-sila-asparagine (**Asn-Si**), β-sila-glutamine (**Gln-Si**), β-sila-isoleucine (**Ile-Si**), β-sila-leucine (**Leu-Si**), β-sila-methionine (**Met-Si**), β-sila-phenylalanine (**Phe-Si**), β-sila-proline (**Pro-Si**), β-sila-serine (**Ser-Si**), β-sila-threonine (**Thr-Si**), β-sila-tryptophan (**Trp-Si**), β-sila-valine (**Val-Si**), and their permethylsila-analogs (designated "**-Si**^{Me}") — were each calculated as both the uncharged ("-u") molecules [R–CH(COOH)(NH_2)] and the zwitterionic ("-z") molecules [R–CH(COO^-)(NH_3^+)]. In addition, β-sila-serine and β-sila-threonine were investigated as the methoxy-analogs (**Ser**^{OMe}**-Si**, **Ser**^{OMe}**-Si**^{Me}, and **Thr**^{OMe}**-Si**; **Thr**^{OMe}**-Si**^{Me} ≡ **Ser**^{OMe}**-Si**^{Me}) because of the potential reactivity[28] of the Si–OH moiety in these molecules.

Note that although the parent α-amino acids leucine and isoleucine are isomers, and **Leu-Si** and **Ile-Si** are isomers, **Leu-Si**^{Me} and **Ile-Si**^{Me} are not isomers: Only a single methylation was required for the former, while two methyl groups needed to be added to the silicon center for the latter. The uncharged and zwitterionic β-sila-isoleucines (**Ile-Si**-u and **Ile-Si**-z) were calculated to be more stable than the corresponding isomeric uncharged and zwitterionic β-sila-leucines (**Leu-Si**-u and **Leu-Si**-z).

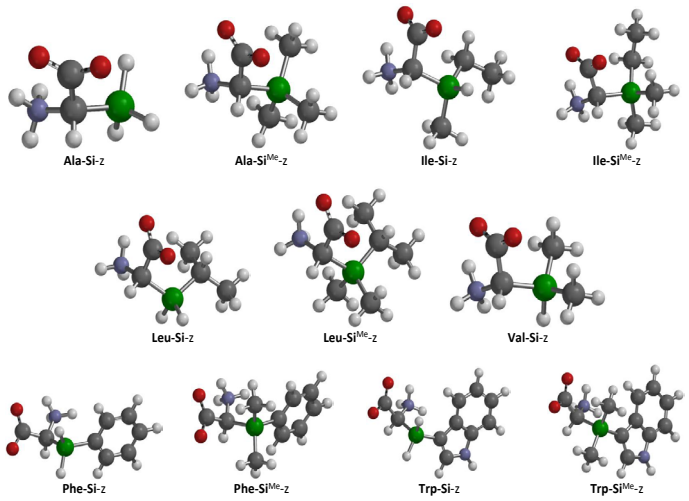


Figure 1. The zwitterionic β-sila-α-amino acids and their permethylsila-analogs with neutral alkyl or aryl side-chains (Val-Si^{Me}-z ⟶ Ala-Si^{Me}-z). Atom key: H, light gray; C, dark gray; N, blue; O, red; Si, green.

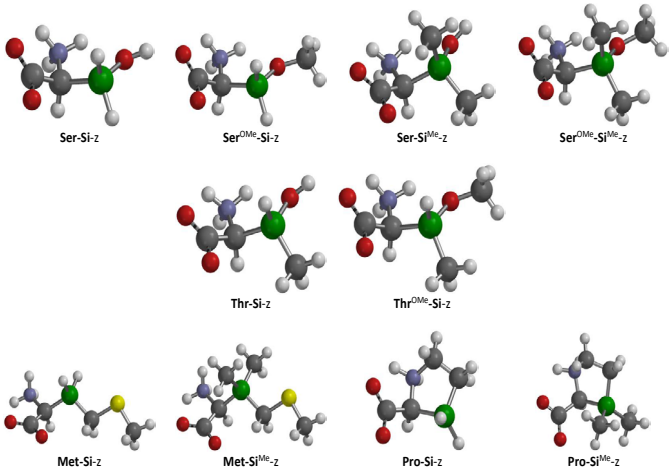


Figure 2. The zwitterionic β-sila-α-amino acids and their permethylsila- and/or methoxy-analogs with oxygen- or sulfur-containing side-chains, including the zwitterionic β-sila-prolines (Thr-Si^{Me}-z ⟶ Ser-Si^{Me}-z and Thr^{OMe}-Si^{Me}-z ⟶ Ser^{OMe}-Si^{Me}-z). Atom key: H, light gray; C, dark gray; N, blue; O, red; S, yellow; Si, green.

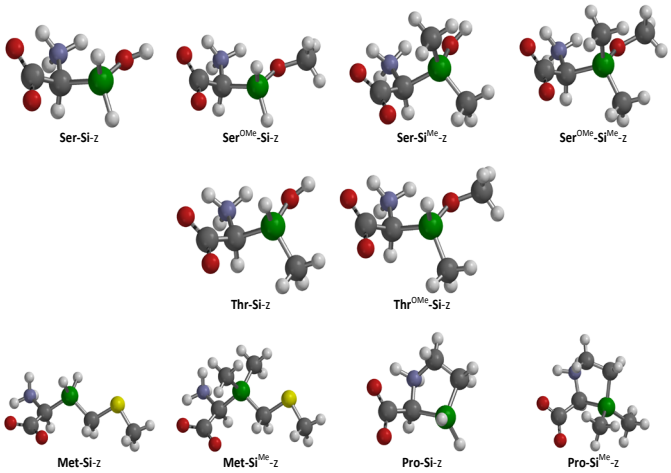


Figure 3. The zwitterionic β-sila-α-amino acids and their dimethylsila-analogs with neutral side-chains that exhibited intramolecular hydrogen bonds. Atom key: H, light gray; C, dark gray; N, blue; O, red; Si, green. The thin lines indicate the positions of the hydrogen bonds.

For each of these twelve β-sila-α-amino acids with neutral side-chains, the zwitterionic form was found to be more stable than the uncharged form, as expected. In Fig. 1, the zwitterionic structures of the β-sila-α-amino acids containing simple alkyl or aryl neutral side-chains are depicted, while Fig. 2 shows the zwitterionic β-sila-α-amino acids with oxygen or sulfur in the neutral side-chains, along with proline. The molecules illustrated in Fig. 3 are those β-sila-α-amino acids containing hetero-atoms in the neutral side-chain which exhibited intramolecular hydrogen bonds: the β-sila-asparagines and β-sila-glutamines.

Interestingly, semiempirical conformational analyses, which were then confirmed at the M06-2X/6-311++G** level of theory, indicated that zwitterionic β-dimethylsila-asparagine could exist as a molecule with an intramolecular hydrogen bond between a hydrogen atom of the positively-charged α-amino group and the amide carbonyl group, **Asn-Si**^{Me}-z-H{H_2NH···OC}, or between the negatively-charged α-carboxylate and the amide NH_2 group, **Asn-Si**^{Me}-z-H{OCO···HNH}. (Note that the hydrogen bond connectivity in the intramolecularly hydrogen-bonded molecules is given in curly brackets.) Although the conformational analyses did not *consistently* indicate the presence of intramolecular hydrogen bonds in other zwitterionic molecules, these results initiated the additional investigations of the potential for intramolecular hydrogen bonding in zwitterionic β-sila-asparagine, **Asn-Si**-z, β-sila-glutamine, **Gln-Si**-z, and β-dimethylsila-glutamine, **Gln-Si**^{Me}-z, for similar structures, and in other β-sila-α-amino acids (*vide infra*). The intramolecular hydrogen bonds between the α-carboxylate and the hydrogen of the amide NH_2-group and those between the α-ammonium and the oxygen of the amide C=O group were calculated to be stable minima on the relevant potential energy surfaces for each of the zwitterionic β-sila-asparagines and the zwitterionic β-sila-glutamines. (However, no hydrogen bonding was calculated to exist between the α-ammonium group and the nitrogen of the amide NH_2 group.) The properties of these hydrogen-bonded molecules are listed in Table 1 and illustrated in Fig. 3.

For both the hydrogen-bonded β-sila-asparagines and β-sila-glutamines, the molecules with the intramolecular hydrogen bond between the positively-charged α-ammonium and the amide carbonyl group were more stable than those molecules with an intramolecular hydrogen bond between the α-carboxylate and the amide –NH_2 groups. Thus, calculations indicated that **Asn-Si**-z-H{H_2NH···OC} was more stable than **Asn-Si**-z-H{OCO···HNH} by 1.7 kJ/mol and **Asn-Si**^{Me}-z-H{H_2NH···OC} was more stable than **Asn-Si**^{Me}-z-H{OCO···HNH} by 3.7 kJ/mol, while **Gln-Si**-z-H{H_2NH···OC} was more stable than **Gln-Si**-z-H{OCO···HNH} by 11.3 kJ/mol and **Gln-Si**^{Me}-z-H{H_2NH···OC} was more stable than **Gln-Si**^{Me}-z-H{OCO···HNH} by 5.4 kJ/mol. The hydrogen-bond distances were shorter and the Mulliken bond orders (in parentheses) greater, in the β-sila-asparagines — **Asn-Si**-z-H{H_2NH···OC}, 186.5 pm (0.10); **Asn-Si**-z-H{OCO···HNH}, 191.2 pm (0.09); **Asn-Si**^{Me}-z-H{H_2NH···OC}, 185.6 pm (0.12); **Asn-Si**^{Me}-z-H{OCO···HNH}, 188.5 pm (0.09) — than in the corresponding β-sila-glutamines — **Gln-Si**-z-H{H_2NH···OC}, 191.7 pm (0.05); **Gln-Si**-z-H{OCO···HNH}, 193.0 pm (0.04); **Gln-Si**^{Me}-z-H{H_2NH···OC}, 192.2 pm (0.05); **Gln-Si**^{Me}-z-H{OCO···HNH}, 191.8 pm (0.04) — although the Mulliken bond orders were calculated to be small for all of these molecules. The shortening of bond distances may be due to the ring-size differences in the intramolecularly-hydrogen-bonded molecules, as the β-sila-asparagines produce six- and seven-membered rings, while the β-sila-glutamines produce seven- and eight-membered rings. In all cases, however, the hydrogen-bonded structures were calculated to be more stable than their non-hydrogen-bonded counterparts.

Other than the presence of the hydrogen bonds and the intramolecular cyclization, the structures and properties of the hydrogen-bonded molecules were calculated to be similar to the structures and properties of those molecules without hydrogen bonds, albeit save for the cyclization. In addition, as the illustrations indicate, the presence of the methyl-groups on silicon (or on oxygen) did not appreciably alter the structures or properties of the molecules, again apart from the cyclization.

The differences in energy between the non-hydrogen-bonded and the more stable intramolecularly-hydrogen-bonded molecules were 9.05 kJ/mol for **Asn-Si**-z-H{H_2NH···OC}, 12.2 kJ/mol for **Asn-Si**^{Me}-z-H{H_2NH···OC}, 19.0 kJ/mol for **Gln-Si**-z-H{H_2NH···OC}, and 13.6 kJ/mol for **Gln-Si**^{Me}-z-H{H_2NH···OC}. While these energy differences cannot be solely ascribed to the hydrogen bond strength, as the values include energy differences between open chain molecules and cyclized molecules in addition to the hydrogen bonds, the energies are reasonable[36] for hydrogen bond strengths.

In three of the optimized molecules with neutral side-chains — **Leu-Si**^{Me}-u, **Pro-Si**-u, and **Pro-Si**^{Me}-u — the Si–C$^{\alpha}$ bonds were calculated to be longer than 194 pm, approaching or exceeding the sum of the tetrahedral covalent radii for silicon and carbon. As previously discussed, such long bonds are likely to be weak and may engender instability or reactivity for these molecules. (Note, however, that the Si–CH_3 bond lengths were, in general, significantly shorter than the Si–C$^{\alpha}$ bonds.) However, all three of these molecules are the uncharged isomers. As the β-sila-α-amino acids were expected and were calculated to exist in solution as the zwitterions (which do not exhibit such long bonds), these results did not indicate that these β-sila-α-amino acids could not exist as zwitterionic species in solution.

Heavily-protected derivatives[37] of **Ala-Si**^{Me}-u have been synthesized, and the crystal structure of one such molecule, the ethyl ester of **β-**(*t*-butyldimethylsi**la**)-*N*-tosyl-α-alanine — termed ethyl **α-**(*t*-butyldimethylsi**lyl**)-α-*N*-tosylaminoacetate in the article), — was determined. The solid-state bond lengths of 191.9 pm (Si–C$^{\alpha}$), 146.9 pm (C$^{\alpha}$–N), 151.2 pm (C$^{\alpha}$–C), 119.8 (C=O), and 132.7 (C–OEt) generally compared well with the calculated values reported herein in a simulated aqueous environment, especially as the comparison is somewhat tenuous between a solid-state molecule and a molecule in a simulated aqueous environment. The γ-sila-leucine derivative 3-trimethylsilylalanine has been synthesized[38] using an

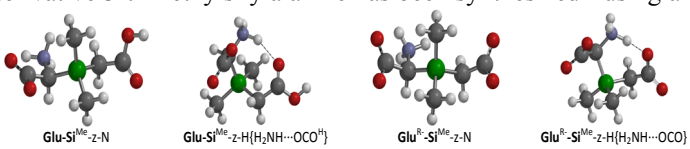


Figure 4. The β-sila-α-amino acids with anionic or potentially anionic side-chains, and their permethylsila-analogs. Atom key: H, light gray; C, dark gray; N, blue; O, red; Si, green. The thin dashed lines indicate the positions of the hydrogen bonds.

enzymatic route, and the molecule and related derivatives were used[39] as inhibitors of thermolysin. Earlier, more general synthetic routes to these molecules[40,41] have also been reported.

β-Sila-α-Amino Acids Containing Potentially-Anionic or Anionic Side-Chains

The properties of the two β-sila-α-amino acids bearing potentially-anionic or anionic side-chains are listed in Table 2, and the zwitterionic forms are illustrated in Fig. 4. Before the discovery of the intramolecularly-hydrogen-bonded structures for the β-sila-asparagines and the β-sila-glutamines, only the neutral uncharged, zwitterionic, and charge isomer systems, along with the corresponding systems with a negative charge on the side-chain, were considered for β-sila-aspartic acid and β-sila-glutamic acid and their permethylsila-analogs. However, hydrogen bonding in the β-sila-asparagines and β-sila-glutamines meant that intramolecular hydrogen bonding might also exist in the β-sila-aspartic acids and β-sila-glutamic acids. Thus, such hydrogen-bonded molecules were examined as the zwitterionic species both with and without a negative charge on the side-chains, *i.e.*, for side-chains containing either the deprotonated carboxylate moiety or the protonated carboxylic acid group, respectively.

None of these β-sila-α-amino acids have been investigated either computationally or experimentally. As was the case for the neutral-side-chain β-sila-α-amino acids, the zwitterionic form was, for all molecules, more stable than the uncharged form. In addition, the charge isomers were calculated to be less stable than the corresponding zwitterionic forms, as was observed previously for the β-sila-α-amino acids with neutral side-chains. The zwitterionic, non-hydrogen-bonded species (**Asp-Si**-z-N, **Asp-Si**^{Me}-z-N, **Glu-Si**-z-N, and **Glu-Si**^{Me}-z-N) were more stable than the uncharged species (**Asp-Si**-u, **Asp-Si**^{Me}-u, **Glu-Si**-u, and **Glu-Si**^{Me}-u), but less stable than the zwitterionic hydrogen-bonded species.

In three instances, the Si-C$^{\alpha}$ bond distance was calculated to exceed 194.9 pm, the sum of the covalent radii for carbon and silicon, in the charge isomers **Asp**$^{R-}$**-Si**^{Me} (195.6 pm), **Glu**$^{R-}$**-Si** (196.2 pm), and **Glu**$^{R-}$**-Si**^{Me} (195.8 pm); the remaining charge isomer, **Asp**$^{R-}$**-Si** exhibited a slightly shorter Si–C$^{\alpha}$ bond distance of 193.8 pm. The Mulliken bond orders for the Si-C$^{\alpha}$ bonds in these charge isomers were **Asp**$^{R-}$**-Si**, 0.50; **Asp**$^{R-}$**-Si**^{Me}, 0.57; **Glu**$^{R-}$**-Si**, 0.41; and **Glu**$^{R-}$**-Si**^{Me}, 0.49, indicating comparatively weak bonds. However, since each of these charge-isomers is calculated to be less stable than the zwitterionic forms, if prepared, the isomers would be expected to convert to the zwitterionic forms which exhibit shorter (and, hence, stronger) bonds.

For the zwitterionic β-sila-aspartic acids and β-sila-glutamic acids, there are two hydrogen-bonded possibilities, these being between the positively-charged α-ammonium group and the side-chain protonated carboxylic acid group or to the deprotonated carboxylate moiety. In the former instance, the hydrogen bond distances and Mulliken bond orders (in parentheses) for the β-sila-aspartic acids and β-sila-glutamic acids — **Asp-Si**-z-H{$H_2NH\cdots OCO^H$}, 196.2 pm (0.09); **Asp-Si**^{Me}-z-H{$H_2NH\cdots OCO^H$}, 193.7 pm (0.11); **Glu-Si**-z-H{$H_2NH\cdots OCO^H$}, 191.1 pm (0.10); **Glu-Si**^{Me}-z-H{$H_2NH\cdots OCO^H$}, 193.0 pm (0.07) — were calculated to be in line with the distances and bond orders calculated for the β-sila-asparagines and the β-sila-glutamines.

The intramolecularly hydrogen-bonded molecules in which the side-chain carboxylate with a negative charge participated in the hydrogen bond exhibited significantly shorter hydrogen bond distances than was observed for hydrogen bonds to the neutral side-chain carboxylic acid. For these cases involving the side-chain carboxylate group, the hydrogen bond distances and Mulliken bond orders (in parentheses) in which the hydrogen bond was between the positively-charged α-amino group and the negatively-charged side-chain carboxylate moiety — **Asp**$^{R-}$**-Si**-z-H{$H_2NH\cdots OCO$}, 168.3 pm (0.16); **Asp**$^{R-}$**-Si**^{Me}-z-H{$H_2NH\cdots OCO$}, 164.6 pm (0.19); **Glu**$^{R-}$**-Si**-z-H{$H_2NH\cdots OCO$}, 155.1 pm (0.22); and **Glu**$^{R-}$**-Si**^{Me}-z-H{$H_2NH\cdots OCO$}, 155.4 pm (0.21) — were calculated to be much stronger. Such a result may not be remarkable, however, as a hydrogen bond between a neutral oxygen-containing moiety such as a protonated –CO_2H group and an ammonium-like cation would be expected to be weaker than a hydrogen bond between a carboxylate anion and such a cation.

The strength of these hydrogen bonds may be approximated by comparing the electronic energies of the hydrogen-bonded and non-hydrogen-bonded structures, although such a comparison, as mentioned earlier, does not account for ring strain or other factors that may occur in converting an open chain molecule into a cyclic hydrogen-bonded molecule. That being stated, the energy differences were as follows (calculated by subtracting the energy of the hydrogen-bonded molecule from the non-hydrogen-bonded molecule, since both energies are expressed as negative values, to result in an overall positive value for the approximated hydrogen bond strength):

***E*[Asp-Si-z-N] – *E*[Asp-Si-z-H{$H_2NH\cdots OCO^H$}]**	0.001509 au	3.96 kJ/mol
***E*[Asp-Si^{Me}-z-N] – *E*[Asp-Si^{Me}-z-H{$H_2NH\cdots OCO^H$}]**	0.008474 au	22.2 kJ/mol
***E*[Asp^{R-}-Si-z-N] – *E*[Asp^{R-}-Si-z-H{$H_2NH\cdots OCO$}]**	0.002271 au	5.96 kJ/mol
***E*[Asp^{R-}-Si^{Me}-z-N] – *E*[Asp^{R-}-Si^{Me}-z-H{$H_2NH\cdots OCO$}]**	0.009375 au	24.6 kJ/mol
***E*[Glu-Si-z-N] – *E*[Glu-Si-z-H{$H_2NH\cdots OCO^H$}]**	0.005331 au	14.0 kJ/mol
***E*[Glu-Si^{Me}-z-N] – *E*[Glu-Si^{Me}-z-H{$H_2NH\cdots OCO^H$}]**	0.013425 au	35.2 kJ/mol
***E*[Glu^{R-}-Si-z-N] – *E*[Glu^{R-}-Si-z-H{$H_2NH\cdots OCO$}]**	0.004074 au	10.7 kJ/mol
***E*[Glu^{R-}-Si^{Me}-z-N] – *E*[Glu^{R-}-Si^{Me}-z-H{$H_2NH\cdots OCO$}]**	0.012713 au	33.4 kJ/mol

While these are weak hydrogen bonds[42] — even the hydrogen bonds with energies greater than ~20 kJ/mol — such cyclic structures are expected to have a significant effect on the β-sila-aspartic acids and β-sila-glutamic acids in solution, and most especially for the molecules with negatively-charged side-chains.

β-Sila-α-Amino Acids Containing Cationic or Potentially-Cationic Side-Chains

The five β-sila-α-amino acids derived from the common α-amino acids with cationic or potentially-cationic side-chains are β-sila-arginine (**Arg-Si**), β-sila-cysteine (**Cys-Si**), β-sila-histidine (**His-Si**), β-sila-lysine (**Lys-Si**), and β-sila-tyrosine (**Tyr-Si**). Because the β-sila-derivatives of histidine represent a somewhat more complicated class of β-sila-α-amino acids than do the remaining four β-sila-α-amino acids in this group, the properties of these molecules are detailed separately in Table 3 and the zwitterionic forms illustrated in Fig. 5. [The commonly-used "pseudo-systematic" Greek lettering for the atom positions on the imidazole ring of histidine was used herein even though such identification is somewhat idiosyncratic: The N and C atoms attached to C$^{\gamma}$ are both designated delta (N$^{\delta}$ and C$^{\delta}$), while the remaining two atoms

are both labeled epsilon (N^{ε} and C^{ε}).]

For β-sila-histidine (which could generically be designated **His-Si** based on the three-letter designation for histidine), three different structures needed to be investigated because of the presence of the imidazole ring: The β-sila-histidine molecules in which the imidazole ring was expected to be fully-protonated and positively-charged — **Hip**$^{R+}$**-Si** — occur (in the parent histidine without a β-silicon atom) in aqueous solutions below about pH 6.0. Either of the two remaining structures could potentially occur above pH 6.0, in which the single proton on the neutral imidazole ring could be bound either to N^{δ} (**Hid-Si**) or to N^{ε} (**Hie-Si**). Thus, the parent designation **His-Si** was not used for any of the β-sila-histidines or their dimethylsila-analogs.

Unlike some of the β-sila-α-amino acids previously described, none of the β-sila-histidines exhibited silicon-carbon bond lengths that approached or exceeded the sum of the tetrahedral covalent radii for silicon and carbon (194.9 pm). However, intramolecular hydrogen bonding was calculated to be possible in the β-sila-histidines: The intramolecular hydrogen bonds occurred between the negatively-charged α-carboxylate group and the N^{δ}–H group of the imidazole ring for **Hid-Si**, **Hid-Si**^{Me}, **Hip**$^{R+}$**-Si**, and **Hip**$^{R+}$**-Si**^{Me} and between the unprotonated N^{δ} and the protonated α-amino group for **Hie-Si** and **Hie-Si**^{Me}. Because of the distances and molecular geometries involved, hydrogen bonding to the ε-nitrogen position, either with the unprotonated N^{ε} in **Hid-Si** and **Hid-Si**^{Me}, or with the protonated N^{ε}–H in **Hip**$^{R+}$**-Si**, **Hip**$^{R+}$**-Si**^{Me}, **Hie-Si**, and **Hie-Si**^{Me}, were determined calculationally to be impossible.

The hydrogen bonds in the unprotonated β-sila-histidines and their dimethylsila-analogs exhibited relatively long bond lengths and relatively weak Mulliken bond orders (in parentheses): **Hid-Si**-z-H{OCO···HN$^{\delta}$}, 195.2 pm (0.04); **Hid-Si**^{Me}-z-H{OCO···HN$^{\delta}$}, 191.3 pm (0.05); **Hie-Si**-z-H, 190.1 pm (0.10){H_2NH···N$^{\delta}$}; and **Hie-Si**^{Me}-z-H{H_2NH···N$^{\delta}$}, 189.9 pm (0.09). However, the protonated β-sila-histidines and their dimethylsila-analogs exhibited slightly shorter hydrogen bond distances, albeit with little change in Mulliken bond orders: **Hip**$^{R+}$**-Si**-z-H{H_2NH···N$^{\delta}$}, 180.1 pm (0.06); and **Hip**$^{R+}$**-Si**^{Me}-z-H{H_2NH···N$^{\delta}$}, 172.4 pm (0.08).

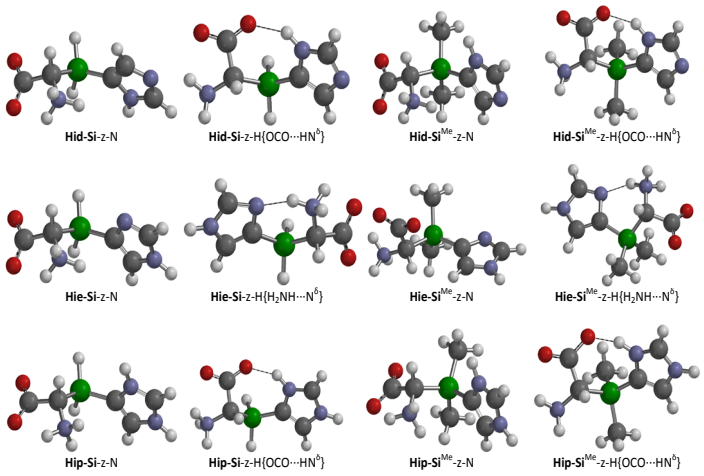


Figure 5. The β-sila-α-amino acids derived from the histidines, and their dimethylsila-analogs. Atom key: H, lighray; C, dark gray; N, blue; O, red; Si, green. The thin dashed lines indicate the position of the hydrogen bonds.

As was described earlier, the strength of these β-sila-histidine hydrogen bonds may be approximated by comparing the electronic energies of the hydrogen-bonded and non-hydrogen-bonded structures. These energy differences were as follows:

***E*[Hid-Si-z-N] – *E*[Hid-Si-z-H{OCO···HN$^{\delta}$}]**	0.005223 au	13.7 kJ/mol
***E*[Hid-Si^{Me}-z-N] – *E*[Hid-Si^{Me}-z-H{OCO···HN$^{\delta}$}]**	0.003115 au	8.18 kJ/mol
***E*[Hie-Si-z-N] – *E*[Hie-Si-z-H{H_2NH···N$^{\delta}$}]**	0.003501 au	9.19 kJ/mol
***E*[Hie-Si^{Me}-z-N] – *E*[Hie-Si^{Me}-z-H{H_2NH···N$^{\delta}$}]**	0.006833 au	17.9 kJ/mol
***E*[Hip^{R+}-Si-z-N] – *E*[Hip^{R+}-Si-z-H{OCO···HN$^{\delta}$}]**	0.007135 au	18.7 kJ/mol
***E*[Hip^{R+}-Si^{Me}-z-N] – *E*[Hip^{R+}-Si^{Me}-z-H{OCO···HN$^{\delta}$}]**	0.007345 au	19.3 kJ/mol

In these molecules, there did not appear to be a pattern in the approximated strengths of the hydrogen bonds, although those formed in the β-sila-histidines with the protonated, positively-charged side-chains (**Hip**$^{R+}$**-Si**-z-H{OCO···HN$^{\delta}$} and **Hip**$^{R+}$**-Si**^{Me}-z-H{OCO···HN$^{\delta}$}) were the shortest and strongest. Although the sila- and dimethylsila-hydrogen-bonded structures were similar (see Fig. 5), the zwitterionic β-sila-histidines without any charge on the side-chains exhibited reversed hydrogen bond energies: **Hid-Si**-z-H{OCO···HN$^{\delta}$} exhibited a stronger hydrogen bond than did **Hid-Si**^{Me}-z-H{OCO···HN$^{\delta}$}, while **Hie-Si**^{Me}-z-H{H_2NH···N$^{\delta}$} exhibited a stronger hydrogen bond than did **Hie-Si**-z-H{H_2NH···N$^{\delta}$}. The sila- and dimethylsila-histidines with positively-charged side-chains had nearly the same hydrogen bond energies. The underlying reasons for these differences are not currently fully understood.

The β-sila-histidines with neutral side-chains may be directly compared for their calculated stabilities, as these molecules contain the same number and kind of atoms. Thus, **Hid-Si**-u is more stable than **Hie-Si**-u, but **Hie-Si**-z-N is more stable than **Hid-Si**-z-N; the same is true for the β-dimethylsila-histidines. Neutral histidine is preferentially protonated[43] at N^{ε} rather than at N^{δ}; **Hie-Si**-z-N was calculated to be more stable than **Hid-Si**-z-N. However, for the dimethyl analogs, the reverse was found: **Hid-Si**^{Me}-z-N was calculated to be more stable than **Hie-Si**^{Me}-z-N. The reasons for this reversal are unclear.

Because of the multiplicity of potentially hydrogen-bonded

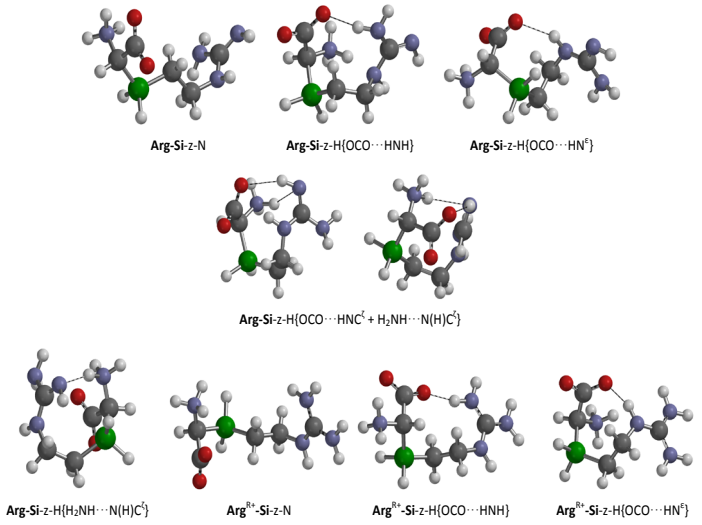


Figure 6a. The β-sila-arginines. Atom key: H, light gray; C, dark gray; N, blue; O, red; Si, green. The thin dashed lines indicate the positions of the hydrogen bonds; two views of **Arg-Si**-z-H{OCO··· HNζ + H_2NH···N(H)Cζ} are presented for clarity.

structures possible for the β-sila-arginines, and those calculated to exist, the structures and properties of the β-sila-arginines are listed separately in Table 4 and the zwitterionic forms, including the hydrogen-bonded systems, are illustrated in Fig. 6a (β-sila-arginines) and 6b (β-dimethylsila-arginines). As observed for the previously-described β-sila-amino acids, the zwitterionic forms were calculated to be more stable than the uncharged molecules, and the "charge isomers" were less stable than the zwitterionic forms. No molecules exhibited unusually-long bonds.

Hydrogen-bonded structures in the neutral side-chain β-sila-arginines were thought to be possible with the α-carboxylate group accepting a hydrogen bond from the side-chain $–NH_2$ group or either of the two independent –NH groups, or with the α-ammonium group donating a hydrogen bond to any of the three nitrogen atoms of the neutral side-chain. Similarly, for the cationic side-chain β-sila-arginines, hydrogen-bonded structures were thought to be possible with the α-carboxylate group accepting a hydrogen bond from either of the equivalent $–NH_2$ groups or from the –NH group, or with the α-ammonium group donating a hydrogen bond to a nitrogen atom of either of the two equivalent $–NH_2$ groups or to the nitrogen atom of the –NH group.

The preceding analysis indicated that up to ten different intramolecularly hydrogen-bonded β-sila-arginines (and up to an additional ten for β-dimethylsila-arginine) could exist. However, calculations demonstrated that only six of these possibilities were true minima on each of the relevant potential energy surfaces. Thus, no hydrogen bonding was calculated to exist for the potential donation of a hydrogen bond from the α-ammonium group to N^{ε} regardless of whether the side-chain was neutral, *i.e.*, **Arg-Si**-z-H{$H_2NH\cdots N^{\varepsilon}$} and **Arg-Si**^{Me}-z-H{$H_2NH\cdots N^{\varepsilon}$}, or positively-charged, *i.e.*, **Arg**$^{R+}$**-Si**-z-H{$H_2NH\cdots N^{\varepsilon}$} and **Arg**$^{R+}$**-Si**^{Me}-z-H{$H_2NH\cdots N^{\varepsilon}$}. In addition, no hydrogen bonding was calculated to exist between the α-ammonium group and the nitrogen of the NH_2 group when the side-chain carried a positive charge, *i.e.*, **Arg**$^{R+}$**-Si**-z-H{$H_2NH\cdots NH_2$} and **Arg**$^{R+}$**-Si**^{Me}-z-H{$H_2NH\cdots NH_2$}, or when the side-chain was neutral, *i.e.*, **Arg-Si**-z-H{$H_2NH\cdots NH_2$} and **Arg-Si**^{Me}-z-H{$H_2NH\cdots NH_2$}.

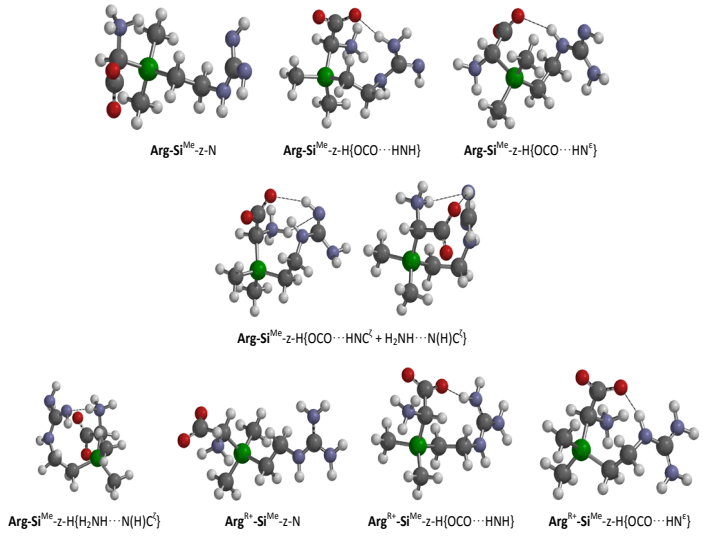


Figure 6b. The β-dimethylsila-arginines. Atom key: H, light gray; C, dark gray; N, blue; O, red; Si, green. The thin dashed lines indicate the positions of the hydrogen bonds; two views of **Arg-Si**^{Me}-z-H{$OCO\cdots HNC^{\zeta} + H_2NH\cdots N(H)C^{\zeta}$} are presented for clarity.

and **Arg-Si**^{Me}-z-H{$H_2NH\cdots NH_2$}.

Interestingly, when hydrogen bonding was investigated between the α-carboxylate group and the HNC^{ζ} group of the neutral side-chain of the molecules, a second hydrogen bond was also calculated to exist in both β-sila-arginine, *i.e.*, **Arg-Si**-z-H{$OCO\cdots HNC^{\zeta} + H_2NH\cdots N(H)C^{\zeta}$}, and in β-dimethylsila-arginine, *i.e.*, **Arg-Si**^{Me}-z-H{$OCO\cdots HNC^{\zeta} + H_2NH\cdots N(H)C^{\zeta}$}. No structure could be identified containing the {$OCO\cdots HNC^{\zeta}$} hydrogen bond without the {$H_2NH\cdots N(H)C^{\zeta}$} hydrogen bond also being present. However, the reverse was not true, as both **Arg-Si**-z-H{$H_2NH\cdots N(H)C^{\zeta}$} and **Arg-Si**^{Me}-z-H{$H_2NH\cdots N(H)C^{\zeta}$}, each without a second hydrogen bond, were stable minima on their respective potential energy surfaces, and these latter systems exhibited the strongest hydrogen bonds of their type (*vide infra*).

None of the β-sila-arginines or β-dimethylsila-arginines exhibited Si–C bonds that approached or exceeded the sum of the tetrahedral covalent bond radii for silicon and carbon. In addition, as was observed for the previously-discussed β-sila-α-amino acids, the zwitterionic forms of these β-sila-α-amino acids were calculated to be more stable than the uncharged forms, and more stable than the charge isomer forms of these molecules. While not all of the intramolecularly hydrogen-bonded structures were calculated to be stable minima on the relevant potential energy surfaces, the cyclic hydrogen-bonded structures which were calculated to exist were more stable than the comparable acyclic, non-hydrogen-bonded structures, as expected.

The four most stable intramolecularly hydrogen-bonded systems, with hydrogen bond lengths (and Mulliken bond orders, in parentheses), were **Arg-Si**-z-H{$H_2NH\cdots N(H)C^{\zeta}$}, 181.4 pm (0.11); **Arg**$^{R+}$**-Si**-z-H{$OCO\cdots HNH$}, 180.8 pm (0.06); **Arg-Si**^{Me}-z-H{$H_2NH\cdots N(H)C^{\zeta}$}, 188.0 (0.13); and **Arg**$^{R+}$**-Si**^{Me}-z-H{$OCO\cdots HNH$}, 179.6 (0.07). When these hydrogen-bonded systems were compared to the corresponding non-hydrogen-bonded molecule, the energy differences so derived were as follows:

***E*[Arg-Si-z-N] – *E*[Arg-Si-z-H{$H_2NH\cdots N(H)C^{\zeta}$}]**	0.010470 au	27.5 kJ/mol
***E*[Arg^{R+}-Si-z-N] – *E*[Arg^{R+}-Si-z-H{$OCO\cdots HNH$}]**	0.007345 au	19.3 kJ/mol
***E*[Arg-Si^{Me}-z-N] – *E*[Arg-Si^{Me}-z-H{$H_2NH\cdots N(H)C^{\zeta}$}]**	0.008047 au	21.1 kJ/mol
***E*[Arg^{R+}-Si^{Me}-z-N] – *E*[Arg^{R+}-Si^{Me}-z-H{$H_2NH\cdots N(H)C^{\zeta}$}]**	0.009293 au	24.4 kJ/mol

Although these approximated hydrogen bond strengths are remarkably similar (albeit including energy differences due to comparisons between open-chain and a cyclic molecules), the calculated energies are still considered weak[41] hydrogen bonds.

Interestingly, the two hydrogen-bonded structures calculated to exhibit two hydrogen bonds each, the β-sila-arginine, **Arg-Si**-z-H{$OCO\cdots HNC^{\zeta} + H_2NH\cdots N(H)C^{\zeta}$}, and the β-dimethylsila-arginine, **Arg-Si**^{Me}-z-H{$OCO\cdots HNC^{\zeta} + H_2NH\cdots N(H)C^{\zeta}$}, are each less stable than the hydrogen-bonded molecules listed above that contain only one hydrogen bond, even though the {$H_2NH\cdots N(H)C^{\zeta}$} hydrogen bond appears in all four molecules. Apparently, this inconsistency derives from the additional strain introduced when both the α-carboxylate and the α-ammonium groups accept from and donate to the (same) HNC^{ζ} group, compared to the apparently reduced strain when the only hydrogen bond originates through the

α-carboxylate accepting a hydrogen bond from the HNᶻ group.

The properties of the remaining three β-sila-α-amino acids with cationic or potentially-cationic side-chains are listed in Table 5 and the zwitterionic forms are illustrated in Fig. 7 (the β-sila-cysteines and -tyrosines) and Fig. 8 (the β-sila-lysines). However, the β-sila-cysteines were not all able to be optimized. Thus, the β-sila-cysteines bearing a positive charge on the side-chain (**Cys**$^{R+}$**-Si**-u, **Cys**$^{R+}$**-Si**-z, and **Cys**$^{R+}$**-Si**), and the analogous β-dimethylsila-cysteines (**Cys**$^{R+}$**-Si**^{Me}-u, **Cys**$^{R+}$**-Si**^{Me}-z, and **Cys**$^{R+}$**-Si**^{Me}) — *i.e.*, all those β-sila-cysteines containing the moiety $-SH_2^+$ — could not be optimized, strongly indicating that the β-sila-cysteines and the β-dimethylsila-cysteines will not be protonated on the side-chains of the molecules in aqueous solutions. (Cysteine itself appears unlikely to be protonated[44] in a neutral aqueous solution such as is simulated here.) In addition, no hydrogen-bonded β-sila-cysteines were observed, even when the systems were calculated from initial trial geometries favoring a structure containing a hydrogen bond.

For those β-sila- and β-dimethylsila-cysteines that were able to be optimized, the Si–S bond lengths were in the range of ~214–218 pm. For sulfur, the tetrahedral covalent radius has been determined[34] to be 104.2 pm, leading to the sum of the tetrahedral covalent radii for Si–S bonds to be 221.8 pm. In silicon-sulfur cluster[45] compounds, Si–S bonds range between 214 pm and 216 pm, while in $Si(SC_6H_5)_4$, for example, the Si–S bond lengths in the solid state[46] are 212.3 pm and 212.9 pm. Thus, bond lengths suggest that **Cys-Si**-u, **Cys-Si**-z, **Cys-Si**^{Me}-u, and **Cys-Si**^{Me}-z are realistic synthetic targets (although as was observed for the β-sila-α-amino acids previously discussed, only the zwitterionic forms were calculated to exist in aqueous solution). Further, the calculated Mulliken bond orders for the Si–S bonds in these four molecules are 1.22, 1.26, 1.46, and 1.62, respectively, indicating that the Si–S bonds are expected to be strong.

None of the β-sila- or β-dimethylsila-cysteines, -lysines, or -tyrosines exhibited Si–C bonds that approached or exceeded the sum of the tetrahedral covalent bond radii for silicon and carbon (194.9 pm). Further, the zwitterionic forms of the molecules were again more stable than the uncharged forms and the charge isomers were less stable than the zwitterionic molecules. While the zwitterionic β-sila-tyrosines did optimize with both neutral and cationic side-chains (the pK_a of the side chain of the parent tyrosine is ~10.1), no possibility for intramolecular hydrogen bonding was observed for the β-sila- and β-dimethylsila-tyrosines due to the geometric constraints of the arene ring.

Interestingly, two of the hydrogen-bonded β-sila-lysines, **Lys-Si**-z-H{OCO···HNᶻ} and **Lys-Si**^{Me}-z-H{OCO···HNᶻ}, while optimized successfully with hydrogen bonds, were calculated to be less stable than the non-hydrogen-bonded open-chain molecules. These two hydrogen-bonded molecules exhibited hydrogen bond donation from the free, neutral amine group of the side-chain to the α-carboxylate group. However, those hydrogen-bonded β-sila-lysines in which the α-carboxylate group accepts a hydrogen bond from the positively-charged ammonium group of the side-chain or in which the α-ammonium group donates a hydrogen bond to the neutral amine of the side-chain were more stable than the corresponding open-chain molecules. The lowered stability of **Lys-Si**-z-H{OCO···HNᶻ} and **Lys-Si**^{Me}-z-H{OCO···HNᶻ} with respect to the non-hydrogen-bonded **Lys-Si**-z-N and **Lys-Si**^{Me}-z-N might, as was mentioned earlier, be due to the ring size in the hydrogen bonded molecules. For **Lys-Si**-z-H{OCO···HNᶻ} and **Lys-Si**^{Me}-z-H{OCO···HNᶻ}, the ring size is nine [–Si–Cᵅ–C=O···H–N–C–C–C–] while in the hydrogen-bonded molecules which were calculated to be more stable than non-hydrogen-bonded open chain molecules, **Lys-Si**-z-H{H_2NH···Nᶻ} and **Lys-Si**^{Me}-z-H{{H_2NH···Nᶻ }, the ring size is eight [–Si–Cᵅ–N–H···N–C–C–C–]. Cyclononane is reported[47] to exhibit a slightly higher ring

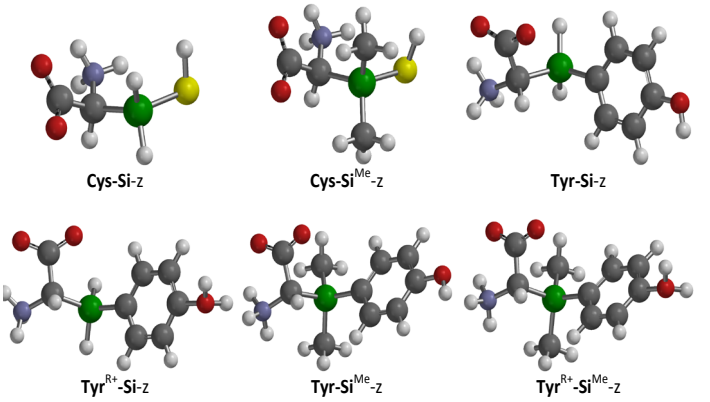


Figure 7. The β-sila-α-amino acids derived from cysteine and tyrosine, and their dimethylsila-analogs. Atom key: H, light gray; C, dark gray; N, blue; O, red; S, yellow; Si, green.

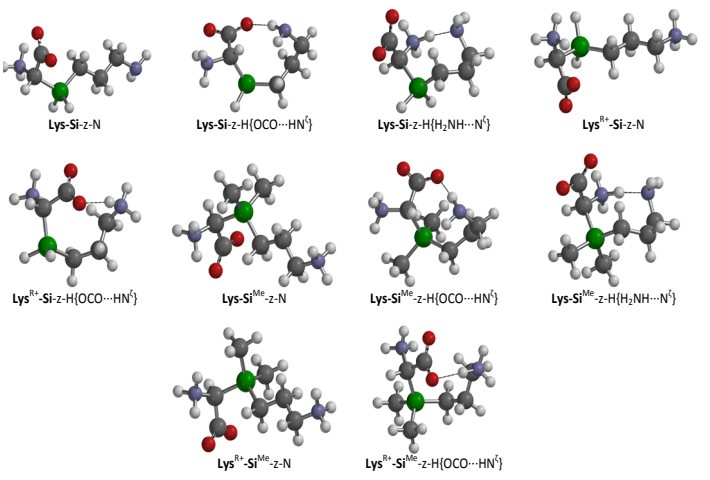


Figure 8. The β-sila-α-amino acids derived from lysine, and their dimethylsila-analogs. Atom key: H, light gray; C, dark gray; N, blue; O, red; Si, green. The thin dashed lines indicate the positions of the hydrogen bonds.

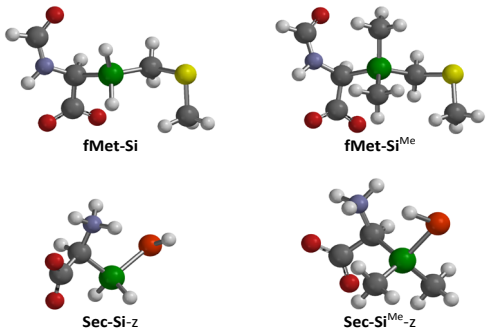


Figure 9. The β-sila-α-amino acids derived from *N*-formylmethionine and selenocysteine, and their dimethylsila-analogs. Atom key: H, light gray; C, dark gray; N, blue; O, red; Si, green; S, yellow; Se, orange.

strain than does cycloöctane.

β-Sila-α-Amino Acids Derived from the Three Uncommon α-Amino Acids

The structures and properties of the β-sila-derivatives of the three uncommon α-amino acids are detailed in Table 6 (β-sila-*N*-formylmethionines, **fMet-Si**, and β-sila-selenocysteines, **Sec-Si**) and Table 7 (β-sila-pyrrolysines, **Pyl-Si**); none of these three β-sila-α-amino acids have been reported in the literature. The zwitterionic forms of these molecules are illustrated in Fig. 9 (β-sila-*N*-formylmethionines and β-sila-selenocysteines), Fig. 10a (β-sila-pyrrolysines), and Fig. 10b (β-dimethylsila-pyrrolysines). Uniquely among the β-sila-α-amino acids, whether the nineteen common or three uncommon, the β-sila-*N*-formylmethionines cannot exist as a zwitterion in aqueous solution. These β-sila-α-amino acids, however, can carry a negative charge on the α-carboxylic acid group, but the *N*-formyl group is not protonated under aqueous conditions. Also, note that the uncharged and anionic forms of the β-sila-*N*-formylmethionines cannot be compared energetically because the uncharged form contains one hydrogen atom more than is contained in the anionic form.

As was observed previously for some of the uncharged forms of the β-sila-α-amino acids, the uncharged form of β-dimethylsila-α-amino acids, the uncharged form of β-dimethylsila-*N*-formylmethionine, **fMet-Si**^{Me}-u, was calculated to have a Si–C$^{\alpha}$ bond length of 194.3 pm (with a Mulliken bond order of 0.65). Even though a direct comparison of this molecule to the negatively-charged **fMet**$^{O-}$**-Si**^{Me} (or between **fMet-Si**-u and **fMet**$^{O-}$**-Si**) is not possible because the molecules contain different numbers of atoms, the negatively-charged forms are likely to be the forms to exist in an aqueous solution. In addition, no intramolecular hydrogen-bonded structures were observed for these molecules.

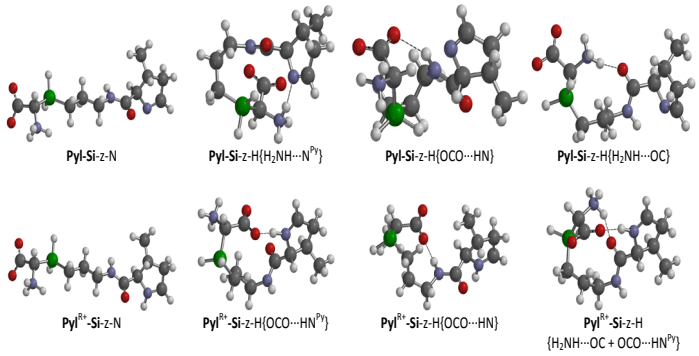


Figure 10a. The β-sila-α-amino acids derived from pyrrolysine. Atom key: H, light gray; C, dark gray; N, blue; O, red; Si, green. The thin dashed lines indicate the positions of the hydrogen bonds.

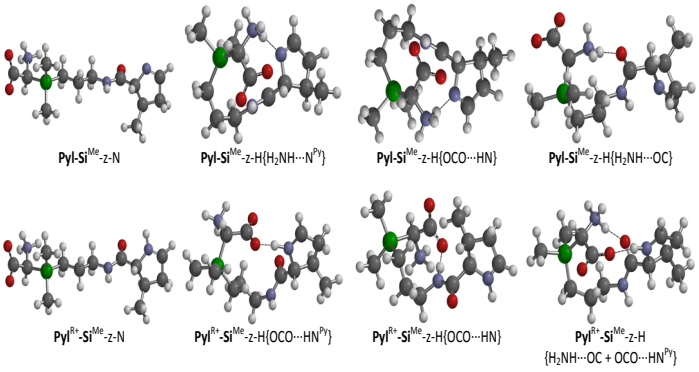


Figure 10b. The β-dimethylsila-α-amino acids derived from pyrrolysine. Atom key: H, light gray; C, dark gray; N, blue; O, red; Si, green. The thin dashed lines indicate the positions of the hydrogen bonds,

None of the zwitterionic molecules in Tables 6 or 7 were calculated to exhibit unusually-long Si–C bonds. However, the Si–Se distances in the β-sila- and β-dimethylsila-selenocysteines deserve some discussion. As discussed earlier, the zwitterionic β-sila- and β-dimethylsila-cysteines could not be optimized with a cationic side-chain, and the same was calculated to be true for the β-sila- and β-dimethylsila-selenocysteines. However, unlike the zwitterionic β-sila- and β-dimethylsila-cysteines with a neutral side-chain, which were calculated to exhibit Si–S bonds consistently shorter than the sum of the tetrahedral covalent radii for silicon and sulfur, the β-dimethylsila-selenocysteines were calculated to have Si–Se distances greater than the sum of the tetrahedral covalent radii[34] for silicon and selenium. These molecules were able to be optimized despite the long Si–Se bonds.

In **Sec-Si**-z, the Si–Se bond length was calculated to be 230.3 pm, with a Mulliken bond order of 1.08; the sum of the tetrahedral covalent radii[34] for silicon and selenium is 232.1 pm. However, in **Sec-Si**^{Me}-z, the Si–Se bond length was calculated to be 232.3 pm, with a Mulliken bond order of 0.91. While such results were not observed with other β-sila- and β-dimethylsila-α-amino acids, steric hindrance from the methyl groups on silicon may have caused the lengthening of the Si–Se bond. In β-sila- and β-dimethylsila-cysteine, the Si–S bond was 2 pm longer in the dimethyl analog, as was observed here (although the Si–S bonds in β-sila- and β-dimethylsila-cysteine were significantly stronger, with a Mulliken bond order of 1.26 in both cases). The calculated bond length is only slightly longer than the maximum calculated from the tetrahedral covalent radii in β-dimethylsila-selenocysteine, and only slightly weaker than a "full" single bond. Thus, while the β-sila- and β-dimethylsila-selenocysteines could not be optimized with cationic side-chains, evidence suggests that the molecules could be viable synthetic targets with a neutral side-chain, despite the long Si–Se bond in the dimethyl analog.

The last of the β-sila-derivatives of the three uncommon α-amino acids, β-sila-pyrrolysine, is a derivative of β-sila-lysine in which a 3-methyl-1-pyrroline-2-carboxylic acid moiety has been attached to lysine through an amide bond to the amine side-chain of lysine (see Fig. 10a and 10b). None of the β-sila-pyrrolysines have been reported in the literature. As free 1-pyrroline[48] has a pK_a of 6.8, the β-sila-pyrrolysines needed to be investigated with and without a cationic side-chain and, like lysine, also had the potential for intramolecular hydrogen bonding: Hydrogen bonds could potentially occur, when the side-chain is uncharged, between the α-carboxylate and the amide hydrogen ({OCO···HN}), and between the α-ammonium and either the amide carbonyl ({H_2NH···OC}) or the pyrroline nitrogen ({H_2NH···N^{Py}}). When the side-chain is positively-charged due to protonation at the pyrroline nitrogen, hydrogen bonds were investigated between the α-carboxylate and the amide hydrogen ({OCO···HN}) or the hydrogen attached to the pyrroline nitrogen ({OCO···HN^{Py}}), and between the α-ammonium and the amide carbonyl ({H_2NH···OC}). Each of these possibilities was investigated for both the parent and dimethyl-substituted molecules.

For the zwitterionic β-sila-pyrrolysines, the three hydrogen-bonded systems were calculated to be more stable than the non-hydrogen-bonded system; **Pyl-Si**-z-H{$H_2NH\cdots N^{Py}$} was the most stable of these hydrogen-bonded systems, with an H···N distance of 184.3 pm and a Mulliken bond order of 0.03. Of the zwitterionic β-sila-pyrrolysines with a protonated pyrroline side-chain, again the three cyclic hydrogen-bonded systems were calculated to be more stable than the open-chain non-hydrogen-bonded molecule. For these systems, however, the most stable hydrogen-bonded system was **Pyl**$^{R+}$**-Si**-z-H{$OCO\cdots HN^{Py}$}, again with a hydrogen bond interaction with the pyrroline ring, but in this case through the protonated nitrogen with a very short O···H distance of 149.5 pm and a Mulliken bond order of 0.22. Although not the most stable of the hydrogen-bonded systems, a side-chain protonated molecule calculated to contain two hydrogen bonds, **Pyl**$^{R+}$**-Si**-z-H{$H_2NH\cdots OC$ + $OCO\cdots HN^{Py}$} was also a minimum; no side-chain positively-charged molecule containing only the {$H_2NH\cdots OC$} hydrogen bond, *i.e.*, **Pyl**$^{R+}$**-Si**-z-H{$H_2NH\cdots OC$}, could be identified.

For the zwitterionic β-dimethylsila-pyrrolysines, the three analogous cyclic hydrogen-bonded systems were again calculated to be more stable than the non-hydrogen-bonded system; **Pyl-Si**^{Me}-z-H{$H_2NH\cdots N^{Py}$} was the most stable of these hydrogen-bonded systems, with an H···N distance of 182.7 pm and a Mulliken bond order of 0.04. Of the zwitterionic β-dimethylsila-pyrrolysines with a protonated pyrroline side-chain, again the three cyclic hydrogen-bonded systems were calculated to be more stable than the open-chain non-hydrogen-bonded molecule. For these systems, the most stable hydrogen-bonded system was **Pyl**$^{R+}$**-Si**^{Me}-z-H{$OCO\cdots HN^{Py}$}, again with a hydrogen bond interaction with the pyrroline ring, but in this case through the protonated nitrogen with a very short O···H distance of 147.9 pm and a Mulliken bond order of 0.22. Although not the most stable of the hydrogen-bonded molecules, a side-chain protonated molecule calculated to contain two hydrogen bonds, **Pyl**$^{R+}$**-Si**^{Me}-z-H{$H_2NH\cdots OC$ + $OCO\cdots HN^{Py}$} was also a minimum with the dimethyl systems; no side-chain positively-charged molecule containing only the {$H_2NH\cdots OC$} hydrogen bond, *i.e.*, **Pyl**$^{R+}$**-Si**^{Me}-z-H{$H_2NH\cdots OC$}, could be identified, just as was the case for the unmethylated molecules.

Closing notes

A unique dipole moment is undefined for species which carry a formal positive or negative charge[49] because the dipole moment for such a molecule depends upon the location and orientation of the charged molecule in space. However, some of the neutral species described above were calculated to have unusually large dipole moments. Such dipole moments generally appear in the "charge isomer" molecules, and, for these molecules, eight charge isomers were calculated to exhibit dipole moments in excess of 30 D. These eight molecules (with dipole moments in parentheses) were **Arg**$^{R+}$**-Si** (30.29 D), **Arg**$^{R+}$**-Si**^{Me} (32.12 D), **Lys**$^{R+}$**-Si** (31.55 D), **Lys**$^{R+}$**-Si**^{Me} (31.85 D), **Tyr**$^{R+}$**-Si** (42.01 D), **Tyr**$^{R+}$**-Si**^{Me} (41.09 D), **Pyl**$^{R+}$**-Si** (56.38 D), and **Pyl**$^{R+}$**-Si**^{Me} (57.03 D). While α-chymotrypsin has been reported to exhibit an experimentally-measured permanent dipole moment[50] of ~480 D(!), the values for these eight charge isomer molecules (and, to a lesser extent, the large dipole moments of other molecules calculated herein as well) deserve some discussion.

The simplest explanation for these large dipole moments may be through a correlation between the shortest distance from the α-carboxylate group to the side-chain positively-charged group and the magnitude of the dipole moment, as has been described[49] elsewhere. As Chart 1 demonstrates, such a correlation appears to be present with a high correlation.

These molecules are the charge isomer systems, and were calculated to be less stable than their zwitterionic counterparts. Thus, it is unlikely that such high dipole moment molecules would be able to be synthesized, as the molecules would be expected to convert to the corresponding zwitterionic molecule (at least without the inclusion of substituents to sterically prevent such conversions). However, the magnitude of the calculated dipole moment appears to be directly related to the distance between the positive and negative charges in these molecules.

The energies of the highest occupied molecular orbitals (HOMO) and of the lowest unoccupied molecular orbitals (LUMO), and the energy separation between these energy levels, have been reported for the β-sila-α-amino acids discussed herein. Large separations between these energy levels have been reported to be related to the kinetic stability of fullerenes[51] and polycyclic aromatic hydrocarbon[52] systems. However, in the β-sila-α-amino acids, the LUMO-HOMO separation values vary over a range of ~6.8 eV to ~9.2 eV, again lending credence to the viability of these molecules, should appropriate synthetic methods be found. In addition, the specific molecular orbital energies of the LUMO and HOMO levels may allow these molecules to exhibit useful electron donation (from the HOMO to a good electron acceptor) or acceptance (to the LUMO when that molecular orbital was calculated to exhibit a negative electronic energy) with other molecules or surfaces. Of course, such reactivity may also lead to decomposition pathways for the molecules.

Conclusions

The β-sila-α-amino acids and their β-dimethylsila-analogs

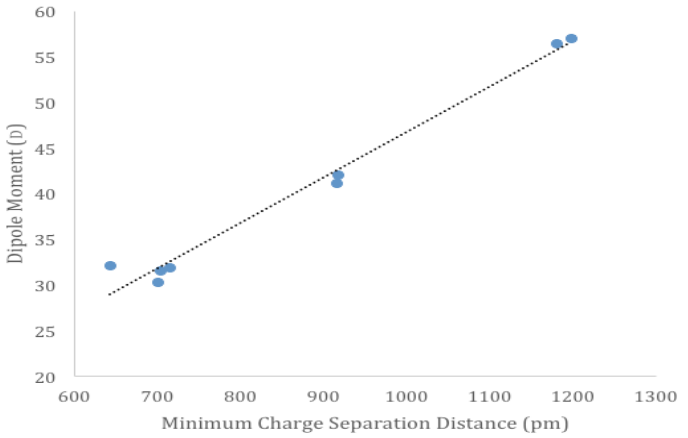


Chart 1. The correlation between the calculated permanent dipole moment and the charge separation between the α-carboxylate and the side-chain positively-charged groups of charge isomer molecules. The line has m = 0.0497 D/pm, b = 3.0382 D, and R^2 = 0.9813.

have been calculated to be both stable molecules as uncharged species containing an α-carboxylic acid and an α-amine, but, more importantly, as zwitterionic species containing an α-carboxylate group and an α-ammonium group in a simulated aqueous solution. Some of these molecules are also calculated to exhibit intramolecular hydrogen bonding. Thus, the incorporation of silicon into α-amino acids may very well hold promise for, potentially, the possibility of silicon-based life or, at a minimum, an alternative to the amino acid life as currently understood. Methylation at the silicon atom, while adding steric bulk to the molecules, does not appear to appreciably alter the structures or electronic properties of these molecules.

While beyond the scope of this report, calculations could now be carried out with the same parameters as those described herein to expand the potential for the incorporation of silicon into proteins. Steps that could be taken include the investigation of the structures and properties of dipeptides (or polypeptides) containing some or all of the β-sila- or β-dimethylsila-α-amino acids in concert with the parent, unaltered α-amino acids to further compare their structures and properties. Initial tests could include the comparison of dipeptides containing two α-amino acids with dipeptides containing either one or two β-sila-α-amino acids to determine whether the dipeptides can be optimized and thus be viable synthetic targets.

Supplementary Information

Supplementary information, including all input and output files, is available on request from the corresponding author, <rwz7001@humboldt.edu>.[53]

Acknowledgements

The National Science Foundation's Course, Curriculum, and Laboratory Improvement Program, Adaptation and Implementation Section (NSF-9950344), is acknowledged for generous support for the initial purchase of computers and computational chemistry software, upon which all of this work is based. Matching funds and additional support were received at various times from the Department of Chemistry, the College of Natural Resources and Sciences, the Office for Research and Graduate Studies, the Office of the Vice-President for Academic Affairs, the Office of the President, and the Sponsored Programs Foundation of Humboldt State University (now California State Polytechnic University, Humboldt).

References

1. Betts, J. G.; Desaix, P.; Johnson, E.; Johnson, J. E.; Korol, O.; Kruse, D.; Poe, B.; Wise, J. A.; Womble, M.; Young, K. A. *Anatomy and Physiology*, Openstax (Rice University): Houston, Texas (**2017**), section 3.4, pp 110–115.
2. Vauquelin, L. N.; Robiquet, P. J. *Ann. Chim.* **1806**, *57*, 88–93.
3. McCoy, R. H.; Meyer, C. E.; Rose, W. C. *J. Biol. Chem.* **1935**, *112*, 283–302.
4. Marcker, K.; Sanger, F. *J. Mol. Biol.* **1964**, *8*, 835–840.
5. Stadtman, T. C. *Science (Washington, D. C.)* **1974**, *183*, 915–922.
6. (a) Srinivasan, G.; James, C. M.; Krzycki, J. A. *Science (Washington, D. C.)* **2002**, *296*, 1459–1462.
(b) Hao, B.; Gong, W.; Ferguson, T. K.; James, C. M., Krzycki, J. A.; Chan, M. K. *Science (Washington, D. C.)* **2002**, *296*, 1462–1466.
7. Petkowski, J. J.; Bains, W.; Seager, S. *Life* **2020**, *10*, 84 (47 pp).
8. (a) Rémond, E.; Martin, C.; Martinez, J.; Cavelier, F. *Chem. Rev.* **2016**, *116*, 11654–11684.
9. Mortensen, M.; Husmann, R.; Veri, E.; Bolm, C. *Chem. Soc. Rev.* **2009**, *38*, 1002–1010.
10. (a) Patai, S.; Rappoport, Z. (eds.) *The Chemistry of Organic Silicon Compounds, Parts 1 and 2, Volume 1*, John Wiley and Sons: Chichester, England (**1989**).
(b) *Chem. Rev.* **1995**, *95*(5): Special issue on silicon chemistry.
(c) Rappoport, Z.; Apeloig, Y. (eds.) *The Chemistry of Organic Silicon Compounds, Volume 2*, John Wiley and Sons: Chichester, England (**1998**).
(d) Rappoport, Z.; Apeloig, Y. (eds.) *The Chemistry of Organic Silicon Compounds, Volume 3*, John Wiley and Sons: Chichester, England (**2001**).
11. Jacob, D. T. *Silicon* **2016**, *8*, 175–176.
12. Nieto-Domínquez, M.; Nikel, P. I. *ChemBioChem* **2020**, *21*, 2551–2571.
13. Haas, R.; Nikel, P. I. *Trends Biotech.* **2022**, doi: 10.1016/j.tibtech.2022.06.004.
14. Sieburth, S. McN.; Somers, J. J.; O'Hare, H. K. *Tetrahedron* **1996,** *52*, 5669–5682.
15. Czapla, M. *Int. J. Quantum Chem.* **2018**, *118*, e25488 (8 pp).
16. Kondratyev, M. S.; Shcherbakov, K. A.; Samchenko, A. A.; Degtyareva, O. V.; Terpugov, E. L.; Khechinashvili, N. N.; Komarov, V. M. *Biophysics* **2022**, *67*, 157–164.
17. Abersfelder, K.; White, A. J. P.; Rzepa, H. S.; Scheschkewitz, D. *Science (Washington, D. C.* **2010**, *327*, 564–566.
18. Clabo, D. A., Jr.; Schaefer, H. F., III. *J. Chem. Phys.* **1986**, *84*, 1664–1669.
19. Ivanov, A. S.; Boldyrev, A. I. *J. Phys. Chem. A* **2012**, *116*, 9591–9598.
20. Nakamura, T.; Mesuda, A.; Kudo, T. *Organometallics* **2020**, *39*, 3041–3049.
21. Sax, A. F.; Kalcher, J.; Janoschek, R. *J. Comput. Chem.* **1988**, *9*, 564–577.
22. Pakiari, A. H.; Jamshidi, Z. *J. Mol. Struct. THEOCHEM* **2004,** *685*, 155–161.
23. Steiner, T. *Angew. Chem. Int. Ed.* **2002**, *41*, 48–76
24. *Spartan '18*, version 1.4.5 (**2020**), Wavefunction, Inc., 18401 Von Karman Avenue, Suite 370, Irvine, CA 92612.
25. (a) Walker, M.; Harvey, A. J. A.; Sen, A.; Dessent, C. E. H. *J. Phys. Chem*. **2013**, *117*, 12590–12600.
(b) Pallavi, L.; Tonannaver, J.; Tonannaver, Jayashree. *J. Mol. Struct.* **2020**, *1211*, 128085 (12 pp).
26. Zhao, Y.; Truhlar, D. G. *Theor. Chem. Acc.* **2008**, *120*, 215–241.
27. Cossi, M.; Rega, N.; Scalmani, G.; Barone, V. *J. Comput. Chem.* **2003**, *24*, 669–681.
28. Sieburth, S. McN.; Fensterbank, L. *J. Org. Chem.* **1993**, *58*, 6314–6318.
29. Kim, J. K.; Siburth, S. McN. *J. Org. Chem.* **2012**, *77*, 2901–2906.
30. Mastryukov, V. S.; Golubinskii, A. V.; Vilkov, L. V. *Zh. Strukt. Khim. (J. Struct. Chem.)* **1980**, *21*, 48–52.
31. Belyakov, A. V.; Zavgorodnii, V. S.; Mastryukov, V. S. *Zh.*

Strukt. Khim. (J. Struct. Chem.) **1989**, *30*, 34–41.
32. Buschmann J.; Lentz, D.; Luger, P.; Röttger, M. *Acta Cryst. Sect. C* **2000**, *C56*, 121–122.
33. Beagley, B.; Monaghan, J. J.; Hewitt, T. G. *J. Mol. Struct.* **1971**, *8*, 401–411.
34. Pyykkö, P. *Phys. Rev. B* **2012**, *85*, 024115 (7 pp).
35. Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1841–1846.
36. Steiner, T. *Angew. Chem. Int. Ed.* **2002**, *41*, 48–76.
37. Bolm, C.; Kasyan, A.; Drauz, K.; Günther, K.; Raabe, G. *Angew. Chem. Int. Ed.* **2000**, *39*, 2288–2290.
38. Yamanaka, H.; Fukui, T.; Kawamoto, T.; Tanaka, A. *Appl. Microbiol. Biotechnol.* **1996**, *45*, 51–55.
39. Ishikawa, H.; Yamanaka, H.; Kawamoto, T.; Tanaka, A. *Appl. Microbiol. Biotechnol.* **1999**, *53*, 19–22.
40. Kobayashi, T.; Nishino, S.; Miura, M.; Hirano, K. *Org. Lett.* **2022**, *24*, 1418–1422.
41. Rugeri, B.; Audi, H.; Jewula, P.; Koudih, R.; Malacea-Kabbara, R.; Vimont, D.; Schulz, J.; Fernandez, P.; Jugé, S. *Eur. J. Org. Chem.* **2017**, 5399–5409.
42. Emsley, J. *Chem. Soc. Rev.* **1980**, *9*, 91–124.
43. Hernández, B.; Pflüger, F.; Adenier, A.; Kruglik, S. G.; Ghomi, M. *Phys. Chem. Chem. Phys.* **2011**, *13*, 17284–17294.
44. Kostco, O.; Ahmed, M. *Phys. Chem. Chem. Phys.* **2021**, *23*, 8847–8853.45. Ackerhans, C.; Roesky, H. W.; Noltemeyer, M. *Organometallics* **2001**, *20*, 1282–1284, and references therein.
46. Shibao, R. K.; Keder, N. L.; Eckert, H. *Inorg. Chem.* **1990**, *29*, 4163–4166.
47. Dragojlovic, V. *ChemTexts* **2015**, *1*(3), 14 (30 pp).
48. Zhang, X.; Chingin, K.; Zhong, D.; Liang, J.; Ouyang, Y.; Chen, H. *Sci. Rep.* **2017**, *7*, 7675 (8 pp).
49. *Spartan '18 for Windows, Macintosh, and Linux: Tutorial and Users Guide* (**2017**), Wavefunction, Inc., 18401 Von Karman Avenue, Suite 370, Irvine, CA 92612, pg 47.
50. Antosiewicz, J.; Porschke, D. *Biochemistry* **1989**, *28*, 10072–10078.
51. Aihara, J.-i. *Theor. Chem. Acc.* **1999**, *102*, 134–138.
52. Aihara, J.-i. *J. Phys. Chem. A* **1999**, *103*, 7487–7495.
53.Coudert, F.-X. *Chem. Mater.* **2017**, *29*, 2615–2617.