

# Segmented contracted basis sets for atoms H through Xe: Sapporo-(DK)- $n$ ZP sets ( $n = D, T, Q$ )

Takeshi Noro · Masahiro Sekiya · Toshikatsu Koga

Received: 27 September 2011 / Accepted: 1 November 2011 / Published online: 4 February 2012  
© Springer-Verlag 2012

**Abstract** For the 54 atoms from H to Xe, compact yet accurate segmented Gaussian-type basis sets have been constructed for all electron calculations. Non-relativistic  $n$ ZP (Sapporo- $n$ ZP) sets for Li–Xe and relativistic  $n$ ZP (Sapporo-DK- $n$ ZP) sets for K–Xe are developed ( $n = D, T, Q$ ), which efficiently incorporate valence and core electron correlations. Test calculations at the coupled-cluster level of theory are performed for spectroscopic constants of 12 hydrides of  $s$ - and  $d$ -block atoms and 12 diatomics of  $p$ -block atoms in their ground states. For all molecules, the calculated spectroscopic constants approach to the experimental values smoothly as the basis set quality increases.

**Keywords** Segmented basis set · Core–valence correlation · Relativistic all electron calculation

## 1 Introduction

During the last decade, correlating contracted (C) Gaussian-type function (GTF) sets in segmented form have been

developed for the 103 atoms H through Lr of the periodic table [1–12]. They were determined to represent the space spanned by important accurate atomic natural orbitals (NOs) generated from large-scale configuration interaction (CI) calculations of atoms and were designed to incorporate significant contributions to the electron correlation efficiently. We constructed non-relativistic correlating sets for the atoms up to Xe and relativistic sets for the heavier atoms in which the relativistic effects are considered through the third-order Douglas–Kroll–Hess (DK) approximation [13, 14]. In spite of their compactness, these correlating sets, named Natural Orbital based Segmented Contracted Gaussian (NOSeC) basis sets, show high performance in molecular calculations when used together with literature segmented Hartree–Fock (HF) GTF sets. The NOSeC sets combined with Tatewaki–Koga (TK) HF CGTF sets [15–19], named (DK)-TK/NOSeC-(C)V- $n$ ZP ( $n = D, T, Q$ ), are available through the web site <http://setani.sci.hokudai.ac.jp/sapporo/>.

In this paper, we construct core–valence correlating basis sets for the second to fifth period  $p$ -block elements and the first and second  $d$ -block elements. In these atoms, especially in heavy atoms, the core–valence correlation effects as well as the relativistic effects are important to obtain reliable spectroscopic constants. In the InI molecule, for example, the coupled-cluster singles and doubles with non-iterative triples correction (CCSD(T)) calculation with DK-TK/NOSeC-V-QZP gave a nontrivial deviation of 0.07 Å in the bond length from the experimental value, which cannot be considered sufficiently small and is expected to be reduced by considering the core–valence correlation. The core–valence correlation effects are taken into account in the cc-pCV $n$ Z basis sets for the second and third period atoms [20, 21], and  $d$ -block atoms of the fourth and fifth periods [22, 23]. For the  $s$ -block atoms from Li to Ca, Iron

---

### Present Address:

T. Noro (✉)  
Department of Chemistry, Faculty of Science,  
Hokkaido University, Sapporo, Hokkaido 060-0810, Japan  
e-mail: tashi@mail.sci.hokudai.ac.jp

M. Sekiya  
Department of Intercultural Studies, Tomakomai Komazawa  
University, Tomakomai, Hokkaido 059-1292, Japan

T. Koga  
Applied Chemistry Research Unit, Graduate School  
of Engineering, Muroran Institute of Technology,  
Muroran, Hokkaido 050-8585, Japan

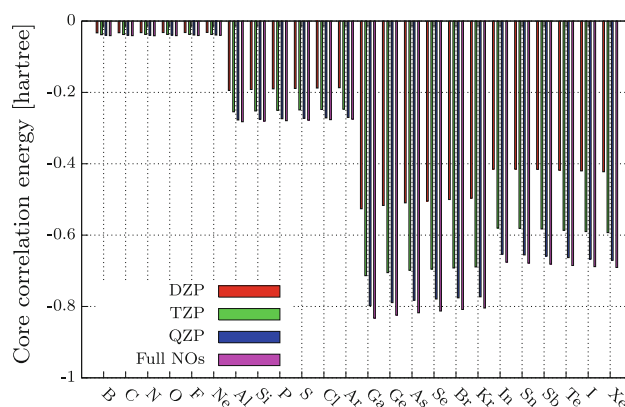
et al. [24] proposed core–valence cc-type basis sets. For the  $p$ -block atoms, the ANO basis sets were developed for the second to sixth period elements where the relativistic effects were considered through the DK approximation [25]. Dyal also constructed relativistic core–valence sets for the fourth to sixth period elements, where the DZP, TZP, and QZP basis sets were optimized in Dirac–Hartree–Fock calculations [26–28]. In this paper, we consider the  $n's$ ,  $n'p$ , and  $n'd$  subshells as the core for  $p$ -block elements, and the  $n's$  and  $n'p$  subshells for  $d$ -block elements, where  $n'$  stands for the principle quantum number of the second outermost shell. For  $d$ -block atoms, the (DK)-TK/NOSeC-CV- $n$ ZP sets do not consider the correlation effects of the  $n's$  and  $n'p$  subshells, which are important in the excitations of the  $n'd$  electrons. Although for  $s$ -block atoms the (DK)-TK/NOSeC-CV- $n$ ZP sets include the correlation effects among electrons in the  $n's$  and  $n'p$  subshells, we decided to refine them by eliminating some similar primitive Gaussians in CGTFs which may cause linear dependency problems.

The next section outlines our computational procedures and presents new basis sets named Sapporo-(DK)- $n$ ZP. Atomic tests and molecular applications to 24 diatomics are given in Sects. 2 and 3, respectively. The following symbols are used throughout this paper: [ ] for CGTFs, ( ) for primitive Gaussians, { } for contraction patterns of CGTFs where powers imply repetition of the same size.

## 2 Method

### 2.1 Standard accurate NO sets

In order to obtain accurate NOs, we performed CI calculation using well-tempered primitive sets by Huzinaga et al. [29, 30] extended by adding higher azimuthal quantum number ( $l$ ) functions as follows: (12s8p8d8f8g) for the second period atoms, (23s16p9d9f9g) for the third period atoms, (26s20p16d9f9g9h) for the fourth period atoms, and (28s23p20d9f9g9h) for the fifth period atoms. For the  $p$ -block atoms, we performed separate CI calculations for core and valence electrons. In the core CI calculations, we considered the correlation of the  $1s$  electrons for Li–Ne, of the  $2s$  and  $2p$  electrons for Na–Ar, and of the  $n's$ ,  $n'p$ , and  $n'd$  electrons for K–Kr ( $n' = 3$ ) and Rb–Xe ( $n' = 4$ ). In the valence CI calculations, we only correlated  $(n' + 1)s$  and  $(n' + 1)p$  electrons. For the  $d$ -block atoms, on the other hand, we took into account the correlation among the  $n's$ ,  $n'p$ ,  $n'd$ , and  $(n' + 1)s$  electrons simultaneously. Non-relativistic CI calculations were performed for the second to fifth period atoms, and relativistic CI calculations, which consider scalar relativistic effects through the DK Hamiltonian [14] with a Gaussian nucleus model [31],



**Fig. 1** Core correlation energies of  $p$ -block atoms

were performed for the fourth and fifth period atoms. In all atomic calculations, we used the ATOMCI program package [32].

Standard  $n$ ZP NO sets were constructed by truncating the accurate NOs generated from the CI wave functions. For the size of  $n$ ZP set, we follow the definition of standard cc-pV $n$ Z basis sets for each subshell. In the Ga atom, for example, we consider  $3s$ ,  $3p$ , and  $3d$  as core subshells, and  $4s$  and  $4p$  as valence subshells. We first define a minimal set of [4s3p1d] for occupied atomic orbitals. In the DZP set, [1s1p1d1f] and [1s1p1d] are considered as correlating sets of core and valence shells, respectively. Then, the total size of DZP is [6s5p3d1f]. Similarly, we define the standard sizes of TZP and QZP to be [8s7p5d3f1g] and [10s9p7d5f3g1h], respectively.

In Fig. 1, we show the core correlation energies of the DZP, TZP, and QZP sets for the  $p$ -block atoms. In the second period atoms, the DZP sets yield about 75% of the total core correlation energies by the full NOs. In the third period atoms, we find the core correlation energy is almost saturated at the TZP set. Thus, we construct the DZP core correlating sets for the second period  $p$ -block atoms, and the DZP and TZP core correlating sets for the third period  $p$ -block atoms. For the fourth and fifth period atoms, on the other hand, the core correlation energies are considerably larger than those of the second and third period atoms, and the QZP set gives nontrivial improvement over the TZP set. Thus, we generate all the DZP, TZP, and QZP core correlating sets for the fourth and fifth period  $p$ -block atoms.

### 2.2 Determination of core correlating basis functions

In the previous papers [1–12], we reported general-purpose correlating basis sets, which are intended to be combined with an arbitrary HF basis set. In this work, however, we develop core correlating functions to complement a particular segment-type basis sets, which contain occupied and

valence-correlating CGTFs. The procedure is summarized as follows:

1. Start from the minimal-type HF sets of Tatewaki-Koga's non-relativistic sets [15–18] or the relativistic DK sets constructed in the previous work [19]. Decontract occupied  $n'$  and  $n' + 1$  subshell orbitals of minimal-type HF functions. Add the valence-correlating basis sets NOSeC-V- $n$ ZP for the  $p$ - and  $d$ -block atoms, and the core-valence-correlating sets of NOSeC-CV- $n$ ZP for the  $s$ -block atoms ( $n = D, T, Q$ ).
2. Optimize higher  $l$  primitive GTFs, which describe the correlation effects among core electrons, to represent the standard accurate NO sets mentioned in the previous sub-section. If the numbers of decontracted  $s$  and  $p$  functions are deficient in the standard size of the  $n$ ZP set, extra  $s$  or  $p$ -type primitive GTFs are supplemented and optimized.
3. Check the combined basis set from the procedures 1 and 2 so that analogous primitives do not appear in different CGTFs.

The procedure is performed independently for each  $l$  in the determination of the present basis sets. For the required optimization, we use the same procedure as our previous works [1], where the contraction coefficients and exponents

were optimized as nonlinear parameters using the conjugate directions algorithm [33].

### 2.3 New DZP, TZP, and QZP basis sets

Following the procedure mentioned in the previous sub-section, we developed non-relativistic DZP, TZP, and QZP sets for Li–Xe, and non-relativistic and relativistic DZP, TZP, and QZP sets for K–Xe, which are referred to as Sapporo-(DK)- $n$ ZP sets ( $n = D, T, Q$ ).

For the second and third period  $p$ -block atoms, the sizes of the TZP and QZP sets are smaller than the standard sizes, because small core correlating sets were sufficient for these atoms as mentioned in Sect. 2.1. In QZP sets of the fourth and fifth period atoms, we removed one each of  $s$ -,  $p$ -, and  $d$ -type CGTF functions, because the contributions from these functions were found to be small. In Ga, for example, the size of QZP reduces to [9s8p7d5f3g1h] after the elimination of [1s1p1d]. We summarize the sizes of the present CGTF sets in Table 1.

In Table 2, we compare the sizes of Sapporo-TZP and available cc-CVTZ sets. The main difference of these sets is the description of occupied orbitals: The segmented contractions are used in the present sets, while the general contractions in cc-CVTZ. The numbers of CGTFs are

**Table 1** Sizes of Sapporo CGTF sets

Period	Atoms	Core	Valence	DZP	TZP	QZP
1	H, He		1s	[2s1p]	[3s2p1d]	[4s3p2d1f]
2	Li, Be	1s	2s	[5s2p]	[6s4p2d]	[7s5p3d1f]
	B–Ne	1s	2s, 2p	[4s3p1d]	[6s4p2d1f]	[7s5p3d2f1g]
3	Na, Mg	2s, 2p	3s	[5s4p1d]	[7s6p3d1f]	[8s7p5d3f1g]
	Al–Ar	2s, 2p	3s, 3p	[5s4p2d]	[7s6p4d2f]	[8s7p5d3f1g]
4	K, Ca	3s, 3p	4s	[6s5p1d]	[8s6p3d1f]	[9s7p4d2f1g]
	Sc–Zn		3s, 3p, 3d, 4s	[6s4p3d1f]	[8s6p4d2f1g]	[9s7p5d3f2g1h]
	Ga–Kr	3s, 3p, 3d	4s, 4p	[6s5p3d1f]	[8s7p5d3f1g]	[9s8p7d5f3g1h]
5	Rb, Sr	4s, 4p	5s	[7s6p2d]	[9s7p4d1f]	[10s8p5d2f1g]
	Y–Cd		4s, 4p, 4d, 5s	[7s5p4d1f]	[9s7p5d2f1g]	[10s8p6d3f2g1h]
	In–Xe	4s, 4p, 4d	5s, 5p	[7s6p4d1f]	[9s8p6d3f1g]	[10s9p7d5f3g1h]

**Table 2** Comparison of sizes of CGTFs, contraction patterns, and total numbers  $N_{\text{GTF}}$  of primitive GTFs between Sapporo-TZP and cc-CVTZ

Period	Atoms	Sapporo-TZP	$N_{\text{GTF}}$	cc-CVTZ	$N_{\text{GTF}}$
		CGTF		CGTF	
2	Li, Be	[6s4p2d] = {41 <sup>5</sup> /1 <sup>4</sup> /21}	16	[6s5p3d1f] = {9 <sup>2</sup> 1 <sup>4</sup> /31 <sup>4</sup> /1 <sup>3</sup> /1}	33
	B–Ne	[6s4p2d1f] = {421 <sup>4</sup> /31 <sup>3</sup> /21/2}	21	[6s5p3d1f] = {8 <sup>2</sup> 1 <sup>4</sup> /31 <sup>4</sup> /1 <sup>3</sup> /1}	31
3	Na, Mg	[7s6p3d1f] = {621 <sup>5</sup> /41 <sup>5</sup> /21 <sup>2</sup> /1}	27	[7s6p4d2f] = {(13) <sup>3</sup> 1 <sup>4</sup> /8 <sup>2</sup> 1 <sup>4</sup> /1 <sup>4</sup> /1 <sup>2</sup> }	69
	Al–Ar	[7s6p4d2f] = {62 <sup>2</sup> 1 <sup>4</sup> /41 <sup>5</sup> /31 <sup>3</sup> /2 <sup>2</sup> }	33	[7s6p4d2f] = {(13) <sup>3</sup> 1 <sup>4</sup> /7 <sup>2</sup> 1 <sup>4</sup> /1 <sup>4</sup> /1 <sup>2</sup> }	67
4	K, Ca	[8s6p3d1f] = {7421 <sup>5</sup> /721 <sup>4</sup> /21 <sup>2</sup> /1}	36	[8s7p5d2f] = {(18) <sup>4</sup> 1 <sup>4</sup> /(12) <sup>3</sup> 1 <sup>4</sup> /41 <sup>4</sup> /1 <sup>2</sup> }	126
	Sc–Zn	[8s6p4d2f1g] = {7421 <sup>5</sup> /721 <sup>4</sup> /621 <sup>2</sup> /21/2}	46	[9s8p6d3f2g] = {(19) <sup>6</sup> 1 <sup>3</sup> /(15) <sup>5</sup> 1 <sup>3</sup> /7 <sup>3</sup> 1 <sup>3</sup> /1 <sup>3</sup> /1 <sup>2</sup> }	224

**Table 3** Correlation energies in hartree

Z	Atom	Correlation energy		
		DZP	TZP	QZP
2	He	-0.034446 (96.4)	-0.039348 (98.4)	-0.040856 (99.4)
3	Li	-0.036596 (95.8)	-0.041559 (96.3)	-0.042412 (95.6)
4	Be	-0.078279 (99.9)	-0.084865 (99.0)	-0.085938 (97.5)
5	B	-0.092315 (94.0)	-0.100025 (93.2)	-0.102075 (93.0)
6	C	-0.109904 (95.2)	-0.123527 (94.7)	-0.127217 (94.3)
7	N	-0.125177 (95.5)	-0.147296 (95.7)	-0.153173 (95.1)
8	O	-0.159438 (95.5)	-0.199293 (96.5)	-0.210938 (96.4)
9	F	-0.193048 (94.7)	-0.248401 (96.7)	-0.266164 (96.9)
10	Ne	-0.227331 (93.8)	-0.296606 (96.7)	-0.320950 (97.2)
11	Na	-0.213380 (98.1)	-0.264158 (94.6)	-0.282338 (94.3)
12	Mg	-0.235847 (101.1)	-0.290504 (96.4)	-0.311658 (95.6)
13	Al	-0.237108 (96.0)	-0.301912 (97.5)	-0.308555 (92.7)
14	Si	-0.253815 (96.4)	-0.323325 (97.9)	-0.330627 (93.1)
15	P	-0.267857 (96.4)	-0.344830 (97.9)	-0.353333 (93.4)
16	S	-0.292959 (96.6)	-0.388445 (98.5)	-0.401451 (94.4)
17	Cl	-0.317006 (96.4)	-0.427604 (98.7)	-0.446041 (95.0)
18	Ar	-0.340172 (96.1)	-0.462347 (98.9)	-0.487049 (95.4)
19	K	-0.180622 (97.3)	-0.233490 (96.6)	-0.253846 (96.7)
20	Ca	-0.211762 (99.4)	-0.270827 (97.7)	-0.291981 (97.0)
21	Sc	-0.238633 (83.9)	-0.284211 (84.6)	-0.323489 (91.9)
22	Ti	-0.256309 (84.6)	-0.324864 (87.7)	-0.367664 (93.7)
23	V	-0.286151 (88.1)	-0.361978 (88.7)	-0.413010 (94.7)
24	Cr	-0.321186 (92.0)	-0.409390 (89.8)	-0.467311 (95.2)
25	Mn	-0.342875 (92.8)	-0.436597 (90.0)	-0.499535 (95.3)
26	Fe	-0.377258 (91.8)	-0.505567 (91.7)	-0.573956 (95.8)
27	Co	-0.406762 (90.5)	-0.562379 (92.2)	-0.638334 (95.8)
28	Ni	-0.448976 (91.4)	-0.622759 (92.7)	-0.706600 (95.9)
29	Cu	-0.534169 (93.8)	-0.730060 (93.2)	-0.823968 (95.9)
30	Zn	-0.508159 (87.8)	-0.740172 (92.7)	-0.839659 (95.3)
31	Ga	-0.529458 (92.3)	-0.751550 (98.4)	-0.841530 (99.1)
32	Ge	-0.536190 (92.7)	-0.766724 (99.1)	-0.852812 (99.2)
33	As	-0.541120 (92.9)	-0.778750 (99.3)	-0.866023 (99.4)
34	Se	-0.556547 (93.0)	-0.811817 (99.2)	-0.903950 (99.3)
35	Br	-0.571181 (93.0)	-0.840792 (99.1)	-0.937928 (99.2)
36	Kr	-0.585217 (93.0)	-0.864946 (99.1)	-0.968608 (99.2)
37	Rb	-0.143646 (98.6)	-0.194723 (97.3)	-0.214968 (97.0)
38	Sr	-0.170514 (100.4)	-0.228135 (98.5)	-0.250313 (97.7)
39	Y	-0.205385 (84.6)	-0.245987 (84.9)	-0.282954 (93.1)
40	Zr	-0.231232 (87.7)	-0.287853 (89.2)	-0.322229 (94.1)
41	Nb	-0.238713 (86.8)	-0.316739 (90.1)	-0.357367 (94.7)
42	Mo	-0.266107 (89.9)	-0.349403 (90.9)	-0.393566 (94.8)
43	Tc	-0.303579 (92.2)	-0.391763 (92.3)	-0.437789 (95.2)
44	Ru	-0.324609 (90.4)	-0.452344 (93.4)	-0.508413 (95.9)
45	Rh	-0.357863 (91.8)	-0.500183 (94.1)	-0.562523 (96.1)
46	Pd	-0.411095 (93.9)	-0.572238 (94.9)	-0.643507 (96.6)
47	Ag	-0.424676 (94.5)	-0.593435 (95.2)	-0.671474 (97.1)
48	Cd	-0.455792 (95.8)	-0.614647 (93.5)	-0.704894 (96.4)
49	In	-0.437236 (95.4)	-0.621641 (99.3)	-0.689097 (98.4)
50	Sn	-0.449919 (95.8)	-0.638402 (99.4)	-0.708519 (98.7)

**Table 3** continued

Z	Atom	Correlation energy		
		DZP	TZP	QZP
51	Sb	-0.460252 (96.2)	-0.655272 (99.6)	-0.728039 (98.8)
52	Te	-0.479291 (96.3)	-0.691820 (99.4)	-0.770178 (98.8)
53	I	-0.497085 (96.4)	-0.722824 (99.4)	-0.807495 (98.8)
54	Xe	-0.513447 (96.5)	-0.747089 (99.3)	-0.839486 (98.8)

Values in parenthesis are reproduction percentage. The symbol Z denotes atomic number

similar in both sets, but the numbers of primitive GTFs are remarkably different. For the fourth period atoms, for example, the number of primitive GTFs of cc-CVTZ is four times larger than that of the Sapporo sets. Thus, the present sets are definitely compact compared with cc-type basis sets.

### 3 Atomic tests

The correlation energies calculated using the present non-relativistic basis sets are compiled in Table 3. For comparison, we show the reproduction percentage of correlation energies relative to those by accurate NOs of the standard size. For the *p*-block atoms, the correlation energies in the table are the sum of the core correlation energies among *n's*, *n'p*, and *n'd* electrons and the valence correlation energies among (*n' + 1*)*s* and (*n' + 1*)*p* electrons obtained by separate CI calculations. The present basis sets reproduce more than 90% of the correlation energies among the *n's*, *n'p*, and (*n' + 1*)*s* electrons for *s*-block atoms, *n's*, *n'p*, *n'd*, (*n' + 1*)*s*, and (*n' + 1*)*p* electrons for *p*-block atoms, and *n's*, *n'p*, *n'd*, and (*n' + 1*) electrons for *d*-block atoms, with minor exceptions for the DZP and TZP sets of early transition atoms Sc–V and Y–Nb.

### 4 Molecular applications

To test the quality of Sapporo-(DK)-*n*ZP basis sets, we have carried out self-consistent field (SCF) and CCSD(T) calculations to obtain spectroscopic constants of 12 hydrides of the *s*- and *d*-block atoms (LiH, BeH, NaH, MgH, KH, CaH, RbH, SrH, CrH, CuH, MoH, and AgH) and 12 diatomic molecules of the *p*-block atoms (BF, CO, N<sub>2</sub>, AlCl, SiS, P<sub>2</sub>, GaBr, GeSe, As<sub>2</sub>, InI, SnTe, and Sb<sub>2</sub>). The non-relativistic Sapporo-*n*ZP were used for H and Li–Cl, while the relativistic Sapporo-DK-*n*ZP sets for Ca–I. In all calculations, the relativistic effects were considered through the third-order DK approximation. All molecular calculations were performed using the Molcas6.4, Molcas7.4 [34], and Molpro2010 [35] program systems.

## 4.1 Diatomic hydrides

The resultant spectroscopic constants of the 12 hydrides are shown in Table 4 along with available experimental data. For *s*-block hydrides, the deviation from the experimental value decreases monotonically as the quality of the basis set increases, and the calculated values reach satisfactory agreement with experimental values at the QZP set, where the maximum deviations are 0.01 Å in the bond length  $r_e$ , 35 cm<sup>-1</sup> in the vibrational frequency  $\omega_e$ , and 0.1 eV in the dissociation energy  $D_0$ . For *d*-block hydrides, on the other hand, a different tendency is observed. As the quality of the basis set increases, the deviation from the experiment decreases, but we do not reach satisfactory agreement even at the QZP set. For example, the deviations in CrH are 0.02 Å in  $r_e$ , 96 cm<sup>-1</sup> in  $\omega_e$ , and 0.17 eV in  $D_0$ . This is not due to the basis set quality, but the limitation of CCSD(T) method based on a single reference theory. We carried out the complete active space SCF approach with additional

dynamic correlation effects using multiconfigurational second-order perturbation theory (CASPT2) calculations with QZP set for CrH, where 3*d*, 4*s*, and 4*p* of Cr and 1*s* of H are included in a complete active space and the 3*s* and 3*p* electrons of Cr are correlated. The new results are 1.652 Å, 1678 cm<sup>-1</sup>, and 2.190 eV for  $r_e$ ,  $\omega_e$ , and  $D_0$ , respectively. The deviations of  $r_e$  and  $\omega_e$  from the experiment reduce to 0.003 Å and 22 cm<sup>-1</sup>. Thus, the present basis sets have sufficient quality to obtain reliable spectroscopic constants.

In Table 5, we compare the timing data of CASPT2 calculations on CrH using Sapporo-DK-QZP and cc-pwCVQZ-DK [22] basis sets. Both sets gave almost the same values in  $r_e$ ,  $\omega_e$ , and  $D_0$ , but we found considerable difference in timings, especially at generation of integral and SCF steps. Large cost in integral generation using cc-pwCVQZ-DK is due to the long expansion of occupied orbitals as shown in Table 2. The cc-pwCVQZ-DK basis set has six times larger primitive GTFs and need five times longer CPU times in integral generation.

**Table 4** Spectroscopic constants of diatomic hydrides by CCSD(T)

Basis	$r_e/\text{Å}$	$\omega_e/\text{cm}^{-1}$	$D_0/\text{eV}$	$r_e/\text{Å}$	$\omega_e/\text{cm}^{-1}$	$D_0/\text{eV}$
	LiH			BeH		
DZP	1.623	1339	2.012	1.365	1991	1.550
TZP	1.599	1383	2.292	1.356	2013	1.862
QZP	1.590	1413	2.344	1.340	2075	2.033
Exptl. <sup>1</sup>	1.5957	1405.65	2.42871	1.3426	2060.78	2.03
	NaH			MgH		
DZP	1.892	1171	1.572	1.735	1461	0.874
TZP	1.889	1189	1.828	1.728	1543	1.210
QZP	1.887	1187	1.859	1.727	1536	1.252
Exptl. <sup>1</sup>	1.8874	1172.2	1.88	1.7297	1495.20	1.34
	KH			CaH		
DZP	2.285	938	1.348	2.080	1202	1.066
TZP	2.253	983	1.674	2.028	1282	1.542
QZP	2.249	980	1.724	2.016	1281	1.613
Exptl. <sup>1</sup>	2.242	983.6	1.8	2.0025	1298.34	1.70
	RbH			SrH		
DZP	2.420	831	1.583	2.237	1094	1.339
TZP	2.377	926	1.616	2.160	1181	1.544
QZP	2.369	932	1.674	2.153	1181	1.621
Exptl. <sup>1</sup>	2.367	936.9		2.1456	1206.2	1.66
	CrH			CuH		
DZP	1.636	1771	2.036	1.449	1961	2.529
TZP	1.633	1759	2.121	1.448	1998	2.712
QZP	1.632	1752	2.102	1.450	1990	2.726
Exptl.	1.655 <sup>1</sup>	1656.0 <sup>2</sup>	1.93 <sup>3</sup>	1.46263 <sup>1</sup>	1941.26 <sup>1</sup>	2.7 <sup>1</sup>
	MoH			AgH		
DZP	1.724	1788	2.171	1.613	1782	2.140
TZP	1.716	1781	2.327	1.609	1786	2.281
QZP	1.715	1770	2.327	1.612	1783	2.293

<sup>1</sup> See Ref. [36]

<sup>2</sup> See Ref. [37]

<sup>3</sup> See Ref. [38]

**Table 5** Timing data in seconds for CASPT2 calculations of CrH by Molcas7.4 on 3.33 GHz Intel Core 2 Duo E8600 CPU

Basis set	Sapporo-DK-QZP [9s7p5d3f2g1h]	cc-pwCVQZ-DK [10s9p7d4f3g2h]
Integral	7.5	38.9
SCF	3.7	14.9
CASSCF	34.0	111.8
CASPT2	28.4	43.0

#### 4.2 Diatomic molecules of *p*-block atoms

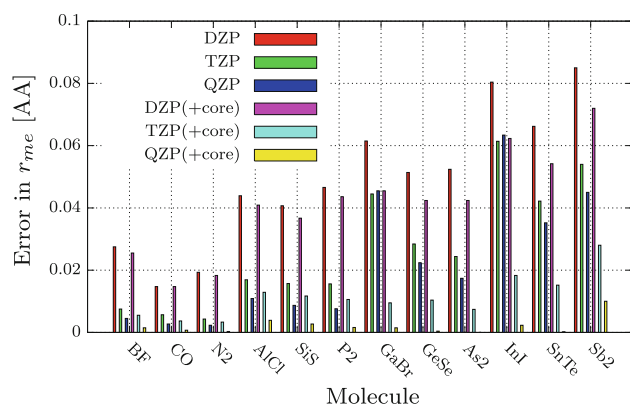
In Table 6, we show the calculated spectroscopic constants of the 12 diatomic molecules of the *p*-block atoms and compare them with available experimental data. In Fig. 2, the deviations from experimental values are exemplified for  $r_0$ . In order to show the effects of correlation among core electrons, we plot the results with

and without the core correlation effects, as denoted by  $nZP(+core)$  and  $nZP$  in the legend. For all molecules, smooth convergences on spectroscopic constants are found as the quality of the basis set increases when the core correlation is considered. At the QZP set, we reach a reasonable agreement with the experiment, where the deviations are smaller than 0.01 Å in  $r_e$ , 10 cm<sup>-1</sup> in  $\omega_e$ , and 0.2 eV in  $D_0$ . The effects of core correlations are remarkable in  $r_e$  of diatomics of the fourth and fifth period atoms. In GeSe, for example, the deviation of calculated  $r_e$  without the core correlation is 0.02 Å, while the inclusion of the core correlation reduces it to 0.0004 Å. We found analogous effects of core correlation for other molecules. For  $\omega_e$ , on the other hand, the inclusion of the core correlation increases the deviations from the experiment for some molecules, such as GaBr and Sb<sub>2</sub>, but the deviations at the QZP set are still reasonably small.

**Table 6** Spectroscopic constants of *p*-block diatomic molecules by CCSD(T)

Basis	$r_e/\text{Å}$	$\omega_e/\text{cm}^{-1}$	$D_0/\text{eV}$	$r_e/\text{Å}$	$\omega_e/\text{cm}^{-1}$	$D_0/\text{eV}$
	BF			CO		
DZP	1.288	1343	7.223	1.143	2142	10.377
TZP	1.268	1405	7.681	1.132	2157	10.860
QZP	1.264	1404	7.775	1.129	2170	11.023
Exptl. <sup>1</sup>	1.2625	1402.1	7.8	1.128323	2169.8358	11.09
	N <sub>2</sub>			AlCl		
DZP	1.116	2323	8.590	2.171	470	4.747
TZP	1.101	2355	9.361	2.143	481	5.060
QZP	1.098	2365	9.584	2.134	482	5.212
Exptl. <sup>1</sup>	1.09768	2358.57	9.759	2.130113	481.30	5.12
	SiS			P <sub>2</sub>		
DZP	1.966	726	5.545	1.937	749	3.871
TZP	1.941	743	6.011	1.904	777	4.587
QZP	1.932	752	6.234	1.895	784	4.858
Exptl. <sup>1</sup>	1.929321	749.64	6.4	1.8934	780.77	5.033
	GaBr			GeSe		
DZP	2.398	261	4.008	2.177	387	4.566
TZP	2.362	268	4.248	2.145	402	4.990
QZP	2.354	268	4.369	2.135	406	5.156
Exptl.	2.35248 <sup>1</sup>	263.0 <sup>1</sup>	4.4 <sup>2</sup>	2.134629 <sup>1</sup>	408.7 <sup>1</sup>	5.0 <sup>2</sup>
	As <sub>2</sub>			InI		
DZP	2.145	409	2.983	2.816	171	3.379
TZP	2.110	431	3.601	2.772	176	3.566
QZP	2.099	437	3.832	2.756	178	3.697
Exptl.	2.1026 <sup>1</sup>	429.55 <sup>1</sup>	3.96 <sup>1</sup>	2.75365 <sup>1</sup>	177.1 <sup>1</sup>	3.6 <sup>2</sup>
	SnTe			Sb <sub>2</sub>		
DZP	2.577	247	3.613	2.548	257	2.172
TZP	2.538	258	3.951	2.504	274	2.690
QZP	2.523	262	4.134	2.486	279	2.925
Exptl.	2.522814 <sup>1</sup>	259.5 <sup>1</sup>	3.93 <sup>2</sup>	2.476 <sup>3</sup>	269.9 <sup>1</sup>	3.09 <sup>1</sup>

<sup>1</sup> See Ref. [36]<sup>2</sup> The experimental values of Ref. [36] were corrected for the *J*-averaged dissociation limits<sup>3</sup> See Ref. [39]



**Fig. 2** Error in calculated  $r_c$  in Å

In this paper, we do not examine the basis set superposition error (BSSE), because we already reported that the effects of BSSE on the spectroscopic constants are quite small for the (DK)-TK/NOSec-(C)V- $n$ ZP sets in the previous papers [11, 19].

## 5 Summary

We developed all electron non-relativistic and relativistic segmented basis sets for the atoms H through Xe. To describe the core correlation, we constructed correlating CGTF sets for the second to fifth period  $p$ -block and the first and second  $d$ -block atoms. These correlating CGTF sets are optimized to represent accurate atomic NOs generated from large-scale CI calculations under the condition that the minimal-type HF sets (non-relativistic and relativistic TK HF CGTF sets) and the valence-correlating sets (NOSec-CV- $n$ ZP sets) exist. For  $s$ -block atoms, the (DK)-TK/NOSec-CV- $n$ ZP sets were re-optimized to remove analogous primitives in CGTFs. The resultant basis sets are named Sapporo-(DK)- $n$ ZP ( $n = D, T, Q$ ) and give more than 90 % of the core and valence correlation energies produced by accurate NOs of the standard size.

Test calculations for 12 hydrides of  $s$ - and  $d$ -block atoms and 12 diatomic molecules of the  $p$ -block atoms were performed at the CCSD(T) level of theory. For all molecules, smooth convergences of calculated spectroscopic constants to experimental values are found as the quality of the basis set increases when the core correlation is considered. At the QZP set, we reach a reasonable agreement with the experiment, for example, in the case of the  $p$ -block diatomic molecules, the deviations are smaller than 0.01 Å in  $r_c$ , 10  $\text{cm}^{-1}$  in  $\omega_e$ , and 0.2 eV in  $D_0$ .

The present sets are available at the web site <http://setani.sci.hokudai.ac.jp/sapporo/>, where basis sets are provided in appropriately formatted forms for popular electronic program packages such as Gaussian, Gamess,

Molpro, Molcas, Turbomole, Dirac, Nwchem, and Alchemy2. The Sapporo-(DK)- $n$ ZP sets have been also implemented in the Gamess program package and can be used by a simple key word in input data.

**Acknowledgments** This work was supported in part by a Grant-in-Aid for Scientific Research from the Ministry of Education, Culture, Sports, Science, and Technology of Japan.

## References

- Noro T, Sekiya M, Koga T (1997) *Theor Chem Acc* 98:25
- Sekiya M, Noro T, Koga T, Matsuyama H (1998) *J Mol Struct (Theochem)* 451:51
- Noro T, Sekiya M, Koga T, Matsuyama H (2000) *Theor Chem Acc* 104:146
- Sekiya M, Noro T, Osanai Y, Koga T (2001) *Theor Chem Acc* 106:297
- Osanai Y, Sekiya M, Noro T, Koga T (2003) *Mol Phys* 101:65
- Noro T, Sekiya M, Koga T (2003) *Theor Chem Acc* 109:85
- Noro T, Sekiya M, Osanai Y, Miyoshi E, Koga T (2003) *J Chem Phys* 119:5142
- Osanai Y, Noro T, Miyoshi E, Sekiya M, Koga T (2004) *J Chem Phys* 120:6408
- Sekiya M, Noro T, Miyoshi E, Osanai Y, Koga T (2006) *J Comput Chem* 27:463
- Noro T, Sekiya M, Osanai Y, Koga T, Matsuyama H (2007) *J Comput Chem* 28:2511
- Noro T, Sekiya M, Koga T (2008) *Theor Chem Acc* 121:289
- Sekiya M, Noro T, Koga T, Saito SL (2010) *J Comput Chem* 31:497
- Hess BA (1986) *Phys Rev A* 33:3742
- Nakajima T, Hirao K (2000) *J Chem Phys* 113:7786
- Tatewaki H, Koga T (1996) *J Chem Phys* 104:8493
- Tatewaki H, Koga T, Takashima H (1997) *Theor Chem Acc* 96:243
- Koga T, Tatewaki H, Matsuyama H, Satoh Y (1999) *Theor Chem Acc* 102:105
- Koga T, Yamamoto S, Shimazaki T, Tatewaki H (2002) *Theor Chem Acc* 108:41
- Noro T, Sekiya M, Koga T, Saito SL (2009) *Chem Phys Lett* 481:229
- Woon DE, Dunning TH Jr (1995) *J Chem Phys* 103:4572
- Peterson KA, Dunning TH Jr (2002) *J Chem Phys* 117:10548
- Balabanov NB, Peterson KA (2005) *J Chem Phys* 123:064107
- Peterson KA, Figgen D, Dolg M, Stoll H (2007) *J Chem Phys* 126:124101
- Iron MA, Oren M, Martin JML (2003) *Mol Phys* 101:1345
- Roos BO (2003) *Mol Phys* 101:87
- Dyall KG (1998) *Theor Chem Acc* 99:366
- Dyall KG (2002) *Theor Chem Acc* 108:335
- Dyall KG (2006) *Theor Chem Acc* 115:441
- Huzinaga S, Miguel B (1990) *Chem Phys Lett* 175:289
- Huzinaga S, Klobukowski M (1993) *Chem Phys Lett* 212:260
- Visscher L, Dyall KG (1997) *At Data Nucl Data Tables* 67:207
- Sasaki F, Sekiya M, Noro T, Ohtsuki K, Osanai Y (1993) *ME-TECC-94. STEF, Cagliari*
- Powell MJD (1964) *Comput J* 7:155
- Karlstrom G, Lindh R, Malmqvist PA, Roos BO, Ryde U, Varyazov V, Widmark PO, Cossi M, Schimmelpfennig B, Neogrady P, Seijo L (2003) *Comp Mat Sci* 28:222
- Werner HJ, Knowles PJ, Knizia G, Manby FR, Schütz M et al (2010) *MOLPRO*, version 2010.1, a package of *ab initio* program

36. Huber KP, Herzberg G (1979) Molecular spectra and molecular structure IV. Constants of diatomic molecules. Van Nostrand Reinhold, New York
37. Bauschlicher CW Jr, Ram RS, Bernath PF, Parsons CG, Galehouse D (2001) J Chem Phys 115:1312
38. Chen YM, Clemmer DE, Armentrout PB (1993) J Chem Phys 98:4929
39. Sontag H, Weber R (1982) J Molec Spectrosc 91:72