
6-31G* Basis Set for Third-Row Atoms

VITALY A. RASSOLOV,¹ MARK A. RATNER,¹ JOHN A. POPLE,¹
PAUL C. REDFERN,² LARRY A. CURTISS²

¹Department of Chemistry, Northwestern University, Evanston, Illinois 60208

²Chemistry and Materials Science Division, Argonne National Laboratory, Argonne, Illinois 60439

Received 10 August 2000; accepted 2 October 2000

ABSTRACT: Medium basis sets based upon contractions of Gaussian primitives are developed for the third-row elements Ga through Kr. The basis functions generalize the 6-31G and 6-31G* sets commonly used for atoms up to Ar. A reexamination of the 6-31G* basis set for K and Ca developed earlier leads to the inclusion of 3d orbitals into the valence space for these atoms. Now the 6-31G basis for the whole third-row K through Kr has six primitive Gaussians for 1s, 2s, 2p, 3s, and 3p orbitals, and a split-valence pair of three and one primitives for valence orbitals, which are 4s, 4p, and 3d. The nature of the polarization functions for third-row atoms is reexamined as well. The polarization functions for K, Ca, and Ga through Kr are a single set of Cartesian d-type primitives. The polarization functions for transition metals are defined to be a single 7f set of uncontracted primitives. Comparison with experimental data shows good agreement with bond lengths and angles for representative vapor-phase metal complexes. © 2001 John Wiley & Sons, Inc. J Comput Chem 22: 976–984, 2001

Keywords: basis sets; *ab initio*; third-row elements; Hartree–Fock

Introduction

The improvements in computational power and efficiency of computational techniques allow *ab initio* computation of properties of compounds containing heavier elements. Such computations require all electron basis sets of heavier elements of the same quality as those available for the first and second rows of the periodic table. Recently,

we extended a popular 6-31G basis set to third row atoms potassium through zinc.¹ The valence space was chosen as 4s, 4p shells for K and Ca, and 4s, 4p, and 3d for the transition metals. The polarization functions were chosen to be a single uncontracted Gaussian shell of the angular momentum higher than that of the valence orbitals.

In the present work we reexamine the valence space selection for K and Ca, redefine the polarization functions for the whole row, and extend the 6-31G basis set to the rest of the third row elements. We compare the newly developed 6-31G set with those of Blaudeau, McGrath, Curtiss, and Radom for K and Ca,² and Binning and Curtiss for Ga

Correspondence to: V. A. Rassolov; e-mail: rassolov@chem.nwu.edu

Contract/grant sponsor: Chemistry Division of the NSF

Contract/grant sponsor: U.S. Department of Energy, Division of Materials Science; contract/grant number: W-31-109-Eng.-38

through Kr.³ These basis sets are, unfortunately, also labeled as "6-31G" in the literature, despite having a different structure than 6-31G sets for the first two atomic rows of the periodic table. We propose to label these basis sets "6-31G*(C)" and "6-31G(C)," for the sets with and without polarization functions, respectively.

The Choice of Valence and Polarization Spaces

The 6-31G notation implies that each of the inner atomic shells is represented by a single six-Gaussian contraction. For historical reasons, p and s shells have identical exponents of the primitive functions, with different contraction coefficients. Each valence shell is represented by two functions, also with identical exponents in s and p spaces. One of the functions is a three-Gaussian contraction, and the other one is a single uncontracted Gaussian.

The p orbitals play a major role in chemistry of alkali and alkali earth elements. Therefore, the 4p shell is included among the valence shells for K and Ca. The d functions, however, also play a major role in the chemistry of calcium. One of the important examples of this is the structure of CaF₂

TABLE I.
The UHF Energies of Atoms with 6-31G and 6-31G* Basis Sets.

Atom	6-31G	6-31G*
K(² S)	-599.119954	-599.119959
Ca(¹ S)	-676.708958	-676.708965
Sc(² D)	-759.674203	-759.676936
Ti(³ F)	-848.327855	-848.333065
V(⁴ F)	-942.787478	-942.792254
Cr(⁵ D)	-1043.191883	-1043.194127
Mn(⁶ S)	-1149.722055	-1149.722055
Fe(⁵ D)	-1262.266962	-1262.268758
Co(⁴ F)	-1381.197761	-1381.200897
Ni(³ F)	-1506.609605	-1506.612283
Cu(² D)	-1638.639638	-1638.640759
Zn(¹ S)	-1777.482753	-1777.482753
Ga(² P)	-1922.895670	-1922.945263
Ge(³ D)	-2074.989222	-2075.037823
As(⁴ S)	-2233.859508	-2233.905143
Se(³ D)	-2399.478837	-2399.526779
Br(² P)	-2572.039558	-2572.087679
Kr(¹ S)	-2751.638332	-2751.683898

molecule. Its ground state configuration is bent (for review and discussion see ref. 4). We have found that to describe the bending on the SCF level, it is necessary to include more than one d-function into the basis set. In our earlier study, we also found that more than one d-function is required to describe the equilibrium geometry of the CaO molecule.¹ Therefore, the 3d shell must be included into the valence space of calcium. This is contrary to our earlier version of an 6-31G basis set, where d-functions were included only for transition metals.

The d-functions are less important for potassium chemistry. However, for consistency we treat all of the third row in the uniform way. The valence

TABLE II.
6-31G Basis Set for Potassium.

α	d_s	d_p
3.159442(4)	1.828010(-3)	
4.744330(3)	1.399403(-2)	
1.080419(3)	6.887129(-2)	
3.042338(2)	2.369760(-1)	
9.724586(1)	4.829040(-1)	
3.302495(1)	3.404795(-1)	
6.227625(2)	-2.502976(-3)	4.094637(-3)
1.478839(2)	-3.315550(-2)	3.145199(-2)
4.732735(1)	-1.226387(-1)	1.351558(-1)
1.751495(1)	5.353643(-2)	3.390500(-1)
6.922722	6.193860(-1)	4.629455(-1)
2.768277	4.345878(-1)	2.242638(-1)
1.184802(1)	1.277689(-2)	-1.221377(-2)
4.079211	2.098767(-1)	-6.900537(-3)
1.763481	-3.095274(-3)	2.007466(-1)
7.889270(-1)	-5.593884(-1)	4.281332(-1)
3.503870(-1)	-5.134760(-1)	3.970156(-1)
1.463440(-1)	-6.598035(-2)	1.104718(-1)
7.168010(-1)	-5.237772(-2)	3.164300(-2)
2.337410(-1)	-2.798503(-1)	-4.046160(-2)
3.867500(-2)	1.141547	1.012029
1.652100(-2)	1.000000	1.000000
α_d	d_d	
9.029078	8.747672(-2)	
2.285045	3.795635(-1)	
0.663892	7.180393(-1)	
1.79389(-1)	1.000000	

Numbers a(N) in parenthesis here and in the following tables represent $a \times 10^N$.

TABLE III.
6-31G Basis Set for Calcium.

α	d_s	d_p
3.526486(4)	1.813501(−3)	
5.295503(3)	1.388493(−2)	
1.206020(3)	6.836162(−2)	
3.396839(2)	2.356188(−1)	
1.086264(2)	4.820639(−1)	
3.692103(1)	3.429819(−1)	
7.063096(2)	2.448225(−3)	4.020371(−3)
1.678187(2)	3.241504(−2)	3.100601(−2)
5.382558(1)	1.226219(−1)	1.337279(−1)
2.001638(1)	−4.316965(−2)	3.367983(−1)
7.970279	−6.126995(−1)	4.631281(−1)
3.212059	−4.487540(−1)	2.257532(−1)
1.419518(1)	1.084500(−2)	−1.289621(−2)
4.880828	2.088333(−1)	−1.025198(−2)
2.160390	3.150338(−2)	1.959781(−1)
9.878990(−1)	−5.526518(−1)	4.357933(−1)
4.495170(−1)	−5.437997(−1)	3.996452(−1)
1.873870(−1)	−6.669342(−2)	9.713636(−2)
1.032271	−4.439720(−2)	−4.298621(−1)
3.811710(−1)	−3.284563(−1)	6.935829(−3)
6.513100(−2)	1.163010	9.705933(−1)
2.601000(−2)	1.000000	1.000000
α_d	d_d	
10.11067	8.747672(−2)	
2.558769	3.795635(−1)	
0.743420	7.180393(−1)	
2.00878(−1)	1.000000	

TABLE IV.
6-31G Basis Set for Gallium.

α	d_s	d_p
8.828461(4)	1.736921(−3)	
1.325606(4)	1.331136(−2)	
3.019649(3)	6.571709(−2)	
8.514222(2)	2.287932(−1)	
2.729997(2)	4.781507(−1)	
9.342593(1)	3.549154(−1)	
1.877680(3)	2.316733(−3)	3.896102(−3)
4.474374(2)	3.090570(−2)	3.066136(−2)
1.451401(2)	1.264173(−1)	1.344509(−1)
5.484977(1)	−1.429714(−2)	3.470761(−1)
2.244351(1)	−6.132855(−1)	4.635435(−1)
9.286622	−4.703598(−1)	2.039435(−1)
8.005681(1)	−5.056378(−3)	−6.947816(−3)
2.757856(1)	6.117037(−2)	−2.938902(−2)
1.171717(1)	2.575692(−1)	5.377307(−2)
5.054113	−2.150754(−1)	3.764511(−1)
2.172525	−7.213703(−1)	4.923913(−1)
9.041840(−1)	−2.785244(−1)	2.073613(−1)
1.112438	1.970334(−1)	−9.151867(−3)
3.287220(−1)	−2.497645(−1)	3.111786(−1)
1.305520(−1)	−8.749447(−1)	7.436549(−1)
4.758900(−2)	1.000000	1.000000
α_d	d_d	
3.911406(1)	8.790043(−2)	
1.061218(1)	3.915600(−1)	
3.273033	6.956990(−1)	
9.156600(−1)	1.000000	

space includes 4s, 4p, and 3d shells. As in earlier studies, we use a set of six Cartesian components {xx,yy,zz,xy,yz,zx} for d-functions. This brings the total number of 6-31G functions for every third row element to 29.

The polarization functions increase the angular flexibility of the basis set in the region of chemical bond. We have used a Cartesian representation of a single f primitive to describe the polarization function for transition metals in our earlier study.¹ Compared with the spherical harmonics representation, this increases the size of the basis set by three basis functions. Such an increase is particularly damaging to the speed and storage requirements of the correlated methods. Additionally, three extra p-type functions can adversely affect the conver-

gence of the calculations. We, therefore, recommend a spherical harmonics representation of polarization f-functions for transition metals Sc through Zn. The total number of 6-31G* functions for these elements is 36.

The bonding of nontransition third row elements involves mainly s and p type functions, despite the importance of d functions in geometry optimization.⁵ Therefore, the natural choice of the polarization functions for these elements is d type. Alternatively, one can define the polarization function to be of higher orbital angular momentum than any valence function. This choice leads to polarization f functions for K, Ca, and Ga through Kr. In this work we investigate both types of polarization functions.

TABLE V.
6-31G Basis Set for Germanium.

α	d_s	d_p
9.428132(4)	1.732993(-3)	
1.415642(4)	1.328181(-2)	
3.224935(3)	6.557319(-2)	
9.094821(2)	2.283712(-1)	
2.917149(2)	4.778104(-1)	
9.989074(1)	3.557135(-1)	
<hr/>		
2.016629(3)	2.299186(-3)	3.872605(-3)
4.806599(2)	3.068823(-2)	3.051218(-2)
1.560616(2)	1.262906(-1)	1.338971(-1)
5.907914(1)	-1.105405(-2)	3.462496(-1)
2.422346(1)	-6.103659(-1)	4.635741(-1)
1.004418(1)	-4.755387(-1)	2.047879(-1)
<hr/>		
8.728112(1)	-5.330845(-3)	-6.893957(-3)
3.028230(1)	5.874495(-2)	-2.954252(-2)
1.285367(1)	2.598349(-1)	5.042291(-2)
5.587437	-1.926917(-1)	3.699366(-1)
2.438461	-7.190570(-1)	4.933147(-1)
1.040324	-2.995181(-1)	2.116445(-1)
<hr/>		
1.344960	2.338815(-1)	-1.976804(-2)
4.436620(-1)	-2.189617(-1)	3.028906(-1)
1.760820(-1)	-9.242006(-1)	7.562828(-1)
6.466500(-2)	1.000000	1.000000
<hr/>		
α_d	d_d	
4.463105(1)	8.431036(-2)	
1.220184(1)	3.847726(-1)	
3.823423	7.003323(-1)	
1.108831	1.000000	

TABLE VI.
6-31G Basis Set for Arsenic.

α	d_s	d_p
1.005955(5)	1.726750(-3)	
1.510482(4)	1.323462(-2)	
3.440884(3)	6.535848(-2)	
9.703961(2)	2.278042(-1)	
3.112852(2)	4.774525(-1)	
1.066284(2)	3.567619(-1)	
<hr/>		
2.166679(3)	2.271761(-3)	3.832156(-3)
5.165414(2)	3.033475(-2)	3.023558(-2)
1.678674(2)	1.259057(-1)	1.328632(-1)
6.364638(1)	-6.687172(-3)	3.447648(-1)
2.613673(1)	-6.065306(-1)	4.640368(-1)
1.085439(1)	-4.823144(-1)	2.064824(-1)
<hr/>		
9.506989(1)	-5.587423(-3)	-6.816583(-3)
3.318087(1)	5.632506(-2)	-2.970303(-2)
1.406773(1)	2.625835(-1)	4.704335(-2)
6.153288	-1.718349(-1)	3.645042(-1)
2.721712	-7.175645(-1)	4.945157(-1)
1.185334	-3.184598(-1)	2.149830(-1)
<hr/>		
1.615315	2.645372(-1)	-2.574061(-2)
5.513300(-1)	-1.952737(-1)	3.072764(-1)
2.227620(-1)	-9.595400(-1)	7.537368(-1)
8.292300(-2)	1.000000	1.000000
<hr/>		
α_d	d_d	
5.030227(1)	8.144711(-2)	
1.384166(1)	3.792908(-1)	
4.393458	7.040401(-1)	
1.310755	1.000000	

Computational Methods

The values of all primitive exponents and contraction coefficients are obtained by minimization of total SCF energy. We used the ATOM-SCF program,⁶ modified by us to allow variation of exponents and contraction coefficients at the same time, and to use the modified Powell method of conjugate gradients^{7,8} to carry out the minimization of parameters. This program uses spin-restricted open shell Hartree-Fock theory (ROHF), and pure d-functions (based on spherical harmonics). Thus, our parameters are optimized for pure functions, even though the final basis is specified in Cartesian representation. This procedure is identical to the one used to obtain 6-31G set for the elements K through Zn.¹

Therefore, the s and p functions of K and Ca are identical to those published earlier. The optimization of 6-31G set for the elements Ga through Kr was performed on high spin, high angular momentum atomic states, namely Ga(²P), Ge(³D), As(⁴S), Se(³D), Br(²P), and Kr(¹S).

The optimization of d-functions for potassium and calcium was first performed on the excited states K(²D) and Ca(³D). The procedure yielded very diffuse d-functions. To get a more realistic valence-type d function, the d-functions of scandium were scaled with the square of the ratio of nuclear charges. The latter approach yields basis sets with optimized molecular geometries closer to experimental results and lower molecular energies. Therefore, the final 6-31G basis contains the scandium-like d functions for calcium and

TABLE VII.
6-31G Basis Set for Selenium.

α	d_s	d_p
1.070273(5)	1.722646(−3)	
1.607076(4)	1.320324(−2)	
3.661226(3)	6.520494(−2)	
1.032673(3)	2.273787(−1)	
3.313339(2)	4.771451(−1)	
1.135470(2)	3.575553(−1)	
2.313540(3)	2.261924(−3)	3.818409(−3)
5.516849(2)	3.019493(−2)	3.015145(−2)
1.794401(2)	1.258828(−1)	1.325614(−1)
6.813044(1)	−4.373809(−3)	3.443419(−1)
2.803062(1)	−6.043277(−1)	4.639237(−1)
1.166572(1)	−4.861200(−1)	2.068198(−1)
1.015754(2)	−5.752618(−3)	−6.942389(−3)
3.561545(1)	5.675608(−2)	−3.014441(−2)
1.513135(1)	2.651243(−1)	4.776411(−2)
6.646923	−1.670582(−1)	3.663827(−1)
2.972805	−7.188737(−1)	4.940086(−1)
1.316707	−3.221907(−1)	2.100109(−1)
1.846991	2.823156(−1)	−2.653920(−2)
6.471590(−1)	−2.129616(−1)	3.357291(−1)
2.579870(−1)	−9.545384(−1)	7.301815(−1)
9.410700(−2)	1.000000	1.000000
α_d	d_d	
5.618544(1)	7.904963(−2)	
1.554808(1)	3.746449(−1)	
4.989394	7.071645(−1)	
1.523844	1.000000	

TABLE VIII.
6-31G Basis Set for Bromide.

α	d_s	d_p
1.137182(5)	1.717696(−3)	
1.707444(4)	1.316744(−2)	
3.889576(3)	6.504553(−2)	
1.097096(3)	2.269505(−1)	
3.520624(2)	4.768357(−1)	
1.207002(2)	3.583677(−1)	
2.471138(3)	2.243687(−3)	3.790182(−3)
5.893838(2)	2.994853(−2)	2.995979(−2)
1.918738(2)	1.256009(−1)	1.318228(−1)
7.295339(1)	−9.832786(−4)	3.432708(−1)
3.005839(1)	−6.013141(−1)	4.642345(−1)
1.252927(1)	−4.913983(−1)	2.079387(−1)
1.096411(2)	−5.975683(−3)	−6.907483(−3)
3.858948(1)	5.542122(−2)	−3.041432(−2)
1.637818(1)	2.681200(−1)	4.602725(−2)
7.221836	−1.543606(−1)	3.650689(−1)
3.263697	−7.206306(−1)	4.949232(−1)
1.465499	−3.316437(−1)	2.090394(−1)
2.103651	3.029029(−1)	−2.826714(−2)
7.547050(−1)	−2.152659(−1)	3.503065(−1)
3.005140(−1)	−9.633941(−1)	7.182446(−1)
1.090710(−1)	1.000000	1.000000
α_d	d_d	
6.225514(1)	7.704229(−2)	
1.731284(1)	3.707384(−1)	
5.607915	7.097628(−1)	
1.746486	1.000000	

potassium. The UHF energies obtained using the Q-Chem⁹ program with these basis sets are given in Table I, with basis sets themselves given in Tables II–IX. They can be also downloaded through the Northwestern University Chemistry Department FTP server.¹⁰

The value of the polarization function exponent was optimized by minimizing the energy of various molecules with geometries optimized at the Hartree–Fock level using the 6-31G basis set. The values of optimal f- and d-functions for different compounds are given in Table X, and the UHF energies are shown in Table I.

The optimized value of f polarization exponent for most compounds is around 0.5, with the clear exception of potassium compounds. The d-orbitals do not actively participate in forming potassium

bonds at the SCF level, and therefore, do not require f-functions for the polarization. The value of optimized f-exponent corresponds instead to some very diffuse Rydberg state, as in the case of CaF₂ molecule. We recommend using 0.5 value for all polarization exponents. It is interesting to note that this value is less than 0.8 recommended for lighter transition metals. This is the reflection of greater participation of d-functions in the bonding of transition metals.

The optimized polarization d-exponents are closely related to the values of outermost valence d-functions, with the ratio between them of about 4 (Table X). For some potassium and calcium compounds there was no optimal d-function more diffuse than the outermost valence d. For the final basis set we use polarization functions with the expo-

TABLE IX.
6-31G Basis Set for Krypton.

α	d_s	d_p
1.205524(5)	1.714050(−3)	
1.810225(4)	1.313805(−2)	
4.124126(3)	6.490006(−2)	
1.163472(3)	2.265185(−1)	
3.734612(2)	4.764961(−1)	
1.280897(2)	3.591952(−1)	
2.634681(3)	2.225111(−3)	3.761911(−3)
6.284533(2)	2.971122(−2)	2.977531(−2)
2.047081(2)	1.253926(−1)	1.311878(−1)
7.790827(1)	1.947058(−3)	3.425019(−1)
3.213816(1)	−5.987388(−1)	4.644938(−1)
1.341845(1)	−4.958972(−1)	2.087284(−1)
1.175107(2)	−6.157662(−3)	−6.922855(−3)
4.152553(1)	5.464841(−2)	−3.069239(−2)
1.765290(1)	2.706994(−1)	4.480260(−2)
7.818313	−1.426136(−1)	3.636775(−1)
3.571775	−7.216781(−1)	4.952412(−1)
1.623750	−3.412008(−1)	2.086340(−1)
2.374560	3.251184(−1)	−3.009554(−2)
8.691930(−1)	−2.141533(−1)	3.598893(−1)
3.474730(−1)	−9.755083(−1)	7.103098(−1)
1.264790(−1)	1.000000	1.000000
α_d	d_d	
6.853888(1)	7.530705(−2)	
1.914333(1)	3.673551(−1)	
6.251213	7.120146(−1)	
1.979236	1.000000	

nents four times smaller than the outermost valence ones. The values are given in Table X.

Performance

We have performed a number of geometry optimization calculations with the 6-31G, 6-31G(f), and 6-31G(d) basis sets for the molecules for which there is experimental data on their geometric parameters in the gas phase. The results are summarized in Tables XI and XII. All single states are calculated at the spin-restricted Hartree–Fock (RHF) level. For K_2 , As_2 , BrO^- , and KrF_2 the spin-unrestricted (UHF) calculations yield lower energy. The geometry optimization at the UHF level with 6-31G(f) basis set gives the bond distance of 4.9478 Å for K_2 , and

TABLE X.
Dependence of the Optimized Polarization Exponent for Various Molecules.

Molecule	f Exp.	Final f	d Exp.	Final d
KH	0.1823			
KOH	0.0514			
KF	0.0673	0.5	0.04485	
KCl	0.0587			
Ca(CH ₃) ₂	0.5223			
CaO	0.4731	0.5	0.0502	
CaF ₂	0.09843			
GaCl	0.2935	0.5	0.2235 (4.09)	0.2289
GeH ₄	0.446		0.2778 (3.99)	
GeO	0.567	0.5	0.2738 (4.05)	0.2772
AsH ₃	0.5605		0.3270 (4.01)	
As ₂	0.2521	0.5	0.3150 (4.16)	0.3277
SeH ₂	0.5525	0.5	0.3762 (4.05)	0.3810
BrOH	0.6155		0.4365 (4.00)	
BrCH ₃	0.5853	0.5	0.4342 (4.02)	0.4366
KrF ₂	0.4920	0.5	0.4525 (4.37)	0.4948

All geometries are optimized at the 6-31G level. The number in parentheses shows the ratio of the diffuse exponent to the exponent of the outermost valence function.

2.2993 Å for As_2 . Both values are significantly larger than the corresponding RHF distances. Two other molecules KrF_2 and BrO^- are unstable with respect to their dissociation at the UHF level. This instability is similar to that of the F_2 . The BrO molecule has a metastable structure at the UHF level, and is also unstable with respect to dissociation. At the same time, Br_2 , $BrCl$, and BrF molecules are stable and bound at the RHF level of theory with 6-31G and 6-31G(d) sets, unlike the F_2 molecule. We have no simple explanation for such behavior of Hartree–Fock approximation of molecular binding. This subject requires further study.

To study the importance of d-functions in the geometry optimization, the K and Ca containing molecules are placed in a separate table. The relative root-mean-square deviations (RRMS) deviations are given at the bottom of the table. The values in parentheses show RRMS for bond lengths only. The BrO molecule is excluded from the calculation, as it would give infinite RRMS for 6-31G set.

First of all, the need to include d-functions in the valence space of K and Ca is obvious. The RRMS drops by almost a half when diffuse d-functions

TABLE XI.
Calculated and Experimental Equilibrium Hartree–Fock Geometries for Molecules with K and Ca atoms.

Molecule	Parameter	6-31G*(C)	6-31G*(old)	6-31G(f)	6-31G	6-31G(d)	STO-3G	pVDZ	Expt.
K ₂ ^a	r(KK)	4.1803	4.202	4.1968	4.2003	4.2017	3.4633	4.2131	3.9051
KOH ^b	r(KO)	2.1844	2.244	2.2637	2.1831	2.2107	2.1575	2.2716	2.196
	r(OH)	0.9441	0.944	0.9503	0.9502	0.9440	0.9885	0.9456	0.96
KCl ^a	r(KCl)	2.7314	2.777	2.8097	2.7522	2.7430	2.673	2.7662	2.6667
KF ^a	r(KF)	2.1423	2.1893	2.2153	2.1450	2.1468	2.2865	2.2383	2.1716
KH ^a	r(KH)	2.2420	2.3287	2.3747	2.2901	2.3121	2.0808	2.3786	2.242
CaH ^a	r(CaH)	1.9899	2.1294	1.9959	2.0489	2.0581	2.1794	2.3307	2.0025
CaS ^a	r(CaS)	2.3363	2.4882	2.3537	2.3735	2.3588	2.5573	2.5544	2.3178
CaF ^a	r(CaF)	1.9357	2.0355	1.9591	1.9755	1.9618	3.9855	2.0724	1.967
CaCl ^a	r(CaCl)	2.4879	2.5893	2.5032	2.5375	2.5218	2.5826	2.6036	2.4390
CaCl ₂ ^b	r(CaCl)	2.5006	2.5571	2.5055	2.5239	2.5153	2.5573	2.5787	2.483
CaO ^b	r(CaO)	1.7763	2.038	1.8135	1.8251	1.8224	2.7526	2.1832	1.8221
CaF ₂ ^c	r(CaF)	1.9868	2.0866	1.9997	2.0079	2.0012	2.209	2.0698	2.1
	∠(FCaF)	147.19	180	155.96	153.486	153.11	180	180	135
RRMS		3.6(2.8)%	10.4(5.5)%	5.4(3.6)%	4.7(3.1)%	4.7(3.1)%	32.4(32.4)%	12.3(8.7)%	

The "6-31G*(old)" column is calculated with 6-31G* basis set without d-functions in valence space, and d-function polarization, as reported in ref. 1. The 6-31G(f) sets have valence d functions optimized with the excited states. The 6-31G and 6-31G(d) bases have valence d functions rescaled from Sc. Bond lengths are given in angstroms, angles are in degrees. The last row gives relative root mean square deviations for all values, and only for bond lengths in parentheses.

^a Experimental geometry from ref. 15.

^b Experimental geometry from ref. 16.

^c Experimental geometry from JANAF tables (ref. 17).

are included in the basis, as seen by comparing 6-31G*(old) and 6-31G(f) cases. It is further reduced when the d-functions in the basis are obtained by rescaling those of scandium. The inclusion of polarization functions does not significantly affect the overall agreement with experimental geometries.

The calculations on the molecules containing Ga through Kr elements show the importance of d-type polarization functions. Addition of f-type functions to the 6-31G set improves the agreement with the experimental geometries for a chosen set of molecules by 30%, while the use of d-functions result in the 51% improvement. We have attempted to define the d-type polarization functions for transition metals as well. The calculation on a set of model compounds used in ref. 1 shows that the replacement of f-type polarization functions by the d-type significantly worsens the agreement with the experimental geometries. We, therefore, define the polarization functions differently for the transition metals and nontransition metal elements. The transition metal polarization functions polarize the bonds formed by d-type functions, and, therefore, must be of f-type. For Ga through Kr the valence d-functions are too tight to play a significant role in

polarizing the molecular bonds, and polarization d functions are needed.

The performance of the new 6-31G* basis set for gas phase geometry optimization is compared with a smaller STO-3G¹¹ and a comparable-size VDZ + polarization¹² sets for the molecules containing K and Ca in Table XI. Both of these sets do not contain valence d-type functions for K and Ca. This leads to poor optimized geometries, especially for calcium containing molecules. The 6-31G and 6-31G* basis sets performs slightly worse than the old basis set by Blaudeau et al.²

We compare 6-31G* set for the molecules containing Ga through Kr with the old 6-31G*,³ STO-3G,¹¹ pVDZ,¹² and cc-pVDZ¹³ basis sets for gas phase geometries optimizations in Table XII. With the exception of a small STO-3G set, all basis sets have similar performance in geometry optimization, as shown by overall RRMS. Much of the error comes from correlation effects. The MP2(full) geometry optimization with 6-31G* set calculated with Gaussian 98 program¹⁴ on the set of molecules given in Table XII gives overall RRMS of 1.05%, compared to 1.33% for the 6-31G*(C) set. This MP2 geometry optimization accuracy is comparable to the 6-31G* performance for second-row elements.

TABLE XII.
Calculated and Experimental Equilibrium Hartree-Fock Geometries for Molecules with Ga through Kr Elements.

Molecule	Point Group	Parameter	6-31G* (C)	6-31G	6-31G(f)	6-31G(d)	STO-3G	pVDZ	cc-pVDZ	Expt.
GaH ^a		r(GaH)	1.6743	1.6943	1.6893	1.6821	1.5028	1.6744	1.6633	1.663
GaO ^a		r(GaO)	1.8017	1.8087	1.8379	1.7915	1.6104	1.7906	1.8037	1.743
GaCl ^a		r(GaCl)	2.2403	2.3364	2.2583	2.2458	1.9302	2.2275	2.2486	2.202
GeH ₄ ^b	<i>T_d</i>	r(GeH)	1.5353	1.5353	1.5352	1.5369	1.4306	1.5376	1.5349	1.514
GeO ^a		r(GeO)	1.6041	1.6220	1.5851	1.5947	1.4542	1.5929	1.6023	1.625
As ₂ ^a		r(AsAs)	2.0489	2.1114	2.0892	2.0660	1.9785	2.0583	2.0606	2.103
AsH ^a		r(AsH)	1.5298	1.5351	1.5315	1.5235	1.464	1.5236	1.5201	1.535
AsH ₂ ^c	<i>C_{2v}</i>	r(AsH)	1.5264	1.5251	1.5249	1.5176	1.4601	1.5201	1.5176	1.518
		∠(HAsH)	92.2159	92.8186	93.192	92.426	92.113	92.671	92.791	90.7
AsH ₃ ^c	<i>C_{3v}</i>	r(AsH)	1.5134	1.5151	1.5151	1.5123	1.4567	1.5158	1.5132	1.511
		∠(HAsH)	93.167	95.040	95.520	93.917	93.987	94.111	94.325	92.1
SeH ^a		r(SeH)	1.4709	1.4735	1.4735	1.4628	1.4466	1.4664	1.4617	1.475
SeH ₂ ^c	<i>C_{2v}</i>	r(SeH)	1.4665	1.4652	1.4655	1.4579	1.4401	1.4629	1.4579	1.46
		∠(HSeH)	92.2947	93.635	94.128	92.680	92.491	92.889	93.063	90.6
Br ₂ ^a		r(BrBr)	2.2859	2.4062	2.3697	2.2825	2.2858	2.2986	2.297	2.281
BrO ^a		r(BrO)	1.7785	∞	1.7989	1.7783	1.8508	1.7713	1.7946	1.717
BrO ^{-d}		r(BrO)	1.8768	2.0405	1.9401	1.8894	2.0043	1.8735	1.9131	1.814
BrH ^{+a}		r(BrH)	1.4468	1.4471	1.4486	1.4366	1.6641	1.4483	1.441	1.448
BrH ^a		r(BrH)	1.4193	1.4229	1.4227	1.4129	1.4117	1.4196	1.4122	1.414
BrB ^a		r(BrB)	1.9047	1.9876	1.9086	1.9174	1.891	1.8939	1.9121	1.888
BrF ^a		r(BrF)	1.7374	1.8050	1.7378	1.7399	1.7704	1.7423	1.7568	1.759
BrNa ^a		r(BrNa)	2.5113	2.5967	2.5415	2.5454	2.3245	2.5345		2.502
BrCl ^a		r(BrCl)	2.1435	2.2728	2.1832	2.1384	2.1733	2.1610	2.1526	2.136
BrOH ^c	<i>C_s</i>	r(BrO)	1.8090	1.8697	1.8152	1.8084	1.8297	1.8021	1.8154	1.834
		r(OH)	0.9505	0.9571	0.9506	0.9507	0.9985	0.9524	0.9499	0.961
		∠(BrOH)	104.985	108.519	105.536	105.300	107.025	106.789	104.81	102.3
BrCH ₃	<i>C_{3v}</i>	r(BrC)	1.9476	1.9912	1.9610	1.9471	1.9058	1.9398	1.9466	1.934
		r(CH)	1.0764	1.0741	1.0764	1.0767	1.0868	1.0858	1.0845	1.082
		∠(HCH)	111.196	111.556	111.168	111.032	108.922	110.830	111.197	111.2
KrF ₂ ^c	<i>D_{∞h}</i>	r(KrF)	1.8342	1.9028	1.8233	1.8155	1.9214	1.8225	1.8329	1.875
RRMS			1.6(1.6)%	3.7(3.7)%	2.6(2.5)%	1.8(2.7)%	5.9(6.3)%	1.8(1.5)%	2.0(1.9)%	

Bond lengths are given in angstroms, angles are in degrees. The last row gives relative root-mean-square deviations for all values, and only for bond lengths in parentheses. Much of the error in the 6-31G(d) geometry optimization is due to correlation, as explained in the text.

^a Experimental geometry from ref. 15.

^b Experimental geometry from ref. 18.

^c Experimental geometry from ref. 16.

^d Experimental geometry from ref. 19.

^e Experimental geometry from ref. 20.

^f Experimental geometry from ref. 21.

The 6-31G basis is uniformly defined for the entire third row. It contains 29 basis functions. The 6-31G* basis for K, Ca, and Ga through Kr includes an additional shell of 6d functions, making it equivalent to 6-31G(d) set. The total number of 6-31G* basis functions for these elements is 35. The 6-31G* basis for Sc through Zn is actually 6-31G(f) set, as it uses 7f polarization functions, with the total number of 36 basis functions.

Conclusions

The 6-31G* polarization basis set has been extended to elements Ga through Kr. The analysis of the valence space of calcium leads to a modification of previously published 6-31G basis sets for K and Ca, and a modification of the nature of polarization functions for the elements K through Zn.

Now, the 6-31G basis for the entire third row has 4s, 4p, and 3d functions in the valence space. The different nature of chemical bonding for transition and nontransition elements does not permit the uniform definition of polarization functions without the loss of accuracy. The polarization functions for K, Ca, and Ga through Kr elements are a single set of Cartesian d-functions. The polarization functions for Sc through Zn are a single set of 7f spherical harmonic polarization functions. Thus, the polarization functions labeled by the "star" imply a single set of Gaussian functions with the orbital angular momentum higher than that of the functions forming molecular bonds at the Hartree-Fock level. The Hartree-Fock calculations on a sample set of molecules show that the agreement with experimental geometries is satisfactory, and the performance is comparable to the slightly larger set of Binning and Curtiss,³ also labeled "6-31G*." We propose to change the notation for that set to "6-31G*(C)," along with similar set for K and Ca.² The new basis set is consistent with 6-31G* set for first- and second-row elements.

References

1. Rassolov, V. A.; Pople, J. A.; Ratner, M. A.; Windus, T. L. *J Chem Phys* 1998, 109, 1223.
2. Blaudeau, J.-P.; McGrath, M. P.; Curtiss, L. A.; Radom, L. *J Chem Phys* 1997, 107, 5016.
3. Binning, R. C.; Curtiss, L. A. *J Comput Chem* 1990, 11, 1206.
4. DeKock, R. L.; Peterson, M. A.; Timmer, L. K.; Baerends, E. J.; Vernooijs, P. *Polyhedron* 1990, 9, 1919.
5. Coulson, C. A. *Isr J Chem* 1973, 11, 683.
6. Roothaan, C. C. J.; Bagus, P. S. *Atom-SCF*, 1963, Roos, B.; Salez, C.; Veillard, A.; Clementi, E. Eds.; IBM Research Report RJ518, 1968.
7. Brent, R. P. *Algorithms for Minimization without Derivatives*; Prentice-Hall: Englewood Cliffs, NJ, 1973, Chapt. 7.
8. Press, W. H.; Teukolsky, S. A.; Vetterling, W. T.; Flannery, B. P. *Numerical Recipes in FORTRAN*; Cambridge University Press: New York, 1994, 2nd ed.
9. White, C. A.; et al. *Q-Chem Inc.*, Version 1.2; 1998.
10. The 6-31G and 6-31G* basis sets are available for anonymous downloading at [ftp.chem.nwu.edu](ftp://ftp.chem.nwu.edu), or via world wide web at <ftp://ftp.chem.nwu.edu/pub/basis>.
11. Pietro, W. J.; Levy, B. A.; Hehre, W. J.; Stewart, R. F. *Inorg Chem* 1980, 19, 2225.
12. Schafer, A.; Horn, H.; Ahlrichs, R. *J Chem Phys* 1992, 97, 2571.
13. Wilson, A. K.; Woon, D. E.; Peterson, K. A.; Dunning, T. H., Jr. *J Chem Phys* 1999, 110, 7667.
14. Frisch, M. J.; et al. *Gaussian 98*; Gaussian, Inc.: Pittsburgh, PA, 1998.
15. Huber, K.; Herzberg, G. *Molecular Spectra and Molecular Structure 4; Constants of Diatomic Molecules*; Van Nostrand: Princeton, NJ, 1979.
16. Madelung, O.; Kuchitsu, K., Eds. *Landolt-Bornstein, New Series, Group II*; Springer: Berlin, 1992.
17. Chase, M. W., Jr.; Davies, C.; Downey, J. R., Jr.; Frurip, D. J.; McDonald, R. A.; Syverud, A. N. *J Phys Chem Ref Data* 1985, 14, Suppl. 1.
18. Ohno, K.; Matsumura, H.; Endo, Y.; Hirota, E. *J Mol Spectrosc* 1986, 118, 1.
19. Gilles, M. K.; Pollak, M. L.; Lineberger, W. C. *J Chem Phys* 1992, 96, 8012.
20. Koga, Y.; Takeo, H.; Kondo, S.; Sugie, M.; Matsumura, C.; McRae, G. A.; Cohen, E. A. *J Mol Spectrosc* 1989, 138, 467.
21. Graner, G. *J Mol Spectrosc* 1981, 90, 394.